solutions réseaux

Réseaux d'entreprise par la pratique

JEAN-LUC **MONTAGNIER**

JEAN-LUC **MONTAGNIER**

*Ingénieur consultant spécialisé dans les réseaux et les télécoms, **Jean-Luc Montagnier** assure des missions de conception d'architectures, d'audit et d'assistance à maîtrise d'ouvrage, aussi bien auprès de PME que de grands comptes. Jean-Luc Montagnier est l'auteur de* Pratique des réseaux d'entreprise, *l'un des ouvrages de référence en langue française sur les réseaux.*

Devenez expert en réseaux par l'exemple

Vous êtes chargé de mettre en place un réseau d'entreprise ou de faire évoluer un réseau existant. Que vous soyez architecte réseau, responsable informatique ou étudiant, cet ouvrage vous apporte une expertise complète, au travers d'exemples de configuration et d'études de cas, sur des réseaux allant d'une dizaine à plusieurs milliers de postes.

Depuis l'installation du système de câblage jusqu'à la sécurisation d'un réseau IP, en passant par les réseaux sans fil, les réseaux étendus, ou encore la qualité de service, l'ouvrage détaille pas à pas les protocoles réseau les plus répandus en entreprise.

Couvre les réseaux Wi-Fi.

Au sommaire : Architectures et Protocoles

Architectures LAN et WLAN : réseaux d'étage et fédérateurs, choix, positionnement et configuration des commutateurs et des bornes Wi-Fi • Ethernet 10/100/1000bT, spanning tree 802.1d, GARP, VLAN, GVRP • 802.11a/b/g, Bluetooth 802.15, LMDS, MMDS, WEP, TKIP, WPA, 802.11i/h/f/e

Administration des réseaux : ping, traceroute, analyseur réseau, station d'administration, configuration automatique des postes de travail, plan d'adressage et de nommage, positionnement et configuration des serveurs de noms • IPv4, IPv6, ICMP, ARP • DHCP, BootP, SNMP, MIB • DNS

Architectures MAN et WAN : dimensionnement des liaisons, choix des technologies, liaisons de secours, installation et configuration des routeurs et des commutateurs, services opérateurs, VPN de niveau 2 et 3 • SDH, WDM, Gigabit, WMAN 802.16 • LS, Frame Relay, ATM • PPP, RNIS, proxy ARP, ADSL, HDSL, Glite • OSPF, BGP, VRRP, MPLS

Gestion avancée des flux IP • **Caractéristiques des flux multimédias :** codec audio G.7xx, codec vidéo H.26x, MPEG, MP3 • **routage multicast :** IGMP, GMRP, PIM, MOSPF, DVMRP • **qualité de service :** choix des files d'attente, modèles de classification et de marquage, WFQ, WRED, ECN, 802.1p IP Precedence, Diffserv, Intserv, RSVP, COPS • **VoIP et ToIP :** configuration des passerelles VoIP et des PABX IP, H.323, Q.931, H.245, H.225, T.120, RTP/RTCP

Systèmes de câblage : caractéristiques et propriétés des câbles, cheminements, locaux techniques, règles d'ingénierie, CTTP, réflectométrie • UTP/FTP/STP catégorie 5/6/7, fibre optique multimode et monomode, 62,5/125 et 50/125, RJ45, MT-RJ, ST, SC • NEXT, FEXT, ACR, ELFEXT, EMC...

Sécurité : vulnérabilités, attaques, chiffrement, authentification, signature, certificats, architecture à contrôle de flux, préconisations pour le filtrage des protocoles les plus sensibles • NAT, PAT, SSL, IPsec, firewall stateful inspection et relais applicatifs, token-cards

Normes et standards : RFC, ITU, IEEE, EIA/TIA, organismes de l'Internet, sites web de référence

www • editions-eyrolles • com

EYROLLES

Code éditeur : G11258
ISBN : 2-212-11258-0

9 782212 112580

Conception : Nord Compo

40 €

Réseaux d'entreprise par la pratique

solutions réseaux

Réseaux d'entreprise par la pratique

JEAN-LUC **MONTAGNIER**

EYROLLES

Préface

N'en déplaise aux nostalgiques des odeurs d'encriers et des plumes d'oies, des belles images de « Jour de fête », des joies et des ambiances fébriles, quand la lettre tant attendue – partie déjà depuis plusieurs jours ! – arrivait enfin, la Correspondance a cédé la place à la Télécommunication, le seul papier au multimédia.

L'échelle du temps de la communication s'est incroyablement comprimée. Souvenez-vous qu'hier encore des compagnies colombophiles existaient dans nos armées alors, que pendant la guerre du Golfe, les anti-missiles américains recevaient leurs données de tir depuis les États-Unis pour intercepter les cibles ennemies repérées par les avions radars quelques secondes seulement après leur lancement.

Tout autant que le temps et conséquence directe de l'évolution fantastique des technologies de l'information, l'espace s'est également virtuellement rétréci. Le monde entier apparaît désormais à portée de main, à portée de « clic ».

Si toutes les entreprises ne sont évidemment pas mondiales, l'économie dans laquelle elles évoluent l'est totalement. Le rythme de l'activité économique est à l'image d'Internet et du haut débit, il est « haute vitesse ». La culture de l'instantanée s'est imposée ! Peu de chance pour celles qui ne peuvent pas suivre.

On l'aura compris, dans un tel environnement les missions du responsable télécommunication d'entreprise sont de toute première importance. D'un service de second plan, hier rattaché aux services généraux et occupant quelques modestes bureaux dans la périphérie immédiate du répartiteur téléphonique, les télécommunications sont dorénavant organisées en direction, souvent associées avec l'informatique. Elles justifient des efforts financiers importants de l'entreprise tant pour le recrutement de ses cadres et techniciens que pour les investissements souvent coûteux dans les équipements réseau. De tels efforts lui sont nécessaires pour rester dans la course.

Toute proportion gardée, cette évolution historique inéluctable est, si on peut dire, rattrapée par la conjoncture géopolitique qui marque ce début de siècle. Les menaces émergentes conduisent ainsi les entreprises à bâtir de véritables plans de secours, voire à développer le concept de « *homeland security* ». Là encore, les réseaux sont un élément clé des mesures de protection décidées. En effet, bien souvent, ces mesures prévoient, aux fins de protection, un éclatement important des structures et des systèmes informatiques, plaçant naturellement les réseaux au cœur de celles-ci. Il est à noter que nous retrouvons là les motivations qui ont conduit, il y a longtemps, le *Department Of Defence* américain à la création du protocole IP.

Véritables systèmes d'irrigation de l'activité économique mondiale, les réseaux de télécommunications et plus largement encore les systèmes d'information sont une cible de choix pour des pirates ou guerriers des temps modernes. Sans aucun esprit de provocation nous pouvons affirmer que « nous sommes en guerre » mais, et c'est bien là tout le paradoxe et la dangerosité de la situation, nous ne le savons pas ou n'avons pas pleinement conscience de la menace qui pèse.

Évidemment, il ne s'agit pas de la guerre avec un grand « G », de celle dont les images d'horreur venues du bout du monde inondent nos journaux télévisés et nos magazines. Pas du tout, la guerre dont il est question est beaucoup moins spectaculaire, tout à fait silencieuse. Elle se déroule au quotidien, sans qu'aucune goutte de sang ne soit jamais versée. Qui s'en plaindrait ?

Les enjeux associés à cette guerre sont très différents de ceux généralement attachés aux conflits militaires. Pour autant, ils n'en sont pas moins importants. En effet, il s'agit - pour ne citer que les principaux - de la sauvegarde de notre patrimoine industriel et intellectuel, de la préservation de notre identité culturelle et, en conséquence, du maintien de notre capacité à porter au-delà de nos frontières, à influer sur les évènements et à tenir notre rang de grande puissance économique.

Les menaces sont nombreuses, puissantes et multiformes. Sans sombrer dans une paranoïa démesurée, force est de constater, qu'observées avec le filtre de la menace qui pèse, plusieurs situations parfaitement connues et admises sont pour le moins inquiétantes.

Prenons comme seule illustration de ces situations – mais il en existe bien d'autres qu'il serait trop long de relater ici - les réseaux de communication de données qu'il s'agisse du réseau Internet ou des réseaux de communication des entreprises au sens le plus large. Ceux-ci, nous l'avons vu, sont aujourd'hui parfaitement indispensables à la vie économique de toute nation développée et de plus en plus nécessaires à la vie sociale. L'ensemble de ces réseaux interconnectés constitue un énorme maillage à l'échelle mondiale ; à chaque nœud de ce maillage, on trouve un « routeur », machine informatique avec un système d'exploitation particulier, qui assure le « juste » aiguillage des communications. Ne nous sommes-nous jamais posé la question de la menace constituée par le fait que le système d'exploitation de ces routeurs, par lesquels – excusez du peu – transitent toutes nos communications des plus banales au plus stratégiques, est presque toujours le même, celui d'un constructeur bien connu, et que ces routeurs sont à une immense majorité issus de ce même constructeur ?

Bien sûr, il n'est pas sérieux de vouloir tout bouleverser dans l'ordre établi en matière de marché des NTIC et plus largement de toute l'industrie du savoir qui, il faut bien le constater, est assez largement entre les mains des acteurs nord américains. En revanche, il apparaît

plus que jamais nécessaire de sensibiliser et au-delà de mobiliser tous les responsables télécommunication et tous les responsables informatique, sur lesquels reposent désormais des enjeux fondamentaux pour les entreprises qui les emploient, afin qu'ils prennent pleinement conscience de ces menaces pour que celles-ci ne soient pas oubliées dans les choix qu'ils seront amenés à faire.

Les réseaux de télécommunication sont donc totalement indispensables à la vie économique, leur fiabilité et leur performance doivent être à la hauteur des enjeux qu'ils supportent, ils constituent une cible privilégiée pour un ensemble de menaces.

L'ouvrage de Jean-luc Montagnier trouve ainsi pleinement sa place au cœur de cette problématique. Il apportera aux architectes et responsables de télécommunications les réponses claires et documentées aux questions essentielles que ceux-ci sont amenés à se poser dans leur quotidien ou dans les choix plus stratégiques qu'ils sont conduits à faire. Un volet complet traitant de la sécurité des réseaux abordera plus spécifiquement les réponses techniques qu'il convient d'adopter.

D'un ouvrage de référence, cette nouvelle édition devient un véritable ouvrage d'actualité et pour longtemps encore, soyons-en sûr.

Olivier Koczan

EADS Defence and Security Systems

Table des matières

PREMIÈRE PARTIE
LES RÉSEAUX LOCAUX

CHAPITRE 1

Installer son premier réseau local **3**

Le contexte 4

Les choix de base 5

Quel réseau ? 5

Quelle topologie ? 5

De quoi a-t-on besoin ? 7

De cartes réseau 7

De cordons de raccordement 9

D'un concentrateur 10

De logiciels de communication 13

Comment faire fonctionner tout cela ? 15

Installer les cartes réseau et les drivers 15

Configurer les adresses IP 17

Installer les concentrateurs et y raccorder les PC 19

CHAPITRE 2

Mettre en place un système de câblage 21

Quelle est la démarche à suivre ? ... 22
L'avant-projet ... 22
L'étude d'ingénierie ... 26
Quel type de câble ? ... 26
Cuivre ou fibre optique ? ... *27*
Coaxial ou paires torsadées ? ... *28*
Le choix de la paire torsadée en distribution ... 28
Écranté ou non ? Blindé ou non ? ... *28*
Catégories 5, 6 ou 7 ? ... *29*
Le choix de la fibre optique entre les locaux techniques ... 32
Multimode ou monomode ? ... *32*
62,5/125 ou 50/125 ? ... *32*
Le câble contenant les fibres ... *33*
Le coaxial et la paire torsadée pour la vidéo ... 34
Selon quels critères choisir le type de câble ? ... *34*
Quel type de prise ? ... 36
L'aménagement des locaux techniques ... 36
Les baies ... 36
Le cheminement des cordons de brassage ... 37
L'organisation du local ... 37
Le cahier des charges ... 38
Le suivi du chantier et la recette ... 40

CHAPITRE 3

Architecture des réseaux locaux 47

Les choix de base ... 48
Quel type de réseau choisir ? ... 48
Quel débit retenir ? ... 48
Quel format d'équipement ? ... 50

Concentrateur ou commutateur ? ... 53
L'architecture ... 56
Mise en place d'un réseau local d'étage ... 56
Extension du réseau d'étage ... 57
Conception d'un réseau d'immeuble ... 59
Mise en place d'un réseau fédérateur ... 59
Quel débit et quelle technologie ? ... 61
Suivre l'évolution des besoins ... 63
Assurer la continuité de service ... 65
L'échange de trames Ethernet ... 68
Échange de trames sur un segment Ethernet ... 68
Échange de trames entre différents segments Ethernet ... 70

CHAPITRE 4

L'alternative du sans fil 77

Introduction ... 78
Les principes de transmission radio ... 79
Le signal radio ... 79
La modulation du signal ... 80
Le multiplexage des canaux de fréquence ... 81
Le codage ... 82
Propriétés des ondes radio ... 83
Exemples d'application ... 83
Revue des réseaux sans fil ... 84
Les applications ... 84
Quel WLAN choisir ? ... 86
Contraintes réglementaires ... 86
Performances ... 88
Architectures et principes de fonctionnement ... 89
Installer un WLAN ... 93
Choisir la bonne antenne ... 94
Les WLAN en entreprise ... 96

Extension du réseau 97
Le WLAN à la maison 98
La configuration des postes client 99
Les paramètres système *99*
Les paramètres de communication *99*
Les paramètres propres au protocole 802.11b *100*
La configuration des points d'accès 101
Comment sécuriser son réseau ? 102
La sécurité de base : WEP 102
WPA et 802.11i 104
La boucle locale radio 106
Les micro-ondes point à point 107

DEUXIÈME PARTIE
LES RÉSEAUX IP

CHAPITRE 5
Démarrer son réseau IP **111**
Le plan d'adressage IP 112
La démarche 112
Les principes de base 114
Impact sur l'Internet 115
Les sous-réseaux IP 115
Méthode d'affectation des réseaux LAN 117
Méthode d'affectation des réseaux WAN 120
Méthode d'affectation des stations au sein des réseaux 121
L'encapsulation des protocoles 122
L'adressage 124
Le multiplexage 125
La résolution d'adresse 129
L'échange de données entre applications 130

CHAPITRE 6

Administrer son réseau IP 135

Les utilitaires de base 136
Le ping 136
Le traceroute 137
Observer ce qui se passe sur son réseau 140
Piloter son réseau 142
Quelle station d'administration ? 142
Pour quelle utilisation ? 143
Configurer automatiquement ses PC 149
Quelle utilisation de DHCP ? 149
Comment configurer un serveur DHCP ? 151
Définir les pools d'adresses 151
Définir les options à distribuer 153
Configurer les routeurs 156
Installer plusieurs serveurs 157
Vérifier la configuration de son PC 157

CHAPITRE 7

La gestion des noms 161

Présentation du DNS 162
Les composants du DNS 162
Élaborer un plan de nommage 162
Définir l'arborescence DNS 163
Standardiser le nommage des objets 166
Configurer les serveurs DNS 168
Configurer le fichier cache 171
Configurer un serveur primaire 172
Activer la résolution de nom 174
Activer le routage de la messagerie 174
Du bon usage des alias 175

Configurer un serveur racine 175
Configurer un serveur secondaire 177
Configurer un serveur cache 177
Déléguer l'autorité à un autre serveur 178
Les domaines de résolution inverse 178
Le fichier d'initialisation 179
Configurer les clients DNS 180
Vérifier le fonctionnement du DNS 181

CHAPITRE 8

En route vers la sixième dimension d'IP **185**

Genèse d'IPv6 186
Les besoins 186
Les solutions 186
Comparaison avec IPv4 187
Remarque sur la terminologie *187*
Configuration des nœuds 190
Fonctionnement standard des nœuds 194
Fonctionnement standard des nœuds 195
La découverte des nœuds voisins (RFC 2461 et 3122) 197
Les options utilisées dans les messages de découverte 197
La résolution d'adresse 198
La résolution inverse d'adresse 199
La découverte des routeurs 200
La vérification de l'état des voisins 201
La redirection 202
L'auto-configuration des adresses (RFC 2462) 202
L'affectation automatique des adresses 202
La détection d'adresse dupliquée 203
La découverte du MTU (RFC 1981) 204
La gestion des groupes de diffusion (RFC 2710 et 2711) 204

Migrer de la v4 à la v6 205

TROISIÈME PARTIE
LES RÉSEAUX ÉTENDUS

CHAPITRE 9

Mettre en place sa première interconnexion 209

Le contexte 210

Les choix de base 210

Quel protocole de niveau 2 ? 211

Quel équipement réseau ? 212

Quel opérateur ? 212

De quoi avons-nous besoin ? 213

D'une liaison entre les deux sites 213

À quel débit ? *213*

Avec quel support de transmission ? *214*

Avec quel service opérateur ? *215*

De routeurs 216

De câbles 217

Connecter un PC au routeur 219

Configurer le routeur 220

Affecter les adresses IP *220*

Activer le routage *221*

Configurer les postes de travail 223

Tester le réseau 226

Réduire l'overhead 226

Mettre en place une liaison de secours 227

Installation d'un accès de base T0 230

Sécurisation de la liaison 231

Gestion du débordement 232

CHAPITRE 10

Architecture des réseaux étendus 233

Les solutions disponibles sur le marché 234
Les liaisons point à point 234
Les réseaux opérateur 235
L'accès aux réseaux des opérateurs 236
Les services proposés par les opérateurs 236
Les choix du client 237
Les lignes spécialisées 240
Équipements installés chez l'utilisateur 241
Les technologies DSL 242
Les applications du xDSL 246
Les boucles SDH 247
Le multiplexage WDM 250
Dimensionner les liaisons 250
Identifier les flux 251
Les flux de type conversationnel *252*
Les flux de type transactionnel *253*
Les flux de type transfert de fichiers *253*
Les flux client-serveur *254*
Estimer la volumétrie 255
Volumétrie liée à la messagerie *256*
Volumétrie liée aux transferts de fichiers *256*
Volumétrie liée aux applications transactionnelles site central *256*
Volumétrie liée aux applications transactionnelles intranet *256*
Volumétrie liée à d'autres services *257*
Rassembler toutes les données *257*
Tenir compte des temps de réponse 260

CHAPITRE 11

Bâtir un réseau de transport 261

LS, Frame Relay ou ATM ? 262
Qualité de service et facturation 266
Débit garanti *266*
Connecter un routeur au réseau de transport 268
Si le routeur ne supporte pas Frame Relay *270*
Si le routeur supporte Frame Relay *271*
Gérer les circuits virtuels 272
Combien de circuits virtuels ? 273
Correspondance entre adresses IP et DLCI *276*
Configurer les SVC 277
Gérer la qualité de service 278
Les sous-interfaces 280
Mettre en place un réseau ATM 281
Qualité de service et facturation 282
La gestion du trafic : TMS 4.0 (ATM Forum af-tm-0056.000) *282*
Les classes de service ATM Transfer Capabilities (ITU I.371 et TMS) *283*
Connecter le routeur au réseau de transport 284
Si le routeur ne dispose pas d'interface ATM *285*
Si le routeur supporte ATM *286*
Configurer les SVC 290
Gérer la qualité de service 292
Les paramètres décrivant les classes de service (ITU I.356 et ATM Forum TMS 4.0) 293
L'adressage ATM 295
L'adressage Frame Relay 296
Interopérabilité entre Frame Relay et ATM 296

CHAPITRE 12

Commutation et routage LAN / MAN / WAN 297

Mettre en place un réseau fédérateur 298

Quels équipements ? 299
Routeur ou commutateur de niveau 3 ? 300
Quelle architecture ? 301
Configurer les VLAN 302
Mettre en place un réseau de campus 305
L'adressage et le routage IP 307
La redondance du routage 308
La rencontre du LAN et du WAN 310
Le routage sur le WAN - OSPF 311
Configurer le routage OSPF 311
Redondance en cas de panne 313
Ajustement des paramètres 315
Modifier le coût des routes 315
Limiter la diffusion des routes 316
Modifier la fréquence des échanges 316
Forcer l'élection du routeur désigné 316
Les performances d'OSPF 317
Le routage entre systèmes autonomes - BGP 317
La commutation sur le WAN - MPLS 320

QUATRIÈME PARTIE
LA GESTION AVANCÉE DES FLUX IP

CHAPITRE 13
Les flux multimédias 327
Les caractéristiques des flux multimédias 328
Choisir un codec audio 331
Caractéristiques 331
Performances 332
Qualité 333
Les codec vidéo 334
Principes de compression des images 335

Les problèmes posés par les transmissions audio et vidéo 338
Estimation du temps de transit 339
Le transport des données multimédias 340

CHAPITRE 14

La qualité de service sur IP 343

Améliorer les performances du réseau 344
Affecter des priorités aux files d'attente 344
Agir sur les files d'attente 346
L'algorithme FIFO – Un fonctionnement simple 346
Gérer les congestions 346
Prévenir les congestions 348
Réguler le trafic 348
Quelle file d'attente choisir pour son réseau ? 350
Gérer la qualité de service 351
La qualité de service selon la préséance 352
Maintenir la qualité de service 353
La qualité de service selon DiffServ 355
Configuration des routeurs 355
Configuration des commutateurs de niveau 2 360
Configuration des commutateurs de niveau 3 361
Définir une règle de marquage 361
Définir une règle de policing 362
Définir une règle de classification 363
Associer une politique à un port 363
Affecter des valeurs au champ DSCP 364
Configuration des postes de travail 365
La qualité de service selon IntServ 367
La réservation des ressources 367
La description de la qualité de service 374
Les classes de service 374
Déployer une politique de qualité de service 376

La qualité de service sur MPLS 377

CHAPITRE 15

Le routage des flux multimédias 379

La diffusion sur un réseau IP 380

La gestion des groupes de diffusion 382

Le routage des flux multicasts 386

Le routage à l'aide de DVMRP 386

Le routage à l'aide de MOSPF 391

Le routage à l'aide de PIM 396

Principe de PIM-SM 397

Principe du calcul des routes 397

Principe du routage 399

Routage sur les liaisons WAN 399

Quel protocole choisir ? 401

Architecture adaptée au protocole DVMRP 403

Architecture adaptée au protocole MOSPF 403

Architecture adaptée au protocole PIM 404

Contrôler la diffusion sur son réseau 405

Limiter les flux multicasts sur le réseau local 408

CHAPITRE 16

La téléphonie et la vidéo sur IP 409

Présentation des protocoles multimédias 410

Les composants d'un système H.323 411

L'établissement d'une communication 414

Interconnecter les PABX *via* IP 416

Mettre en place un gatekeeper 420

La voie vers le tout IP 425

Configurer le PABX et la passerelle VoIP 427

Déclarer les terminaux téléphoniques 428

Assurer la qualité de service 430
Transporter les flux multimédias 431
Le transport des flux audio et vidéo via RTP et RTCP 432
Optimiser les flux multimédias 435
Compression des en-têtes 435
Échanger des données multimédias 437

CINQUIÈME PARTIE
LA SÉCURISATION DES RÉSEAUX IP

CHAPITRE 17
Protéger son réseau et ses données 441
L'importance d'une politique de sécurité 442
Les vulnérabilités des protocoles IP 443
Telnet 443
FTP 443
DNS 444
HTTP 445
Netbios 445
SNMP 445
RPC (Remote Procedure Calls) 446
NFS 446
ICMP 447
Les différents types d'attaques 447
Les attaques de type refus de service (denial of service) 447
Quelques exemples d'attaques par refus de service 448
Les attaques par inondation SYN (syn flooding) 448
La dissimulation d'adresses (spoofing) 448
Les attaques par tunnel 449
Le vol de session (session stealing, splicing ou hijacking) 449
Le rebond 449
Les chevaux de Troie 450

Les vers 450
Le contrôle de flux 450
Les technologies utilisées par les firewalls 451
Le filtrage de paquet 451
Le « stateful inspection » 452
Le relais applicatif 452
Le relais de circuit 452
Comparaison des technologies 453
Choix d'un firewall 453
Comparaison entre les routeurs et les firewalls 455
Définition d'une architecture à contrôle de flux 455
Acteurs et matrice de confiance 456
Matrice de communication 457
Mécanismes de sécurité supplémentaires 458
Cas du DNS 460
Cas de Netbios 461
Services IP et matrice de flux 462
Matrice de filtrage 463
Rôle et fonctionnement des firewalls 464
Architecture 464
Fonctionnement 465
Relais applicatif 465
Cas des protocoles non relayés 466
Authentification 466
Translation d'adresses 467
Redondance 467
Administration 467
Log et audit 467
Mécanismes de protection 468
Intégrité 468
Chiffrement des données 468
Remarques sur l'administration des firewalls 468

La confidentialité 469
Les algorithmes de chiffrement 469
Efficacité des algorithmes de chiffrement 471
Les protocoles de sécurité 473
Les protocoles d'échange de clés 474
La gestion des certificats 475
La signature numérique 476
L'enveloppe numérique 477
Les contraintes législatives 477
L'authentification 480
Les mécanismes d'authentification 480
Fonctionnement d'une calculette d'authentification 482
Les serveurs d'authentification 482
Exemple d'authentification forte pour les accès distants 483

ANNEXES

Normes et standards **487**

Le câblage 487
Normes CEN relatives au câblage *487*
Normes EIA/TIA relatives au câblage *487*
Normes ITU-T relatives au câblage *487*
Les interfaces physiques 488
Avis de l'ITU-T relatifs aux interfaces physiques *488*
Normes EIA/TIA relatives aux interfaces physiques *488*
Avis de l'ITU-T relatifs aux échanges ETTD-ETCD *488*
Les réseaux locaux 489
Normes IEEE relatives aux réseaux locaux *489*
La famille des protocoles TCP/IP 491
RFC relatives aux protocoles TCP/IP *491*
*Standards originaux du DOD (*Department Of Defense*)* *492*
RFC relatives *aux protocoles de routage IP* *492*
RFC relatives aux applications utilisant TCP/IP *492*

RFC relatives à IP sur Frame-Relay *493*
RFC relatives à IP sur ATM *493*
RFC relatives à PPP *494*
RFC relatives à SNMP *494*
Normes ISO et équivalents ITU-T relatifs à la syntaxe ASN.1 ..... *495*
RFC relatives à IPv6 ..... *495*
RFC relatives à la voix sur IP *496*
Avis de l'ITU-T relatifs à la voix sur IP *496*
RFC relatives à la qualité de service *497*
RFC relatives au routage multicast *497*
Les réseaux RNIS 498
Organisation et nomenclature des normes *498*
Série I.100 : Concepts généraux du RNIS *499*
Série I.200 : Services assurés par le RNIS *499*
Série I.300 : Aspects réseaux du RNIS *499*
Série I.400 - Interfaces usager-réseau *500*
Série I.500 : Interfaces d'interconnexion du RNIS *501*
Série I.600 : Administration du RNIS *501*
Avis de l'ITU-T relatifs aux réseaux ATM *502*
Avis de l'ITU-T et équivalents ANSI relatifs au Frame Relay *502*
Avis de l'ITU-T relatifs aux systèmes de transmission numérique MIC *503*
Avis de l'ITU-T relatifs aux réseaux SDH *503*
Organisation de l'Internet **505**
Quelques chiffres *506*
La gestion de l'Internet 507
Quelques autres organismes d'intérêt général *510*
Les anciens organismes de régulation *510*
Où les contacter *510*
Glossaire **513**

Bibliographie **529**

Sites web **531**

Accès aux RFC ... 531
Câblage ... 531
Internet ... 531
Modem-câble ... 532
Organismes de normalisation ... 532
Organismes de régulation ... 533
Protocoles ... 533
Qualité de service ... 533
Réseaux sans fils ... 534
Revues de presse ... 534
Sécurité ... 535
VoIP ... 535

Index **537**

Table des encarts **547**

Avant-propos

J'ai retrouvé, il y a peu, un ouvrage faisant partie de ceux qui ont abreuvé une génération d'étudiants en réseaux & télécom, un pavé de plus de 900 pages. Il faisait partie d'une de ces bibles immanquables que l'on se devait de lire. Celui qui n'avait pas lu le « Machin » ou le « Truc » passait assurément à côté de quelque chose, et risquait de compromettre ses examens.

Sur les 900 pages dédiées aux réseaux, une seule était consacrée à TCP/IP sous une rubrique bizarrement intitulée « La productique ». Il y était dit que cette architecture était démodée et que, depuis quelques années, tous les utilisateurs de ce type de réseau étaient amenés à évoluer vers l'architecture OSI. Le livre datait de 1987, TCP/IP avait 18 ans et l'Internet explosait… aux États-Unis.

Cela m'a rappelé les quelques temps passés à travailler au cœur de la Silicon Valley en tant que programmeur. J'étais alors en charge de développer des couches logicielles autour de TCP/IP. Un de mes collègues avait affiché à l'entrée de son bureau un manifeste intitulé « Why OSI ? ». Parmi les réponses saugrenues, il y avait celles-ci : « parce que c'est normalisé », « parce qu'il y a 7 couches », « parce que c'est compliqué », etc.

Voilà un des problèmes de l'Europe : d'un côté un centre d'activité qui crée la technologie de demain, de l'autre des commentateurs avertis. Dans l'entreprise, contentons-nous donc d'utiliser au mieux ce qu'on nous propose.

J.-L. Montagnier

Les technologies gagnantes

Les technologies dont il est question ici, sont celles qui sont les plus utilisées dans les entreprises. Celles qui ont supplanté les autres.

Le monde des réseaux a, en effet, longtemps été caractérisé par une quantité innombrable de protocoles plus ou moins exotiques, essentiellement poussés par des constructeurs soucieux de verrouiller leur marché (Digital avec Decnet, IBM avec SNA, Novell avec IPX, etc.) et par des organismes de normalisation sacrifiant aux plaisirs des techniciens (OSI de l'ISO…).

Ainsi, seuls quelques protocoles ont survécu ; ils ont pour point commun d'avoir su s'adapter aux besoins des entreprises tout en présentant le meilleur rapport qualité/prix.

Par exemple, l'Ethernet de l'an 2000 ne ressemble plus à celui des années 70. La famille TCP/IP s'est enrichie de dizaines de protocoles, et le Frame Relay a su intégrer la voix et les données en offrant le minimum de qualité de service nécessaire. Tandis que le RNIS a su répondre à un besoin bien spécifique d'interconnexion.

Tout cela parce ces protocoles reposent sur les technologies ouvertes, simples et qui apportent une réelle plus-value aux entreprises.

Le principe de l'apport minimal

L'histoire montre que les technologies qui se sont imposées sont celles qui répondent à au moins un des trois critères suivants : conservation de l'existant, réponse aux besoins et réduction des coûts.

Par exemple, la technologie conserve-t-elle l'existant ? Répond-elle aux besoins ? Réduit-elle les coûts ? Si l'ensemble des réponses aboutit à une réponse positive, alors oui, elle apporte quelque chose !

La nouvelle technologie peut ne pas conserver l'existant, mais apporter une réelle plus-value ; c'est le cas du Compact Disc qui a remplacé le vinyle en quelques années (meilleure qualité de son, plus grande résistance, etc.).

La nouvelle technologie peut être plus chère mais répondre aux besoins ; c'est le cas des téléphones mobiles. Pour un usage personnel, on a besoin de communiquer, de se sentir important (besoins irrationnels). Pour un professionnel, être joignable à tout moment fait partie de la qualité de service qu'il offre à ses clients (besoin rationnel) : cela lui permet d'être plus réactif et donc de gagner plus d'argent, même si les communications coûtent deux fois plus cher que celles d'un téléphone fixe.

La nouvelle technologie peut remettre en cause l'existant mais réduire les coûts. Si le retour sur investissement est assuré, elle a alors de fortes chances de s'imposer. Ce constant souci de productivité est présent dans toutes les industries.

Face à une nouvelle technologie, la question à se poser est donc : est-ce qu'elle apporte quelque chose ? est-ce qu'elle répond au principe de l'apport minimal ?

Le syndrome de Vinci

Faut-il alors donner du temps à une technologie pour que celle-ci s'impose ?

Par exemple, IP n'a pas connu une diffusion planétaire immédiate : ce protocole s'est imposé rapidement dans les universités et les centres de recherche, puis plus lentement dans le monde des entreprises.

Cela n'est cependant pas un gage de réussite ; il suffit de se souvenir de ce qui est arrivé à l'ATM : technologie censée tout faire, elle s'est avérée trop complexe à mettre en œuvre et reste, aujourd'hui, cantonnée aux réseaux des opérateurs. L'échec dans les réseaux locaux est patent.

Le problème des technologies trop en avance sur leur temps est qu'elles sont soumises au syndrome de Vinci[1] :

- Au début, elles ne répondent pas au principe de l'apport minimal.
- Leur complexité et leur coût freinent leur développement.
- Dix ans après, elles sont dépassées (techniquement ou économiquement) par d'autres technologies ou de plus anciennes qui ont évolué avec leur temps.

C'est ce qui aurait pu arriver à IP si les protocoles OSI avait été plus performants, et c'est ce qui est arrivé à ATM face à un Ethernet tout terrain qui a su s'adapter au fil du temps.

Dans dix ans, on pourra se poser de nouveau la question : Ethernet et IP ont-ils atteint leurs limites par rapport aux besoins du moment ? Si oui, quelle est la solution qui répond au principe de l'apport minimal ?

Un exemple dans le domaine des réseaux

L'Ethernet à 100 Mbit/s, puis le Gigabit Ethernet, se sont imposés très rapidement, parce que d'une part les composants électroniques proviennent de technologies existantes (*Fibre Channel*, etc.) ce qui réduit les coûts, et d'autre part l'existant est préservé.

ATM est sans doute ce qui se fait de mieux en matière de technologie réseau, mais il en fait trop par rapport aux besoins d'aujourd'hui. En faire trop implique de dépenser plus d'argent en R&D, en formation, en fabrication, etc., ce qui a pour conséquence de ralentir la diffusion de ladite technologie.

En réalité lorsqu'ils bénéficient d'une dynamique importante comme Ethernet et IP, les technologies et les produits s'améliorent au fil du temps. La méthode américaine est, en effet, de sortir un produit le plus rapidement possible afin rentabiliser les premiers investissements, les améliorations ne venant que si le marché se développe.

---

[1] Artiste de génie, Léonard de Vinci était également « Premier ingénieur et architecte du Roi, Mécanicien d'état ». Cependant, nombre de ses projets sombraient dans les limbes, parce que trop novateurs ou irréalisables avec les moyens techniques de l'époque. Ses talents d'ingénieur étaient surtout mis à contribution pour réaliser des automates de fêtes foraines et de spectacles.

Guide de lecture

L'ouvrage est structuré en quatre parties : les réseaux locaux, les réseaux IP, les réseaux étendus et la gestion avancée des flux IP.

Cette organisation correspond à une progression technique, qui suit l'évolution d'une architecture réseau, allant de la plus simple à la plus complexe, telle qu'elle peut être rencontrée par les responsables réseaux et télécoms dans le monde de l'entreprise.

Les réseaux locaux

Dans sa première partie, l'ouvrage traite des réseaux Ethernet en partant du plus simple aux architectures les plus complexes, tout en expliquant les techniques mises en œuvre. Ensuite, les technologies liées au câblage, passage obligé pour bâtir des réseaux, sont expliquées dans le détail. Enfin, alternative au câblage, les réseaux Ethernet sans fils sont également abordés sous les aspects techniques et architectures.

Chapitre	Contenu
1. Installer son premier réseau local	Topologies bus/étoile, coaxial 10b2, hubs 10bT
2. Mettre en place un système de câblage	Cuivre UTP/STP, catégories 5/6/7, RJ45 Fibre optique multimode/monomode, ST/SC
3. Architecture des réseaux locaux	Ethernet 10/100/1000bT, couche MAC, spanning tree 802.1d, GARP, switches
4. L'alternative du sans fil	WLAN 802.11a/b/g Wi-Fi, WPAN 802.15 Bluetooth

Les réseaux IP

Dans sa deuxième partie, l'ouvrage traite du protocole IP et son interaction avec Ethernet en termes d'adressage et d'encapsulation. Les outils d'administration sont ensuite expliqués en établissant le lien entre le fonctionnement des protocoles et leur mise en œuvre. La gestion des noms, qui permet de découpler les applications du réseau, occupe une place à part étant donnée l'ampleur du sujet. Enfin, la nouvelle version d'IP, qui commence à apparaître avec les terminaux UMTS et chez les opérateurs, est décrite dans le détail.

Chapitre	Contenu
5. Démarrer son réseau IP	IPv4, adressage IP, interaction IP / Ethernet, TCP, UDP
6. Administrer son réseau IP	Ping, traceroute, ICMP, SNMP, MIB, DHCP, Bootp
7. La gestion des noms	DNS, protocole, architecture et configuration
8. En route vers la sixième dimension d'IP	IPv6, adressage, ICMPv6, protocole et architecture

Les réseaux étendus

Dans sa troisième partie, l'ouvrage traite des réseaux étendus, les WAN. À l'inverse des réseaux locaux, où Ethernet règne sans partage, une profusion de protocoles occupe le terrain. Un premier chapitre présente, sous forme de guide de choix, un panorama de la problématique. Les technologies employées dans les boucles locales sont ensuite décrites tout en faisant la relation avec les offres de service des opérateurs. Les réseaux de transports, utilisés par les opérateurs et les entreprises, sont ensuite détaillés. Enfin, les protocoles réalisant le lien entre le LAN et le WAN sont traités.

Chapitre	Contenu
9. Mettre en place sa première interconnexion	Guide de choix, PPP, RNIS, proxy ARP, routeurs, interfaces WAN
10. Architecture des réseaux étendus	VPN niveaux 2/3/4, HDSL, ADSL, SDH, WDM
11. Bâtir un réseau de transport	LS, Frame Relay, ATM
12. Commutation LAN et routage WAN	VLAN, GVRP, VRRP, OSPF, BGP, MPLS

La gestion avancée des flux IP

Dans sa quatrième partie, l'ouvrage présente des techniques avancées de gestion des flux. Les réseaux IP véhiculant des flux de plus en plus variés, des mécanismes complémentaires et sophistiqués doivent être mis en œuvre. Un premier chapitre expose la problématique posée par les flux multimédias afin d'expliquer la nécessité de ces mécanismes. Ceux-ci sont ensuite décrits dans les deux chapitres suivants : gestion de la qualité de service et routage des flux multimédias. Des exemples de configuration aident à expliquer leur fonctionnement.

Chapitre	Contenu
13. Les flux multimédias	Codec audio G.7xx, codec vidéo MPEG, H.26x
14. La qualité de service sur IP	WFQ, WRED, TOS, Diffserv, Intserv, RSVP, COPS
15. Le routage des flux multimédias	IGMP, DVMRP, MOSPF, PIM-SM, PIM-DM, GMRP
16. La téléphonie et la vidéo sur IP	H.323, Q.931, H.225, H.245, RTP/RTCP, T.120

La sécurisation des réseaux IP

Enfin, et non des moindres, le dernier chapitre traite de la sécurité des réseaux, non seulement sur les plans techniques, mais aussi en termes d'architecture. Trop souvent, en effet, les équipements de sécurité sont ajoutés sans discernement comme une surcouche aux réseaux. Alors qu'une véritable politique de sécurité suivie d'une réflexion permettent de concilier sécurité et communication.

Chapitre	Contenu
17. Sécuriser son réseau	Algorithmes de chiffrement, firewalls, IPsec, SSL

Les autres protocoles

D'autres protocoles sont à venir et pourraient trouver leur place dans les entreprises (et dans cet ouvrage !) : SIP (concurrent de H.323), MEGACO (*Media Gateway Control*), GMPLS (*Generalized MPLS*), EFM (*Ethernet in the First Mile*), 802.17 RPR (*Resilient Packet Ring*) ou encore SNMPv3 (*Simple Network Management Protocol*).

Inversement, des protocoles, qui existent pourtant toujours, ne sont volontairement pas traités : ATM pour son usage dans le LAN, FDDI, Token-Ring, X.25, SNA (et les protocoles associés, tels que DLSW et STUN) ou encore Decnet. Le Token-Ring a, en effet, été supplanté par l'Ethernet, le X.25 par le Frame Relay, IPX par IP, etc. Le SNA n'est présent que dans les sociétés qui ont beaucoup investi dans les *mainframe* IBM. La migration vers IP est cependant bien amorcée.

PREMIÈRE PARTIE

Les réseaux locaux

1

Installer son premier réseau local

La connexion à l'Internet a été l'occasion pour nombre d'entre nous de se frotter aux réseaux en manipulant ses composants : modems, câbles, logiciels de communication, navigateurs et messageries.

De retour au bureau, l'étape suivante consiste à construire son propre réseau pour les besoins de son entreprise, c'est-à-dire un **intranet**. Le réseau offre, en effet, de formidables possibilités de développement : pouvoir vendre des produits au monde entier sans ouvrir de boutiques dans chaque pays, collecter des informations sur des sujets précis, échanger des documents avec ses fournisseurs, etc. Avec les réseaux, nous entrons de plain-pied dans la société de l'information.

Dans une entreprise, le réseau est tout d'abord local, c'est-à-dire limité à un ou plusieurs bâtiments. Commençons donc par là.

Dans ce chapitre, vous apprendrez ainsi :

- quels sont les principes de base d'un réseau local ;
- à choisir les matériels et logiciels pour votre réseau ;
- à installer une carte réseau ;
- à configurer votre PC.

Le contexte

Avec la connexion à l'Internet, vous avez commencé, sans le savoir, à construire les prémisses d'un réseau. Vous en avez utilisé tous les composants : câbles, matériel de connexion (dans notre cas le modem), logiciels TCP/IP, etc. Il s'agissait du type de réseau, parmi les nombreuses autres variantes possibles, qui était le plus approprié pour connecter un seul poste de travail.

Maintenant, le but est de relier plusieurs PC entre eux (de 2 à 10), par exemple, le vôtre à ceux de la secrétaire et du comptable ; ou, chez vous, entre votre bureau, la cave et la cuisine (histoire de s'amuser). Les PC sont distants de quelques mètres.

Il s'agit donc de créer un réseau adapté à ce besoin. Par conséquent, on utilisera des matériels adaptés aux réseaux d'entreprise (à chaque problème sa solution).

Le réseau qui correspond à notre situation est appelé réseau local (LAN, *Local Area Network*). Pour le mettre en place, vous avez besoin :

- d'une série de câbles qui relient les PC entre eux ;
- de cartes réseau qui permettent aux PC de se raccorder à ces câbles, d'envoyer des données et d'en recevoir ;
- de logiciels de communication, tels qu'un pilote pour les cartes réseau et une pile TCP/IP. IP (*Internet Protocol*) est le protocole qui permet aux ordinateurs d'échanger des données sous forme de paquets, tandis que TCP gère les sessions entre ces mêmes ordinateurs.

QU'EST-CE QU'UN RÉSEAU LOCAL ?

Un **LAN** (*Local Area Network*) est un réseau dont la portée est limitée de quelques mètres à plusieurs centaines de mètres. C'est le type de réseau que l'on peut installer chez soi, dans des bureaux ou dans un immeuble. Un LAN, comme tout réseau, repose sur un support de transmission : un câble (en cuivre ou fibre optique) ou, plus rarement, les ondes radio.

Les réseaux locaux les plus répandus sont **Ethernet** (85 %) et **Token-Ring** (15 %). Il existe plusieurs **topologies** pour un LAN :

- En **anneau**. Les PC sont chaînés entre eux, le premier étant connecté au dernier, afin de former l'anneau.
- En **bus**. Les PC sont connectés à un câble qui parcourt tous les bureaux ou toutes les pièces de la maison.
- En **étoile**. Autour d'un équipement spécifique appelé **concentrateur** (couramment appelé *hub* pour Ethernet *et MAU* pour Token-Ring).

La topologie en étoile est la plus courante, tandis que le bus est le moyen le plus simple pour construire un réseau Ethernet.

On retrouve les mêmes briques à assembler que pour une connexion Internet. La carte réseau remplace ici le modem (RTC, ADSL ou câble) qui sert à se connecter à l'Internet. Une telle carte est conçue pour le réseau local (LAN – *Local Area Network*), tandis que le modem est utilisé pour les réseaux longue distance (WAN – *Wide Area Network*).

QUELLES DIFFÉRENCES ENTRE ETHERNET ET TOKEN RING ?

Le principe d'Ethernet repose sur un **bus partagé** : chaque station émet quand elle le souhaite mais, quand deux stations émettent en même temps, il se produit une **collision** matérialisée par la somme des deux signaux véhiculant les deux trames. Dans ce cas, les émissions sont stoppées et au bout d'un laps de temps aléatoire, une autre tentative est faite.

Le principe de Token Ring repose sur un **anneau** : chaque station attend de disposer d'un **jeton** (matérialisé par une trame d'un format particulier) avant d'émettre une trame. Le jeton circule de station en station, formant un anneau.

Le bus partagé à détection de collision et l'anneau à jeton sont deux **méthodes d'accès** à un support de transmission tel qu'un câble.

À l'inverse du bus partagé dont l'accès est **aléatoire**, la technique du jeton est **déterministe** : chaque station parle à tour de rôle au bout d'un laps de temps fixe qui dépend du nombre de stations (le temps pour le jeton de faire le tour de l'anneau). La bande passante est mieux exploitée avec Token-Ring, ce qui le rend plus performant.

L'avantage technique offert par le Token Ring n'est pas utile aux réseaux locaux. De plus, il nécessite des composants électroniques plus complexes et donc plus chers à fabriquer. En résumé, Ethernet est plus simple, plus évolutif et présente le meilleur compromis coût/performances.

Les choix de base

Quel réseau ?

Tout d'abord, quel type de réseau retenir ? Ethernet (gestion des collisions sur un bus) ou Token-Ring (gestion d'un jeton sur un anneau) ?

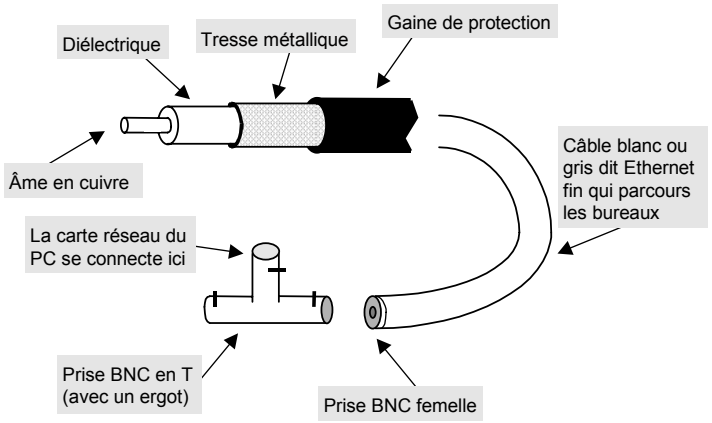
Question performances, les deux se valent, même si, à débit égal, il y a un léger avantage à utiliser Token-Ring. Cependant, Ethernet détient plus de 85 % du marché et a toujours été techniquement en avance sur Token-Ring. Si l'on doit créer soi-même un réseau à partir de rien, autant se lancer dans Ethernet : c'est plus simple et cela coûte moins cher.

Si, dans une entreprise, Token-Ring est déjà bien implanté, on peut envisager de poursuivre dans cette voie. Mais une migration vers Ethernet est toujours envisageable : tout n'est qu'une question de retour sur investissement.

Quelle topologie ?

Historiquement, le bus a été la première topologie pour Ethernet : elle repose sur un câble spécifique en cuivre, appelé câble coaxial, qui parcourt tous les bureaux dans lesquels il y a un PC à connecter.

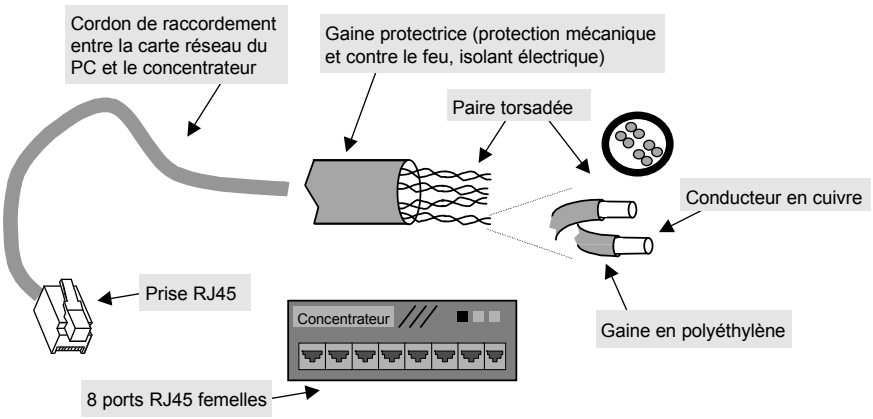
Figure 1-1.
Composant d'un réseau Ethernet en bus.



Le câble est dit « Ethernet fin » par comparaison à une autre variante d'Ethernet, de moins en moins répandue, qui utilise un câble plus épais de couleur jaune.

Aujourd'hui, la topologie la plus répandue est celle de l'étoile qui consiste à relier tous les PC à un équipement central appelé concentrateur (*hub*, en anglais). Le câble est constitué de quatre paires de fils de cuivre torsadées et est terminé par des connecteurs RJ45.

Figure 1-2.
Composant d'un réseau Ethernet en étoile.



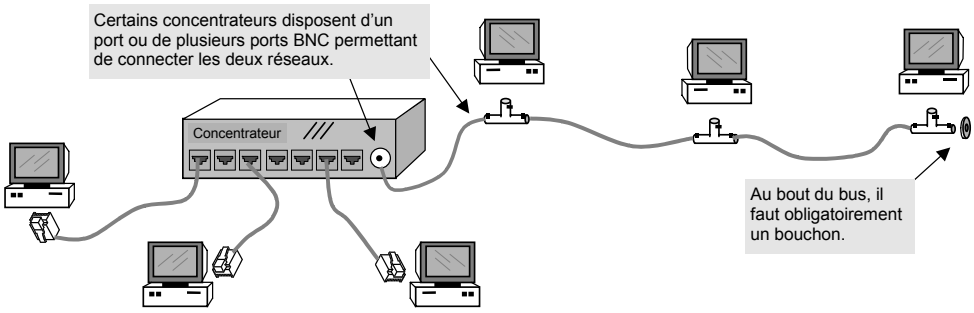
Il existe de nombreuses variantes de câbles de fils de cuivre en paires torsadées selon l'impédance, le diamètre des fils et la nature des protections. Elles seront étudiées au chapitre suivant.

Ethernet	Bus	Étoile
Câble cuivre	Coaxial	Paires torsadées
Connecteurs	BNC	RJ45
Vitesse	Limité à 10 Mbit/s	10 Mbit/s et plus
Modification du réseau	Difficile	Très facile
Remarque	De moins en moins répandu	Nécessite un concentrateur Ethernet
Adapté aux…	Petits réseaux locaux	Petits et grands réseaux locaux

En définitive, l'Ethernet en bus est la solution la plus économique lorsque l'on veut connecter quelques PC qui sont regroupés dans une seule pièce. L'Ethernet en étoile est plus cher puisqu'il nécessite un concentrateur (de 75 à 300 € en entrée de gamme selon le nombre de ports RJ45).

À moins que vous ne disposiez de matériel de récupération de type BNC, la topologie en étoile est conseillée. En effet, elle vous permettra de faire évoluer votre réseau tout en conservant les cartes et le concentrateur.

Figure 1-3.
Réseau Ethernet en étoile et en bus



De quoi a-t-on besoin ?

De cartes réseau

Chaque PC a besoin d'un équipement capable de « parler Ethernet » : c'est le rôle de la carte réseau, dite carte Ethernet, et souvent appelée NIC (*Network Interface Card*). Elle s'insère dans un emplacement (*slot*) du PC qui lui est réservé.

Il existe plusieurs types de cartes Ethernet qui se distinguent par leur connecteur :

- BNC pour l'Ethernet fin en bus ;
- RJ45 pour l'Ethernet en étoile ;
- AUI pour l'Ethernet en bus ou en étoile.

Certaines cartes proposent une combinaison de deux de ces prises, voire les trois.

La prise AUI (*Attachment Unit Interface*) permet de connecter un équipement appelé *transceiver*, qui réalise l'adaptation au câble. Il existe ainsi des *transceivers* de types BNC, RJ45 et en fibre optique. L'acquisition de cette carte (30 € environ) peut être envisagée si votre réseau nécessite plusieurs types de câbles (coaxial, en paire torsadée, en fibre optique), voire un support de transmission radio (très peu répandu). Dans le cas de cartes RJ45 ou BNC, le *transceiver* est intégré à la carte.

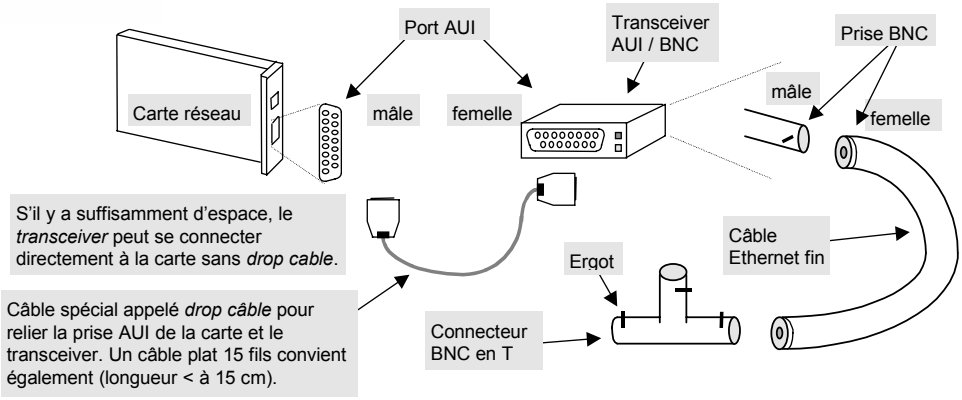
QU'EST-CE QU'UNE CARTE ETHERNET ?

L'ordinateur traite les informations sous forme numérique et sous forme de mots de 32 ou 64 bits (64 éléments d'informations binaires — 0 ou 1).

Une carte réseau Ethernet permet de convertir ces informations en signaux électriques qui sont émis sur le câble. La manière de représenter les bits d'information en signaux s'appelle le **codage**. Pour Ethernet, il s'agit du codage **Manchester**.

La carte envoie ces bits par groupes, appelés **trames Ethernet**. La norme Ethernet spécifie les couches 1 (physique : transmission des signaux par la carte réseau) et 2 (logique : format des trames Ethernet).

Figure 1-4.
Connectique AUI / BNC.



Si vous démarrez avec un réseau en bus, il est conseillé d'acheter une carte équipée de deux connecteurs, un BNC et un RJ45 (la différence de coût est minime). Cela vous permettra de la réutiliser si vous changez de réseau.

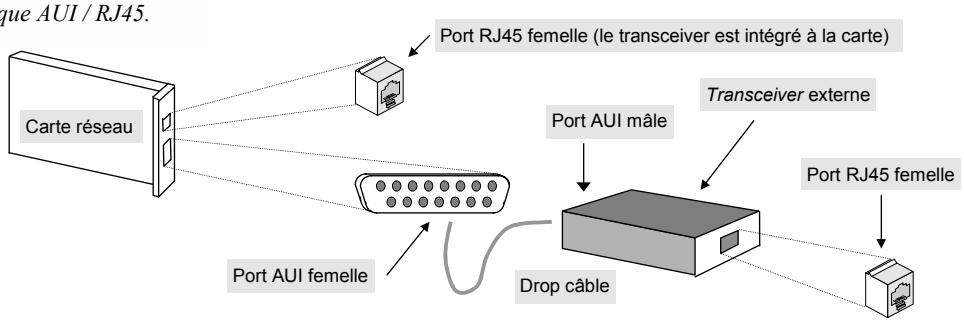
Le coût d'une carte dépend de ses performances, et notamment du bus :

- entrée de gamme avec bus PCI : de 40 à 90 € HT ;
- entrée de gamme avec bus ISA (moins performante que PCI) : de 40 à 120 € HT.

Dans le haut de gamme, les cartes performantes sont celles qui échangent les données avec l'ordinateur *via* un DMA (*Direct Memory Access*), composant électronique spécialisé.

Pour notre réseau, une carte d'entrée de gamme suffira. Les cartes haut de gamme sont plutôt destinées aux serveurs.

Figure 1-5.
Connectique AUI / RJ45.



De cordons de raccordement

Le cordon de raccordement (également appelé cordon de brassage) est nécessaire pour connecter chaque carte réseau au concentrateur. Pour notre réseau, le plus simple est de poser un câblage volant, c'est-à-dire constitué uniquement de cordons de brassage. D'autres types de câblages — plus complexes et plus chers — seront étudiés au chapitre suivant, car au-delà de dix PC le câblage volant devient ingérable et source de problèmes.

Le support de transmission que nous avons choisi est un câble cuivre en paires torsadées dont chaque extrémité est pourvue d'une prise RJ45. Sa longueur ne doit pas excéder cent mètres, selon la qualité du câble et l'environnement électrique. Pensez notamment à éloigner vos câbles de toutes sources de perturbations : appareils électriques tels que la cafetière, le ventilateur, l'aspirateur, le moteur de l'ascenseur, le transformateur de courant, etc.

Il existe de nombreux types de câbles. Toutefois, pour notre premier réseau, le choix n'a guère d'importance. Précisons simplement, et sans entrer dans les détails, qu'il est conseillé d'acheter un câble UTP (*Unshielded Twisted Pair*), 100 Ohms, catégorie 5. C'est le moins cher, il répond à des normes précises et il est parfaitement adapté à nos besoins.

LES CÂBLES CUIVRE EN PAIRES TORSADÉES

Les câbles se différencient, avant tout, par leur **impédance**, exprimée en Ohms (Ω). Les valeurs rencontrées pour les réseaux locaux sont : 100, 120 et 150 Ohms. Plus l'impédance est élevée, meilleure est la qualité du câble (le signal est moins affaibli), mais plus son coût est élevé. Le plus répandu est le **100 Ohms**.

Les câbles se différencient également par la présence ou non de protection contre les perturbations émises par les courants électriques. Il existe des câbles sans protection, dits **UTP** (*Unshielded Twisted Pair*), avec un **écran**, dits **FTP** (*Foilded Twisted Pair*) et avec un **blindage**, dits **STP** (*Shielded Twisted Pair*). Il existe aussi la combinaison **SFTP**.
La prise la plus répandue pour les câbles en paires torsadées est la **RJ45** (*Registered Jack 45*), normalisée ISO 8877.

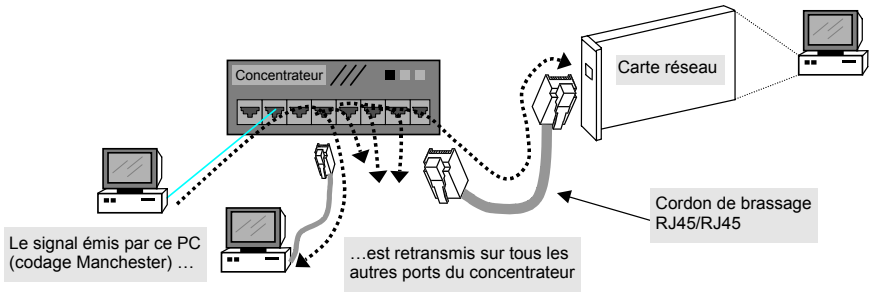
Le câble utilisé entre le PC et le concentrateur doit être droit ; les fils émission et réception ne doivent pas être croisés. Le croisement est, en effet, réalisé dans la prise RJ45 du concentrateur. Pour s'en assurer, il suffit de mettre l'un à côté de l'autre les deux connecteurs RJ45 du câble, orientés dans le même sens, et d'examiner la couleur des huit fils (généralement, le connecteur RJ45 est transparent). Si le câble est droit, les couleurs apparaissent dans cet ordre : bleu, orange, noir, rouge, vert, jaune, marron et gris. Ethernet utilise les fils numérotés 1,2 (émission) et 3,6 (réception).

Si, en revanche, vous connectez deux PC directement (sans concentrateur), le cordon doit être croisé (ce qui est logique).

D'un concentrateur

Le concentrateur est un appareil qui régénère les signaux. En effet, le signal émis par la carte Ethernet s'affaiblit en parcourant le câble et, au-delà de cent mètres, il peut devenir trop faible. Cette distance correspond en fait au maximum autorisé par la norme entre un PC et le concentrateur. Un signal émis par un PC est régénéré sur tous les autres ports du concentrateur (il joue le rôle de répéteur).

Figure 1-6.
Fonctionnement d'un réseau Ethernet en étoile.



Cela nous amène à la remarque suivante : s'il n'y a que deux PC à connecter, un concentrateur est inutile ; les deux cartes réseau peuvent être reliées directement *via* un cordon de brassage n'excédant pas cent mètres. Un câble ne disposant que de deux extrémités (loi physique incontournable de notre univers), la connexion de trois PC ou plus passe obligatoirement par un concentrateur

La plupart des concentrateurs sont dits « intelligents ». Cela signifie qu'ils disposent de mécanismes permettant de détecter les erreurs (signal trop faible, collisions, etc.) et de désactiver le port par lequel ces erreurs ont été détectées afin de ne pas perturber les autres ports (fonction de partitionnement).

Quel concentrateur choisir ? Dans notre cas, un modèle d'entrée de gamme est nettement suffisant. Le critère de choix est alors le nombre de ports RJ45, qui conditionne le nombre de PC à connecter : 4, 5, 6, 8, 12, 16, 24, 32 et 48 ports sont des valeurs couramment proposées.

En dehors du nombre de ports, les concentrateurs se distinguent par différentes fonctions.

Certains sont dits administrables ou *manageables*). Cela signifie qu'ils sont équipés d'un logiciel SNMP (*Simple Network Management Protocol*) qui permet de les administrer à distance. Dans notre cas, cette fonction n'est pas indispensable, d'autant plus que la différence de prix peut atteindre 150 €.

D'autres sont dits empilables (*stackable*) : cela veut dire qu'ils peuvent être chaînés afin d'augmenter le nombre total de ports. Le chaînage est effectué à l'aide d'un bus souvent matérialisé par un câble externe spécifique. Il n'existe aucune norme en la matière, ce qui signifie que vous ne pouvez pas chaîner deux concentrateurs de marque différente (un 3com avec un Dlink, par exemple).

QU'EST CE QU'UN SEGMENT ?

Un concentrateur Ethernet (*hub*) concentre les connexions réseau des PC pour former un **segment** Ethernet.

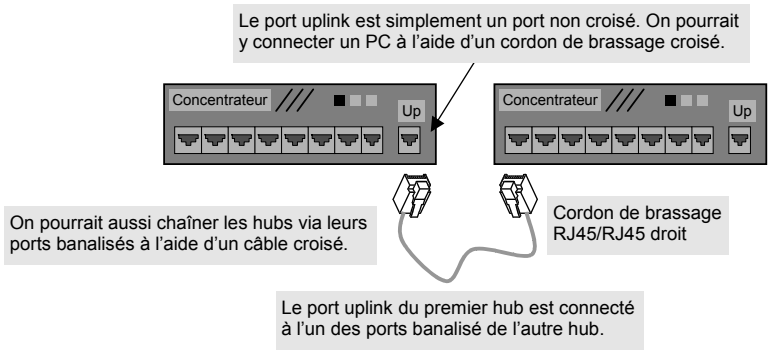
Au sein d'un segment, toutes les trames émises par un PC sont transmises par l'intermédiaire d'un *hub* à tous les autres ports (qu'un PC soit ou non connecté). Cela signifie que si deux PC émettent en même temps, la somme des deux signaux générés excédera la limite permise par la norme, ce qui correspondra à une collision. Un segment délimite donc un **domaine de collision**.

QU'EST CE QUE SNMP ?

Le protocole SNMP (*Simple Network Management Program*) permet d'interroger, de configurer et de surveiller à distance un équipement réseau. Un logiciel serveur, appelé **agent SNMP**, est implanté dans l'équipement à gérer, par exemple un concentrateur. Cet agent répond aux requêtes de clients situés dans les **stations d'administration**.

Dans notre cas, cette fonction n'est pas intéressante, car il existe un autre moyen de chaîner les concentrateurs. Les concentrateurs sont, en effet, couramment équipés d'un port « *uplink* » de type RJ45 et/ou AUI et/ou BNC (attention au choix !) qui permet de chaîner les concentrateurs.

Figure 1-7.
Chaînage des concentrateurs.



Le bus externe permet à une pile de concentrateurs d'être vue comme étant un seul et unique élément, ce qui peut être utile pour l'administration à distance *via* SNMP.

Les *stackables* offrent également d'autres fonctions, telles que la segmentation port par port ou par groupes de ports. Un réseau Ethernet est constitué d'un segment matérialisé par le câble du bus ou le concentrateur de l'étoile. Un concentrateur segmentable permet de créer plusieurs segments Ethernet indépendants : un logiciel interne permet d'affecter un port (parfois cette action n'est possible que par groupes de 4 ou 8 ports) au segment Ethernet 0, et l'autre au segment Ethernet 1. Aucun trafic ne passe entre les deux réseaux ; les PC situés sur des segments différents ne peuvent donc pas communiquer entre eux.

L'intérêt de la segmentation est de créer des réseaux protégés (un pour la comptabilité séparé des autres, par exemple) ou de pallier un problème de charge : s'il y a trop de trafic sur un segment, il est possible de segmenter le réseau en répartissant les PC de part et d'autre en fonction de leur besoin de communication.

Fonctionnalité	Description	Intérêt	Coût pour 8 ports en € HT
Intelligent	Partitionnement des ports	Limite la portée d'un problème	De 75 à 150
Administrable	Gestion à distance *via* SNMP	Intéressant pour les grands réseaux	De 150 à 300
Empilable (+ administrable)	Chaînage *via* un bus propriétaire	Traité comme étant un seul *hub* administrable	De 300 à 460
Autres fonctions justifiant les différences de prix			
Port uplink	Chaînage *via* un port dédié	Augmentation du nombre de ports	± 30
Segmentation	Répartition de la charge sur plusieurs segments	Souplesse d'évolution	± 150
Un ou deux slots d'extension	Ajout de ports en fibres ou autre	Souplesse d'évolution	± 150

Pour notre premier réseau, un concentrateur dépourvu de fonction spécifique (c'est-à-dire non empilable, non administrable, sans segmentation, etc.) convient parfaitement. Il couvre tout à fait les besoins d'un particulier, d'une association, d'un cabinet de profession libérale, etc., c'est-à-dire tous les cas où vous êtes certain que le nombre de PC ne dépassera jamais 3, 16 ou 32 postes. Il suffit d'acheter le *hub* qui offre la bonne modularité.

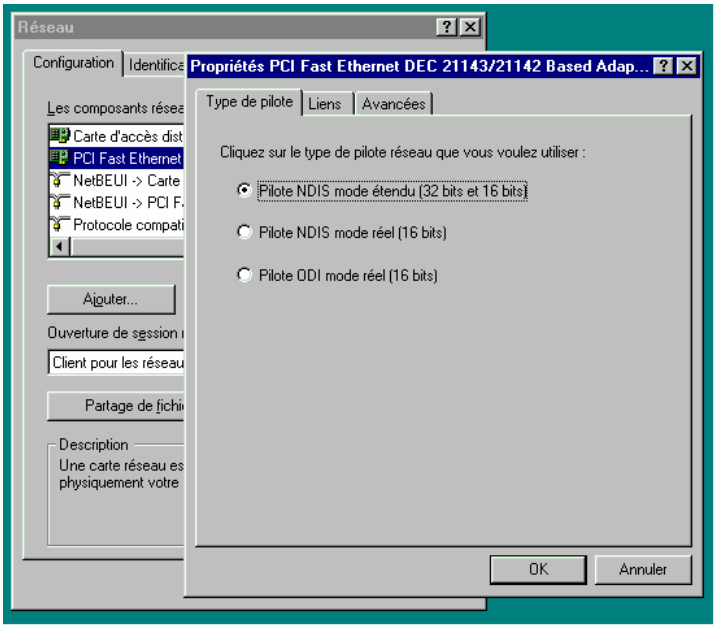
De logiciels de communication

Tout réseau nécessite du matériel et des logiciels, à l'instar d'une connexion à l'Internet qui requiert un *driver* (pour piloter la carte réseau), une pile TCP/IP et au moins un navigateur.

Le pilote est fourni par le constructeur de la carte. Cependant, si ce dernier a passé des accords avec Microsoft, il sera fourni en standard avec Windows. C'est le cas, par exemple, des cartes 3com, Dlink, HP, etc.

La pile TCP/IP est la même que celle utilisée avec l'Internet. En effet, n'importe quel type de carte peut être utilisé avec différentes piles TCP/IP du marché (celles de Windows — Netmanage — de FTP software — WRQ, etc.). Cela est rendu possible grâce à une interface d'accès standardisée sous Windows, appelée NDIS (*Network Driver Interface Specification*).

Par exemple, lors de l'installation de l'accès distant, Windows a installé un driver NDIS pour votre modem — driver fourni par le constructeur ou livré en standard avec Windows — qui dispose d'une interface NDIS. Vous pouvez le vérifier en allant dans le menu « Démarrer→Paramètres→Panneau de configuration→Réseau » qui affiche l'écran suivant (Windows 95).



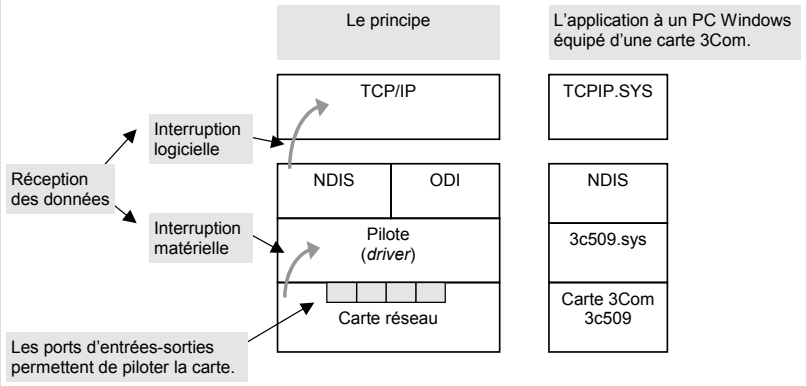
Ainsi, TCP/IP utilise-t-il les mêmes commandes NDIS, quel que soit le périphérique à piloter (une carte réseau, un modem, etc.), grâce à un driver propre à chaque matériel mais qui respecte la même interface logicielle.

LE POINT SUR LES DRIVERS

Le pilotage de chaque carte varie d'un constructeur à l'autre. Il n'existe pas, comme c'est le cas des PC, de standard relatif à la compatibilité matérielle des cartes (il n'y a pas d'« Intel Inside » !). Pour cette raison, chaque carte nécessite un pilote adapté.

À un moment ou à un autre, il faut quand même respecter un standard pour que la même pile TCP/IP puisse dialoguer avec n'importe quel pilote. La standardisation est réalisée au niveau de l'interface d'accès au pilote : la couche TCP/IP donne ainsi des ordres au pilote (du type « envoie une trame ») *via* une interface unique.

Dans le monde des PC, ces interfaces sont appelées **NDIS** (*Network Driver Interface Specification*) par Microsoft et **ODI** (*Open Data-link Interface*) par Novell. Ces deux standards étant bien sûr incompatibles, les cartes sont donc livrées avec deux versions du même driver : l'un avec une interface NDIS, l'autre avec une interface ODI .



Une fois l'interface d'accès au pilote standardisée, n'importe quel logiciel réseau peut être implanté. Le principe consiste à ouvrir un lien au moyen d'interruptions logicielles, de mémoires partagées, de descripteurs de tampons, etc. L'opération s'appelle ***bind*** (liaison), et le lien un **SAP** (*Service Access Point*).

Au-dessus de TCP/IP, on retrouve des applications diverses, comme notre navigateur Internet qui pourra être utilisé sur notre réseau local, que l'on appellera alors intranet. D'autres applications sont possibles en local, à commencer par le partage des fichiers et l'impression. Dans l'environnement Windows, cela est réalisé par un protocole appelé Netbios (*Network Basic Input Ouput System*) qui fonctionne au-dessus de TCP/IP. On ne retrouve pas Netbios sur l'Internet. En effet, d'une part il est spécifique à Microsoft, d'autre part il n'est pas adap-

té à ce type de réseau. Netbios est avant tout conçu pour le réseau local et devra être installé sur le PC en même temps que TCP/IP.

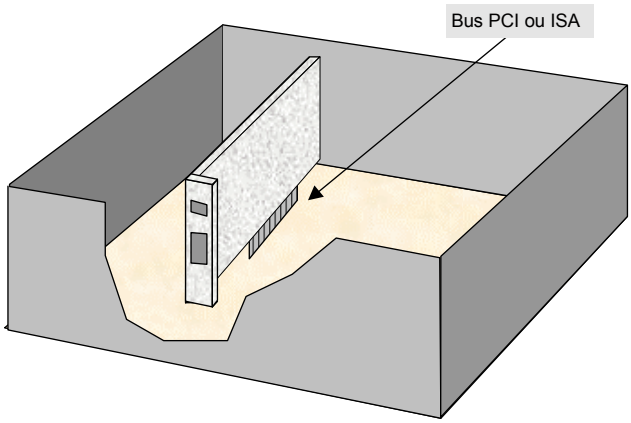
Comment faire fonctionner tout cela ?

Maintenant que vous avez acheté les cartes réseau (une par PC à connecter), un concentrateur (ou plusieurs !) ainsi que les câbles, il ne reste plus qu'à assembler tout cela.

Installer les cartes réseau et les drivers

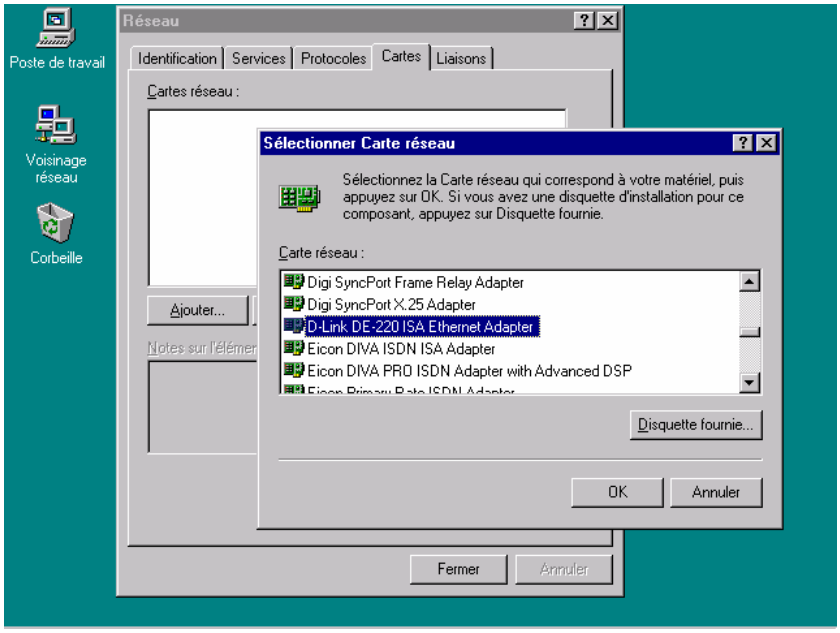
La première chose à faire est d'installer les cartes sur chaque PC. La procédure est standard, mais certaines subtilités — telles que le positionnement de cavaliers (*jumpers*) ou de commutateurs (*switches*) — sont à prendre en compte. Généralement, il n'y a rien à configurer avec Windows 9.x ; ce dernier reconnaît automatiquement la carte et la configure avec les bons paramètres. La documentation livrée avec la carte indique la procédure à suivre.

Figure 1-8.
Insertion d'une carte réseau dans un PC.



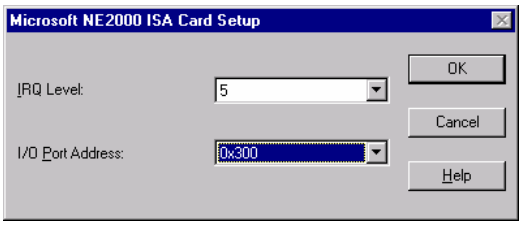
L'étape suivante consiste à installer le driver de la carte.

Si celle-ci n'est pas détectée par la fonction *Plug and Play* de Windows, vous pouvez lancer la procédure en cliquant sur le menu « Démarrer→Paramètres→Panneau de configuration→Réseau ». Vous obtenez alors l'écran présenté à la figure 1-9 (quasi identique sous Windows NT et Windows 9x).



Le programme propose de choisir le driver dans une liste. S'il n'y figure pas, cliquez sur « Disquette fournie… ». Cliquez ensuite sur « OK ».

Selon les cartes, une boîte de dialogue vous demande de spécifier la valeur de certains paramètres.



Ces paramètres concernent :

- les interruptions matérielles, appelées IRQ (*Interruption Request*), utilisées par la carte pour avertir le driver qu'une trame vient d'arriver, par exemple ;
- les ports d'entrée-sortie (*I/O port*, *Input Output*) qui correspondent à des registres (mémoire partagée par la carte et le PC) permettant au driver d'envoyer des commandes à la carte (envoyer une trame, par exemple) ;
- le port DMA (*Direct Memory Access*) si la carte utilise ce mode.

La documentation vous indique la marche à suivre. Généralement, les valeurs par défaut conviennent. Elles doivent être modifiées seulement si vous possédez d'autres cartes qui utilisent les mêmes IRQ et/ou ports I/O.

Cliquez sur « OK » pour terminer l'opération. Le PC affiche alors une série d'écrans et vous demande de réinitialiser l'ordinateur.

Configurer les adresses IP

À la différence de la connexion Internet, il n'existe pas d'ISP pour attribuer automatiquement des adresses IP. À présent, vous êtes chez vous, sur votre réseau, et vous êtes seul maître à bord.

Il faut donc affecter vous-même une adresse IP à chaque poste de travail afin qu'il puisse être identifié de manière unique.

À QUOI SERT L'ADRESSAGE ?

Comme pour le courrier postal, l'adresse permet d'acheminer les trames Ethernet et les paquets IP.

Les réseaux Ethernet utilisent un **adressage plat** : les cartes réseau sont identifiées par une adresse unique, l'adresse MAC (de niveau 2).

Le protocole IP utilise, quant à lui, un **adressage hiérarchique** (de niveau 3) structuré en un **numéro de réseau** et un **numéro de station** au sein de ce réseau (32 bits en tout). L'adresse IP est indépendante de l'adresse MAC : un segment Ethernet peut comprendre plusieurs réseaux IP, et inversement.

Aussi bien au niveau MAC que IP, il existe trois types d'adresses :
- l'adresse **unicast** qui est affectée à une station ;
- l'adresse **multicast** qui désigne un groupe de stations ;
- l'adresse de **broadcast** qui désigne toutes les stations sur un réseau.

Une station est configurée avec une adresse MAC (celle de la carte réseau) et une adresse IP (celle de la pile IP). Des mécanismes spécifiques permettent de réaliser automatiquement la correspondance entre les deux types d'adresses.

Chaque trame Ethernet contient l'adresse MAC de l'émetteur et celle du destinataire. De même, chaque paquet contient les adresses IP de l'émetteur et du destinataire, ce qui permet de les acheminer indépendamment les uns des autres.

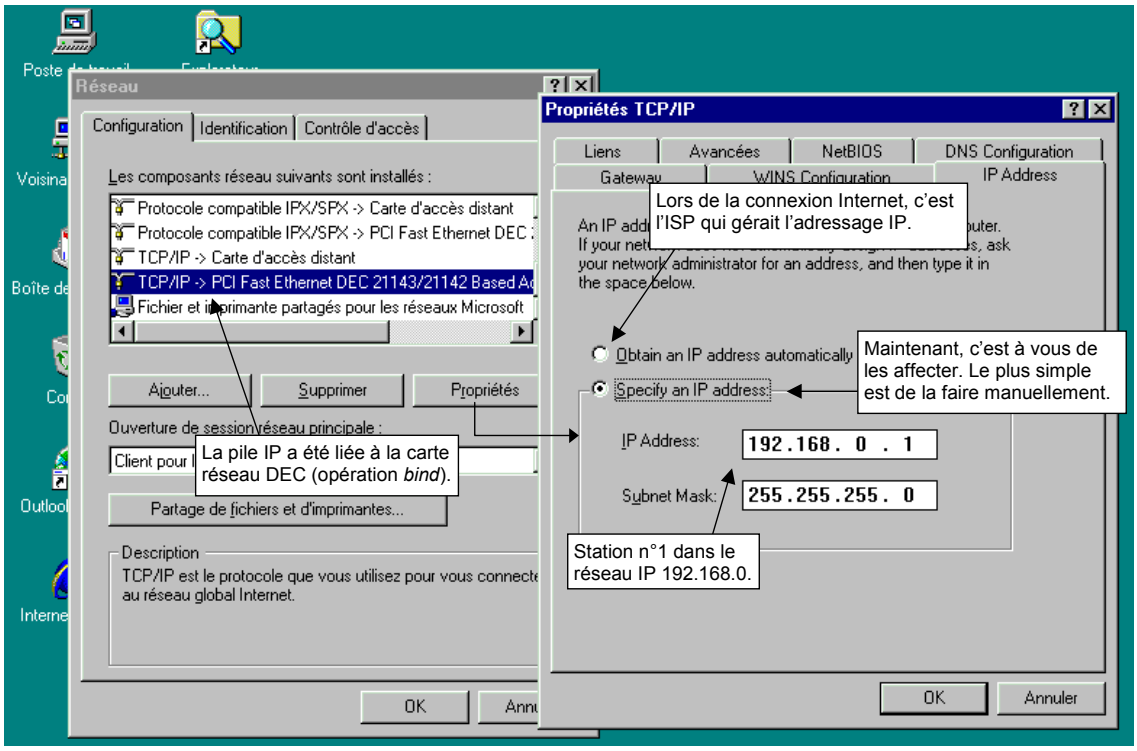
Pour reprendre l'analogie avec les adresses postales, une adresse IP est composée d'un numéro de réseau (le nom d'une rue) et d'un numéro de station au sein de ce réseau (le numéro de votre maison).

Par convention, l'adresse IP s'écrit avec quatre numéros, de 1 à 255, séparés par des points, par exemple 192.162.0.1. Une partie de cette adresse désigne un réseau, l'autre le numéro de

station au sein de ce réseau. Le protocole IP utilise un masque pour distinguer les deux parties. Dans cet exemple, il sera égal à 255.255.255.0, indiquant que les trois premiers chiffres de l'adresse désignent le numéro de réseau, et le dernier celui de la station.

Dans notre cas, il faut s'arranger pour configurer toutes nos stations dans le même réseau logique IP. Nous choisirons donc le réseau 192.168.0 et affecterons à nos PC les numéros compris entre 1 et 254, ce qui donne une plage d'adresses comprise entre 192.168.0.1 et 192.168.0.254.

Sur chaque PC (Windows 9.x), il faut donc aller dans le menu « Démarrer→Panneau de configuration→Réseau » pour configurer ces adresses. Vous obtenez alors l'écran illustré sur la figure ci-après.



Pour l'instant nous n'avons pas besoin d'en savoir plus, car nous avons créé un petit réseau. Le chapitre 5 présente, dans le détail, tous les mécanismes de l'adressage.

Installer les concentrateurs et y raccorder les PC

L'installation des concentrateurs est simple, puisqu'il n'y a aucun paramètre à configurer, ni logiciel à installer. Il suffit de les brancher sur une prise électrique et d'appuyer sur l'interrupteur. Si rien ne se produit, le matériel est en panne ☹.

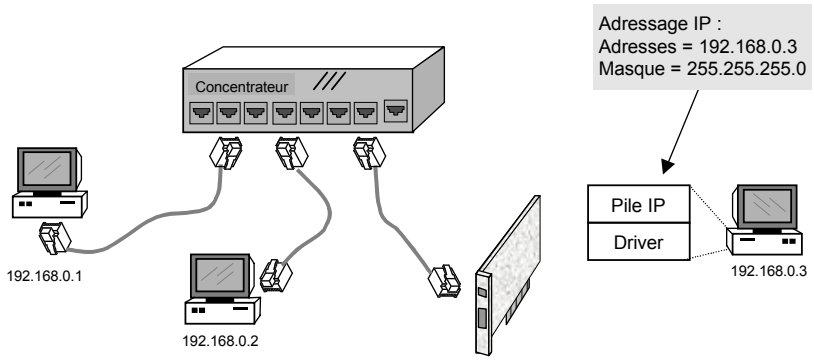
Quelques précautions doivent cependant être prises :

- Utilisez une prise électrique protégée par un disjoncteur dédié aux équipements informatiques afin d'éviter tout parasite provenant d'un autre appareil électrique (cafetière, aspirateur, etc.).
- Placez le concentrateur en hauteur, dans un endroit aéré et éloigné de toute source électrique importante (moteur d'ascenseur, encore la cafetière, etc.).

S'il est administrable, le concentrateur pourra être configuré ultérieurement pour des fonctions spécifiques liées à l'administration SNMP et à la création de segments.

La connexion des PC est également simple : il suffit de raccorder un cordon de brassage au PC et de choisir, au hasard, un des ports du concentrateur.

Figure 1-9.
Un réseau local Ethernet simple.



Comme vous le voyez, ce type d'installation convient à un faible nombre de PC, de préférence regroupés dans un bureau. Rapidement, il devient nécessaire d'organiser le câblage et la mise en place des concentrateurs d'une autre manière. C'est ce que nous allons voir au chapitre suivant.

2

Mettre en place un système de câblage

Le câblage volant, tel qu'il a été installé pour notre petit réseau lors du chapitre précédent, ne peut pas être généralisé à grande échelle. En effet, au delà de dix postes, il devient rapidement source de problèmes : posé à même le sol, il est encombrant, voire gênant (on se prend les pied dedans). Il est de ce fait soumis à une usure plus rapide.

Une première amélioration consiste à le faire circuler dans des parties protégées de l'immeuble : on peut le poser sous un faux plafond ou sous un faux plancher ou encore le faire passer dans une goulotte le long des murs.

Mais, à chaque nouvelle connexion ou à chaque déménagement de PC, il faut déplacer le câble et trouver un nouveau cheminement, ce qui présente des inconvénients majeurs :

- Cette opération est extrêmement difficile, voire impossible si la longueur des câbles est de plusieurs dizaines de mètres.
- Le problème d'usure demeure lorsque les câbles sont déplacés.
- Le cheminement des câbles est difficile à maîtriser : on arrive inévitablement à des situations dans lesquelles les câbles informatiques s'entrecroisent avec les câbles électriques qui sont sources de perturbations importantes. Le réseau peut ne plus fonctionner à cause de cela.

Il est donc impératif de mettre en place un système de câblage permanent (fixe et stable dans le temps) et évolutif (qui s'adapte à tous les besoins présents et futurs). Pour cela, il convient de respecter un certain nombre de règles.

Quelle est la démarche à suivre ?

Le câblage d'un immeuble requiert un certain nombre d'étapes importantes et étalées dans le temps :

- Si l'immeuble existe, un **audit** préalable est nécessaire afin de repérer les locaux, les sources de courants forts, les câbles existants, les cheminements possibles des futurs câbles, etc.
- La phase d'étude et d'expression des besoins, généralement appelée **APS** (avant-projet sommaire), a pour but de déterminer les spécifications fonctionnelles de l'infrastructure (locaux, gaines techniques) et du système de câblage (cuivre, fibre optique, connectique).
- Pour les grandes réalisations, l'APS n'est que l'ébauche de plusieurs scénarios. L'**APD** (avant-projet détaillé) permet alors de choisir la solution en fonction de critères techniques organisationnels et économiques.
- Le dossier de consultation (le cahier des charges) peut être formé de trois documents principaux : le **CCTP** (cahier des clauses techniques particulières), puis éventuellement le **CCTG** (cahier des clauses techniques générales) et, si vous travaillez avec l'administration française, le **CCAP** (cahier des clauses administratives particulières). Cette phase se termine par la sélection d'une entreprise de câblage.
- La phase de **suivi de chantier** nécessite un contrôle régulier et des réunions de coordination.
- La phase de réception (**recette**) consiste à tester et à valider les travaux effectués.

La première tâche est avant tout de repérer les lieux, ou de se contenter d'examiner les plans si l'immeuble n'existe pas encore.

Dans les deux cas, l'objectif est de mettre en place un câblage **systématique**, c'est-à-dire d'équiper entièrement l'immeuble. Si seuls quelques étages sont concernés, la démarche est plus ou moins la même.

Il ne s'agit donc pas de savoir où sera situé tel ou tel utilisateur, mais d'installer des prises partout dans le but de connecter n'importe qui à n'importe quelle prise pour n'importe quel type d'application. On parlera alors d'un **précâblage** multimédia ou **VDI** (voix, données, image).

L'avant-projet

Lors d'une opération de précâblage, il est important de systématiser l'implantation des prises dans tout l'immeuble. Une fois le chantier achevé, tout aménagement complémentaire sera plus délicat, plus long et plus coûteux. Le chantier de câblage est l'occasion unique de réaliser une fois pour toute une infrastructure sans avoir à y revenir avant dix ou quinze ans.

La densité communément admise est d'environ un boîtier VDI pour 7 à 10 m^2 de bureaux, un boîtier pouvant regrouper de deux à quatre prises. Cette densité peut être plus élevée pour certaines applications spécifiques comme les salles de marché : on peut trouver jusqu'à dix prises par position (occupant 3 m^2 environ).

On peut prendre comme repère une travée délimitée par une largeur de fenêtre. Selon les besoins, on pourra installer un boîtier VDI de deux à quatre prises par travées. Généralement, il faut une prise pour le poste de travail informatique, une autre pour le téléphone, et une troisième pour un besoin particulier (une ligne téléphonique directe, une imprimante en réseau, etc.).

Il est important de noter que le boîtier VDI doit se trouver à proximité d'un bloc de prises électriques : cela paraît une évidence, mais il faut penser à se coordonner avec l'entreprise qui réalise les travaux courants forts.

Les locaux concernés sont non seulement les bureaux, mais aussi les locaux collectifs : local photocopieur, cafétéria (on y pose souvent des bornes d'information), salles de réunions, salles de conférences, halls d'entrée (pour les bureaux d'accueil, les locaux des gardiens, etc.).

En outre, il faut aussi prévoir le câblage pour la GTB (gestion technique du bâtiment), bien que celui-ci soit souvent réalisé par une entreprise spécialisée avec laquelle il faudra de toute façon se coordonner. La GTB regroupe des besoins comme la détection incendie, les alarmes, la sécurité d'accès aux locaux, la surveillance, etc.

De même, il y a toute une série d'équipements annexes qui peuvent requérir l'emploi d'une prise :

- les téléphones d'ascenseurs ;
- les lignes directes (celles qui ne passent pas par le PABX) ;
- les bornes de réseau sans fil et de téléphonie sans fil (DECT) ;
- les badgeuses ou pointeuses ;
- etc.

Une fois l'implantation des prises définie, il faut prévoir de la place pour le cheminement des câbles et la création des locaux techniques. Le principe retenu est quasi systématiquement une topologie en étoile : les câbles relient les prises VDI à d'autres prises en local technique. Les différentes normes définissent une longueur maximale de quatre-vingt-dix mètres pour les câbles en cuivre.

De ce fait, il faut prévoir plusieurs locaux techniques au sein de l'immeuble et donc des câbles pour les relier entre eux. Plusieurs facteurs déterminent le nombre et la position des locaux techniques :

- La distance maximale qui est de quatre-vingt-dix mètres.
- La densité des prises : on admet qu'un local peut centraliser jusqu'à 250-350 prises ; au-delà, son exploitation devient complexe (trop de câbles, trop grande concentration d'équipements).
- L'architecture des réseaux informatiques et téléphoniques : de nos jours, ils reposent sur une topologie en étoile avec des équipements installés à chaque étage et d'autres qui ont une fonction fédératrice.

L'INFRASTRUCTURE NÉCESSAIRE À UN SYSTÈME DE CÂBLAGE

Au sein d'un immeuble, de l'espace doit être réservé pour accueillir le système de câblage. Il s'agit essentiellement de **locaux techniques** et de cheminements utilisés pour relier les locaux entre eux.

Les câbles qui relient les prises VDI aux locaux techniques sont appelés **câbles de distribution**. Ceux qui relient les locaux techniques entre eux sont appelés **câbles de rocade**.

Dans les zones de circulation (couloirs, halls d'entrée, etc.), les câbles sont installés dans des **chemins de câbles** métalliques qui servent de support et offrent une protection mécanique et électromagnétique. Dans les bureaux, ces mêmes câbles sont installés dans des **goulottes** ou des tubes noyés dans le béton.

Les câbles de distribution sont généralement horizontaux et cheminent sous les **faux plafonds** et/ou sous les **faux planchers**. Ces derniers ont une fonction essentiellement esthétique et sont constitués de dalles amovibles destinées à en faciliter l'accès.

Les rocades sont verticales ou horizontales, et cheminent sous les faux plafonds, les faux planchers et dans les **gaines techniques** (conduits réservés aux câbles et tuyaux de toute nature).

L'architecture d'un système de câblage suit donc celle des réseaux : on définit ainsi deux niveaux de locaux techniques :

- Les **LTE** (locaux techniques d'étages) qui concentrent les prises VDI et accueillent les équipements de communication de distribution (concentrateurs, commutateurs, etc.).
- Les **LN** (locaux nodaux) qui sont reliés à tous les locaux techniques et accueillent les équipements de communication fédérateurs (PABX, commutateurs fédérateurs, routeurs, etc.).

Un troisième niveau de concentration est parfois nécessaire dans le cas où les distances sur un étage excèdent quatre-vingt-dix mètres. On peut alors trouver la dénomination de LTR (local technique rapproché) ou de LTP (local technique de proximité). Ce type d'architecture n'est cependant pas conseillé, car trop complexe et mal adapté aux architectures réseaux.

Tous ces locaux sont autant d'espaces prélevés sur la superficie utile de l'immeuble. Préparez-vous donc à quelques négociations avec l'architecte (si l'immeuble est à construire) ou avec le responsable des services généraux (s'il existe déjà). Le tableau suivant donne une idée de la surface à réserver à ces locaux techniques.

Local	Superficie moyenne
Local technique d'étage	6 m^2 (3 m x 2 m)
Local nodal	24 m^2 (6 m x 4 m)
Local énergie	8 m^2 (4 m x 2 m)
Local opérateur	9 m^2 (3 m x 3 m)
Salle informatique	28 m^2 (7 m x 4 m)

Bien entendu, ces superficies doivent être ajustées en fonction du nombre de prises à câbler.

LES DIFFÉRENTS TYPES DE LOCAUX TECHNIQUES

Les LTE (**locaux techniques d'étages**) accueillent une à deux baies de câblage (distribution d'étage et rocades) ainsi qu'une à deux baies de communication (équipements de distribution des réseaux téléphonique et informatique). Généralement, un ou deux LTE par étage sont suffisants. Pour des raisons de simplicité, il faut s'arranger pour que les LTE soient tous à l'aplomb les uns des autres.

Le LN (**local nodal**) accueille les baies de câblage (distribution des serveurs et rocades) ainsi que les baies de communication (équipements centraux pour les réseaux téléphonique et informatique). Généralement, il y a deux locaux nodaux dans le bâtiment afin d'offrir une redondance pour le cheminement des câbles. Chaque LTE peut ainsi être relié aux LN *via* deux chemins de câbles différents. Il en est de même entre deux LTE d'un même étage.

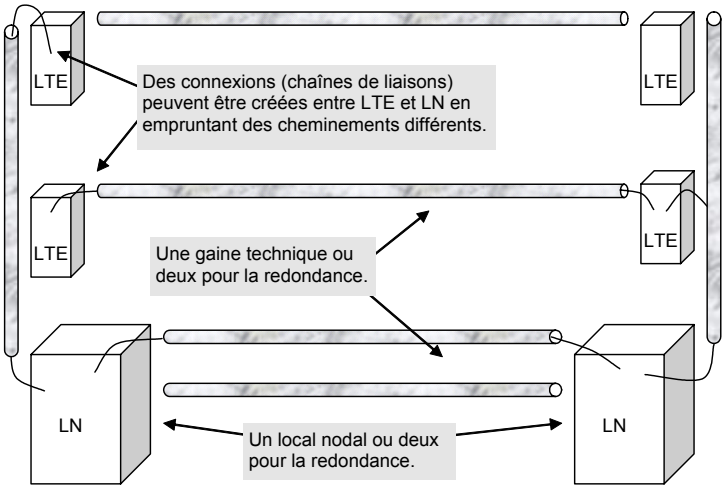
Il est généralement souhaitable de créer un LO (**local opérateur**) réservé aux arrivées télécom des opérateurs afin d'assurer une séparation claire des responsabilités entre lui et le client. Deux LO sont préférables, car la plupart des opérateurs peuvent offrir deux accès physiquement séparés et redondants. Ils doivent jouxter le ou les LN qui hébergent les équipements de communication du client (routeurs, par exemple).

Il faut aussi prévoir un LE (**local énergie**) pour accueillir l'alimentation du PABX (armoire 48v et batteries) ainsi qu'un régulateur de courant/onduleur. Il doit jouxter un local nodal. Là encore, deux locaux énergie offrent un bon niveau de redondance renforcé par deux accès EDF physiquement différents.

Une SI (**salle informatique**) accueille une à deux baies de câblage (distribution) ainsi que des serveurs informatiques. Elle doit de préférence être dédiée afin de mieux contrôler l'accès aux locaux et de séparer les responsabilités entre les équipes système et réseau. De la même manière, deux salles informatiques permettent de limiter les dégâts en cas de sinistre, et de répartir les serveurs en cluster.

Pour les petits sites, il est plus économique et plus simple de regrouper les fonctions de LN, LO et SI au sein d'un même local.

Figure 2-1.
Gaines et locaux techniques pour un système de câblage.



LES COMPOSANTS D'UN SYSTÈME DE CÂBLAGE

Les parties visibles d'un système de câblage sont les **prises utilisateur**, également appelées prises VDI (voix, données, images), installées dans les bureaux. Elles sont regroupées par blocs de 2 à 4, appelés **boîtiers VDI**. Une densité courante est d'un bloc VDI pour 7 à 10 m^2 de bureau.

Les prises utilisateur sont reliées en étoile à un local technique par l'intermédiaire d'un **câble** (en cuivre ou en fibre optique). Le local technique concentre 100 à 350 **câbles de distribution**, chacun se terminant par une prise identique à celle installée du côté utilisateur. Ces **prises de distribution** sont regroupées dans des **panneaux de brassage** fixés dans des **baies.**

Les prises sont reliées aux équipements informatiques et téléphoniques par l'intermédiaire de **cordons de brassage** de même nature que les câbles.

Les prises, câbles, cordons et panneaux de brassage doivent tous être issus du même constructeur afin de bénéficier de sa garantie (généralement dix à quinze ans).

Ensuite, il faut évaluer la puissance électrique consommée par les équipements informatiques. On pourra même prévoir des disjoncteurs séparés, un par baie ou pour un groupe d'équipements.

Il faut enfin prévoir une climatisation dans chaque local technique, et donc évaluer la dissipation calorifique des équipements (exprimée en Watts ou en BTU – *British Thermal Unit*). Ces valeurs sont données par les constructeurs de tout équipement informatique.

Tous ces besoins seront regroupés dans un document appelé **APS** (avant-projet sommaire) et communiqués aux corps d'état concernés (architecte, électricien, société de climatisation, etc.).

L'étude d'ingénierie

La phase d'expression des besoins est suivie d'une étude permettant d'arrêter un certain nombre de choix importants :

- Quel type de câble utiliser ?
- Quel type de prise choisir en bureau ? En local technique ?
- Où faire passer les câbles ? Comment placer les prises ?
- Où positionner les locaux techniques ? Comment les aménager ?

Quel type de câble ?

Éternel débat que celui du choix des câbles, chaque constructeur ayant des arguments en faveur de son produit. De nombreuses combinaisons techniques viennent compliquer le choix. Pour résoudre ce dilemme, un certain nombre de questions sont à se poser, et dans le bon ordre.

Cuivre ou fibre optique ?

Les réseaux Ethernet fonctionnent sur cuivre à 10 Mbit/s et à 1 Gigabit. L'avantage de la fibre optique est qu'elle permet de s'affranchir des contraintes de distance (plusieurs centaines de mètres au minimum contre quatre-vingt-dix mètres pour le cuivre). Cela tient à l'atténuation du signal, beaucoup plus important sur un câble en cuivre.

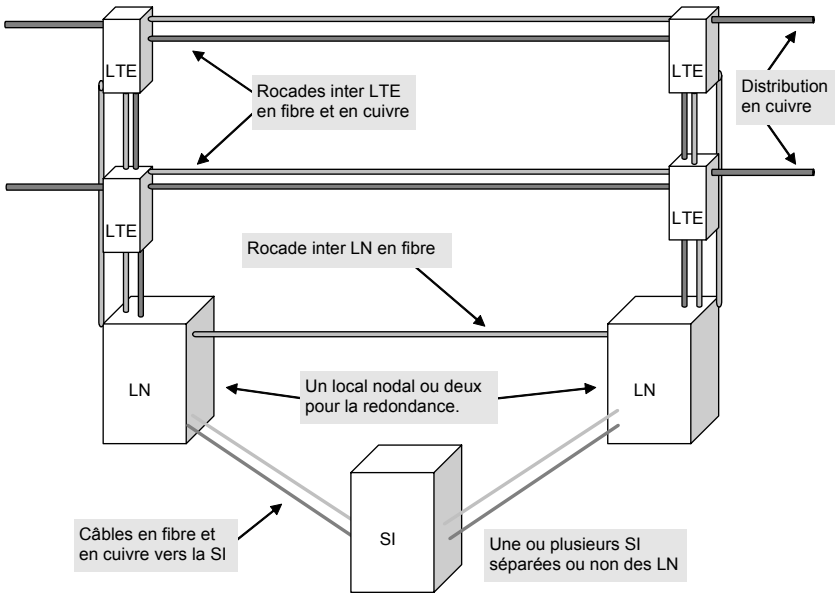
En revanche, le coût global d'un système de câblage en fibre optique est plus élevé que l'équivalent en cuivre. En effet, l'ingénierie nécessaire pour poser des câbles optiques (raccordement des connecteurs et tests) est plus complexe et plus coûteuse qu'avec des câbles en cuivre. De plus, les composants tels que les connecteurs SC et les tiroirs optiques sont également beaucoup plus chers que les prises RJ45 et les panneaux de brassage.

À titre d'indication, un système de câblage en fibre optique coûte en moyenne 60 % plus cher que l'équivalent en câble de cuivre SFTP catégorie 5E.

Il faut ajouter à cela le coût des équipements actifs (les commutateurs et cartes Ethernet), deux fois plus chers en version fibre optique, et pour une densité de ports deux fois moins élevée que leur équivalent en cuivre.

En conclusion, le câble cuivre sera privilégié pour la distribution, et la fibre optique pour la connexion entre les locaux techniques. Cette répartition des rôles offre, en outre, plus de souplesse pour positionner les LTE qui doivent être à moins de quatre-vingt-dix mètres de toutes les prises qu'ils irriguent.

Figure 2-2.
Architecture de câblage type.



Pour ajouter plus de sécurité, on peut envisager de doubler les liaisons entre les LTE d'un même étage ainsi qu'entre les LN, chacune d'entre elles passant alors par deux gaines techniques différentes.

Attention, cependant, les installations téléphoniques classiques requièrent encore des connexions en cuivre entre les postes et le PABX central. Les contraintes de distance étant moins fortes (quelques centaines de mètres), il faut donc envisager du câble en cuivre, dit multipaire, entre les LTE et les locaux nodaux.

De nos jours, on privilégiera une architecture téléphonique identique à celle du réseau local avec des unités déportées dans chaque étage (de type Voice Hub, tels que proposés par Alcatel) et raccordées en fibre optique à un petit PABX central. Si vous optez pour cette solution, vous n'avez plus besoin de câbles multipaires entre le local nodal et les LTE.

Coaxial ou paires torsadées ?

Nous l'avons vu au chapitre précédent, le câble coaxial (50 et 75 Ohms) n'est plus utilisé pour les réseaux locaux. Il pourra cependant être posé pour les besoins spécifiques de la vidéo (voir plus loin).

En revanche, la paire torsadée est le standard pour l'informatique et la téléphonie ; elle peut également être utilisée pour la distribution vidéo. Tous ces équipements (concentrateurs, commutateurs, PABX, etc.) sont, en effet, équipés de prises RJ45.

Le choix de la paire torsadée en distribution

Quelle impédance : 100, 120 ou 150 Ohms ?

Le 150 Ohms n'est qu'un artefact des réseaux Token-Ring IBM qui ne s'est jamais imposé car trop coûteux. Le 120 Ohms se voulait un compromis entre coût et performances entre le 100 et le 150 Ohms, mais ne s'est imposé qu'en France. Le 100 Ohms est le plus répandu, car il est moins cher et est soutenu par les Américains, ATT en tête.

De plus, tous les équipements informatiques sont américains et donc pourvus de connecteurs RJ45 de 100 Ohms. Cependant, un câble 120 Ohms peut y être connecté sans problème, l'affaiblissement, résultant de l'adaptation d'impédance, étant largement compensé par les meilleures performances du câble 120 Ohms.

En conclusion, les 100 et 120 Ohms conviennent tous deux, avec un avantage pour le premier qui est meilleur marché.

Écranté ou non ? Blindé ou non ?

Tout câble informatique véhiculant un courant alternatif, dit courant faible, génère un champ électromagnétique. Inversement, toute source électrique (câble, moteur d'ascenseur, néon à ballast électromagnétique, etc.) véhiculant un courant alternatif, dit courant fort, génère un champ électromagnétique pouvant perturber les câbles informatiques.

La première protection consiste à torsader les paires, de manière à balancer les signaux électriques. Une protection supplémentaire peut être offerte soit par un écran consistant en une

feuille d'aluminium entourant toutes les paires ou chaque paire, soit par un blindage consistant en une tresse métallique, soit par les deux (blindage collectif et écran par paire).

Il faut cependant noter que, d'une part les techniques de codage utilisées par le Fast Ethernet et le Gigabit rendent ces protocoles plus résistants aux perturbations, et d'autre part que la qualité des torsades confère aux câbles des performances conformes aux catégories 6 et 7. Ces deux éléments militent en faveur de l'UTP, câble non écranté et non blindé, qui revient légèrement moins cher que les versions blindées.

Indiquons cependant que le coût du câble ne représente qu'une faible part (environ 20 %) du coût total d'un chantier de câblage, la plus grosse part étant celle de la main d'œuvre pour la pose. Mais, les câbles écrantés et/ou blindés nécessitent la mise en place d'une terre informatique qui grève le budget de petites réalisations.

En conclusion, le câble UTP convient pour des réalisations de petite et grande taille ; le câble STP (avec un écran collectif) est conseillé pour bâtir un réseau évolutif vers les hauts débits, au-delà du Gigabit ; et le SFTP (avec un écran par paire et un blindage collectif) offre une garantie supplémentaire, mais non nécessaire, si vous disposez du budget requis.

Catégories 5, 6 ou 7 ?

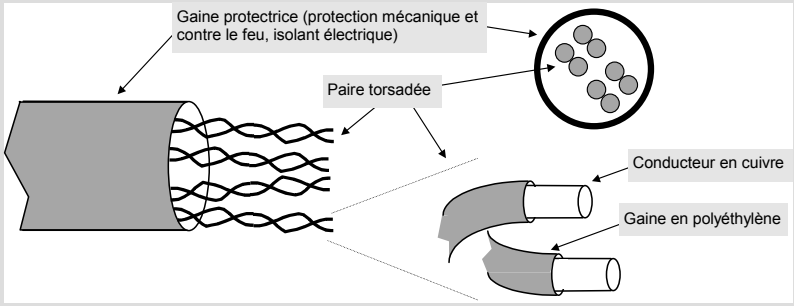
Actuellement, la catégorie 5 (100 MHz de fréquence maximale) est la plus répandue et supporte le Gigabit Ethernet sur ses quatre paires (250 Mbit/s par paire). La catégorie 6 permet de doubler le débit (250 MHz) et sa normalisation est stable. En revanche, le câble catégorie 7 n'a toujours pas de connecteur normalisé.

Câble	Caractéristiques	État de la norme
Cat 5 / Classe D	100 MHz, RJ45, 10bT, 100bT et Gigabit si testé TSB-95	TIA/EIA-568-A Ratifié en octobre 1995
Cat 5E	100 MHz, RJ45 jusqu'à 1 Gbit/s	TIA/EIA-568-B Ratifié en mai 2001
Cat 6 / Classe E	250 MHz, RJ45 jusqu'à 2,5 Gbit/s au moins	TIA/EIA-568-B.2-1 Ratifié en juin 2002
Cat 7 / Classe F	600 MHz, prise non définie jusqu'à 10 Gbit/s	Spécifications en cours

En conclusion, le câblage catégorie 5, actuellement le plus répandu, supporte le 100bT et même le Gigabit à condition qu'il soit testé selon de nouveaux critères (norme TSB-95), tels que la paradiaphonie cumulée. Pour de nouvelles installations, on préférera cependant les câbles certifiés catégorie 6 supportant d'entrée de jeu le Gigabit.

LE POINT SUR LA PAIRE TORSADÉE (TIA/EIA-568 ET EN50173)

Ce type de câble en cuivre comporte huit fils appariés par deux. Les deux fils de chaque paire sont torsadés selon un **pas de torsade** précis, de même que les quatre paires au sein du câble. Le but de cet arrangement est de limiter les interférences produites par chaque fil sur les autres.

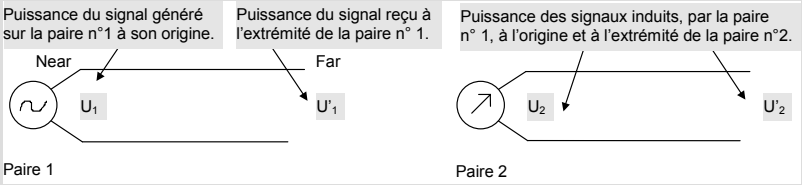


La qualité d'un câble cuivre dépend bien sûr de la qualité du matériau, mais aussi des éléments suivants :

- du **diamètre** des fils, exprimé en AWG (*American Wire Gauge*) :
  22 AWG = 0,63 mm ; 24 AWG = 0,5 mm (le plus courant) ; 26 AWG = 0,4 mm ;
- de l'**impédance** caractéristique, exprimée en Ohm, qui représente la résistance du câble, le 100 Ω étant la plus courante (les 120 et 150 Ω peuvent encore être rencontrés) ;
- de la manière de **torsader** les fils d'une paire et les paires entres elles ;
- de sa **protection** contre les champs électromagnétiques : sans protection (**UTP**, *Unshielded Twisted Pair*) autre que celle offerte par la torsade, avec un écran collectif (**FTP**, *Foilded Twisted Pair*) – consistant en une feuille d'aluminium entourant les paires, avec un blindage collectif (**STP**, *Shielded Twisted Pair*) – consistant en une tresse métallique entourant les paires, ou avec les deux (**SFTP**, *Shielded & Foilded Twisted Pair*) – consistant en un blindage collectif et un écran par paire.

La norme **TIA/EIA-568** précise les paramètres à mesurer pour vérifier la qualité des câbles (TSB 36) et des prises RJ45 (TSB40) :

- l'**atténuation**, exprimé en décibels pour cent mètres, qui mesure l'affaiblissement du signal (U'_1-U_1) ;
- la **paradiaphonie** NEXT (*Near End Cross Talk*), exprimée en décibels, qui mesure la différence entre la quantité de signal U_1 produite par une paire et celle U_2 qui est engendrée sur une autre paire ;
- la **paradiaphonie cumulée** (*Powersum Next*), exprimée en décibels, qui mesure la somme des paradiaphonies engendrées sur une paire par toutes les autres paires ;
- le **rapport signal/bruit** ACR (*Attenuation to Crosstalk Ratio*), exprimé en décibels, qui mesure la différence entre la puissance d'un signal induit U'_2 et celle d'un signal atténué (ACR = NEXT - atténuation).



La meilleure qualité est obtenue pour une faible atténuation ainsi que pour une paradiaphonie et un rapport signal/bruit élevés. L'affaiblissement est proportionnel à la racine carrée de la fréquence.

• • •

LE POINT SUR LA PAIRE TORSADÉE (SUITE)

La norme **TIA/EIA-568** (TSB 95) précise cinq paramètres supplémentaires qui permettent de s'assurer du fonctionnement du Gigabit en full duplex sur les câbles :

- le **taux de réflexion** (*Return loss*), exprimé en décibels, qui mesure l'affaiblissement résultant de l'adaptation d'impédance ;
- la **télédiaphonie** FEXT (*Far-End Crosstalk)*, exprimée en décibels, mesure la différence entre la quantité de signal U_1 produite par une paire et celle U'_2 qui est engendrée sur une autre paire ;
- le **signal de couplage** ELFEXT (*Equal Level Far-End Crosstalk*), qui exprime l'ACR mesuré à l'extrémité du câble (ELFEXT = FEXT – atténuation) ;
- le **signal de couplage cumulé** (*Powersum ELFEXT*), et l'**ACR cumulé** (*Powersum NEXT*), exprimés en décibels.

Quatre **catégories** de câbles sont définies, **5**, **5E**, **6** et **7** (les catégories 1 à 4 ne sont pas utilisées en informatique) selon la **fréquence** maximale du signal pouvant être véhiculé : cat 5 et 5E à 100 MHz, cat 6 à 250 MHz et cat 7 à 600 MHz. Plus la fréquence est élevée, plus le débit du réseau le sera. Par exemple, le 100bT fonctionne à 62,5 MHz sur deux paires, et le Gigabit à 100 MHz sur quatre paires. Les normes précisent les valeurs minimales ou maximales des paramètres pour différentes fréquences de fonctionnement.

La norme européenne **EN50173** reprend le même principe, mais définit les valeurs pour une chaîne de liaison comprenant un câble de 90 mètres et deux cordons de brassages de 5 mètres chacun. On parle alors de **classes D**, **E**, **F** et **G**.

Le tableau suivant présente quelques valeurs de référence spécifiées par la norme TIA/EIA-568-B.2-1 pour la catégorie 6.

Fréquence	Atténuation	Next	PS-Next	ACR	PS-ACR	ELFEXT	PS-ELFEXT	Return Loss
MHz	Max dB/100m	Min dB	Min dB	Min dB/100m	Min dB/100m	Min dB/100m	Min dB/100m	Min dB
10	6,0	59,3	57,3	53,3	51,3	50,0	47,0	25,0
62,5	15,5	47,4	45,4	31,9	29,9	34,1	31,1	21,5
100	19,7	44,3	42,3	24,4	22,4	27,8	27,0	20,1
250	33	38,3	36,3	5,3	3,3	22,0	19,0	17,3

Le tableau suivant présente les valeurs de référence correspondant à la norme CENELEC EN-50173, classe E (câbles, connecteurs et cordons de brassage catégorie 7).

Fréquence	Atténuation	Next	PS-Next	ACR	PS-ACR	ELFEXT	PS-ELFEXT	Return Loss
MHz	Max dB	Min dB	Min dB	Min dB	Min dB	Min dB	Min dB	Min dB
10	6,5	56,6	54,0	50,1	47,5	43,2	40,2	19,0
62,5	16,9	43,4	40,6	26,5	23,7	27,3	24,3	14,1
100	21,7	39,9	37,1	18,2	15,4	23,3	20,2	12,0
250	36,0	33,1	30,2	-2,9	-5,8	17,2	14,2	8,0

Il faut noter que tous les types de câble (UTP, FTP, STP, SFTP, 100 ou 120 Ohms) peuvent être conformes aux normes catégorie 5, 6 ou 7. Les deux ne sont pas liés.

Le choix de la fibre optique entre les locaux techniques

Généralement, le choix de la fibre optique se justifie essentiellement pour des questions de distance, au-delà de la limitation à quatre-vingt-dix mètres de la paire torsadée.

Cependant, un système de câblage en fibre optique coûte quasiment le même prix qu'un équivalent en cuivre catégorie 7. Le choix de ce support peut donc être pris en considération pour la distribution, d'autant plus que la connectique en bureau, de types SC ou MT-RJ, est désormais de bonne qualité.

La fibre optique offre également un gage de pérennité pour le support des hauts débits.

Multimode ou monomode ?

Une fibre monomode offre de meilleures performances mais coûte plus cher que la multimode, de même que les cartes réseau correspondantes. De plus, la multimode convient à presque tous les usages pour la mise en place de réseaux locaux au sein d'un bâtiment ou sur un campus.

Toutefois, si les distances sont réellement importantes, le choix de la monomode s'impose. Elle pourra être envisagée pour connecter d'autres sites à hauts débits dans le cadre d'une boucle optique sur un campus ou avec un opérateur.

62,5/125 ou 50/125 ?

Cette question porte sur les diamètres du cœur et de la gaine optique de la fibre, exprimés en microns. Un cœur de plus petit diamètre affaiblit moins le signal et permet donc de le véhiculer sur de plus grandes distances.

Un câble…	… supporte le Gigabit sur
Cuivre cat 5, 5E, 6, 7	90 mètres
Multimode 62,5/125	300 m à 850 nm 550 m à 1 300 nm
Multimode 50/125	550 m à 850 nm et à 1 300 nm
Monomode	3 km à 1300 nm

Cependant, la fibre optique qui convient à tous les usages actuels et pour laquelle les équipements actifs sont le plus répandus est la multimode 62,5/125. C'est également la moins chère. Tous les réseaux Ethernet, du 10 Mbit/s au Gigabit, fonctionnent avec une longueur d'onde de 850 nm ou, plus rarement, de 1 300 nm.

LA COMPATIBILITÉ ÉLECTROMAGNÉTIQUE (EMC)

Quand un câble est exposé à un champ électromagnétique normal, un **courant** est **induit** sur chacune des paires du câble en cuivre. Sa puissance varie entre **1 et 50 mv** pour un câble UTP catégorie 5, et entre **0 et 0,5 mv** pour un câble FTP. Elle dépend de la qualité du câble (mais pas de sa longueur) ainsi que de la fréquence et de la puissance du signal perturbateur. Les tenants de l'UTP tablent sur le pas de torsade des paires pour limiter les effets électromagnétiques, tandis que les tenants du FTP misent, en plus, sur un écran (une feuille d'aluminium collectif), les STP sur un blindage.

Il est à noter que l'inverse est vrai : le câble ne doit pas rayonner au point de générer des interférences sur les autres équipements électroniques.

Vous pouvez vous-même faire l'expérience de ce phénomène avec le tuiter de votre chaîne hi-fi : même débranché, le câble qui le relie à l'ampli est capable de capter des émissions radio de manière suffisamment puissante pour activer ce petit haut-parleur. Il est même possible d'entendre la radio en branchant un simple écouteur téléphonique au niveau du panneau de brassage !

Jusqu'à présent, ce phénomène ne perturbait pas les réseaux locaux, mais, de nos jours, les fréquences utilisées avoisinent les 100 MHz, ce qui correspond très exactement à la gamme de fréquences des radios FM et des talkies-walkies.

Aujourd'hui, les câbles catégorie 7 sont prévus pour fonctionner jusqu'à 600 MHz, et il est probable que, dans le futur, les fréquences continuent d'augmenter pour avoisiner celles du téléphone DECT (1 800 MHz) et du GSM (900 et 1 800 MHz).

Il est donc important que les câbles aient une bonne performance EMC évaluée en mesurant l'**atténuation de couplage** (AC). Les valeurs précises sont en cours de normalisation :

AC < 40 dB	la qualité EMC est mauvaise
41 < AC < 50	minimum requis pour les câbles UTP catégorie 5 (100 MHz)
51 < AC < 60	minimum requis pour les câbles FTP catégorie 5 (100 MHz)
61 < AC < 70	minimum requis pour les câbles FTP catégorie 6 (250 MHz)
71 < AC < 80	bon câble pour un FTP / câble médiocre pour un SFTP
81 < AC < 90	minimum requis pour les câbles FTP et SFTP catégorie 7 (600 MHz).

Ainsi, le Gigabit Ethernet peut fonctionner sur un câble UTP catégorie 5 en respectant l'EMC uniquement si l'AC est supérieur à 50 dB.

Le câble contenant les fibres

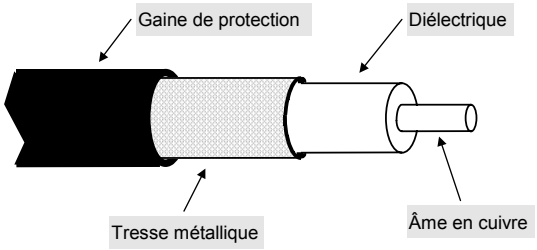
Les qualités d'un câble en fibre optique dépendent de ses caractéristiques optiques mais aussi mécaniques : gaine rigide ou souple, traitée anti-rongeurs et isolations (thermique, incendie, corrosion) qui conditionnent sa duré de vie (dix à quarante ans).

Au sein d'un bâtiment, il conviendra de choisir un câble dit d'intérieur, souple, tandis que, pour les connexions entre bâtiments, on choisira un câble dit d'extérieur, plus rigide, dont la gaine extérieure peut même être métallique.

Le coaxial et la paire torsadée pour la vidéo

Le mode de diffusion le plus répandu pour la vidéo est actuellement le câble coaxial (organisation en bus), car la plupart des équipements sont pourvus de ce type de prise. Un seul câble parcourt alors tout l'immeuble et véhicule plusieurs dizaines de canaux vidéo.

Figure 2-3.
Le câble coaxial.



En revanche, la diffusion sur câbles de cuivre à paires torsadées (organisation en bus-étoile) tend à se généraliser, car elle permet de banaliser le système de câblage et donc de profiter de sa souplesse en termes de reconfiguration et d'évolutivité.

Enfin, la diffusion vidéo sur IP (norme H.323) fait désormais partie de l'offre des constructeurs de matériels vidéo tels que Tonna (gamme de produits Viscable++). Or, qui dit IP dit réseau local Ethernet et donc paire torsadée.

Selon quels critères choisir le type de câble ?

Dans le premier cas, la connexion aux équipements est simple : en bureau, un cordon coaxial relie la prise à une télévision ou à une carte dans un PC (de type WinPC) ; à son extrémité, le câble est connecté à la régie vidéo située dans le local nodal.

Dans le second cas de figure, la connexion est plus coûteuse, car elle requiert l'installation d'équipements intermédiaires dans les LTE. L'architecture ressemble alors à celle mise en place pour le réseau local et la téléphonie.

Il est envisageable de connecter directement la régie vidéo aux prises utilisateur en brassant les prises de distribution aux câbles de rocades jusqu'à l'endroit où est située la régie. Cette solution nécessite néanmoins beaucoup de câbles en cuivre.

En définitive, le câble coaxial est adapté à des besoins ponctuels de diffusion vidéo (moins d'une centaine de postes de travail, des salles de conférence, etc.), tandis que le câble à paires torsadées est bien mieux adapté à des gros besoins, tels que ceux nécessités dans le monde de l'audiovisuel.

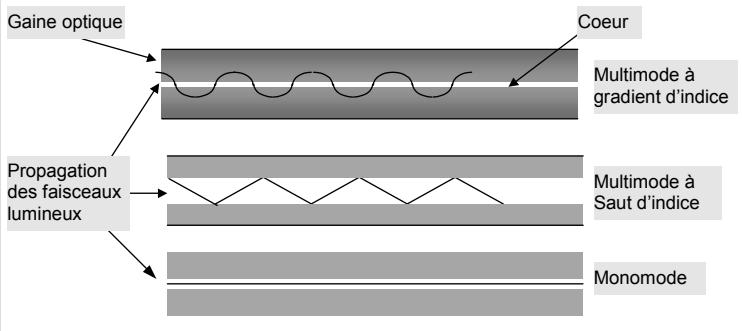
Par contre, avec la généralisation de la vidéo sur IP, le câble coaxial risque bien de disparaître au profit de la paire torsadée.

LE POINT SUR LA FIBRE OPTIQUE (TIA/EIA 492AAAA)

Ce type de câble véhicule des ondes lumineuses au sein d'une fibre caractérisée par sa **gaine optique** et son **cœur**, se différenciant par leur indice de réfraction.

La qualité d'une fibre dépend de trois paramètres :

- du mode de **propagation** de la lumière (**multimode** ou **monomode**) ;
- du **diamètre** de la gaine et du cœur (62,5/125 µ ou 50/125 µ pour la multimode) ;
- de leur **composition** (verre de silice, plastique ou composite).



La qualité du signal dépend en plus :

- de la **longueur d'onde** émise (850 et 1 300 nm pour les multimodes ; 1 310 et 1 550 nm pour les monomodes) ;
- de la source lumineuse : une diode électroluminescente LED (*Light-Emitting Diode*) ou laser ILD (*Injection Laser Diode*).

Les tableaux suivants présentent les performances comparées en fonction de ces paramètres.

	Multimode à saut d'indice	**Multimode à gradient d'indice**	**Monomode**
Source lumineuse	LED ou laser	LED ou laser	Laser
Bande passante	20 à 200 MHz/km	200 MHz à 1,5 GHz/km	3 à 50 GHz/km
Diamètre du cœur	de 50 à 125 µ	de 50 à 125 µ	de 2 à 8 µ
Diamètre de la gaine	de 125 à 440 µ	de 125 à 440 µ	de 15 à 60 µ

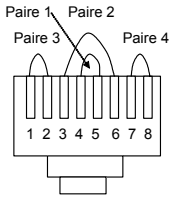
Type de fibre	**Composition Cœur / Gaine**	**Affaiblissement en dB/km** à 850 nm	à 1 300 nm	à 1 500 nm
Multimode à saut d'indice	Verre de silice / verre de silice	2	0,5	0,2
	ou plastique	2,5	---*	---*
	ou verre composite	3,4	---*	---*
Multimode à gradient d'indice	Verre de silice / verre de silice	2	0,5	0,2
	ou verre composite	3,5	1,5	---*
Monomode	Verre de silice / verre de silice	2	0,5	0,2

*--- Affaiblissement trop important, non utilisable

Quel type de prise ?

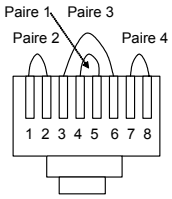
Les standards étant bien établis, les choix sont ici plus limités :

- **RJ45** pour la paire torsadée (la prise RJ11 du téléphone peut s'insérer dans une prise RJ45 femelle, mais pas l'inverse) ;



Fil	Paire	Couleur
1	3	blanc/vert
2	3	Vert
3	2	blanc/orange
4	1	bleu
5	1	blanc/bleu
6	2	orange
7	4	blanc/marron
8	4	marron

TIE/EIA T568A



Fil	Paire	Couleur
1	2	blanc/orange
2	2	orange
3	3	blanc/vert
4	1	bleu
5	1	blanc/bleu
6	3	vert
7	4	blanc/marron
8	4	marron

TIE/EIA T568B

- **SC** ou MT-RJ pour la fibre optique (attention, on trouve encore du ST) ;
- **BNC** pour le coaxial.

Pour des questions de simplicité, les mêmes types de prise sont posés en bureau et en local technique. Cela permet d'utiliser les mêmes cordons de brassage.

L'aménagement des locaux techniques

L'aménagement d'un local technique est plus complexe qu'il n'y paraît, car il n'y a pas de solution universelle. Ce qui doit présider à sa conception est la facilité d'utilisation, à savoir l'accès aux équipements actifs et la facilité de brassage.

Ce qui complique la tâche, c'est que les locaux dédiés à l'informatique sont généralement de petite dimension. Ils doivent néanmoins accueillir le câblage d'étage ainsi que les équipements actifs.

Les baies

La hauteur utile d'une baie est généralement de 36 ou **42 U** (un U équivalant à 4,44 cm) ; ses largeur et profondeur peuvent varier entre 800 × 800 cm, 600 × 800 cm ou 600 × 600 cm. La taille de 800 × 800 a ma préférence, car elle offre suffisamment d'espaces latéraux pour y faire passer des cordons de brassage et suffisamment de profondeur pour y loger tous types d'équipements actifs.

Dans tous les cas, elle doit être équipée de rails crénelés fixés sur les montants droit et gauche, de manière à offrir une largeur de **19 pouces** (48,26 cm). Les rails doivent être fixés à l'avant et à l'arrière, en retrait de **10 à 15 cm** par rapport aux façades. Cet espace permettra de fermer la porte lorsque tous les cordons de brassage seront installés. Détail pratique, mais qui est parfois oublié…

Les baies peuvent être dédiées au câblage ou mixtes câblage/équipements, accueillant, par exemple, un panneau de brassage dans leur partie haute et les équipements actifs dans leur partie basse.

Le cheminement des cordons de brassage

Si le local contient plusieurs baies, de nombreux cordons de brassage seront nécessaires pour raccorder les équipements aux panneaux de brassage : autocommutateurs, routeurs, concentrateurs, etc. Afin de maintenir une installation avec le minimum de cordons emmêlés, il est essentiel de simplifier la tâche des exploitants.

L'utilisation du faux plancher est déconseillée, car on y laisse toujours s'accumuler un sac de nœuds bien caché ; d'autre part, soulever les dalles est toujours une opération fastidieuse. Bien souvent, elles ne peuvent se soulever aisément, car il y a toujours un équipement posé dessus, à cheval entre deux dalles.

Il est, en revanche, préférable de faire circuler les cordons de brassage dans les flancs des baies équipées de guides câbles ainsi que dans un chemin de câble fixé en hauteur, à l'arrière de ces dernières (attention à ne pas en sous-estimer la largeur). Les cordons le moins souvent manipulés circuleront dans le chemin de câble, tandis que les cordons utilisés pour la distribution (connexion des prises utilisateurs aux équipements) chemineront dans les guides câbles.

Afin de faciliter les opérations d'exploitation (brassage, installation d'équipements, etc.), il convient également de réserver un dégagement de 80 cm au moins en face avant et en face arrière des baies.

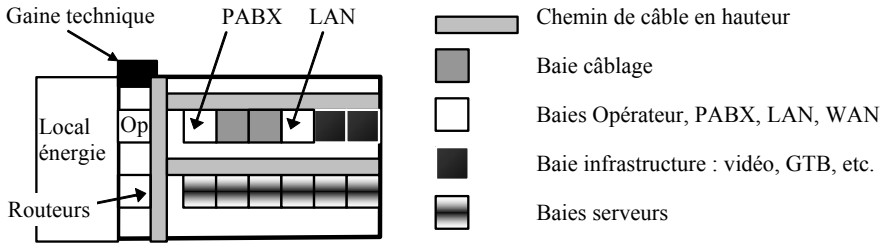
Détail non moins pratique, les luminaires seront disposés de manière à éclairer les zones de dégagement entre les baies.

L'organisation du local

Une des règles de base est de positionner les équipements actifs à proximité des panneaux de brassage, afin de limiter la longueur des cordons et également les sources de « sacs de nœuds ».

Par ailleurs, tous les équipements susceptibles d'être connectés à des lignes télécom (autocommutateurs et routeurs) doivent de préférence être situés à proximité de l'arrivée de ces lignes dans la baie de l'opérateur. Il faut dans tous les cas prévoir des câbles de déport (avec ferme CAD et/ou panneaux RJ45) entre les deux types de baies, surtout si le local opérateur est distinct du local informatique.

Figure 2-4. *Agencement du local nodal.*



Le local nodal peut très bien être situé dans la salle informatique ou être séparé pour des questions d'organisation, les exploitants réseaux ne devant pas avoir accès aux serveurs, et inversement pour les exploitants système.

Si les serveurs sont proches des équipements actifs, ils pourront être directement connectés à ces derniers à l'aide des cordons circulant dans les guides câbles et/ou le chemin de câble.

Si, en revanche, ils sont éloignés, ou si la salle informatique est distincte du local technique, un panneau de distribution (prises RJ45 et/ou SC) devra être installé à proximité de chaque serveur, dans une baie mixte câblage/serveur, et relié par un câblage à un panneau de distribution analogue dans le local technique.

Le cahier des charges

La rédaction de ce document synthétise les données collectées lors des phases précédentes.

Le cahier des charges a pour but d'expliquer ce qui est attendu et de guider les soumissionnaires dans leur réponse.

Dans le cadre de réalisations plus importantes, le document peut être scindé en un **CCTG** (cahier des clauses techniques générales) et un **CCTP** (cahier des clauses techniques particulières).

Le CCTG définit les engagements attendus de la part de l'entreprise :

- ses responsabilités techniques ;
- la nécessité de coordination avec d'autres corps d'état ;
- le respect d'un calendrier de réalisation ;
- le maintien du site dans un bon état de propreté si l'immeuble est déjà occupé, et en particulier l'évacuation des gravats à sa charge ;
- les garanties de qualité ;
- le respect du plan hygiène et sécurité.

Le CCTP doit décrire l'existant (l'infrastructure d'immeuble et le câblage, s'il existe), étayé par des plans (plan de masse, étage type, sous-sol). Il doit ensuite fournir tous les éléments techniques qui permettront aux soumissionnaires de répondre. Ses données sont en fait une synthèse de l'étude réalisée dans les phases précédentes. Il s'agit de décrire :

- le cheminement des câbles (dans ses principes généraux) ;
- les règles d'espacement par rapport aux sources de courants forts (câbles électriques, moteurs d'ascenseur, alimentations à coupure, tubes néon, etc.) ;
- les règles d'ingénierie que l'entreprise devra impérativement respecter concernant les câbles, les chemins de câble, la connexion des prises, le raccordement à la terre informatique, etc.

Dans une autre partie, le CCTP décrit les prestations attendues, à savoir :

- le percement des murs, si nécessaire ;
- la fourniture et la pose de tous les composants requis : chemins de câbles, goulottes, panneaux de brassage, baies, prises, boîtiers VDI dans lesquels viennent s'insérer les prises, etc.
- le raccordement à la terre informatique et, si nécessaire, la réalisation de la terre informatique à partir du puits de terre jusqu'à la distribution.

Des détails qui ont leur importance :

- la documentation des travaux réalisés tels que le cheminement exact des câbles et la position exacte des prises reportés sur les plans ;
- les fiches de tests de chaque prise ;
- l'étiquetage des prises avec des étiquettes gravées autocollantes (et non pas des Dimos ou du papier qui s'effacent ou se décollent au bout de quelques mois).

Et enfin, la nature des composants fournis et installés :

- pour le câblage cuivre : types de câbles de distribution et de rocades, types de connecteurs, etc. ;
- pour le câblage optique : types de câbles de distribution et de rocades, types de tiroirs optiques, etc. ;
- les types de baies : dimensions, avec ou sans portes, etc. ;
- les types de boîtiers VDI : nombre et types des prises (RJ45, CD, etc.) ;
- le descriptif des tests sur chaque prise et les valeurs à mesurer lors des tests réflectométriques (voir plus loin).

Afin de comparer facilement les réponses, un modèle de bordereau de prix tel que celui présenté ici pourra être joint.

Désignation des ouvrages	**U**	**Qté**	**P.U. HT**	**P.T. HT**
Lot 1 – Câblage cuivre (fourniture, pose et raccordement y compris toute sujétion)				
Cheminements entre la salle informatique et le LN-A				
Percements	ens			
Chemins de câble	m			
Raccordement à la terre électrique	ens			

Le reproche que certains peuvent faire à une description aussi détaillée est qu'elle est du ressort de l'entreprise, que c'est son métier. Certes, mais le but est ici de bien définir le **niveau de prestation** que l'on attend. Si on ne le fait pas, les soumissionnaires feront des réponses minimales :

- avec des composants de faible qualité ;
- en omettant certains composants qui peuvent paraître accessoires, comme l'étiquetage des prises ou des colliers de fixation ;
- en calculant la longueur des câbles au plus juste (en les faisant passer en ligne droite sans respecter les contraintes d'écartement des sources de courant fort) ;
- sans prendre en compte la coordination avec les autres entreprises ;
- etc.

Ce type de réponse paraîtra attractif sur le plan financier, mais passera sous silence de nombreux aspects importants.

Il peut également être tentant de confier la réalisation du câblage à une société spécialisée en électricité. Cela serait une erreur, sauf si bien sûr elle dispose des compétences requises en courant faible. Car le câblage informatique n'a rien à voir avec l'électricité : ce sont deux métiers différents qui font appel à des expertises sans aucun rapport entre elles.

Enfin, la certification de l'entreprise pour le système de câblage proposé est un gage de qualité : non seulement elle offre la garantie du constructeur pendant dix à quinze ans mais, de plus, elle signifie que les techniciens ont suivi une formation spécifique de la part du constructeur sur le type de matériel proposé.

En fin de chantier, un représentant du constructeur vérifie (en plus de la recette dont nous parlerons plus loin) la qualité de l'installation, et appose son certificat de garantie. Celui-ci assure la remise en état, pendant dix à quinze ans, selon le constructeur, de n'importe quel composant défectueux (câble, connecteur, panneau de brassage). Si l'entreprise disparaît dans l'intervalle, le constructeur prend le relais ou désigne une autre société.

Le suivi du chantier et la recette

En cours de réalisation, il est nécessaire d'organiser un point hebdomadaire avec les représentants de la société de câblage : l'objectif est de contrôler l'avancement des travaux, de résoudre certains problèmes techniques, de préciser des détails comme le principe d'étiquetage des prises, etc.

L'ORGANISATION D'UN CHANTIER DE CÂBLAGE

Le **maître d'ouvrage** est le donneur d'ordre, celui qui paie, c'est-à-dire vous, le client.

Le **maître d'œuvre** est l'exécutant, le responsable des travaux ; il rend des comptes au maître d'ouvrage.

Le **soumissionnaire** est l'entreprise qui répond à l'appel d'offres ; le terme entreprise désigne l'**entreprise de câblage** qui a été retenue en tant que maître d'œuvre du projet **courants faibles,** par opposition aux câblages **courants forts** qui concernent l'électricité, généralement réalisée par une autre entreprise.

L'entreprise de câblage désigne un **chef de chantier** qui coordonne le travail des ouvriers sur site ; elle est parfois l'interlocuteur du maître d'ouvrage. Dans le cas de réalisations importantes, un **conducteur de travaux** est désigné en tant qu'interlocuteur.

De son côté, le maître d'ouvrage est souvent assisté d'un consultant qui, dans le cadre d'une intervention d'**assistance à maîtrise d'ouvrage**, assure le lien entre l'utilisateur exprimant des besoins généraux et le monde du câblage avec sa spécificité et son vocabulaire.

Une **réunion de chantier** est régulièrement organisée afin de coordonner et de suivre l'avancée des travaux. Elle a aussi pour but de coordonner les activités de l'entreprise de câblage avec d'autres **corps d'états** (électricien, société en charge de la climatisation, société en charge des faux plafonds, etc.). Sur le terrain, l'entreprise de câblage doit assurer cette coordination.

Il est également indispensable de procéder à des visites régulières du site (une à deux fois par semaine selon l'état d'avancement des travaux). L'objectif est ici de contrôler la qualité des travaux en cours afin de procéder à d'éventuelles rectifications avant la fin du chantier. Mieux vaut détecter le plus en amont possible tout problème pouvant nécessiter la reprise des travaux supposés achevés.

Le chantier terminé, il est nécessaire de procéder à sa recette. Celle-ci comprend la vérification exhaustive, qualitative et quantitative de l'ensemble des composants installés, ainsi que l'analyse des documents remis (cahier de tests, plans, etc.). Il s'agit notamment :

- de valider le cahier de tests fourni par l'entreprise ;
- de réaliser des tests complémentaires sur un échantillon de prises avec le même réflectomètre fourni par l'entreprise de câblage ;
- d'effectuer le contrôle physique de l'installation avec le câbleur.

Le câbleur doit ainsi tester toutes les prises avec un réflectomètre et fournir tous les résultats. Ceci permet de s'assurer des performances du système de câblage posé (câbles + prises) et de vérifier que les mesures sont dans les normes. C'est également l'occasion pour le câbleur de réparer d'éventuelles prises défectueuses.

Prise n° 22 du 1er étage

Fiche de test - Câblage Class D

Ici, un câble et ses prises catégorie 5 sont testés et comparés à la classe D de la norme EN50173.

ID circuit :	1.022	Date : 01/09/97	
Résultat :	Passe	type de câble : Cat 5 SFTP	
Propriétaire :	Société X	NVP :	
Niveau :		Taille :	
No de série :	38T95KB0205	Fabricant :	
Version SW :	V04.10	Connecteur :	
Inj. Série # :	38T95K00879	Utilisateur :	
Extrémité #1 :		Extrémtité #2 :	
Bâtiment:		Bâtiment:	
Armoire:		Armoire:	
Bâti:		Bâti:	
Carte:	porte:	Carte:	porte:

L'impédance du câble est garantie par le constructeur à ±15%...

Pour la paire de fils 1/2, la plus mauvaise valeur est de 4,7db à 97 Mhz, ce qui est bien au delà de la norme (22,8 db).

Test		Résultat attendu		Résultat de test			
Câblage		local : 12345678S Dist. : 12345678S		local : 12345678S Dist. : 12345678S			
		min	max	Paire12	Paire36	Paire45	Paire78
Longueur	m	3.0	100.0	23.5	24.0	23.6	23.3
Délai	ns	0	32767	109	111	109	108
Impédance	ohms	85.0	115.0	105.0	102.0	102.0	98.0
Résistance	ohms	0.0	40	4.7	4.9	4.0	4.4
Capacité	pF	40	5600	1100	1102	1108	1084
Affaibliss	dB			4.7	4.9	4.8	4.8
@Fréquence	MHz			97.0	100.0	97.0	100.0
Limite	dB	Class D		22.8	23.2	22.8	23.2

Comb. de paires		12/36	12/45	12/78	36/45	36/78	45/78
Perte de PARA(REFLEC)	dB	39.1	41.7	36.8	37.1	37.4	40.2
@Fréquence 20.0-100.0	MHz	77.5	68.5	100.0	96.5	97.0	76.5
Limite : Class D	dB	25.8	26.5	24.0	24.2	24.2	25.8
Marge actif (REFLEC)	dB	34	[illegible]	[illegible]	[illegible]	32.6	36.3
@Fréquence 20.0-100.0	MHz	100	[illegible]	[illegible]	[illegible]	97.0	76.0
Limite : Class D	dB	4.0	4.1	1.0	1.9	4.7	9.7
Perte de PARA(INJ)	dB	37.8	40.2	38.9	36.9	38.9	37.6
@Fréquence 20.0-100.0	MHz	[illegible]	[illegible]	.5	96.0	73.5	97.0
Limite : Class D	dB	[illegible]	[illegible]	.4	24.3	26.1	24.2
Marge actif (INJ)	dB	[illegible]	[illegible]	.3	32.2	35.1	32.8
@Fréquence 20.0-100.0	MHz	77.0	93.0	69.0	96.0	73.0	97.0
Limite : Class D	dB	9.5	5.6	11.4	4.9	10.4	4.7

Mesure dans les deux sens : côté instrument de mesure et côté injecteur de signal

Marge = le delta entre la norme et ce qui est mesuré.

La fiche de test présentée ci-dessus comporte deux parties :

- Une description des paramètres saisis manuellement par la personne qui a procédé aux tests. Parmi les champs importants, on trouve « ID Circuit » qui correspond au numéro de la prise tel qu'on le trouve sur l'étiquette et les plans d'implantation.
- Deux tableaux indiquant les caractéristiques et les performances du câble mesurées par le réflectomètre. Ces valeurs ne sont, en théorie, pas modifiables manuellement.

Le premier tableau montre les caractéristiques du câble pour chaque paire. Les plus importantes sont la longueur, l'impédance et l'affaiblissement.

La première colonne indique les valeurs limites paramétrées dans le réflectomètre au moment du test. Les colonnes suivantes indiquent les valeurs mesurées pour chaque paire : 1-2, 3-6, 4-5 et 7-8. Les valeurs les plus importantes à vérifier lors de la recette sont :

- La longueur : celle-ci varie de quelques centimètres d'une paire à l'autre. Ceci est dû à la précision de l'appareil de mesure, au pas de torsade légèrement différent, et éventuellement à la courbure du câble tout au long de son cheminement.
- Le délai de propagation et l'impédance (dépend du type de câble). Les valeurs sont fournies par le constructeur avec une marge de ±15%.
- La fréquence à laquelle le plus mauvais affaiblissement a été mesuré. Par exemple, pour la paire 1-2, on trouve un affaiblissement de 4.7 dB à 97 MHz alors que la limite à ne pas dépasser, imposée par la norme Classe D, est de 22,8 dB (valeur calculée par le réflectomètre pour la longueur de 23.5 mètres).

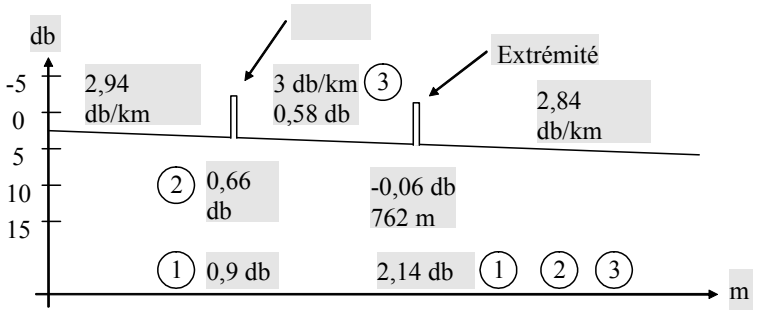
Le deuxième tableau indique les mesures de paradiaphonie pour chaque combinaison de paires. Là encore, la fiche de test indique la fréquence à laquelle les plus mauvaises valeurs ont été mesurées. Par exemple, pour la combinaison 12/36, on a une paradiaphonie de 39,1 dB à 77,5 MHz alors que pour la combinaison 12/78, la plus mauvaise valeur est obtenue à 100 MHz.

Les combinaisons de paires à vérifier dépendent du réseau que l'on va installer. Par exemple, il faut considérer les paires 12/36 pour le 10bT, 12/78 pour le 100bT et toutes les paires pour le Gigabit.

Si des anomalies sont constatées, la recette peut donner lieu à des réserves. Les réserves sont levées seulement lorsque l'entreprise de câblage a corrigé les défauts. Le procès verbal de recette peut alors être signé.

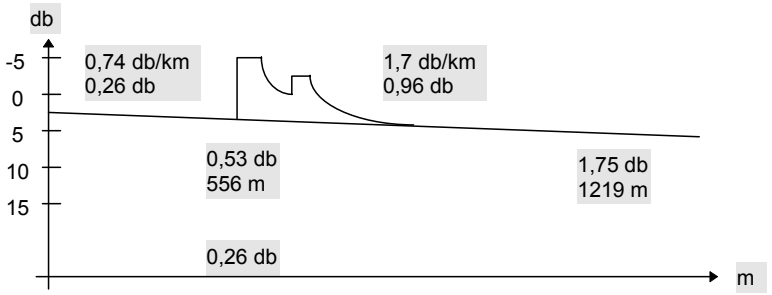
La fibre optique doit également faire l'objet d'un test réflectométrique dans les deux sens et aux deux longueurs d'ondes de référence. La plupart des équipements réseau utilisent, en effet, la longueur d'onde de 850 nm, mais il n'est pas exclu qu'avec le 10 Gigabit, la longueur de 1 300 nm soit la seule possible.

Le schéma suivant présente, à titre d'exemple, une courbe issue des tests réflectométriques d'une fibre optique. La chaîne de liaison comporte le réflectomètre, une bobine d'amorce (un câble d'une fibre optique de plusieurs dizaines de mètres enroulé), la fibre à tester y compris les connecteurs d'origine et d'extrémité, et une bobine de terminaison.



La première valeur, notée « 1 », représente l'affaiblissement de la bobine d'amorce. Les deux pics de la courbe correspondent aux connecteurs ST bornant la section de fibre à mesurer. L'affaiblissement et la longueur des trois sections (bobine d'amorce, fibre optique et bobine de terminaison) sont ainsi mesurés. La valeur du brin à tester est appelée bilan et est exprimée en décibels (symbole dB). Dans l'exemple ci-dessus, la fibre a une longueur de 194 mètres et son bilan est de 1,18 dB.

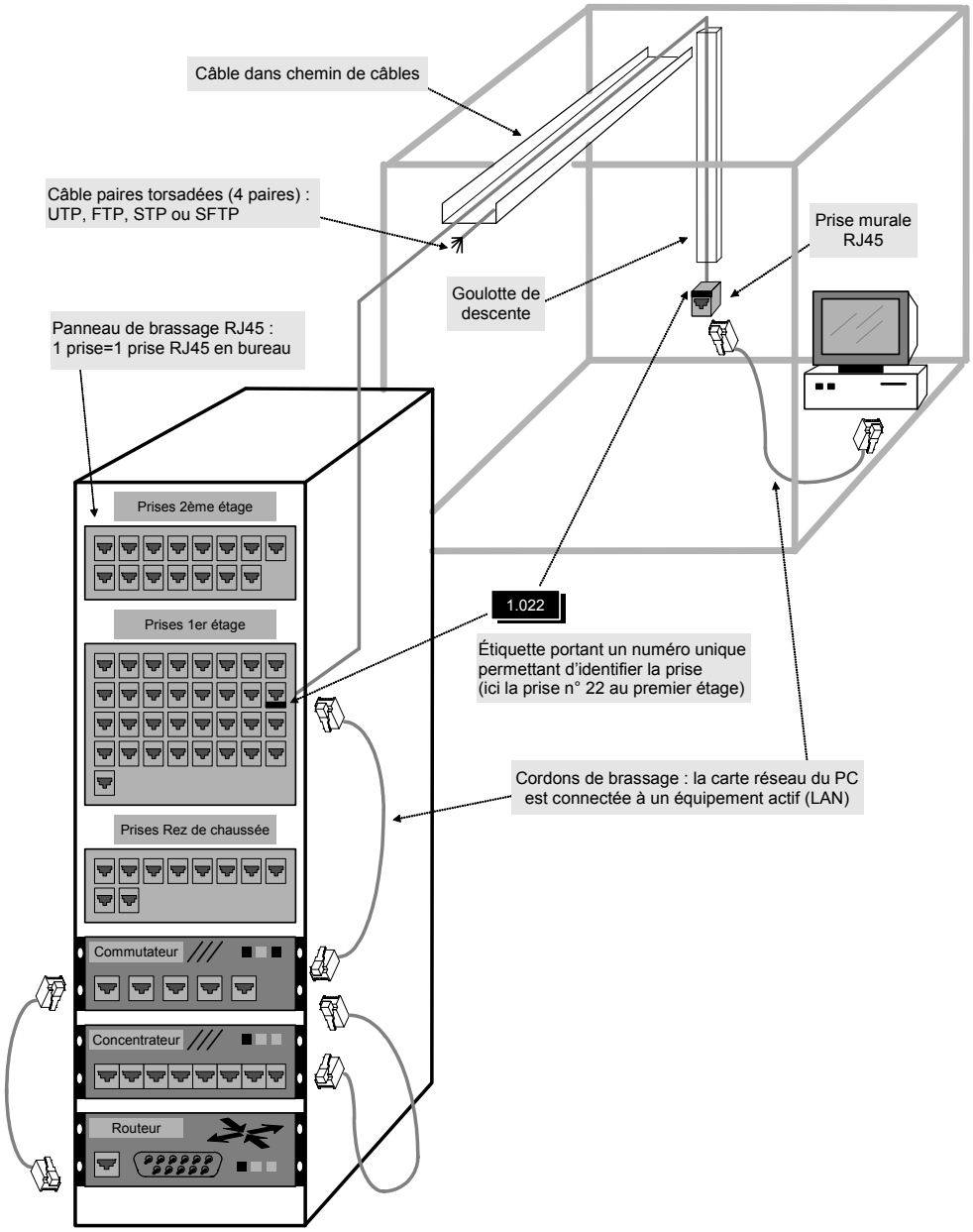
Le schéma suivant montre, en revanche, la courbe obtenue par la mesure d'une fibre défectueuse.



L'interprétation des résultats est toujours délicate. Dans le cas présent, il se peut que la fibre soit légèrement coupée ou présente un rayon de courbure trop faible (un angle droit par exemple).

Par ailleurs, un photomètre permettra de mesurer avec précision la longueur de chaque fibre et de s'assurer ainsi que leur longueur ne dépasse pas les normes imposées par Ethernet (cf. chapitre 3)

La recette terminée, vous pouvez utiliser votre câblage pour constituer votre réseau local et éventuellement votre réseau téléphonique tel que le montre le schéma suivant.


Câble dans chemin de câbles
Câble paires torsadées (4 paires) :
UTP, FTP, STP ou SFTP
Prise murale
RJ45
Goulotte de
descente
Panneau de brassage RJ45 :
1 prise=1 prise RJ45 en bureau
Prises 2ème étage
Prises 1er étage
1.022
Étiquette portant un numéro unique
permettant d'identifier la prise
(ici la prise n° 22 au premier étage)
Cordons de brassage : la carte réseau du PC
est connectée à un équipement actif (LAN)
Prises Rez de chaussée
Commutateur
Concentrateur
Routeur

3

Architecture des réseaux locaux

Le système de câblage étant maintenant prêt à l'usage, on peut alors commencer à y installer un réseau.

La conception d'une architecture réseau est élaborée en fonction du nombre de postes de travail à connecter. On peut considérer les cas suivants :

- petit réseau : moins de 200 postes dans un même bâtiment ;
- réseau moyen : de 200 à 800 postes dans un même bâtiment ;
- gros réseau : plus de 800 postes dans un même bâtiment ;
- 1ère variante : plusieurs bâtiments contenant un nombre varié de postes de travail ;
- 2ème variante : plusieurs sites contenant un nombre varié de postes de travail.

Bien qu'arbitraires, ces bornes correspondent à des ordres de grandeur et à des sauts technologiques. En effet, plus le nombre de postes est important, plus il faut répondre à un certain nombre de contraintes et d'exigences qui n'apparaissent qu'avec la complexité du réseau.

L'architecture réseau est bien sûre liée au système de câblage, mais celui-ci doit avoir été conçu pour en limiter les contraintes, c'est-à-dire s'adapter à toutes les situations. Le chapitre précédent a aidé à œuvrer en ce sens. Tous les systèmes de câblage sont en étoile, de même que la topologie ; les équipements actifs seront donc positionnés dans les locaux techniques. Ils serviront à connecter les postes de travail aux serveurs.

Dans le cas le plus simple, la conception d'une architecture consiste à choisir et à positionner les équipements actifs, puis à les connecter entre eux en utilisant le câblage.

Les choix de base

Quel type de réseau choisir ?

La création d'un réseau local nécessite de faire des choix, et tout d'abord celui du type : Ethernet, Token-Ring ou un autre ?

Le premier est le plus répandu et le moins cher, alors que, pour le deuxième c'est l'inverse. Le choix ira donc de préférence au premier. L'intérêt de Token-Ring est surtout sa compatibilité avec les équipements grands systèmes IBM (3090, etc.). Si vous avez des contrôleurs 3174, des terminaux 3270 ou des PC devant se connecter en émulation 3270 à un système central IBM, le choix se portera naturellement vers ce type de réseau local. Mais ce n'est pas une obligation, car la plupart des équipements IBM acceptent désormais des cartes Ethernet. Cela étant, les postes de travail peuvent toujours être connectés à un réseau Ethernet, et les systèmes IBM à des réseaux Token-Ring.

D'autres solutions sont envisageables pour construire un réseau local, mais elles sont nettement plus chères. Il s'agit, par exemple, d'ATM (*Asynchronous Transfert Mode*). Ce type de réseau est surtout destiné à d'autres usages que nous verrons au chapitre 11.

Pour toutes ces raisons, le choix d'Ethernet s'impose.

Concernant la topologie, la plus pratique est celle de l'étoile : tous les systèmes de câblage sont, on l'a vu, fondés sur ce principe.

Quel débit retenir ?

La décision suivante concerne le débit du réseau, c'est-à-dire la vitesse de transmission des trames Ethernet, encore appelée bande passante. La norme Ethernet est déclinée en plusieurs variantes : 10 Mbit/s (norme 10bT), 100 Mbit/s (norme 100bT) et 1 Gbit/s (norme 1000bT). De par son coût et son caractère innovateur, le Gigabit Ethernet est réservé aux liaisons entre les équipements de concentration et aux serveurs.

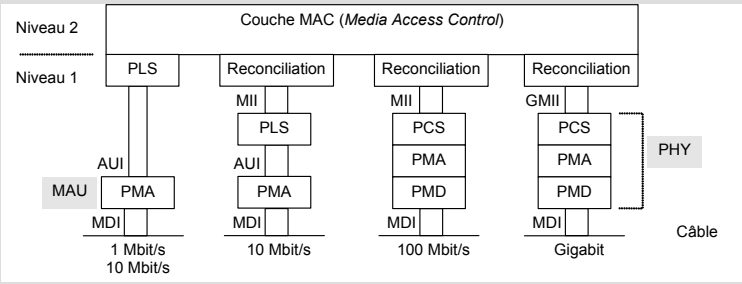
Le choix du débit se fera donc en fonction des coûts, plutôt 10/100 Mbit/s pour les PC et 100/1000 Mbit/s pour les serveurs.

Débit	Utilisation
10 Mbit/s	Postes de travail bureautique
100 Mbit/s	Postes de travail multimédias et serveurs
1 Gbit/s	Pour connecter les équipements réseaux entre eux ainsi que les gros serveurs

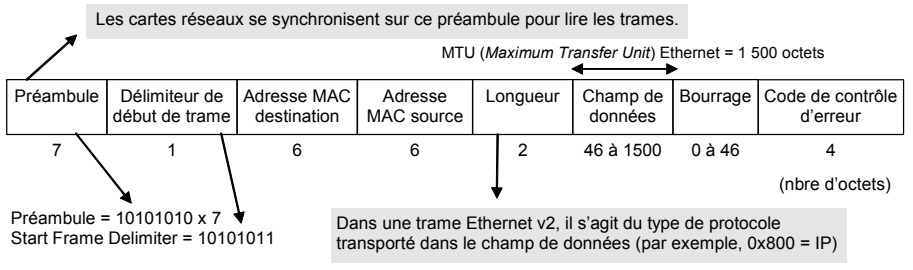
Côté poste de travail, la plupart des cartes réseau fonctionnent à 10 et 100 Mbit/s, et sont au même prix que les cartes 10 Mbit/s. En outre, la plupart des commutateurs offrent des ports à détection automatique de vitesse (port *autosense*).

LE POINT SUR ETHERNET (IEEE 802.3)

La norme 802.3 décrit un protocole de niveau 1 (couche physique) – c'est-à-dire la manière de générer les signaux sur le câble – ainsi que le protocole de niveau 2 (couche liaison). Les spécifications, édictées par l'IEEE (*Institute of Electrical and Electronics Engineers*), reposent sur le modèle suivant :



Au niveau 2, la couche **MAC** (*Medium Access Control*) a pour rôle de transporter les données entre deux points d'un segment Ethernet. Pour ce faire, elle encapsule les données dans une **trame MAC,** identifie émetteurs et destinataires par des **adresses MAC** et assure que les données sont transmises sans erreur, grâce à un code de contrôle.



Chaque carte réseau est identifiée par une adresse MAC unique sur 6 octets (par exemple, 01-00-0A-FB-5D-52). Toute trame MAC émise comporte les adresses de l'émetteur et du destinataire. La carte réseau ne prend en compte que les trames MAC qui lui est destinée. Exception à cette règle : une carte peut émettre une trame spéciale, appelée trame de diffusion (ou trame de ***broadcast***), qui sera lue par toutes les autres cartes (adresse destination dont tous les bits sont à 1 — FF-FF-FF-FF-FF-FF).

La couche MAC est également responsable de gérer l'accès au segment Ethernet quand il est partagé par plusieurs nœuds. Lorsque deux stations émettent simultanément, la somme des deux signaux dépasse un seuil qui est interprété comme étant une **collision**. Dans un tel cas, chaque station arrête d'émettre et attend un temps aléatoire avant d'essayer de réémettre leur trame. Cette méthode de gestion de la contention est appelée **CSMA/CD** (*Carrier Sense Multiple Access / Collision Detection*).

• • •

LE POINT SUR ETHERNET (SUITE)

Afin d'être le plus possible indépendant des méthodes et des supports de transmission, le modèle de référence de la norme IEEE 802.3 repose sur plusieurs sous-couches physiques :

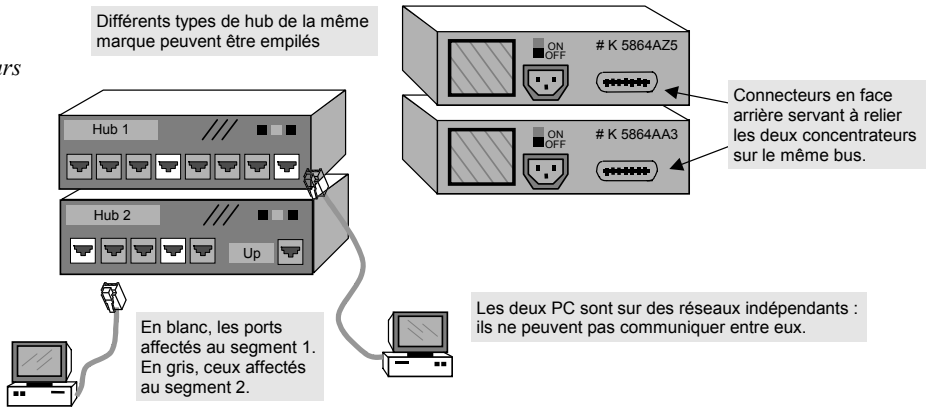
- La sous-couche **réconciliation**, qui a pour rôle d'adapter les bits manipulés par la couche MAC avec les signaux utilisés par les couches physiques sous-jacentes.
- **PLS** (*Physical Layer Signaling*), qui interprète les bits véhiculés sur le câble (données, détection de collision, auto-négociation, etc.).
- **PCS** (*Physical Coding sublayer*), qui code les groupes de bits en symboles, c'est-à-dire en des signaux. Par exemple, le codage 4B/5B consiste à coder 4 bits en un groupe de 5 signaux.
- **PMA** (*Physical Medium Attachment*), qui transmet et reçoit les signaux véhiculés sur le câble en bande de base (par exemple, NRZI). La PMA, associée à un débit, est activée ou désactivée en fonction de la négociation de débit.
- **PMD** (*Physical Medium Dependant*), qui désigne l'électronique qui génère les signaux électriques ou optiques.

Les interfaces **AUI** (*Attachement Unit Interface*), **MII** (*Media Independant Interface*) et **GMII** (*Gigabit MII*), qui sont optionnelles, assurent la liaison entre la couche MAC (logiciel hébergé dans des composants électroniques) et les composants gérant les signaux (PHY) de la carte réseau. Concrètement, elles se présentent sous la forme de connecteurs, pouvant être étendus par un cordon de quelques centimètres, et raccordant la couche MAC à un ***transceiver*** ou MAU (*Medium Attachment Unit*), qui intègrent donc les fonctions de la couche PHY. Les GBIC de Cisco sont des exemples de tels *transceivers*.

Quel format d'équipement ?

Le marché offre le choix d'équipements seuls (*stand alone*), empilables (*stackable*), ou en châssis : des cartes sont insérées dans des emplacements prévus à cet effet (*slots*). Pour compliquer le choix, il existe également des *stackables* avec des *slots* d'extension qui permettent d'ajouter un ou deux ports, dans le but de chaîner l'équipement à un autre.

Figure 3-1.
Les possibilités des concentrateurs empilables.



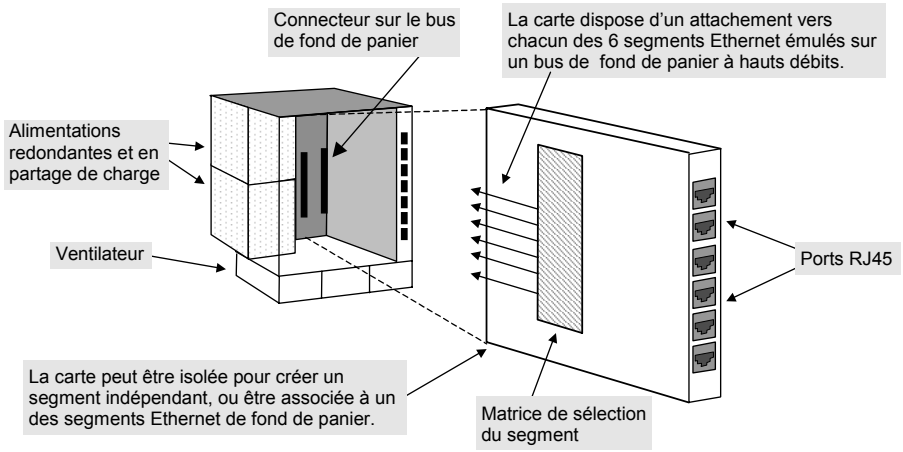
Les modèles *stand alone* visent le marché d'entrée de gamme ; ils sont parfaits pour créer un premier réseau local tel que celui du premier chapitre.

Les modèles empilables sont envisageables dès qu'il y a des possibilités d'extension. Par exemple, votre société dispose de cinquante postes de travail, et vous commencez par en connecter dix dans un premier temps. Vous pouvez alors acheter un hub 12 ports, puis un autre 12 ports plus tard.

Un concentrateur, ou un commutateur, comprend un nombre limité de ports (généralement 8, 16, 24 ou 32). Or, la plupart du temps, un local technique concentre beaucoup plus de postes de travail (jusqu'à plusieurs centaines). La solution est alors de chaîner les *stackables* entre eux *via* un bus spécial, dédié à cet effet ; il s'agit d'un câble externe reliant les équipements entre eux pour n'en faire qu'une unité logique. Attention, le bus étant propriétaire, seuls les équipements d'un même constructeur pourront être chaînés entre eux, généralement de 5 à 8 au maximum.

Une autre solution consiste à installer des châssis, certes plus chers, mais qui offrent de plus grandes capacités d'accueil. Ces équipements permettent de créer plusieurs segments indépendants à des débits différents.

Figure 3-2. *Principes et fonctionnalités d'un châssis.*



Il est possible d'insérer différents types de cartes dans un châssis : concentrateur Ethernet, commutateur Ethernet, carte Token-Ring, FDDI, ATM, etc. Il est également possible de combiner les débits (10, 100 et 1 000 Mbit/s) sur des segments séparés. Mais, attention, il n'est en aucun cas possible de mélanger des débits sur un même segment Ethernet. De même, les réseaux Token-Ring, FDDI et ATM créés seront indépendants.

LES DIFFÉRENTS RÉSEAUX ETHERNET

Il existe actuellement trois déclinaisons d'Ethernet normalisées par l'IEEE (*Institute of Electrical and Electronics Engineers*) : le **10bT** à 10 Mbit/s (norme 802.3), le **100bTx** alias Fast Ethernet à 100 Mbit/s (norme 802.3u) et le **1000bT** alias Gigabit Ethernet (norme 802.3ab).

Le premier chiffre qualifie le débit du réseau Ethernet, la lettre " b " signifie un codage des signaux en bande de base (ex : Manchester) et la lette " T " représente " *Twisted Pair* ", ce qui signifie que le réseau Ethernet fonctionne sur un câblage en cuivre paires torsadées.

Il existe également les mêmes déclinaisons fonctionnant sur fibre optique : **10bF**, **100bFx** et **1000bX** (norme 802.3z). Parmi cette dernière, on distingue le **1000bSX** (S pour *short wavelength*) opérant à 850 nm sur fibre optique multimode et le **1000bLX** (L pour *long wavelength*) opérant à 1 300 nm sur les fibres multimode et monomode.

Ethernet	Codage en ligne	Codage complet	Paires cuivre utilisées	Fréquence du signal / Longueur d'onde	Distance maximale
10bT	Manchester	---	1,2 / 3,6	10 MHz (±15%)	100 m
10bF (FL et FB)	Manchester	---	---	850 nm	2 000 m
100bTx	NRZI	4B/5B	1,2 / 7,8	62,5 MHz	100 m
100bFx	NRZI	4B/5B	---	850 nm	2 000 m
1000bT	PAM5	8B1Q4	Toutes	125 MHz	100 m
1000bSX	NRZ	8B/10B	---	850 et 1 300 nm	220 à 550 m*
1000bLX	NRZ	8B/10B	---	1 300 et 1 550 nm	550 à 5 000 m*

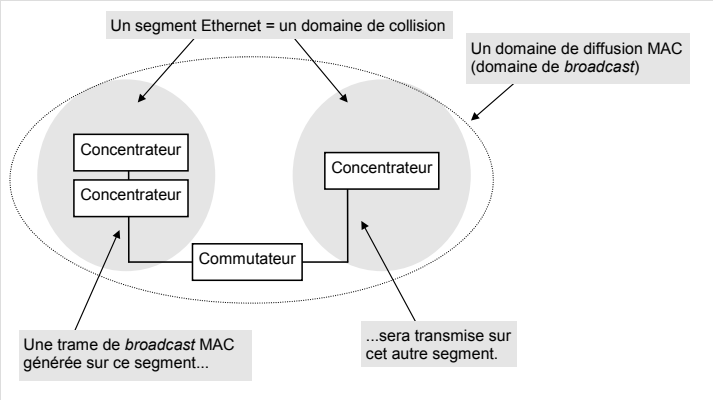
* Selon le type de fibre optique et la qualité des équipements (voir chapitre 12).

Ces variantes utilisent toutes le même principe d'accès au support de transmission (détection de collision en half duplex), les mêmes trames Ethernet et le même adressage MAC (*Medium Access Control*).

Il existe deux types d'équipements : les **concentrateurs** (*hubs*), qui partagent un segment Ethernet entre plusieurs ports, et les **commutateurs** (*switch*), qui créent un segment Ethernet par port.

Certains concentrateurs peuvent être partagés en plusieurs segments indépendants, tandis que le commutateur est capable d'**interconnecter** des segments qui sont physiquement indépendants.

Un **domaine de collision** comprend un ou plusieurs concentrateurs Ethernet ; une trame émise sur un port est régénérée sur tous les autres ports des concentrateurs chaînés (*via* un cordon de brassage) ou empilés (*via* un bus externe par un câble spécifique). Le commutateur ne laisse pas passer les collisions, mais une trame de *broadcast* émise sur un port sera régénérée sur tous les autres ports.



Un PC connecté à un port du commutateur est seul sur le segment Ethernet : aucune collision n'est donc possible, et il dispose de toute la bande passante. Inversement, tous les PC connectés à un concentrateur partagent la même bande passante (10, 100 ou 1000 Mbit/s) et peuvent émettre des trames en même temps, d'où une probabilité de plus en plus importante de collision qui croît avec le nombre de PC.

Le tableau suivant présente quelques éléments de comparaison.

Critère	Empilable	Châssis
Évolutivité	Ajout d'éléments empilables limité à cinq environ	Ajout de cartes limité par le nombre de slots
Segmentation	Limitée à un ou deux segments	Plusieurs segments par port ou par groupe de ports
Capacité de traitement	Bus externe limité à quelques centaines de Mbit/s	Bus de fond de panier de 100 Mbit/s à plusieurs Gbit/s
Alimentation	Une par élément ou, rarement, une pour tous	Une à trois pour l'ensemble du châssis
Redondance d'alimentation	Pas tout le temps	Oui
Création de plusieurs segments indépendants	Oui	Oui
Utilisation	Moins de cent postes en Ethernet, en Token-Ring ou en ATM	Plus de cent postes ; combinaison Ethernet, Token-Ring et ATM possible

En définitive, l'empilable sera choisi pour des faibles densités (moins de cent postes par local technique) ; les châssis seront privilégiés dans les autres cas. D'une manière générale, plus le réseau concentre de postes de travail, plus la fiabilité des équipements doit être importante.

Les châssis seront donc choisis là où le besoin en bande passante est élevé et où la fiabilité est primordiale, c'est-à-dire à des points de concentration stratégiques du réseau.

Concentrateur ou commutateur ?

L'autre décision à prendre consiste à choisir entre les concentrateurs et les commutateurs. Les premiers se contentent de générer le signal, alors que les seconds permettent de créer un segment par port. Ces derniers sont bien sûr plus chers.

Pour une utilisation bureautique du réseau (traitement de texte, comptabilité, base de données, connexion à un serveur, etc.), les concentrateurs suffisent pour connecter les postes de travail car il y a peu de trafic entre eux. Pour améliorer les performances, on peut jouer sur la vitesse (10 ou 100 Mbit/s).

L'utilisation des commutateurs s'envisage dans plusieurs cas de figures :

- lorsqu'on emploie des applications multimédias (voix et vidéo) générant des débits importants et nécessitant des temps de réponse courts ;
- d'une manière générale, lorsque le flux réseau est important et que les temps de réponse sont mauvais ;
- pour interconnecter plusieurs segments Ethernet.

QU'EST-CE QU'UN COMMUTATEUR ?

Un commutateur (*switch*) est un équipement qui offre une bande passante dédiée pour chaque port (10, 100 ou 1 000 Mbit/s par port) alors que le concentrateur partage la bande passante entre tous ses ports. Cela revient à créer un segment Ethernet par port.

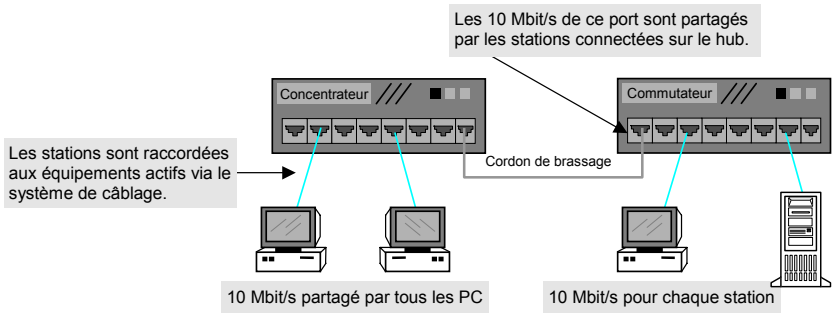
On distingue les *switches* ***cut-through*** (*on the fly*, à la volée) et les *switches* ***store and forward*** (les plus courants aujourd'hui). Les premiers se contentent de régénérer la trame (même les trames erronées et les collisions), tandis que les seconds la stockent en mémoire avant de la régénérer. La méthode ***adaptative cut-through*** combine les deux principes : dès qu'une trame est en erreur, le commutateur bascule en mode *store and forward* pendant un certain temps. La méthode ***fragment free***, la plus performante, lit les 64 premiers octets avant de décider du mode de transmission.

Chaque port du commutateur correspond à un segment Ethernet, c'est-à-dire à un domaine de collision (voir encadré " Le point sur Ethernet ").

Si vous constatez un nombre élevé de collisions lié à une charge réseau importante, vous pouvez, dans un premier temps, segmenter le réseau (c'est-à-dire le couper en deux). Dans ce cas, les PC situés sur un segment ne pourront plus communiquer avec ceux situés sur l'autre.

Au lieu d'acheter un second concentrateur, l'achat d'un commutateur résoudra le problème : les postes seront répartis sur les deux équipements, et les segments ajoutés seront interconnectés.

Figure 3-3.
Concentrateur et commutateur.



Les stations consommant le plus de bande passante seront de préférence connectées au commutateur. C'est le cas des serveurs qui concentrent toutes les connexions des utilisateurs.

Les commutateurs apportent, par ailleurs, de nouvelles fonctionnalités et ils peuvent être préférés aux concentrateurs rien que pour cela, indépendamment de tout problème de charge réseau.

Étant donné qu'un commutateur crée un segment par port, il est possible de combiner les débits (10, 100 et 1 000 Mbit/s) au sein de la même boîte. Il est à noter que seul le mode *store and forward* le permet tandis que le mode *cut-through*, quant à lui, ne bénéficie pas de

LE POINT SUR L'ETHERNET FULL DUPLEX (IEEE 802.3X)

Le mode full duplex conserve le débit nominal de 10, 100 ou 1 000 Mbit/s, mais sépare les canaux émission et transmission, et donc, élimine le besoin de la détection de collision CSMA/CD employée par l'Ethernet classique (en half duplex).

Le mode full duplex n'est possible :

- que sur des liaisons point à point entre un PC et un commutateur ou entre deux commutateurs ;
- qu'avec des câbles qui séparent physiquement les canaux émission et réception, c'est-à-dire le cuivre en paires torsadées et la fibre optique, mais pas les câbles coaxiaux ;
- qu'avec les commutateurs qui sont seuls capables de stocker les données à envoyer lorsque le canal émission est occupé.

L'absence de détection de collision CSMA/CD a une conséquence importante : **la limitation de longueur des câbles imposée par le délai de propagation des collisions disparaît**. Seule l'atténuation du signal, et donc la performance du câble, limite désormais la distance entre deux équipements : en pratique, on atteint des valeurs comprises entre 150 et 200 m au lieu des 100 m habituels, et plusieurs centaines de kilomètres en fibre optique.

cette possibilité car les trames sont commutées à la volée (elles ne sont pas stockées en mémoire).

On trouve sur le marché les formules suivantes :

- ports à détection automatique de vitesse (*port autosense*, 10/100 Mbit/s) ;
- port uplink à 100 Mbit/s ou 1 Gbit/s.

Les cartes réseau autosense 10/100 posent des problèmes avec les commutateurs qui sont également autosense, car il n'y a pas de négociation de débit ; chacun essaie de se caler sur la vitesse de l'autre. Il est donc recommandé de désactiver cette fonction au niveau de la carte et de fixer la vitesse manuellement (configuration à l'aide de Windows).

Les commutateurs permettent également d'augmenter les débits de plusieurs manières :

- avec le mode full duplex entre un PC et un port du commutateur ou entre deux commutateurs ;
- en agrégeant plusieurs ports full duplex du commutateur (technique du *port trunking*) pour le relier à un autre commutateur.

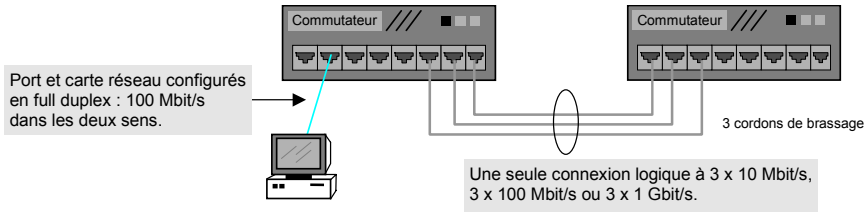


Figure 3-4. *Augmentation des débits grâce aux commutateurs.*

Le mode full duplex ne permet pas d'augmenter la vitesse de transmission, mais consiste à séparer les canaux émission et réception. Le débit global est ainsi augmenté, et peut, théoriquement, être multiplié par deux. Dans la pratique, ce mode est donc intéressant pour les serveurs et les liaisons intercommutateurs qui peuvent avoir à traiter un flux simultané dans les deux sens.

La carte réseau doit également supporter le mode full duplex.

À l'inverse, l'agrégation consiste à créer une seule liaison logique constituée de plusieurs liaisons physiques. Le débit obtenu est alors égal à la somme des débits des ports agrégés. Cette technique, normalisée 802.3ad, est utilisée pour interconnecter deux commutateurs généralement de même marque. Certaines cartes réseau supportent ce mode de fonctionnement.

Fonctionnalité	Concentrateur	Commutateur
Segment	Un seul pour le concentrateur	Un par port
Segmentation	(Optionnel) Réseaux indépendants (c'est-à-dire isolés, ne pouvant pas communiquer entre eux)	Segments interconnectés par la matrice de commutation
Mélange des débits	Non	Oui pour le mode *store and forward*
Port Uplink	Pour chaîner les hubs ; 4 maximum en cascade	Pour chaîner les commutateurs entre eux
Administrable	SNMP en option	SNMP + RMON en option
Autres fonctionnalités	Partitionnement des ports sur erreur	Détection automatique 10/100 Mbit/s Full duplex Agrégation de ports VLAN

Les VLAN (*Virtual LAN*) seront étudiés au chapitre 12. Pour l'instant, nous n'en avons pas besoin.

L'architecture

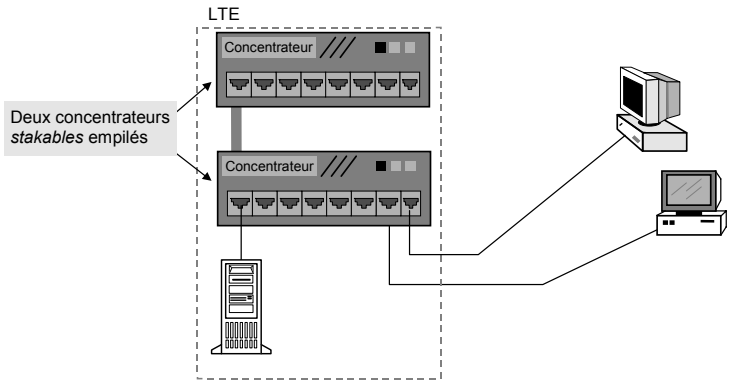
Mise en place d'un réseau local d'étage

Ayant en tête toutes les possibilités des équipements à notre disposition, la conception d'une architecture réseau simple consiste à assembler les concentrateurs et les commutateurs en exploitant au mieux les capacités du câblage.

Partons d'un cas simple : une cinquantaine de PC situés au même étage d'un immeuble quelconque. Les utilisateurs ont juste besoin d'échanger des données entre eux et de partager des applications (un serveur web, une base de données, des traitements de texte, etc.).

Le schéma suivant décrit l'architecture de base qui en résulte.

Figure 3-5.
Réseau local sur un étage.



C'est on ne peut plus simple :

- En vertu de ce qui a été dit aux sections précédentes, deux concentrateurs empilables 10bT ont été installés et chaînés entre eux sur un bus externe.
- Chaque poste est raccordé à un port d'un concentrateur *via* un système de câblage tel que décrit au chapitre précédent : un câblage en cuivre à paires torsadées (avec, bien sûr, des prises RJ45) centré en étoile autour d'un local technique.
- Les serveurs sont ici situés dans le local technique, et chacun d'entre eux est raccordé directement à un port d'un concentrateur *via* un cordon de brassage RJ45/RJ45.

Il n'y a pas de question à se poser.

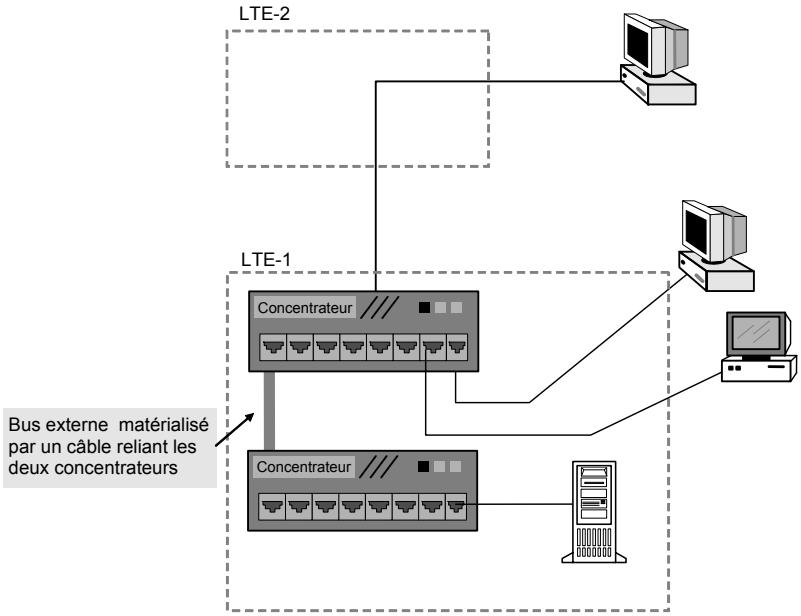
QUELS CORDONS DE BRASSAGE ?

Pour connecter un PC à un hub, un cordon de brassage **droit** (de même nature que le câblage en cuivre à paires torsadées) doit être utilisé. Pour connecter deux hubs entre eux, un cordon **croisé** (paires émission et réception) doit être utilisé, sauf si le port uplink est utilisé.

Extension du réseau d'étage

Maintenant, des utilisateurs, situés à l'étage au-dessus, ont les mêmes besoins.
Pas de problème : il y a moins de dix PC, ce qui nous permet d'utiliser les rocades en cuivre existantes, et de les connecter directement aux concentrateurs qui sont déjà installés.

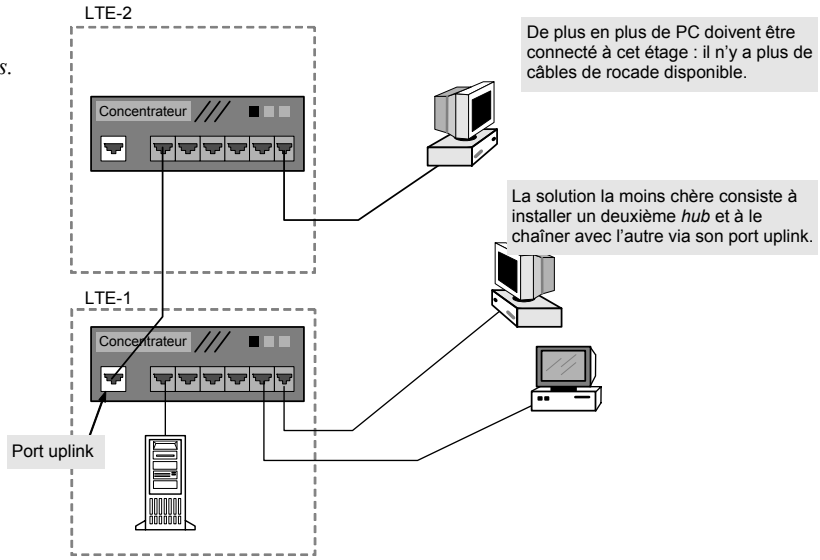
Figure 3-6.
Extension d'un réseau local sur plusieurs étages.



S'il y a davantage de PC, ou, d'une manière générale, s'il n'y a plus de câbles de rocade en nombre suffisant (en général, entre 6 et 12), il suffit de créer un autre réseau sur le même modèle. Le problème est maintenant de connecter les réseaux construits sur les deux étages pour que tout le monde puisse accéder aux mêmes données et aux mêmes applications.

La solution la plus simple consiste à connecter les deux concentrateurs en cascade, *via* un câble de rocade en cuivre.

Figure 3-7.
Un réseau local étendu sur plusieurs étages.



Conception d'un réseau d'immeuble

Maintenant, la situation se corse un peu : le réseau est un succès, il y a de plus en plus de demandes de connexions, la société utilise de plus en plus l'informatique. Il faut maintenant créer des réseaux à chaque étage. La solution précédente, consistant à chaîner les concentrateurs n'est plus applicable, car on est limité par le nombre de cascades possible. De plus, au-delà de cent postes connectés, le réseau Ethernet deviendrait saturé (du fait du nombre de collisions qui augmente avec le nombre d'utilisateurs) et, en définitive, les temps de réponse seraient trop grands.

L'installation de boîtes doit maintenant faire place à une plus grande réflexion, c'est-à-dire à un travail d'architecture. Imaginons donc que nous ayons trois cents utilisateurs répartis sur une demi-douzaine d'étages, soit en moyenne cinquante postes par étage, plus les imprimantes et les serveurs. On se retrouve avec soixante à soixante-dix connexions par étage.

Comme précédemment, on peut installer une pile de concentrateurs à chaque étage pour créer un réseau local d'étage. Là encore, il faut se poser de nouveau les mêmes questions :

- Quel débit : 10bT, 100bT ou Gigabit ?
- Quelle technologie : concentrateurs ou commutateurs ?
- Un seul segment Ethernet ou plusieurs ?

Les réponses à ces questions dépendent avant tout du trafic prévisionnel, des perspectives d'évolution et des performances mises en balance par rapport au coût. Si l'immeuble comprend trente autres étages, on peut supposer qu'il faudra tôt ou tard étendre le réseau. Si, en revanche, l'immeuble n'en comprend que six, on sait que la configuration sera figée pour un bon moment.

L'architecture doit donc être conçue pour couvrir les besoins futurs et non seulement ceux du moment. Elle doit donc être évolutive, c'est-à-dire être bâtie sur des équipements que l'on pourra récupérer (recycler pour d'autres usages).

Mise en place d'un réseau fédérateur

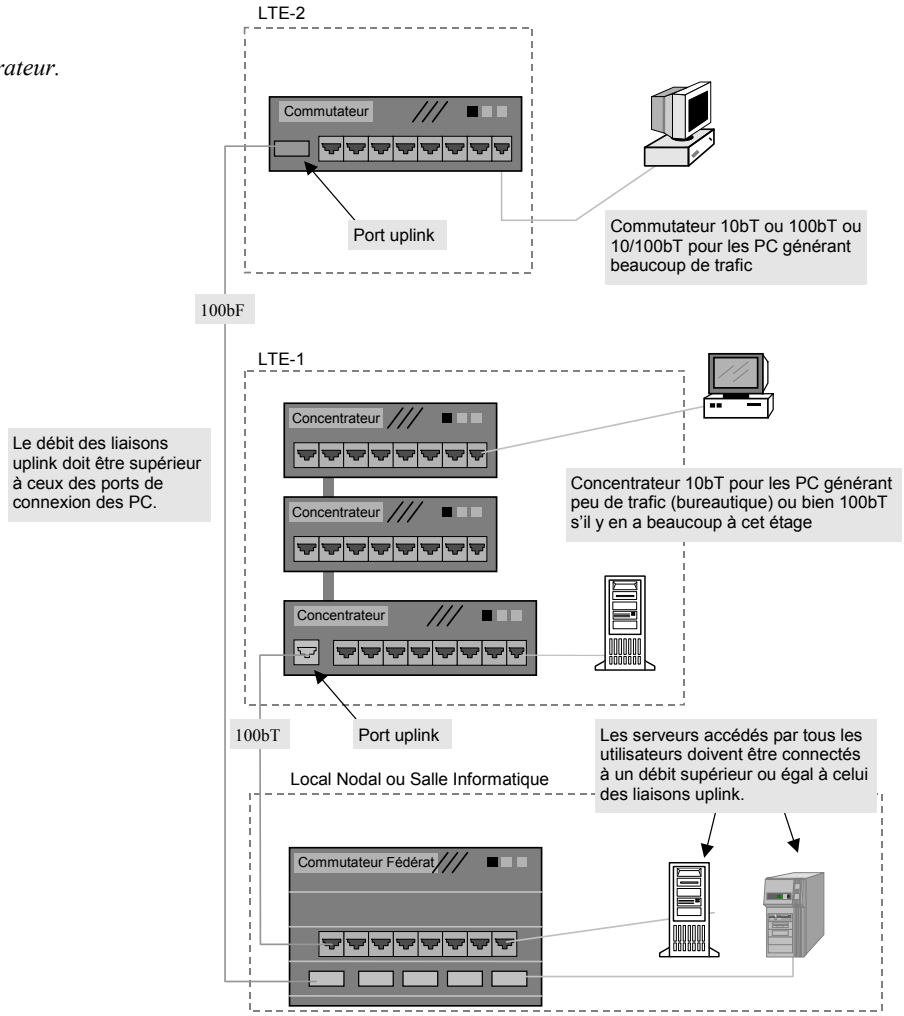
Si les besoins en trafic sont importants (applications multimédias, applications, voix, données et visioconférences), on peut envisager des commutateurs 10bT à tous les étages. Dans la plupart des cas, un débit de 10 Mbit/s suffira, mais la différence de coût étant minime, on peut envisager le 100bT.

La question est maintenant de savoir comment connecter les réseaux locaux entre eux. La solution repose sur la création d'un réseau fédérateur (*backbone*).

On peut imaginer un câble Ethernet (topologie en bus), ou FDDI (anneau), qui parcourt tous les étages et auquel on connecte les concentrateurs. C'est une solution peu évolutive, car le débit est limité à la technologie utilisée (100 Mbit/s pour FDDI). En outre, le Gigabit Ethernet et l'ATM ne sont pas prévus pour une topologie en bus ou en anneau. C'est également une solution peu sûre : le câble étant un élément passif, il n'y a aucun moyen de le superviser à distance.

L'architecture couramment utilisée est de type *collapse backbone* (littéralement, réseau fédérateur effondré). Le principe consiste à concentrer le backbone en un seul point : au lieu d'avoir un réseau qui parcourt tous les étages, le backbone est réalisé dans un commutateur unique. Cela revient à créer une architecture en étoile à deux niveaux, un premier concentrant les PC à chaque étage et un second concentrant les équipements d'étage en un point central, en général la salle informatique ou un local nodal dédié aux équipements réseau.

Figure 3-8.
Conception d'un réseau fédérateur.



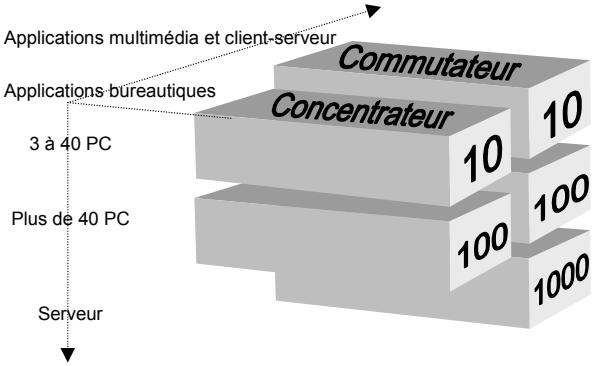
Pour un réseau de taille moyenne (de 200 à 800 utilisateurs), l'équipement central doit être de grande capacité en termes d'accueil et de performances. Le choix se portera donc sur un châssis qui offre une matrice de commutation à haut débit. Le réseau fédérateur n'est alors pas limité à 10 Mbit/s, mais à 100 Mbit/s par étage et à la capacité de la matrice de commutation du commutateur central, généralement plusieurs gigabits.

Quel débit et quelle technologie ?

Le choix du débit du réseau fédérateur dépend de celui utilisé par les PC.

Si les PC sont connectés à un…	Le débit des liens uplink vers le commutateur central doit être au moins égal à…
Concentrateur à 10 Mbit/s	10 Mbit/s
Concentrateur à 100 Mbit/s	100 Mbit/s
Commutateur 10 Mbit/s	100 Mbit/s
Commutateur 100 Mbit/s	1 Gbit/s

Le choix du débit des PC dépend, quant à lui, du volume de trafic généré et du type d'application (du flux bureautique au flux multimédia). Mais le passage du concentrateur au commutateur évite ou retarde l'augmentation du débit, ce qui permet de conserver les cartes réseau existantes dans les PC.



Indépendamment de la charge réseau et du nombre de PC, certaines combinaisons sont plus appropriées que d'autres.

COMMENT FONCTIONNE UN COMMUTATEUR ?

Un commutateur permet d'**interconnecter** plusieurs segments Ethernet. Sur chacun de ses ports, on peut raccorder un concentrateur (plusieurs PC partagent alors la bande passante sur ce port) ou un seul PC (technique de la **microsegmentation**).

Afin de limiter le trafic réseau inutile, les trames Ethernet échangées entre les PC d'un même segment ne sont pas transmises sur les autres segments gérés par le commutateur (voir encart " Le point sur Ethernet ").

Pour savoir quelles sont les trames qui doivent sortir du segment et celles qui ne le doivent pas, le commutateur regarde toutes les adresses MAC sources des trames qui entrent sur chacun de ses ports, et les enregistre dans ses **tables d'adresses** (il existe une table par port). C'est le mécanisme d'**apprentissage**. Si un PC est déplacé d'un port à l'autre ou d'un commutateur à l'autre, son adresse MAC peut se retrouver dans deux tables et créer ainsi des conflits. Pour éviter cela, les adresses sont effacées de la table au bout de 15 à 30 secondes.

Les commutateurs d'entrée de gamme ne disposent que de très peu de mémoire et ne peuvent apprendre qu'une, deux ou quatre adresses MAC par port. Ils sont plutôt dédiés à la microsegmentation. Les commutateurs fédérateurs doivent en revanche disposer de beaucoup de mémoire et être capables d'enregistrer plusieurs milliers d'adresses MAC, car ils fédèrent tous les flux interréseaux locaux (c'est-à-dire intersegment Ethernet).

La **mémoire tampon** doit également être suffisamment importante pour permettre l'adaptation des débits (entre 10 et 100 Mbit/s, et surtout entre 10/100 et 1 Gbit/s).

L'interconnexion des segments est réalisée par une **matrice de commutation** à haut débit capable de supporter la somme des débits des ports (8 × 10 Mbit/s, par exemple). Là encore, les commutateurs fédérateurs doivent comprendre une matrice de commutation très puissante (généralement des ASIC et des processeurs RISC).

Type de trafic	Réseau d'étage	Fédérateur
Applications bureautiques (Word, Excel, bases de données)	Concentrateurs 10bT	Commutateurs 10/100bT
Client-serveur et un peu de multimédia (voix et vidéo)	Commutateurs 10/100bT	Commutateurs 100bT
Applications multimédias intensives	Commutateurs 100bT	Commutateurs Gigabit

Le coût est un autre critère de décision, sans doute le plus important. Lors du choix d'une technologie (concentration ou commutation) et du débit, il faut tenir compte du nombre de cartes réseau pour les PC ainsi que du nombre d'équipements.

Équipement actif 8 ports	Coût en Euros HT
Carte 10bT	De 38 à 120
Cartes 100bT et 10/100bT	De 75 à 185
Concentrateur 10bT	De 75 à 770
Concentrateur 100bT (ou 10/100bT)	De 385 à 1 000
Commutateur 10bT	De 770 à 2 300
Commutateur 100bT (ou 10/100bT)	De 2 300 à 3 000
Concentrateur 1000bT	N'existe pas
Commutateur 1000bT	De 7 600 à 13 800

Les écarts de prix sont dus à des différences dans les fonctionnalités proposées (concentrateur administrable ou non, empilable ou non, avec ou sans slot d'extension, etc.). Un exemple en a été donné au premier chapitre concernant les concentrateurs 10bT. Le choix de la fibre optique fait également monter les prix.

Suivre l'évolution des besoins

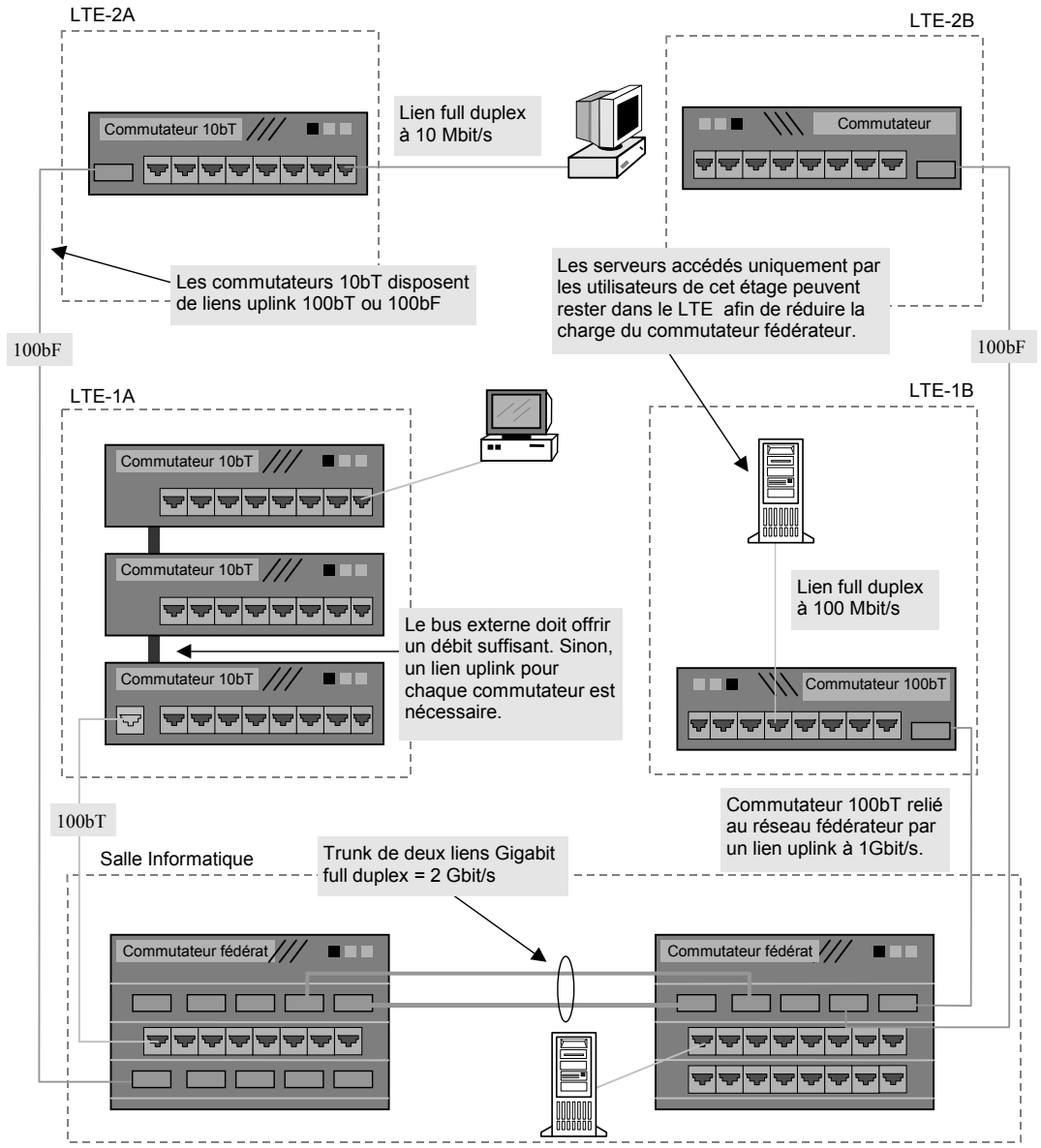
La direction a maintenant décidé de regrouper plusieurs sites sur le nôtre et de louer l'autre aile de l'immeuble. Le nombre d'utilisateurs va ainsi passer de 300 à 500. Ils seront répartis sur 10 étages, soit une moyenne de 70 connexions par étage en comptant les imprimantes, les doubles connexions, etc.

Par ailleurs, les utilisateurs de bases de données client-serveur se plaignent des mauvais temps de réponse. Il s'agit de gros consommateurs de bande passante (gestion électronique de documents, applications décisionnelles de type datawharehouse, etc.) qui perturbe les autres trafics. Les utilisateurs qui se connectent sur des serveurs Unix se plaignent également de ralentissements brusques lors de l'affichage des écrans et du déplacement du curseur à l'écran.

Pour faire face à cette montée en charge, nous avons plusieurs solutions, que nous pouvons combiner.

En premier lieu, les commutateurs peuvent être généralisés. Cela permet de segmenter le réseau afin qu'un flux important ne viennent en perturber un autre. Ensuite, les débits peuvent être augmentés à 100 Mbit/s pour les plus gros consommateurs. Afin de limiter le coût du réseau, les autres utilisateurs peuvent rester connectés à 10 Mbit/s.

Figure 3-9.
Utilisation des capacités des commutateurs pour monter en charge.



Au niveau du réseau fédérateur, la capacité d'accueil doit être augmentée. La sécurité de fonctionnement doit également être améliorée, car la panne de l'équipement central entraînerait l'arrêt de tout le réseau. Le ou les commutateurs empilables doivent donc faire place à un châssis plus performant (en termes de bande passante de la matrice de commutation) et qui offre des alimentations redondantes.

L'ajout d'un second commutateur, identique au premier et donc interchangeable, renforce encore la fiabilité du réseau fédérateur.

Afin de ne pas créer un engorgement entre les deux commutateurs, un lien à très haut débit doit être créé. Cela est réalisé en agrégeant plusieurs ports Gigabit (technique du *port trunking*).

Les serveurs très sollicités et certains postes de travail consommateurs de bande passante peuvent voir leurs performances d'accès réseau accrues par des connexions full duplex.

Assurer la continuité de service

Près de 500 utilisateurs sont connectés. Le réseau devient un élément stratégique du système d'information. En effet, sans lui, les utilisateurs ne peuvent plus travailler : plus de connexion au système central, plus d'accès aux bases de données, plus de messagerie, etc. Bref, il est nécessaire de construire une architecture redondante pour parer aux éventuelles pannes. Celles-ci ne sont pas un mythe. En voici quelques-unes qui arrivent fréquemment :

- panne d'alimentation ;
- coupure d'un câble entre les commutateurs (erreur de manipulation lors du brassage dans les locaux techniques) ;
- débranchement d'un câble, d'un convertisseur ou de tout autre petit boîtier : eh oui, un *transceiver* ou un connecteur mal enfoncés dans la prise tendent non pas à se remettre en place d'eux-mêmes (dommage), mais plutôt à tomber sous l'effet des vibrations et de la pesanteur...

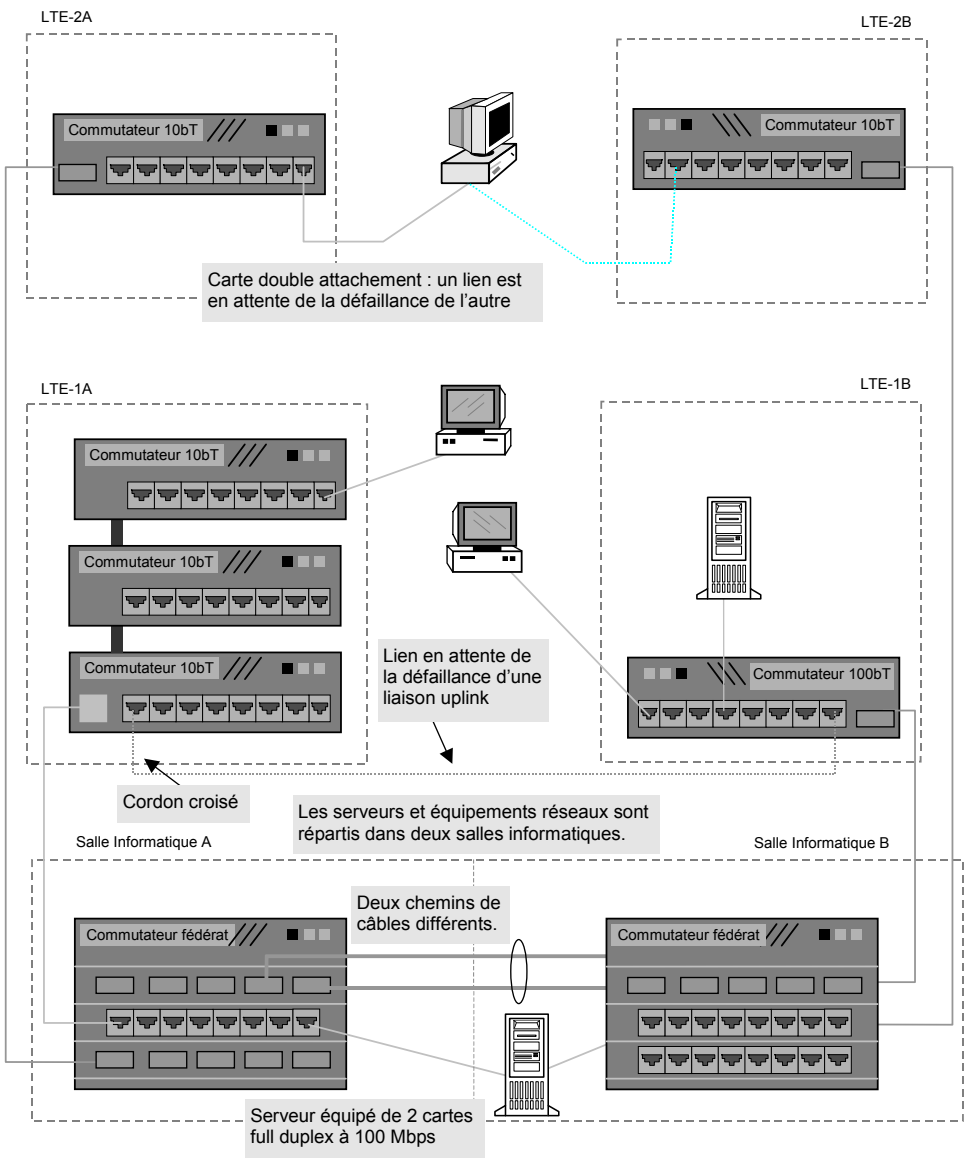
L'incendie ou l'inondation sont plus rares, mais, quand ils se produisent, il n'est pas question de reconstruire un système de câblage et un réseau avant quelques heures, voire plusieurs jours.

La mise en place d'un réseau redondant repose d'une part sur un système de câblage qui offre l'alternative de plusieurs cheminements, d'autre part sur la mise en place d'équipements de secours.

La redondance est généralement limitée au réseau fédérateur, mais, dans des cas critiques (salles de marché, processus industriels en flux tendus, etc.), il peut être nécessaire de l'étendre jusqu'au poste de travail.

On peut alors prévoir d'équiper ces derniers en carte deux ports : quand le premier perd le contact avec le port du concentrateur, il bascule sur le second, connecté à un autre concentrateur situé dans un autre local technique. Côté serveur, on peut envisager la même solution ou encore prévoir deux cartes ou plus, ce qui offre l'intérêt du partage de charge en fonctionnement normal.

Figure 3-10. *Cas d'un réseau redondant.*



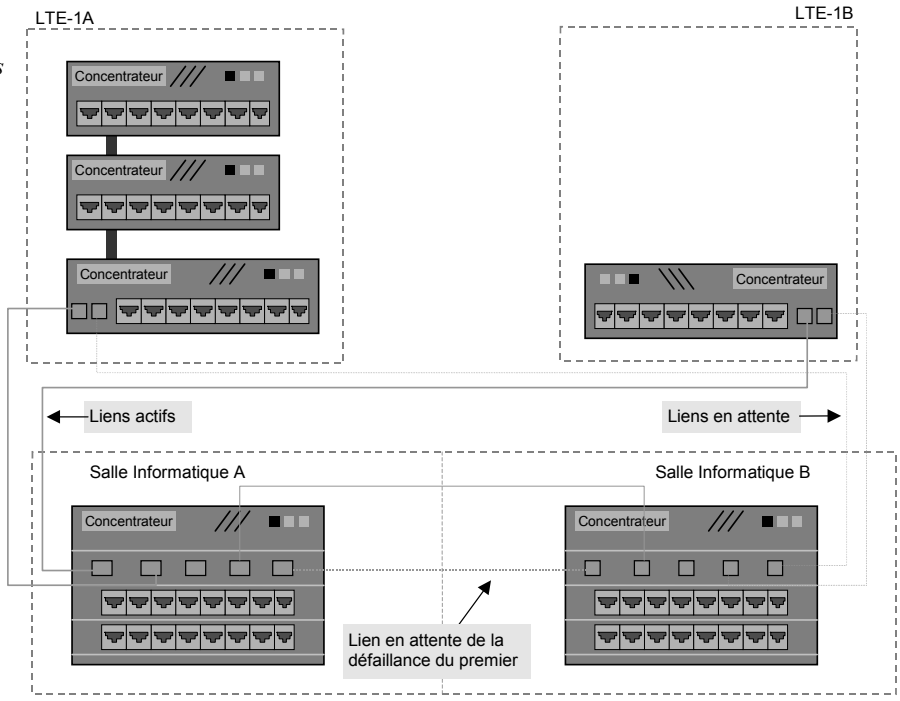
Les équipements fédérateurs sont ici des commutateurs reliés entre eux pour former un seul réseau fédérateur. Cette liaison doit être de débit élevé afin d'absorber tous les flux issus des étages et allant vers un serveur connecté au commutateur fédérateur opposé. On peut envisager ici du Gigabit Ethernet. Les liaisons entre les commutateurs fédérateurs et les équipements d'étage peuvent être réalisées en cuivre ou en fibre optique, le choix dépendant de la distance.

L'agrégation de liens offre ici une protection supplémentaire car chaque lien passe par deux gaines techniques différentes : en cas de perte d'une des liaisons (coupure du câble, incendie, etc.), l'autre reste opérationnelle et permet un fonctionnement en mode dégradé à 1 Gbit/s au lieu de 2 Gbit/s.

Il est à noter que cette architecture redondante n'est possible que parce que les équipements sont des commutateurs. Ceux-ci échangent des informations sur la topologie du réseau et déterminent les routes actives et celles qui ne le sont pas. Le protocole utilisé s'appelle le **spanning tree**.

La même architecture avec des concentrateurs serait impossible car le segment Ethernet ferait une boucle, ce qui est interdit par la norme IEEE. Pour assurer le même niveau de redondance, une configuration équivalente avec des concentrateurs serait donc beaucoup plus complexe et nécessiterait de doubler le nombre de câbles. De plus, seule l'utilisation de la fibre optique permet ce type de redondance.

Figure 3-11.
Redondance avec des concentrateurs Ethernet.

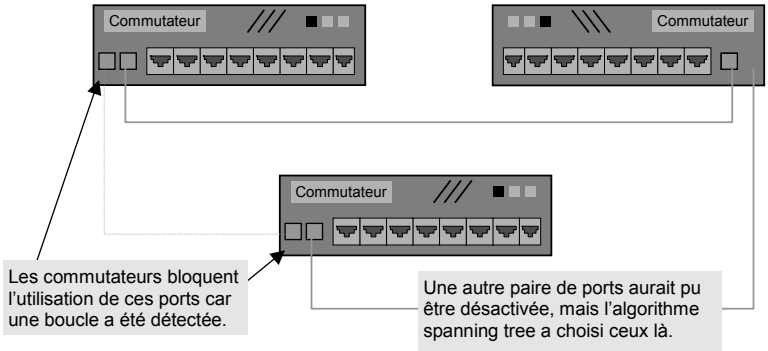


Les liaisons en pointillé sont inactives en fonctionnement normal. En cas de rupture de la liaison principale, le concentrateur d'étage active la seconde liaison en quelques microsecondes.

La liaison entre les deux concentrateurs fédérateurs doit être doublée. En effet, si elle était perdue, on se retrouverait avec deux réseaux isolés ayant le même numéro de réseau IP.

Inversement, avec le spanning tree, les liens sont tous actifs, mais une des routes est invalidée par les commutateurs. En cas de rupture d'un lien, une nouvelle route sera calculée, généralement au bout de trente secondes.

Figure 3-12.
La redondance avec le spanning tree.



L'échange de trames Ethernet

Un réseau local tel qu'Ethernet est constitué de concentrateurs et de commutateurs, deux équipements au comportement bien différent.

Caractéristique	Ethernet partagé	Ethernet commuté
Équipement	Concentrateur	Commutateur
Segment	Un segment partagé par tous les ports	Un segment par port
Architecture matérielle	Composants électroniques ; un ou plusieurs bus Ethernet	ASIC et processeur RISC, matrice de commutation
Architecture logicielle	Petit logiciel pour des options de configuration et pour l'administration SNMP	Logiciel plus complexe, mais traitement essentiellement matériel
Algorithme de routage	Aucun	Spanning tree
Traitement des trames	Aucun traitement (transparent pour les trames Ethernet)	Filtrage et transmission (*forward*) des trames ; apprentissage des adresses MAC
Couche réseau	Niveau 1 = couche physique	Niveaux 1 et 2 = couche physique et liaison
Fonction	Connecter plusieurs PC au sein d'un segment	Interconnecter plusieurs segments

Échange de trames sur un segment Ethernet

Commençons par étudier le fonctionnement au sein d'un segment Ethernet partagé (la trame circule sur un seul segment).

La norme Ethernet spécifie l'utilisation d'adresses physiques liées aux cartes réseau : les adresses MAC.

QU'EST-CE QUE LE SPANNING TREE ?

Lorsqu'un commutateur reçoit une trame, il recherche son adresse de destination dans ses tables d'adresses MAC (une par port). Si elle ne s'y trouve pas, il la transmet sur tous ses ports. Sinon, il l'envoie uniquement sur le port identifié. C'est le mécanisme de **forwarding**.

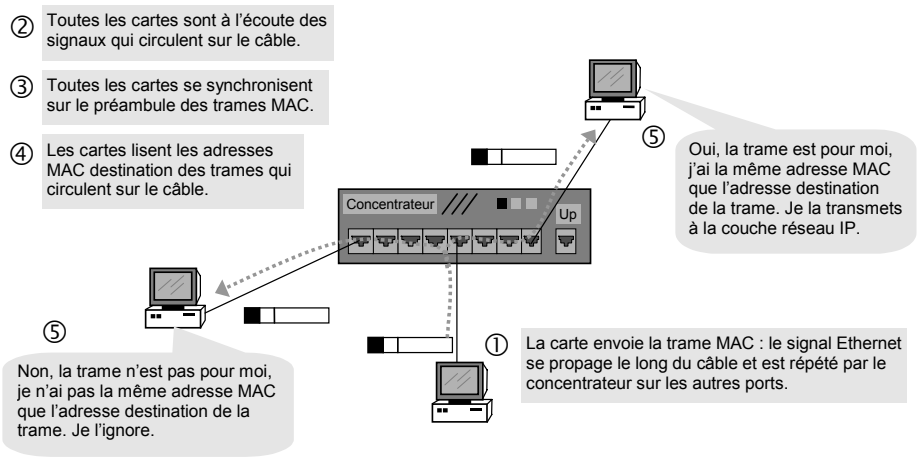
S'il existe plusieurs chemins, par exemple entre un commutateur d'étage et deux fédérateurs, la trame risque de boucler, car chacun des commutateurs transmet la trame (*forward*).

Le spanning tree est un **protocole de routage de niveau 2** associé à un algorithme qui permet d'éviter que les trames bouclent dans le réseau. Les commutateurs s'échangent des trames *spanning tree* et calculent une route en invalidant les chemins multiples susceptibles de créer des boucles au sein du segment Ethernet.

La carte recevant une trame Ethernet ne la prendra en compte que si l'adresse MAC de destination de la trame est identique à celle qui est inscrite dans sa mémoire. La seule exception à cette règle concerne les adresses de broadcast et, éventuellement, les adresses multicast.

En résumé, les cartes sont programmées pour accepter les trames qui leur sont destinées, plus toutes les trames de broadcast ainsi que les trames multicast qui ont été configurées.

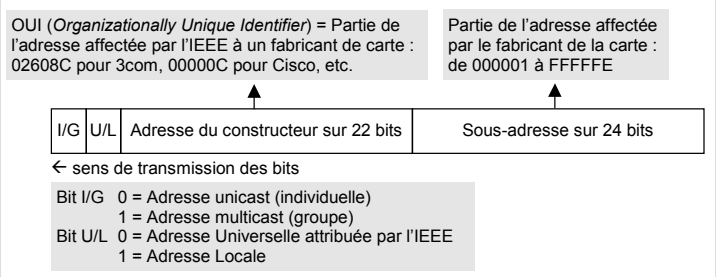
Figure 3-13. *Échange des trames Ethernet.*



Les trames qui sont acceptées sont remises au protocole de niveau 3, qui correspond à l'identifiant trouvé dans le champ « Type », soit 0×0800 pour IP. Celles qui ne sont pas destinées à la station sont ignorées.

LES ADRESSES MAC (IEEE 802.3)

Une adresse MAC est constituée de 6 octets et s'écrit sous forme hexadécimale séparé par des tirets, comme par exemple 02-60-8C-EB-25-D2. L'adresse de broadcast est FF-FF-FF-FF-FF-FF.

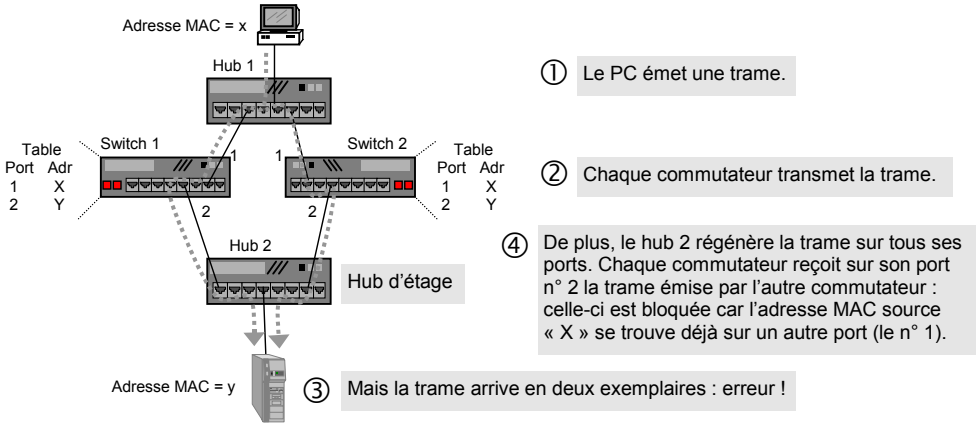


Le schéma montre l'adresse dans l'ordre de transmission des bits, alors qu'en représentation hexadécimale, l'ordre des bits au sein de chaque octet est inversé (cf. exemple au début du chapitre 15). Ainsi, le premier octet " 02 " indique que le bit U/L est à " 1 ". Le site standards.ieee.org/search.html permet de rechercher les OUI affectés par l'IEEE.

Échange de trames entre différents segments Ethernet

On l'a vu, le commutateur est un équipement permettant de segmenter le réseau Ethernet et d'interconnecter ces différents segments. Pour décider si la trame doit passer d'un segment à l'autre, il se base sur son adresse MAC qu'il compare avec celles se trouvant dans ses tables d'adresses en mémoire.

On peut envisager différentes situations dans lesquelles il existe plusieurs chemins possibles entre deux stations.

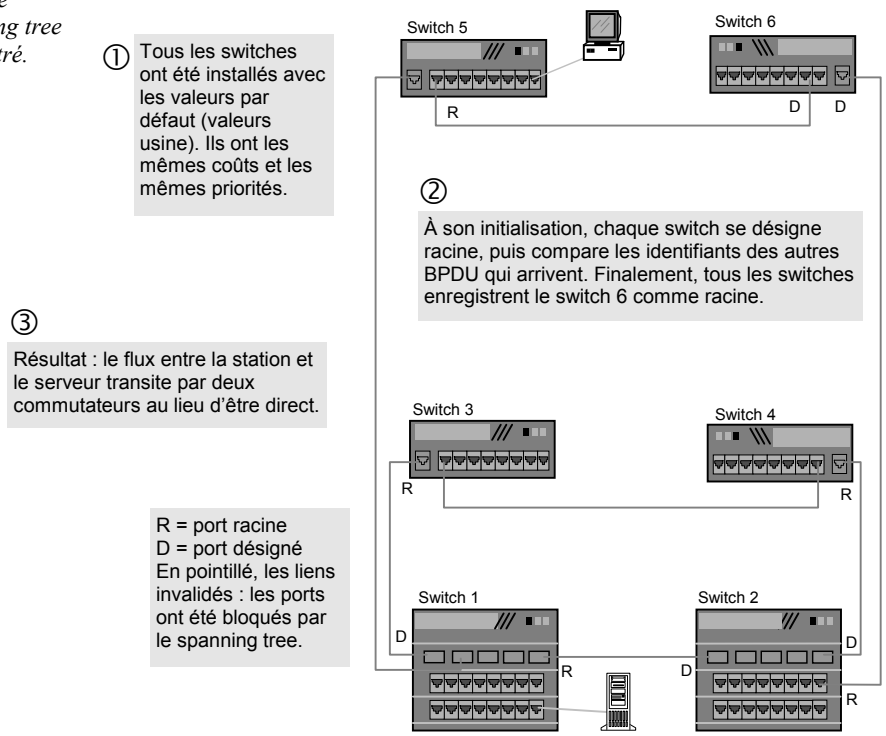


Il faut noter que, si le serveur n'a jamais émis de trame, son adresse MAC « Y » n'est pas encore connue des commutateurs. Ces derniers transmettront alors la trame sur tous les autres ports, dont le port n° 2. De plus, la même situation se produit pour toutes les trames de broadcast et de multicast.

Pour éviter ces problèmes, il faut qu'une des deux routes soit interdite. C'est là qu'intervient le **spanning tree**. Le but de ce protocole est de définir une route unique vers un commutateur désigné racine en se basant sur des coûts et des priorités.

Compte tenu des processus d'élection, il est important de bien paramétrer les coûts (exprimés en nombre de sauts et/ou dépendant du débit du port) ainsi que les priorités. Les valeurs par défaut (fixées en usine) peuvent en effet aboutir à choisir des chemins qui ne sont pas les meilleurs. Prenons le cas de notre réseau local redondant.

Figure 3-14.
Conséquence d'un spanning tree mal paramétré.



LE POINT SUR LE SPANNING TREE (IEEE 802.1D)

L'algorithme du spanning tree consiste à construire un arbre définissant un chemin unique entre un commutateur et sa racine.

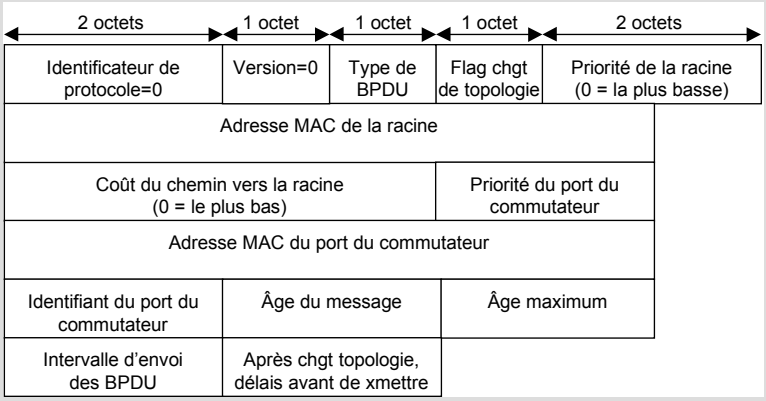
Lors de la construction de l'arbre (suite à un changement de topologie), chaque commutateur émet un **BPDU** (*Bridge Protocol Data Unit*) de configuration sur tous ses ports. Inversement, il retransmet tous les BPDU (éventuellement en les modifiant) qui lui arrivent, et ainsi de suite jusqu'à ce que les BPDU échangés contiennent tous la même valeur.

La première étape de ce processus consiste à élire un **commutateur racine** : c'est celui dont la **priorité** est la plus basse ou, en cas d'égalité, celui dont l'adresse MAC est la plus basse.

Ensuite, chaque commutateur détermine le **port racine** — celui par lequel un BPDU émis par la racine arrive. S'il y en a plusieurs, le port choisi est celui qui a le **coût** de chemin vers la racine le plus bas. Le coût est déterminé par la somme des coûts des ports situés entre le commutateur et la racine. En cas d'égalité, le port choisi est celui qui a la priorité la plus basse ; en cas de nouvelle égalité, c'est celui qui a l'adresse MAC la plus basse.

Enfin, sur chaque segment Ethernet, le commutateur dont le port racine a le coût de chemin vers la racine le plus bas est élu **commutateur désigné**. En cas d'égalité, c'est celui qui a la priorité la plus basse et, en cas de nouvelle égalité, celui qui a l'adresse MAC la plus basse.

En définitive, sur chaque segment Ethernet, un seul chemin vers le commutateur racine sera calculé. Les commutateurs désactivent tous leurs ports qui ne sont ni racines ni désignés.



Afin de détecter les changements de topologie (apparition ou disparition d'un commutateur), la racine envoie régulièrement (toutes les deux secondes par défaut) un BPDU d'annonce (comprenant seulement les trois premiers octets) sur tous ses ports. Les commutateurs transmettent ce BPDU sur leurs ports désignés.

Les BPDU sont envoyés dans des trames Ethernet multicast d'adresse MAC 01:80:C2:00:00:10 transportant les trames LLC identifiées par le SAP 0x42.

Tous les commutateurs ont la même priorité (par exemple 32 768) et le même coût sur chaque port de même débit (par exemple 19 pour les ports à 100 Mbit/s). Le switch 6 a été désigné racine parce que son identifiant (priorité + adresse MAC) était le plus bas. Le même processus de sélection a déterminé les routes menant vers la racine uniquement en se fon-

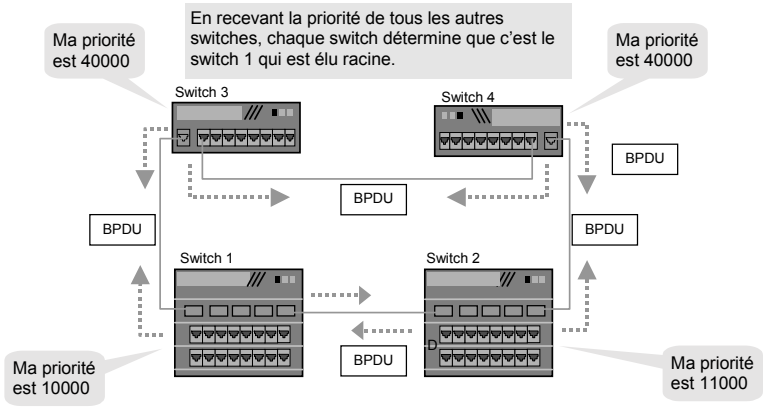
dant sur les valeurs des adresses MAC, puisque toutes les autres valeurs (priorité et coût) sont identiques.

Résultat, certains flux ne sont pas optimisés et peuvent dégrader les performances.

Reprenons les différentes phases de calcul de l'arbre spanning tree. Les commutateurs désignent la racine. Afin d'optimiser les flux, il est préférable que ce soient les commutateurs fédérateurs qui assurent ce rôle. Leur priorité doit donc être abaissée par rapport aux commutateurs d'étage. Étant donné qu'il s'agit de Catalyst 5000, la commande est la suivante :

```
Console> (enable)set spantree priority 10000
VLAN 1 bridge priority set to 10000.
```

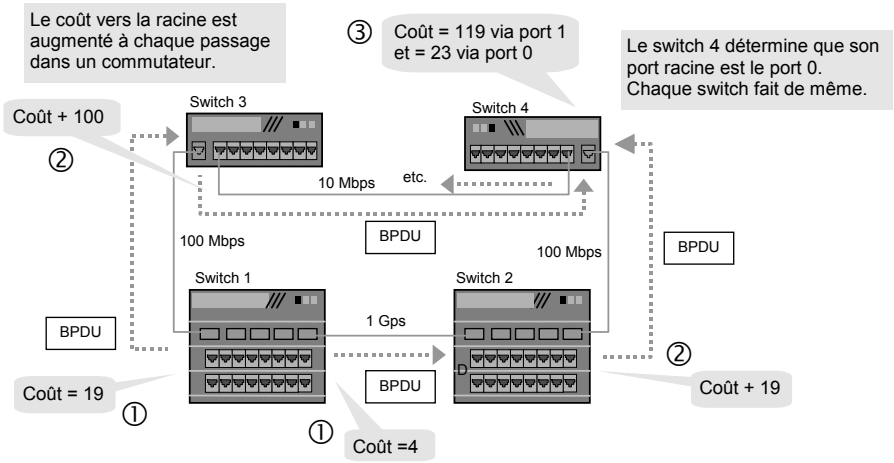
Figure 3-15.
Élection du commutateur racine.



Chaque commutateur choisit ensuite son port racine, celui dont le coût de chemin vers la racine est le plus bas. Sur un Catalyst, le coût de chaque port dépend de son débit : 4 pour 1 Gbit/s, 19 pour 100 Mbit/s, et 100 pour 10 Mbit/s. La commande suivante permet de changer la valeur par défaut :

```
Console> (enable)set spantree portcost 1/1 4
```

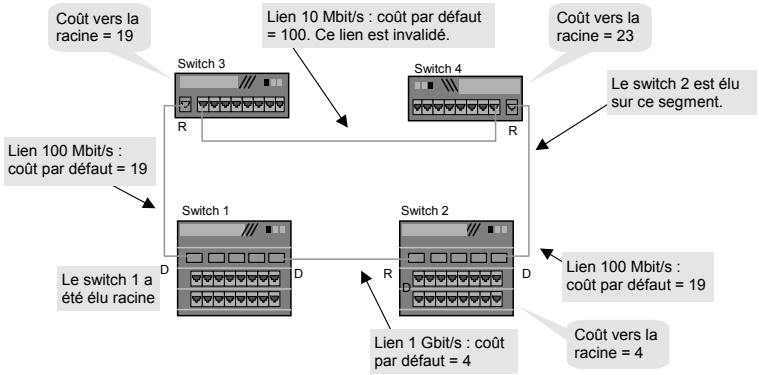
Figure 3-16.
Choix des ports racines.



Sur chaque segment Ethernet, le commutateur désigné est celui dont le port racine a le coût le plus bas. En cas d'égalité, la priorité détermine ce coût. Étant donné que tous les ports ayant un même débit ont le même coût et la même priorité par défaut, le choix s'effectuera en fonction de l'adresse MAC. Pour éviter les mauvaises surprises, il est possible d'abaisser la priorité d'un port pour être sûr qu'il soit désigné en cas de routes multiples :

```
set spantree portpri 1/1 32
```

Figure 3-17. *Élection des commutateurs désignés.*



En définitive, les chemins redondants ne sont pas utilisés, et le partage de la charge entre plusieurs routes n'est pas possible.

Les mêmes BPDU sont envoyés sur tous les ports, même là où il n'y a qu'un PC connecté. Le spanning tree peut donc être désactivé sur ces ports, ce qui présente l'avantage de diminuer (un peu) le trafic et d'éviter que des ajouts sauvages de commutateurs (qui seraient connectés sur ces ports) ne viennent perturber votre réseau.

```
set spantree disable
```

Le commutateur racine émet régulièrement (toutes les deux secondes par défaut) des BPDU pour maintenir l'état du spanning tree. Si le réseau est stable (peu d'incidents et de changements), il est possible d'augmenter cette valeur afin de diminuer le trafic

```
set spantree hello 5
```

On peut s'assurer que, sur le switch 1, le spanning tree s'est stabilisé dans une bonne configuration.

```
Console> (enable) show spantree
VLAN 1
Spanning tree enabled
Designated Root             00-1f-00-40-0b-eb-25-d2
Designated Root Priority    45
Designated Root Cost        0
Designated Root Port        1/1
```

```
Root Max Age   20 sec    Hello Time 2  sec   Forward Delay 20 sec
Bridge ID MAC ADDR          00-40-0b-eb-25-d2
Bridge ID Priority          45
Bridge Max Age 20 sec    Hello Time 2  sec   Forward Delay 20 sec

Port      Vlan  Port-State     Cost   Priority  Fast-Start
--------  ----  -------------  -----  --------  ----------
1/1       1     forwarding         4        32   disabled
1/2       1     forwarding        19        32   disabled
2/1       1     forwarding        19        32   disabled
2/2       1     not-connected     19        32   disabled
2/3       1     not-connected     19        32   disabled
3/1       1     not-connected    100        32   disabled
3/2       1     forwarding       100        32   disabled
```

Le processus de création de l'arbre spanning tree peut durer plusieurs dizaines de secondes. Ce temps est, en réalité, proportionnel au nombre de commutateurs.

Pendant cette phase, aucun commutateur ne traite de trame au cours des 15 premières secondes (valeur par défaut) ; le réseau s'arrête donc de fonctionner chaque fois qu'un commutateur est allumé ou éteint quelque part dans le réseau.

La phase d'apprentissage est encore plus longue lorsque tous les commutateurs s'initialisent en même temps (suite à une panne de courant, par exemple). En effet, les BPDU sont reçus par plusieurs ports dont l'un peut être élu racine, puis invalidé par la suite si un commutateur situé en aval a invalidé sa route. Le temps de stabilisation de l'arbre peut ainsi atteindre plusieurs minutes.

Dans certains cas, notamment sur les commutateurs fédérateurs, il peut être intéressant de diminuer ce temps, surtout si l'architecture réseau est conçue sans aucune boucle.

```
set spantree fwddelay 5
```

Inversement, si ce temps est trop court par rapport au délai de construction de l'arbre, des trames peuvent commencer à circuler et potentiellement être dupliquées dans le cas de routes multiples. Il vaut alors mieux augmenter le paramètre « forward delay » au-delà des 15 secondes par défaut.

La meilleure solution consiste à activer plus rapidement les ports qui ne sont pas concernés par le *spanning tree*, c'est-à-dire ceux sur lesquels est connectée une seule station.

```
set spantree portfast 1/2 enable
```

On peut remarquer que l'échange de BPDU et l'élection d'un commutateur désigné impliquent que chaque port du commutateur soit identifié par une adresse MAC (comme une carte réseau).

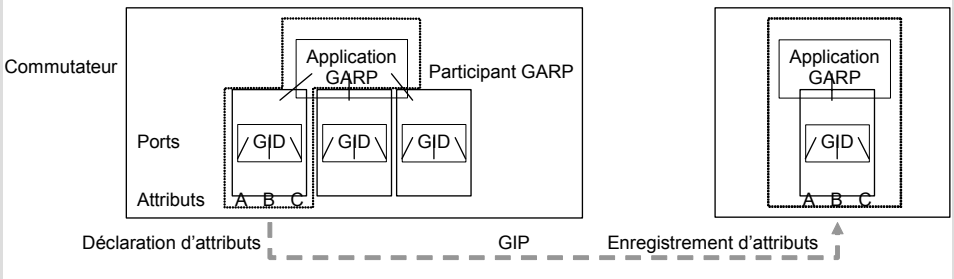
L'alternative à cette optimisation repose sur la nouvelle version du spanning tree, appelée **RSTP** (*Rapid Spanning Tree Protocol*) et issue du groupe de travail IEEE **802.1w**. Le protocole RSTP est compatible avec le spanning tree et sera, à terme, intégré dans le standard 802.1d.

LE POINT SUR GARP (IEEE 802.1D)

La norme 802.d spécifie, en plus du *spanning tree*, le protocole **GARP** (*Generic Attribute Registration Protocol*), qui permet aux commutateurs d'échanger des informations de gestion. Chaque commutateur peut ainsi déclarer aux autres commutateurs appartenir à un groupe possédant certains attributs.

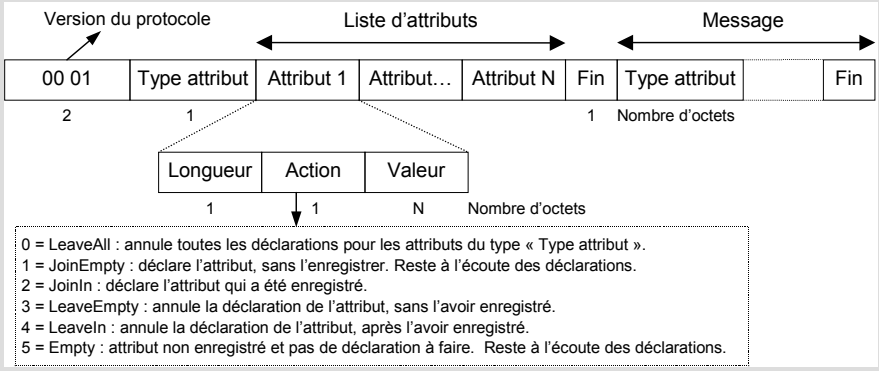
La même norme spécifie une application de GARP à l'enregistrement des groupes multicast, protocole appelé **GMRP** (*GARP Multicast Registration Protocol*). La norme 802.q utilise également GARP pour la propagation dynamique des VLAN entre commutateurs : le protocole est appelé **GVRP** (*GARP VLAN registration Protocol*).

Un commutateur fonctionne avec plusieurs participants GARP, un par application et par port du commutateur. Un tel participant est constitué de l'application et d'une instance du GID (*GARP Information Declaration*), qui gère, pour chaque attribut utilisé par l'application, l'état du participant (membre, observateur…) associé à l'état des attributs (enregistré, déclaré…). Les GID utilisent le protocole GIP (*GARP Information Propagation*) pour échanger ces attributs.



À l'aide de ces composants, un port recevant la déclaration d'un attribut l'enregistre et le déclare (i.e. le propage) aux autres ports. Inversement, l'annulation d'un attribut est propagée vers un port si tous les autres ports ont annulé l'enregistrement de cet attribut.

Le PDU (*Protocol Data Unit*) utilisé par GIP est constitué de messages contenant des listes d'attributs regroupés par types.



Les PDU GARP sont transportés dans des trames LLC identifiées par le SAP 0x42. Les applications GARP sont identifiées par des adresses MAC qui leur sont réservées : 01-80-C2-00-00-20 pour GMRP et 01-80-C2-00-00-21 pour GVRP.

4

L'alternative du sans fil

Pourquoi s'embarrasser de câbles ?

Le système de câblage est, en effet, source de complexité et de coût. Il est également synonyme de rigidité : se déplacer d'un bureau vers une salle de réunion ou d'une pièce à l'autre, nécessite de câbler entièrement les bureaux ou sa maison.

Les concepteurs de réseaux Ethernet proposent ainsi une version sans fil de ce protocole, reposant sur les ondes radio. Les applications étant multiples, de nombreux protocoles se disputent le marché.

Les réseaux Ethernet sans fil répondent aux inconvénients du câblage, mais posent d'autres problèmes inhérents aux ondes radio, telles que les interférences et la confidentialité des informations. L'Ethernet sans fil requiert ainsi ses propres règles d'ingénierie et a donc des limites.

Dans ce chapitre, vous apprendrez ainsi :

- les principes techniques des réseaux sans fil ;
- le fonctionnement des différents protocoles Wi-Fi, etc. ;
- à installer un réseau local Ethernet sans fil.

Introduction

Le monde du sans fil est très vaste puisqu'il couvre de nombreuses applications qui nous sont familières : radio, télévision, téléphone et maintenant transmission de données.

D'un point de vue utilisateur, on peut classer le monde du sans fil en quatre catégories qui se distinguent par la taille de la **cellule**. La cellule désigne la portée de l'onde entre l'émetteur et le récepteur, en d'autres termes la distance maximale entre les deux.

Type de cellule	Large	Macro	Micro	Pico
Taille de la cellule	20 km	1 à 3 km	150 à 100 m	10 à 30 m
Position de l'antenne	Toits, pylônes	Toits, pylônes, façades	Plafonds, murs	Sac, oreillette, PDA
Application	Téléphonie (GSM, UMTS)	Boucle locale radio	Réseau local sans fil	Réseau personnel
Débit des données	Quelques centaines de Kbit/s	Quelques Mbit/s	10 à 50 Mbit/s	Quelques Mbit/s

Les technologies employées sont non seulement adaptées à ces différentes applications, mais également aux conditions réglementaires imposées dans chaque pays. En effet, le domaine des ondes radio couvre un espace de fréquence limité, allant du Kilo-Hertz aux dizaines de Giga-Hertz. Cet espace doit donc être partagé, et les utilisateurs ne manquent pas : militaires, aviation civile, chaînes de télévision, stations de radio, opérateurs télécoms, entreprises et maintenant les particuliers.

Les autorités gouvernementales de chaque pays, l'ART en France (Autorité de régulation des télécommunications), la FCC aux États-Unis (*Federal Communication Commission*), contrôlent cet espace et le gèrent. Des bandes de fréquence sont ainsi attribuées sous forme de licences payantes et limitées dans le temps. Certaines plages sont cependant libérées, c'est-à-dire qu'elles peuvent être utilisées gratuitement, sans licence. Cela est le cas des bandes de fréquence autour de 2,4 GHz et de 5 GHz, celles qui nous intéressent dans ce chapitre consacré aux réseaux locaux sans fil.

Les bandes de fréquence précises utilisées par ces technologies dépendent de celles allouées par les autorités gouvernementales. Elles peuvent donc varier d'un pays à l'autre.

La puissance des émetteurs est également réglementée, puisqu'elle influe sur la taille des cellules. Il est ainsi possible d'utiliser les mêmes fréquences à des endroits distincts, si la portée des ondes est réduite. Nous connaissons tous cela avec les radios FM : d'une région à l'autre, des stations différentes peuvent émettre sur la même longueur d'onde.

Les principes de transmission radio

Cette section n'a pas pour objet de vous faire devenir des experts en transmission hertzienne, mais de présenter les principes de base permettant de comprendre comment fonctionne votre réseau local sans fil.

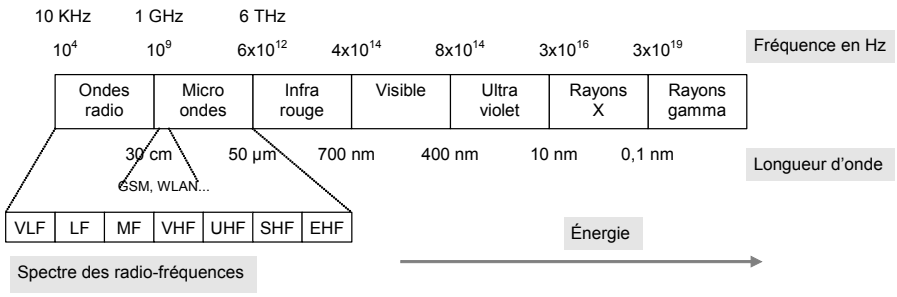
Le signal radio

Un signal radio est une onde électromagnétique se propageant à travers l'espace. Selon sa fréquence, il traverse la matière ou est réfléchi. À ce titre, la lumière se comporte comme une onde électromagnétique. Ces phénomènes bien connus de réflexion, de translucidité, de transparence et d'opacité s'appliquent également aux ondes radio.

De même, la propagation d'une onde électromagnétique dans l'espace peut être visualisée par analogie avec celle d'une onde sur l'eau, provoquée par le jet d'une pierre, l'émetteur de l'onde étant en l'occurrence la pierre. Les principes sont les mêmes, bien que leur nature soit différente.

L'onde électromagnétique est générée par l'oscillation d'un champ électrique qui génère un champ magnétique, qui lui-même oscille et donc génère un champ électrique, et ainsi de suite. La fréquence de l'oscillateur détermine donc celle du signal.

Le schéma suivant permet de se situer dans les gammes de fréquence connues.

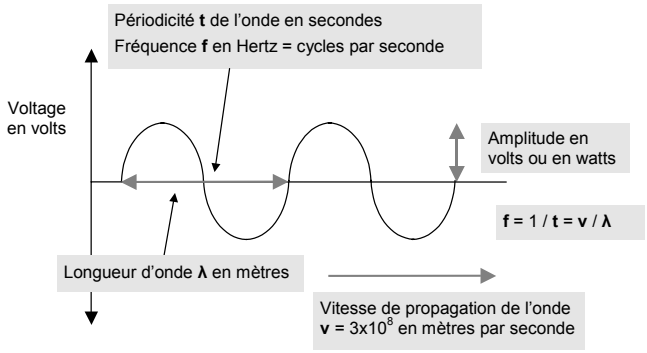


Non seulement ces ondes diffèrent par leurs fréquences, mais aussi par la manière dont elles sont générées, à savoir par transition : du spin (ondes radio), rotationnelle (micro-ondes), vibrationnelle (infrarouge) ou électronique (visible et ultraviolet).

La puissance d'un signal radio se mesure en milliwatts ou en décibels, un ratio logarithmique de la puissance du signal divisé par 1 mW. Tout comme la luminosité diminue avec la distance, la puissance du signal diminue proportionnellement au carré de la distance.

La modulation du signal

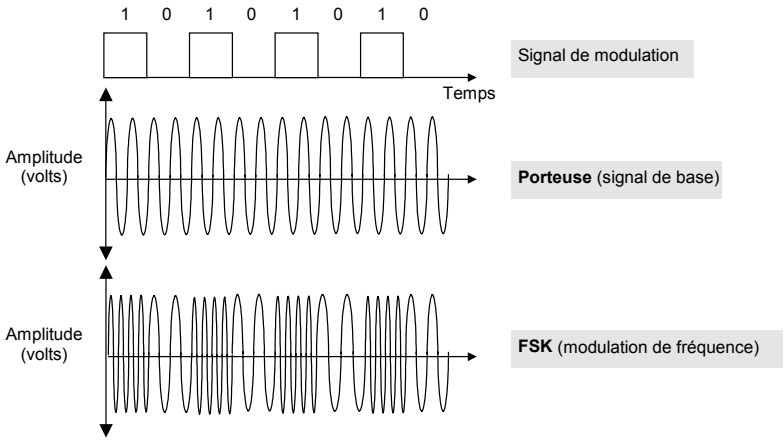
Une onde est classiquement représentée par une sinusoïde et est caractérisée par la longueur entre deux points de référence, par exemple la crête du signal.

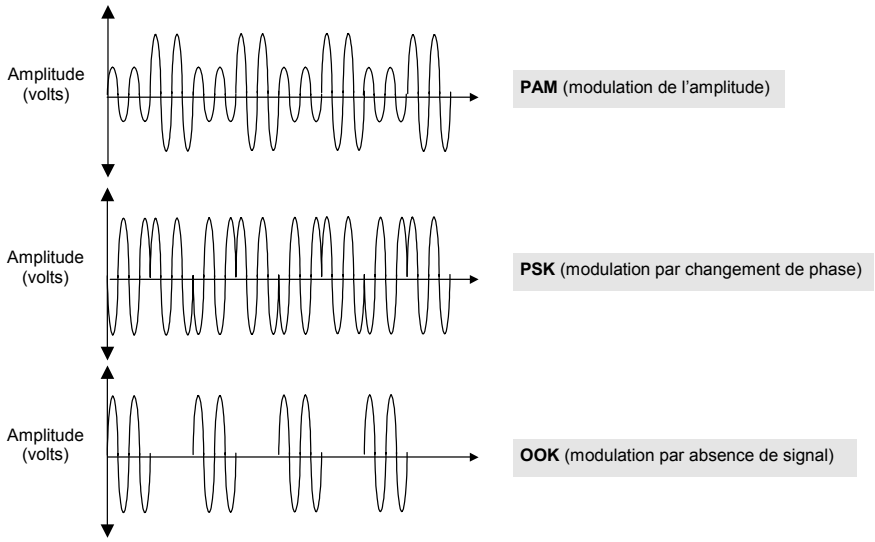


Dans un signal, la porteuse est le signal de référence ne transportant aucune donnée. Les données sont transmises en modulant le signal selon différentes manières : par son amplitude, sa fréquence ou par sa phase.

Pour les signaux analogiques, on parle de modulation de fréquence (FM) ou de modulation d'amplitude (AM). Pour les signaux numériques, on retrouve l'équivalent avec, en plus, des variantes :

- OOK (*On/Off Keying*), à la base du code Morse ;
- FSK (*Frequency Shift Keying*) ;
- PSK (*Phase Shift Keying*) ;
- ASK (*Amplitude Shift Keying*) ou PAM (*Pulse Amplitude Modulation*).





Il est également possible de combiner les modulations ASK et PSK, technique appelée APK (*Amplitude/Phase Keying*) ou QAM (*Quadrature Amplitude Modulation*). Il existe également des variantes à la modulation de fréquence, telle que GFSK (*Gaussian Frequency Shift Keying*), et à la modulation de phase, telle que GMSK (*Gaussian Minimum Shift Keying*) et BPSK (*Binary Phase Shift Keying*).

En outre, le signal peut être transmis à une fréquence donnée (diffusion en **bande étroite**), comme cela est le cas pour la radio, ou être diffusé sur une plage de fréquences (diffusion en **largeur de bande**) comme cela le cas pour les transmissions de données. Dans ce dernier cas, le **canal** désigne ainsi la fréquence au centre de la largeur de bande et la **bande passante** désigne ainsi le **spectre de fréquence** sur laquelle le signal est diffusé. De fait, les canaux doivent être suffisamment espacés afin d'éviter les interférences mutuelles.

L'avantage de la transmission en largeur de bande par rapport à la transmission en bande étroite, est qu'un signal parasite n'altère qu'une partie du spectre de fréquence, et ces conséquences sur le signal originel peuvent même être supprimées en réception grâce à l'application de fonctions mathématiques.

Le multiplexage des canaux de fréquence

À ces techniques de modulation, s'ajoutent des techniques de multiplexage permettant de partager la bande passante entre plusieurs utilisateurs. On trouve ainsi :

- FDMA (*Frequency Division Multiple Accessing*), qui consiste à affecter un canal par utilisateur (la bande de fréquence est découpée en canaux avec une porteuse pour chacun) ;
- TDMA (*Time Division Multiple Accessing*), qui découpe le canal en tranches de temps, une par utilisateur (un canal GSM utilise 3 tranches de temps) ;

- CDMA (*Code Division Multiple Accessing*) qui partage le canal à l'aide d'une adresse affectée à chaque utilisateur (principe retenu pour l'UMTS).

À cela, viennent s'ajouter des techniques de gestion de la bande passante, c'est-à-dire la manière d'utiliser le spectre de fréquences pour transmettre une trame, le but étant d'utiliser au mieux la largeur de bande afin d'augmenter les débits :

- FHSS (*Frequency Hopping Spread Spectrum*), qui émet sur quelques canaux simultanément parmi plusieurs dizaines en en changeant aléatoirement plusieurs fois par seconde ;
- DSSS (*Direct Sequence Spread Spectrum*), qui consiste à émettre sur toute la largeur de bande autour du canal ;
- OFDM (*Orthogonal Frequency Divison Multiplexing*) qui segmente la largeur de bande en sous-fréquences (jusqu'à 52 porteuses), chacune transportant une portion de la trame et étant modulée avec un code différent (BPSK, QPSK, 16-QAM, 64-QAM) selon le débit sélectionné. OFDM est également connu sous le nom de DMT (*Discrete Multi-Tone*).

Alors que DSSS arrose tout le spectre, FHSS n'en utilise qu'une partie en se déplaçant à l'intérieur. Cela rend ce dernier plus robuste aux interférences (la probabilité de perturbation augmente avec la largeur de bande), mais diminue le débit de transmission.

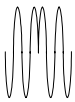
De son côté, OFDM utilise encore mieux le spectre de fréquence, puisque l'espacement entre les sous-porteuses est réduit, de l'ordre de 312 KHz contre 1 MHz entre les canaux FHSS et DSSS.

De ce fait, OFDM convient mieux à un faible nombre de nœuds transmettant beaucoup de données, tandis que DSSS est plus adapté à un réseau comportant beaucoup de nœuds et transmettant moins de données.

Le codage

Les techniques de transmission numériques ne transmettent pas des bits indépendamment les uns des autres, mais des groupes de bits codés par des **symboles**. Dans le cas d'une modulation de phase, un jeu de symboles pourrait être, par exemple :

Bits	Symbole
00	0
01	π/2
10	Π
11	3π/2 (= -π/2)

Un symbole correspond donc à une forme de signal, dépendant de la technique de modulation employée, comme, par exemple  qui peut représenter la valeur 01.

L'intérêt est que la transmission d'un groupe de bits résiste mieux aux interférences : les processeurs spécialisés dans le traitement du signal, les DSP (*Digital Signal Processor*), comparent la forme du signal reçu à une table de référence (en d'autres termes, une table des symboles). Même si une partie du signal est altérée, le DSP sera quand même capable de reconnaître la forme si celle-ci ressemble de manière non ambiguë à un symbole de référence. Les fonctions mathématiques utilisées par ce type de processeur sont les transformations de Fourier rapides (FFT – *Fast Fourier Transform*).

Propriétés des ondes radio

D'une manière générale :

- Plus le débit est élevé, plus la couverture radio est faible.
- Plus la puissance d'émission est élevée, plus la couverture est grande, mais avec une consommation d'énergie plus grande.
- Plus la fréquence radio est élevée, meilleur est le débit, mais plus la couverture est faible.

Par ailleurs, les signaux se comportent différemment dans l'environnement selon leur fréquence :

- Dans la gamme des micro-ondes, les gouttes de pluie atténuent davantage les signaux à haute fréquence que les signaux à basse fréquence.
- Une partie du signal peut traverser les objets et une autre être réfléchie, à l'instar de la lumière sur une étendue d'eau.
- Les ondes basse fréquence, de l'ordre du Kilo-Hertz, sont réfléchies par les couches hautes de l'atmosphère et peuvent donc être reçues au-delà de l'horizon créé par la courbure de la terre.
- Les micro-ondes sont en partie réfléchies par la plupart des objets. Un signal d'origine peut ainsi être réfléchi plusieurs fois. Le récepteur capte donc plusieurs signaux avec de légers décalages de temps.
- Dans le haut du spectre des micro-ondes, au-delà de 5 GHz, les ondes ne pénètrent quasiment pas les objets. L'émetteur et le récepteur ont besoin d'être en vue directe sans obstacle, configuration appelée « *clear line of sight* » dans la littérature anglo-saxonne.

Exemples d'application

La norme **Bluetooth 802.15.1** permet de gérer jusqu'à 7 connexions simultanées de différentes natures : 3 canaux synchrones à 64 Kbit/s (pour la voix et la vidéo) ou 2 canaux synchrones à 64 Kbit/s avec 1 canal asynchrone asymétrique (721 Kbit/s / 57,6 Kbit/s) ou symétrique (432,6 Kbit/s dans les deux sens) pour les données. La norme emploie la technique FHSS sur 5 canaux parmi 79 (4 parmi 23 en France) avec un espacement de 1 MHz entre les canaux. Bluetooth change 1 600 fois de canal par seconde, à comparer avec **Home RF** qui saute de fréquence 50 fois par seconde et qui, comme Bluetooth, combine transmission de données et téléphonie en fournissant 4 canaux DECT et un canal de données à 1,6 Mbit/s.

De son côté, la norme **Wi-Fi 802.11b** emploie la technique DSSS avec des modulations de type PSK, tandis que **Wi-Fi 802.11a** utilise la technique OFDM avec des modulations de type PSK et QAM.

La norme 802.11b utilise le codage Barker pour les débits de 1 et 2 Mbit/s, le codage CCK (*Complementary Code Keying*) pour les débits de 5,5 et 11 Mbit/s.

Revue des réseaux sans fil

Les quelques noms qui viennent d'être cités sont des protocoles normalisés qui définissent les couches physiques et logiques, à savoir : les techniques de gestion de la bande passante et les modulations d'une part, le format des trames et les procédures d'échanges d'autre part.

Selon leur domaine d'application, on peut les ranger dans trois catégories : WPAN pour les micro-réseaux personnels, WLAN pour les réseaux locaux Ethernet, sujet qui nous intéresse dans ce chapitre, et WMAN pour les réseaux à l'échelle d'une ville.

Standard	Nom commercial	Catégorie	Applications
802.11a	Wi-Fi	WLAN	Réseau local
802.11b	Wi-Fi	WLAN	Réseau local
802.11g	Wi-Fi	WLAN	Réseau local
ETSI 300 652	HiperLAN 2	WLAN	Réseau local
HomeRF	HomeRF 2.0	WLAN	Résidentiel et SOHO
IrDA	Infra-rouge	WPAN	Connexion des périphériques aux PC
802.15.1	Bluetooth	WPAN	Connexion des PC, PDA et téléphones
802.15.3	Bluetooth 2	WPAN	Connexion des PC, PDA et téléphones
802.15.4	Zigbee	WPAN	Périphériques à faible consommation
802.16	---	WMAN	Boucle locale radio

La plupart des standards WLAN sont édictés par le groupe de travail 802.11 de l'IEEE, qui spécifie les couches physiques de transmission radio ainsi que la couche logique (niveau 2) appelée MAC (*Media Access Control*), celle que l'on retrouve partout dans le monde Ethernet.

Les technologies relatives au WPAN sont issues des industriels et ont été intégrées, par la suite, dans les travaux de l'IEEE.

Les technologies WMAN, utilisées depuis des décennies mais jamais standardisées, font l'objet depuis peu, de travaux au sein du groupe de travail 802.16 de l'IEEE.

Les applications

L'installation d'un WLAN peut être intéressante à domicile, si vous avez plusieurs utilisateurs, si vous jouez en réseau, etc. Il évite l'installation d'un câblage plus ou moins bricolé.

Dans une entreprise, il permet aux utilisateurs de se déplacer entre leurs bureaux et les salles de réunion. Il permet également de desservir une zone ou un bâtiment dans lesquels les tra-

vaux de câblage sont difficiles ou onéreux. Il permet enfin, d'interconnecter deux bâtiments proches, typiquement sur un campus ou séparés par une cour, sans avoir à tirer un câble.

Dans les lieux publics, il est envisageable de créer des zones de couverture, appelées « *hot spots* » dans les salles d'attente des aéroports, les gares, voire dans le métro. La réglementation pour un usage extérieur étant plus stricte, le développement des *hot spots* dépendra d'une part des autorisations d'organismes de régulation tels que l'ART et la FCC, et d'autre part de la concurrence venant des technologies GPRS et UMTS.

LES STANDARDS DE RÉSEAU SANS FIL

Trois comités de l'IEEE (*Institute of Electrical and Electronics Engineers*) travaillent à trois domaines d'application du sans fil :

- le 802.11, dit WLAN (*Wireless Local Area Network*), pour les réseaux Ethernet ;
- le 802.15, dit WPAN (*Wireless Personal Area Network*), pour les communications entre périphériques, les assistants personnels, les téléphones mobiles, etc. ;
- le 802.16, dit WMAN (*Wireless Metropolitan Area Network*) pour les boucles locales radio.

Ce dernier, étant destiné à des applications radicalement différentes, est traité à la fin de ce chapitre.

Côté Européen, l'ETSI travaille, plus ou moins de concert avec l'IEEE, sur les protocoles Hiperlan (*HIgh PErformance Radio LAN*) et Hiperman (*HIgh PErformance Radio MAN*).

Norme WLAN	Nom commercial	Débit théorique	Portée maximale	Bande de fréquence	Technique de transmission
ETSI EN 300 652	Hiperlan 1	20 Mbit/s	n.c.	5 GHz	GMSK/FSK
ETSI (en cours)	Hiperlan 2	54 Mbit/s	30 m	5 GHz	OFDM
IEEE 802.11a	Wi-Fi	54 Mbit/s	40 m	5 GHz	OFDM
IEEE 802.11g	Wi-Fi	54 Mbit/s	70 m	2,4 GHz	DSSS/OFDM
IEEE 802.11b	Wi-Fi	11 Mbit/s	90 m	2,4 GHz	DSSS
HomeRF 1.0	HomeRF	1,6 Mbit/s	50 m	2,4 GHz	FHSS

Norme WPAN	Nom commercial	Débit théorique	Portée maximale	Bande de fréquence	Technique de transmission
IEEE 802.15.1	Bluetooth	1 Mbit/s	30 m	2,4 GHz	FHSS
IEEE 802.15.3	Bluetooth 2	12 Mbit/s	10 m	2,4 GHz	n.c.
IEEE 802.15.4	Zigbee	250 Kbits/s	75 m	2,4 GHz	DSSS
IrDA	FIR (Fast infrared)	4 Mbit/s	1 m	(850 nm)	optique (LED)

D'autres standards IEEE viennent compléter ces normes de base.

Standard IEEE	Objet
802.11e	Gestion de la qualité de service.
802.11f	IAPP (*Inter-Access Point Prorocol*) : protocole d'échange entre bases (permet le *roaming* multivendeurs).
802.11h	Conformité aux normes européenne relatives à la gestion des bandes de fréquence et à la puissance d'émission.
802.11i	WPA (*Wi-Fi Protected Access*) : sécurisation du WLAN.
802.11j	Convergence des standards 802.11a et Hiperlan 2.

Quel WLAN choisir ?

Pour bâtir un réseau Ethernet sans fil, donc compatible avec votre réseau Ethernet filaire, trois types de produits sont disponibles sur le marché : ceux conformes à la norme 802.11b (débit de 11 Mbit/s), ceux conformes à la norme 802.11a (débit de 54 Mbit/s) et le 802.11g (débit de 54 Mbit/s).

Les normes 802.11a et 802.11b sont incompatibles, car elles fonctionnent à des fréquences différentes. Tous les composants, cartes, bases, etc., doivent respecter la même norme, mais il est possible d'interconnecter ces réseaux *via* un réseau Ethernet filaire. Par ailleurs, certains équipements supportent ces deux normes simultanément. La norme 802.11g présente, quant à elle, l'intérêt d'offrir le même débit que la norme 802.11a tout en étant compatible avec la norme 802.11b.

Il est possible d'installer un réseau WLAN avec des équipements en provenance de constructeurs différents sous certaines conditions :

- Les équipements doivent être certifiés **Wi-Fi** (*Wireless Fidelity*), un label du WECA (*Wireless Ethernet Compatibility Alliance*), organisme chargé de tester l'interopérabilité des équipements (bases, interfaces, etc.).
- Les équipements doivent supporter la norme **802.11f**, qui autorise le déplacement (*roaming*) entre bases issues de constructeurs différents. Cette norme spécifie un protocole d'échange entre les bases, qui permet à une base de prendre en charge les communications précédemment gérées par une autre, lorsque la station choisit de passer d'une cellule à une autre.

D'un point de vue technique, ces normes se distinguent par les fréquences utilisées :

- 2,4 GHz pour le 802.11b, fréquence connue sous le nom de bande ISM (*Industrial Scientific and Medical band*), également utilisée par les fours à micro-ondes, les télécommandes de garage, etc.
- 5 GHz pour le 802.11a, fréquence nettement moins encombrée, ce qui permet d'utiliser des gammes de fréquence plus larges et donc de transmettre à plus haut débit. Cette fréquence est néanmoins soumise à des contraintes réglementaires plus strictes, car elle est utilisée par les militaires pour leurs transmissions satellites.
- 2,4 GHz pour le 802.11g, qui utilise la technique de transmission DSSS pour la compatibilité avec 802.11b et OFDM pour les débits plus élevés.

Contraintes réglementaires

Selon les contraintes réglementaires de chaque pays, le standard 802.11b est autorisé à utiliser un nombre varié de canaux : 11 aux États-Unis, 4 en France, 13 en Europe, 1 au Japon, etc. De même, la FCC aux États-Unis autorise le 802.11a à utiliser 8 canaux. Ces différences influent sur les possibilités d'architecture : avec 11 canaux, 3 WLAN 802.11b peuvent fonctionner sans interférence sur une même zone géographique (technique de la **co-location**), alors qu'en France un seul WLAN peut opérer dans ces mêmes conditions.

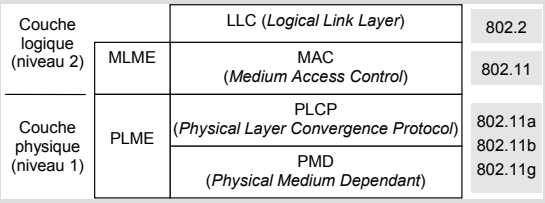
Autorité de réglementation	Bande de fréquences (MHz) autorisées	Bande passante par canal	Puissance d'émission autorisée	Modulation autorisée
FCC (USA)	2400 – 2483,5	83,5 MHz	100 mW	FHSS / DSSS
ETSI (Europe)	2400 – 2483,5	83,5 MHz	100 mW	FHSS / DSSS
ART (France) *	2400 – 2483,5	83,5 MHz	10 mW	FHSS / DSSS
ART (France)	2446,5 – 2483,5	37 MHz	100 mW	FHSS / DSSS
FCC (USA)	5150 – 5260 5260 – 5350	200 MHz	20 mW	OFDM
ART (France)	5150 – 5350	200 MHz	200 mW	OFDM

* Dans certains départements, l'ensemble de la bande est autorisé à l'intérieur des bâtiments ainsi qu'à l'extérieur (décisions de l'ART n° 03-908 du 22/07/03, n° 02-1009 du 31/10/02, n° 02-1008 et n° 02-1088).

Il est à noter que l'utilisation d'une transmission en large bande est un pré requis à l'utilisation d'une fréquence sans licence, car pour une transmission en bande étroite, dont l'énergie est concentrée sur une fréquence, un signal large bande, dont l'énergie est dispersée sur une plage de fréquences, est perçu comme un bruit.

LE MODÈLE DE RÉFÉRENCE DES NORMES IEEE 802.11

Le modèle de référence couvre une partie de la couche logique (architecture et couche MAC) ainsi que la couche physique (transmissions à 1 et 2 Mbit/s), aujourd'hui obsolète. Par la suite, les normes 802.11a, b et g ont introduites les spécifications pour différentes couches physiques (sous-couches PLCP et PMD).



La couche **MAC** a pour rôle de gérer l'accès au support de transmission *via* un adressage (l'adressage MAC classique), une méthode de **contention** (CSMA/CA), une méthode de **réservation** (RTS/CTS) et un format de trame identique quel que soit le support physique. Elle peut également gérer la **fragmentation** et le **réassemblage** des trames si cela est demandé. Dans des milieux perturbés impliquant de nombreuses pertes de données, la retransmission d'une partie de la trame prend, en effet, moins de temps que la retransmission complète. Optionnellement, elle peut gérer des mécanismes de **chiffrement** et d'**authentification** à travers le protocole **WEP** (*Wired Equivalent Privacy*)

La couche **PLCP** réalise l'interface entre les trames MAC et les caractéristiques du support physique (synchronisation, débit, etc.), tandis que la couche **PMD** est propre à chaque technique de transmission (fréquence, codage, modulation, transmission du signal, etc.).

Sur le plan de la gestion, la couche **MLME** (*MAC Sublayer Management Entity*) prend en charge la gestion de l'énergie ainsi que les procédures d'initialisation, d'authentification et d'association, tandis que la couche **PLME** (*Physical Sublayer Management Entity*) permet à tout nœud du réseau d'obtenir les informations relatives aux caractéristiques opérationnelles supportées par une interface : longueur des en-têtes et des préambules, taille maximale des MPDU, temps maximum de basculement entre les modes émission et réception, etc.

Performances

Les performances d'un WLAN sont caractérisées par la portée, c'est-à-dire la distance maximale entre les postes de travail, ainsi que par le débit.

La portée dépend de la puissance émission et du gain de l'antenne, c'est-à-dire de sa sensibilité. Ces paramètres étant imposés par la réglementation, la portée dépend, en définitive, de la qualité des produits et, bien sûr de l'environnement considéré (murs, meubles, personnes).

Le débit dépend, quant à lui, de la norme utilisée. Actuellement, le 802.11g est le mieux placé, d'autant plus qu'il est compatible avec les premières normes 802.11. Le débit diminue avec l'augmentation de la distance entre émetteur et récepteur. La qualité des produits intervient, là encore, dans les performances.

Les tableaux suivants, issus des constructeurs, permettent de donner une idée des performances globales d'un WLAN.

Cisco aironet 802.11b avec une puissance d'émission de 100 mW.

Débit	Intérieur	Extérieur
11 Mbit/s	45 m	243 m
5,5 Mbit/s	61 m	304 m
2 Mbit/s	76 m	457 m
1 Mbit/s	106 m	609 m

Cisco 802.11a avec une puissance d'émission de 100 mW et une antenne de gain de –6dbi.

Débit	Intérieur	Extérieur
54 Mbit/s	21 m	36 m
18 Mbit/s	45 m	213 m
6 Mbit/s	61 m	365 m

Performance comparée des standards Wi-Fi en usage intérieur.

	802.11b	802.11a	802.11g
Portée maximale pour débit maximal	45 m à 11 Mbit/s	12 m à 54 Mbit/s	18 m à 54 Mbit/s
Portée maximale pour débit minimal	91 m à 1 Mbit/s	37 m à 6 Mbit/s	73 m à 6 Mbit/s

Les débits indiqués dans ces tableaux sont les débits de fonctionnement spécifiés par les normes. Si une interface reçoit mal les données à 11 Mbit/s, elle utilisera une vitesse de repli à 5,5 Mbit/s, et ainsi de suite.

En outre, toutes les stations se partageant la même bande passante en half-duplex, les débits réels constatés par les utilisateurs sont inférieurs de 30 à 50 % aux débits de transmission indiqués ci-dessus. Ainsi sur les 11 Mbit/s annoncés par le 802.11b, le débit effectif constaté par l'utilisateur oscille entre 6 et 8 Mbit/s.

Architectures et principes de fonctionnement

Les réseaux WLAN peuvent être bâtis selon deux modèles :

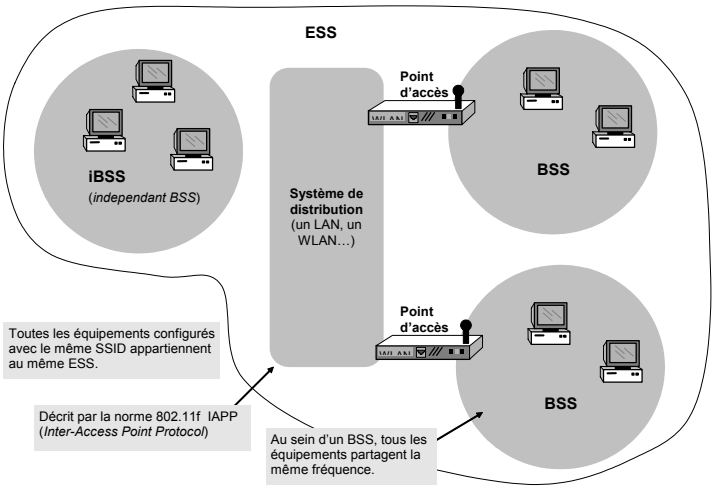
- **ad-hoc**, c'est-à-dire en mode point à point entre équipements (par exemple, PDA-téléphone, PC-téléphone, PC-imprimante) ;
- à **infrastructure**, dans lequel le réseau est géré par des points d'accès (appelés bases ou encore bornes d'accès).

Typiquement, les WPAN sont bâtis sur le modèle ad-hoc.

Le **service de base**, BSS (*Basic Service Set*), repose sur un seul point d'accès raccordant plusieurs équipements.

Une architecture à infrastructure consiste en un ou plusieurs BSS raccordés au réseau Ethernet filaire. Les points d'accès servent alors de pont entre le réseau filaire et le réseau sans fil.

Il est alors possible de mettre en place un **service étendu**, ESS (*Extended Service Set*), permettant aux utilisateurs de se déplacer d'un BSS à l'autre, c'est-à-dire d'être pris en charge successivement par plusieurs bornes. Les bornes forment alors un **système de distribution** qui leur permet d'échanger des informations.



Neuf services sont fournis par cette architecture :

- l'**authentification** des stations sur un ou plusieurs points d'accès, sur la base d'un identifiant ou de l'adresse MAC, qui permet de s'insérer dans le WLAN, sans pour autant pouvoir échanger des données ;
- la **dé-authentification** des stations, permettant à une station de quitter le WLAN ;
- l'**association**, qui permet à une station d'échanger des données *via* un point d'accès auprès duquel elle s'est préalablement authentifiée :
- la **ré-association**, qui permet à la station d'aller d'une base à l'autre, la station transmettant à la nouvelle base des informations sur l'ancienne. La nouvelle base se connecte alors sur l'ancienne pour récupérer des données éventuellement stockées à l'attention de la station qui vient de quitter ladite base ;

- la **dé-sassociation**, qui permet à une station de quitter une base, voire le WLAN ;
- l'**acheminement** des données ;
- la **confidentialité** par un chiffrement simple, RC4 ;
- la **distribution** permettant à une base de savoir où envoyer les données (vers une autre base, une station ou le réseau filaire) ;
- l'**intégration**, qui assure la conversion des trames entre le format WLAN et le format LAN.

Afin de gérer les contentions d'accès à la fréquence qui est partagée par toutes les stations, le protocole 802.11 utilise un mécanisme similaire à celui employé par son cousin filaire, appelé CSMA/CA (*Carrier Sense Multiple Access / Collision Avoidance*). Selon ce principe, une interface n'émet une trame que si aucun trafic n'est détecté sur la bande de fréquence partagée. Dans le cas contraire, l'interface attend un laps de temps aléatoire avant de se mettre de nouveau à l'écoute, et répète le processus.

Afin de minimiser le risque de collision, le mécanisme de réservation suivant peut optionnellement être utilisé : avant d'acheminer une trame à une station, la base lui envoie d'abord une trame RTS (*Requests To Send*). La station doit alors répondre par un CTS (*Clear To Send*) et attendre la réception de la trame. De leur côté, les autres stations recevant ce type de trame retardent d'éventuelles émissions de trame.

Les fréquences radio pouvant par ailleurs être perturbées, le récepteur doit renvoyer à l'émetteur une trame ACK (*Acknowledgement*) pour chaque trame de données reçue. Si l'émetteur ne reçoit pas d'acquittement, il réémettra la trame de données.

Toujours pour réduire l'impact des interférences radio, le standard prévoit de pouvoir échanger des trames courtes. L'inconvénient est que dans un environnement non perturbé, l'overhead protocolaire est plus important qu'avec des trames longues. L'administrateur devra choisir l'un des deux modes lors de la configuration des équipements.

Afin de limiter la consommation électrique des composants 802.11 dans un ordinateur portable, la station peut envoyer à la base une trame PS-Poll (*Power Save Poll*) lui indiquant qu'elle se met en mode veille. Si une trame à destination de la station parvient à la base, celle-ci emmagasinera les données dans une mémoire tampon, le temps que la station se réveille. La station se doit d'envoyer périodiquement une trame PS-Poll.

La taille d'une cellule étant limitée, la mobilité implique de pouvoir passer d'une base à l'autre sans couper les communications. C'est ce qu'on appelle le roaming. Afin de déterminer quelle est la meilleure base, c'est-à-dire celle avec qui la communication est la meilleure, les stations sont à l'écoute des signaux émis par les bases (trames *beacon*). La station, qui s'est authentifiée auprès de plusieurs bases, sélectionne celle dont elle a reçu la trame *beacon* ayant le meilleur rapport signal/bruit, puis utilise le service ré-association avec cette base.

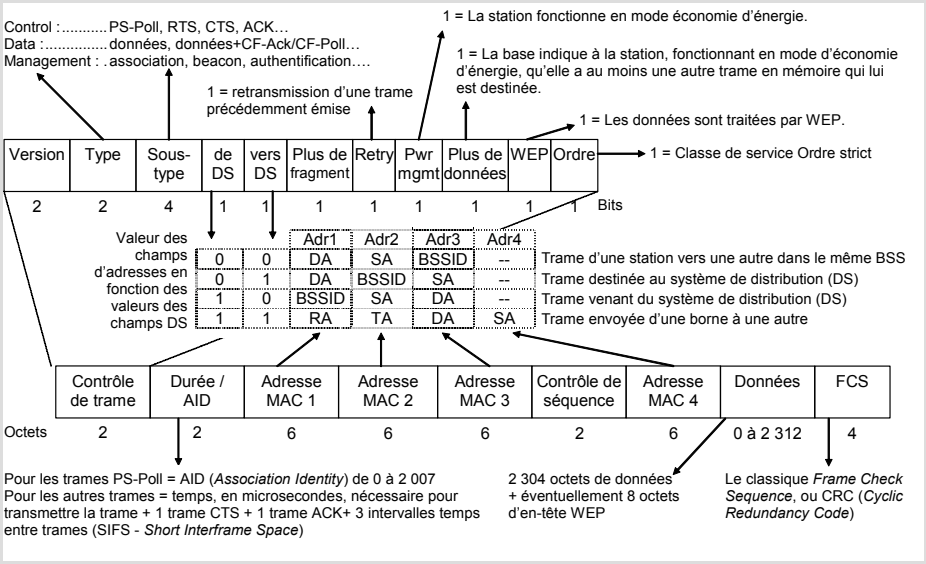
LE POINT SUR LE WLAN (IEEE 802.11, 802.11A, 802.11B, 802.11G)

La couche **MAC** peut fonctionner selon deux modes : **DCF** (*Distribution Coordination Function*), qui définit une méthode d'accès par **contention**, et **PCF** (*Point Coordination Function*), optionnel, qui définit une méthode d'accès par **allocation** de la bande passante par un coordinateur central.

Le mode PCF n'étant pratiquement jamais implémenté, seul le mode DCF est décrit ici. Ce mode repose sur la méthode d'accès **CSMA/CA** (*Carrier Sense Multiple Access / Collision Avoidance*) associé à la réservation du support **RTS/CTS** :

- La station doit respecter un intervalle de temps **ISF** (*Inter Frame Space*) avant de transmettre une trame.
- La station peut attendre un intervalle de temps réduit **SIFS** (*Short Inter Frame Space*) dans les cas suivants : réponse à un polling de la base, envoi de fragments d'une même trame, envoi des trames ACK et CTS.
- La station doit s'assurer que le canal est libre avant toute transmission de trame.
- Si cela n'est pas le cas, elle doit attendre un intervalle de temps **DIFS** (*Distributed IFS*) ajouté d'un délai aléatoire (*back-off window*), avant de réécouter le canal.
- Lorsque le support est libre, la station réserve le canal en envoyant à la borne d'accès une trame **RTS** (*Request to Send*) et attend la réponse **CTS** (*Clear To Send*) avant d'émettre. Ces trames permettent de définir la période de temps réservée pour transmettre la trame de données et, en retour, l'accusé de réception (trame ACK). La fonction RTS/CTS ne peut être utilisée que pour des trames unicasts.

Le format général des trames MAC 802.11 est décrit ci-après.

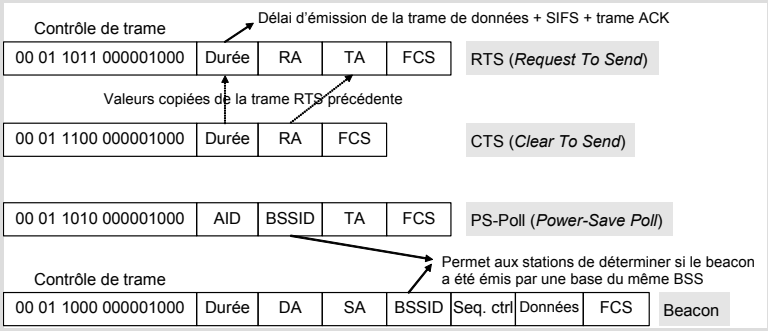


DA (*Destination Address*) = adresse MAC destination
SA (*Destination Address*) = adresse MAC source
TA (*Transmitter Address*) = adresse MAC du point d'accès émetteur
RA (*Receiver Address*) = adresse MAC du point d'accès récepteur
BSSID (*Basic Service Set Identification*) = identifie de manière unique le BSS = adresse MAC de la base

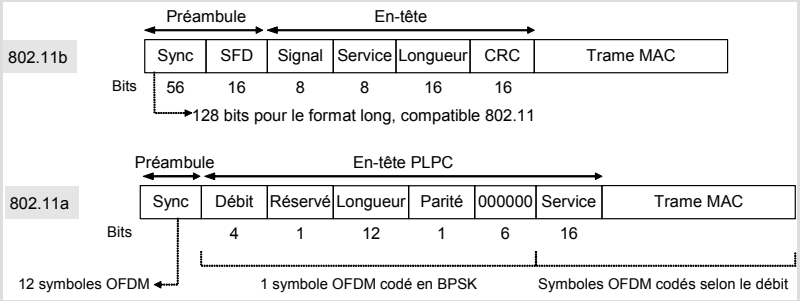
● ● ●

LE POINT SUR LE WLAN (SUITE)

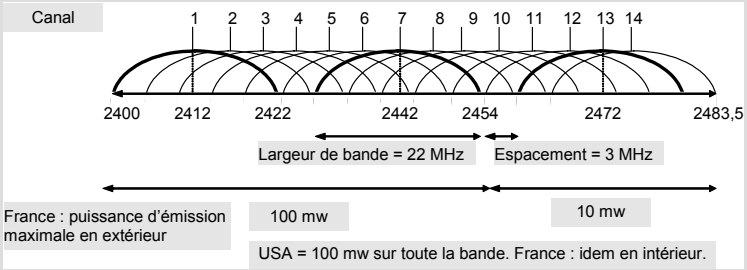
Comme exemples de trames MAC, nous trouvons :



Les trames MAC sont encapsulées dans des PPDU (***PLCP*** *Protocol Data Unit*), trames dont le format est proche de la couche physique sous-jacente (cf. modèle de référence de l'IEEE).



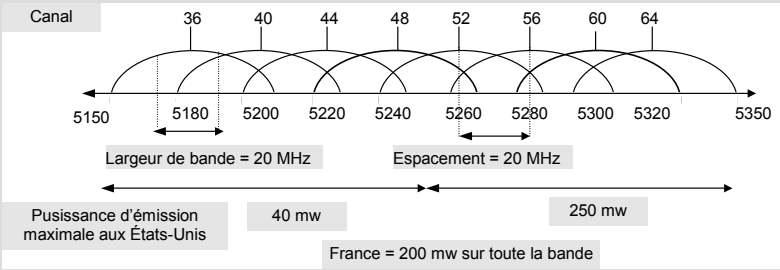
Au niveau de la couche **PMD**, la norme 802.11b définit l'utilisation de 3 canaux parmi les 14 disponibles dans la bande de fréquence autorisée. Le choix de ces canaux dépend de l'autorité de régulation de chaque pays. Trois bornes peuvent donc émettre simultanément dans la même zone géographique sans se gêner, autorisant ainsi la co-localisation de 3 WLAN.



• • •

LE POINT SUR LE WLAN (FIN)

La norme 802.11a définit, quant à elle, l'utilisation de 8 canaux, permettant ainsi la co-localisation de 8 WLAN. Chaque canal est structuré en 52 sous-canaux de 300 KHz de bande passante chacun (48 pour les données et 4 pour la signalisation). La couche PMD 802.11a génère 250 000 symboles par secondes, chacun d'entre eux couvrant les 48 sous-canaux. Le débit résultant varie selon le nombre de bits codés par symbole.



Une autre bande de fréquence est, par ailleurs, prévue par la norme :

- 11 canaux dans la bande 5 470 – 5 725 MHz, en Europe (en cours de négociation avec l'armée en France) ;
- 4 canaux dans la bande 5 725 – 5 825 MHz aux États-Unis.

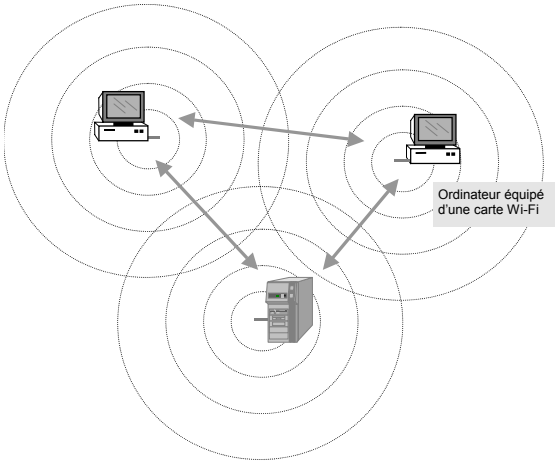
802.11a				802.11b		
Débit en Mbit/s	Modulation	Codage	Bits (*)	Débit en Mbit/s	Modulation	Codage
6	BPSK	OFDM	1 / 48 / 24	1	DBPSK	Barker
9	BPSK	OFDM	1 / 48 / 36	2	DQPSK	Barker
12	QPSK	OFDM	2 / 96 / 48	5,5	DQPSK	CCK
18	QPSK	OFDM	2 / 96 / 72	11	DQPSK	CCK
24	16-QAM	OFDM	4 / 192 / 96			
36	16-QAM	OFDM	4 / 192 / 144			
48	64-QAM	OFDM	6 / 288 / 192			
54	64-QAM	OFDM	6 / 288 / 216			

(*) Nombre de bits codés par sous-fréquence / Nombre de bits codés par symbole / Nombre de bits de données par symbole, les autres bits étant utilisés pour la correction d'erreurs.

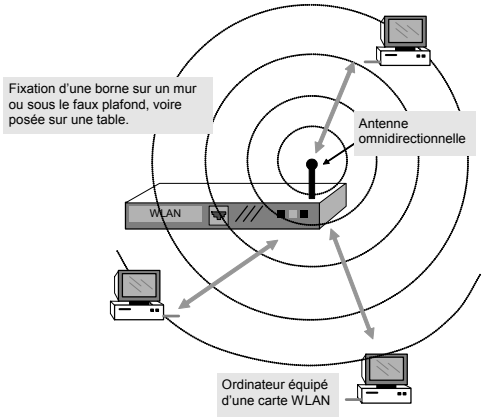
Installer un WLAN

Le plus simple des réseaux sans fil consiste à équiper tous les postes de travail de cartes Wi-Fi qui leur permettent de communiquer directement entre eux. Toutes les applications utilisant le protocole TCP/IP sont alors accessibles : partage de fichiers, partage d'imprimante, jeux en réseau, etc.

Cette architecture, prévue par le standard 802.11, est appelée **ad-hoc**. Les stations WLAN forment alors un iBSS (*Independant Basic Service Set*).



L'étape suivante consiste à ajouter une base, encore appelée borne d'accès ou point d'accès. Tout WLAN mettant en œuvre un point d'accès est appelé, par le standard 802.11, une **architecture à infrastructure**. Le point d'accès et les stations forment alors un BSS (*Basic Service Set*).



L'intérêt de l'architecture à infrastructure est d'étendre la portée du WLAN, puisque maintenant, les stations peuvent être situées à 50 m de part et d'autre du point d'accès, alors que dans le mode ad-hoc, les stations doivent être situées à moins de 50 m les unes des autres. À part cela, les performances sont identiques, car tous les équipements d'un BSS ou d'un iBSS se partagent la même bande de fréquence.

Choisir la bonne antenne

Cinq types d'antenne peuvent être utilisés pour les WLAN :

Type	Forme	Rayonnement	Installation
Verticale	Tige	Omnidirectionnel, cylindrique autour de l'axe de la tige	Salles de réunion, bureaux
Dipôle	Tige surmontée d'une barre horizontale (forme en 'T')	Omnidirectionnel double cylindrique vertical de part et d'autre de l'axe de la tige	Zones longues et étroites comme les couloirs
Sectorielle	Panneau plat ou demi-cylindre	Omnidirectionnel elliptique dans le plan du panneau	Salles de réunion, halls d'entrée, liaisons point à multipoint
Yagi	Barres parallèles liées par une tige perpendiculaire (type antenne TV)	Unidirectionnel elliptique dans l'axe de la tige	Liaisons entre immeubles proches
Parabolique	Parabole (type antenne satellite)	Unidirectionnel elliptique, étroit dans le plan perpendiculaire à la parabole	Liaisons entre immeubles éloignés

L'antenne isotrope, c'est-à-dire rayonnant de la même façon dans toutes les directions, est un modèle théorique. En réalité, l'énergie rayonnée par une antenne est répartie inégalement dans l'espace, certaines directions étant privilégiées et définissant ce qu'on appelle des lobes de rayonnement.

La qualité d'une antenne est appréciée par son **gain** exprimé en dBi (décibels isotropes). Il dépend de la taille de l'antenne, de la fréquence radio, de la capacité du réflecteur à focaliser les ondes radio et de la qualité des matériaux employés.

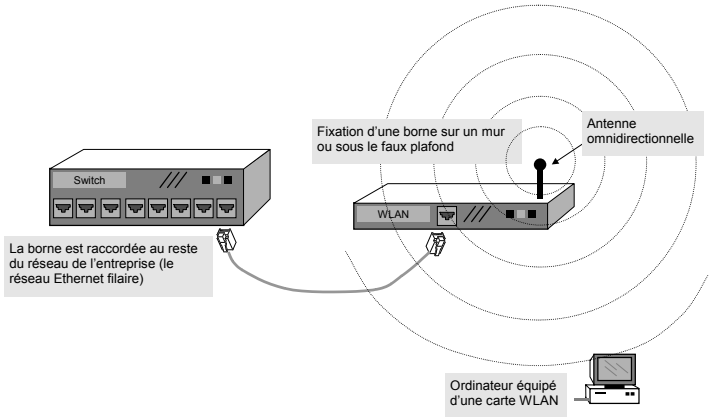
En émission, le gain est le rapport entre l'énergie moyenne rayonnée par l'antenne et celle que rayonnerait une antenne idéale (isotropique) alimentée avec la même puissance. Exprimé en dBi, il mesure la capacité d'une antenne à concentrer les ondes radio dans une direction donnée. Donc, plus une antenne concentre l'énergie dans une direction donnée, plus son gain est élevé.

L'ensemble, puissance de rayonnement multipliée par le gain de l'antenne, définit le PIRE (*puissance isotrope rayonnée équivalente*) – EIRP en anglais (*Equivalent Isotropic Radiated Power*), exprimé en milliwatts. C'est cette puissance qui est limitée par les autorités de régulation (généralement, de 10 à 100 mW dans les bandes de fréquence 2,4 et 5 GHz).

En réception, le gain élevé d'une antenne compense en partie la perte de la puissance du signal. Cette valeur exprime donc sa capacité à amplifier la puissance d'un signal reçu. Les antennes fournies avec les bornes Wi-Fi présentent généralement un très faible gain, de l'ordre de 2.5 dBi. Dans des situations géographiques particulières ou en présence de perturbations électromagnétiques importantes, il convient donc d'installer des antennes de meilleure qualité.

Les WLAN en entreprise

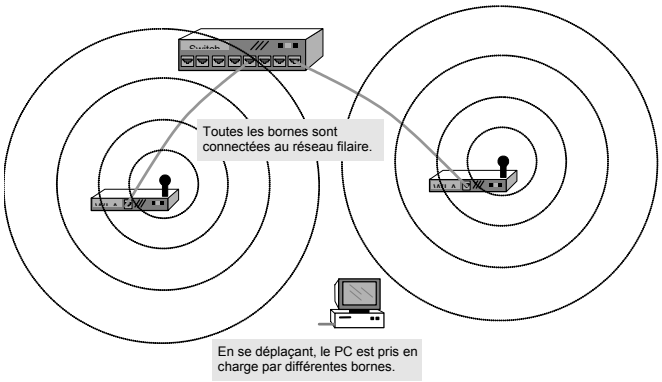
L'intérêt d'une borne est qu'elle permet de raccorder le réseau Ethernet sans fil au réseau Ethernet filaire. Elle intègre alors la fonction de **pont** (*bridge*, en anglais) entre l'Ethernet 802.3 et l'Ethernet 802.11.



Le réseau, qui permet aux BSS de communiquer entre eux, est appelé, par la norme 802.11, un DS (*Distribution System*). L'ensemble des BSS connectés à un DS forme un ESS (*Extended Service Set*).

Ce type d'architecture permet de créer un WLAN sur l'ensemble d'un immeuble, en positionnant les bornes à intervalle régulier. Deux applications sont alors envisageables :

- La création de « hot-spots », c'est-à-dire de zones délimitées telles que les salles de réunion. Les zones de couverture des bornes peuvent être disjointes.
- La création d'un réseau étendu permettant aux utilisateurs de se déplacer dans tout l'immeuble. Les zones de couverture des bornes doivent alors se chevaucher afin de permettre à vos utilisateurs de se déplacer sans perdre la connexion au réseau (service appelé *roaming*). En outre, les bornes doivent être situées dans le même domaine de diffusion MAC (cf. chapitre 3) et être configurées dans le même sous-réseau IP.



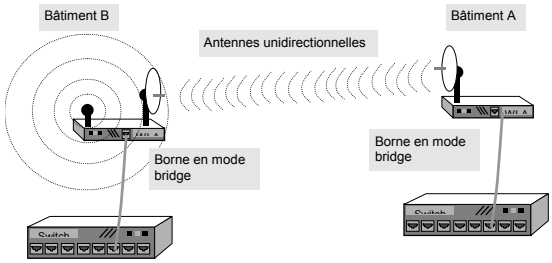
Ce dernier cas de figure est le plus délicat, car il nécessite de réaliser une étude de couverture. Le plus simple est de vous inspirer de la procédure suivante :

- Identifier les environnements types dans votre immeuble (zones de bureaux, open-spaces, zones techniques, halls d'entrées, etc.). Chaque environnement est, en effet, caractérisé par des objets réfléchissant et atténuant les ondes radio (cloisons, mobilier, poteaux, etc.).
- Installer quelques bornes à différents endroits dans chacun de ces environnements types, en respectant les distances maximales indiquées par le constructeur. Les meilleures positions sont en hauteur, fixées à un mur ou sous le faux plafond.
- Se promener avec un PC équipé d'un logiciel affichant la puissance du signal reçu. La plupart du temps, ce type de logiciel est fourni avec la carte WLAN. Il suffit alors de repérer les zones d'ombres, et de déplacer les bornes pour trouver la bonne configuration. Certains équipements proposent d'ailleurs le mode *survey* permettant de réaliser ce type de test.

Avec ces expériences, vous pourrez alors déterminer le nombre de bornes nécessaires par mètre carré et appliquer ces schémas pour chacun des environnements rencontrés dans tout l'immeuble.

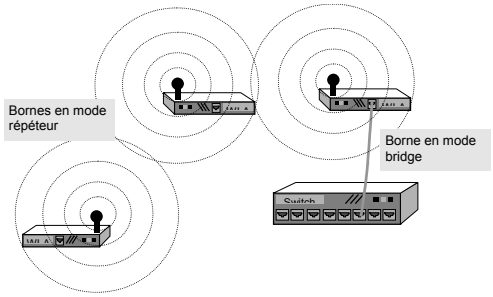
Extension du réseau

La connexion de deux immeubles, séparés de quelques centaines de mètres par un WLAN, peut s'avérer une alternative intéressante à la mise en place d'un câblage. L'utilisation d'antennes unidirectionnelles est alors conseillée afin de canaliser les ondes radio et l'énergie utilisée.



Une autre manière d'augmenter la portée du WLAN est de cascader les bornes grâce à la fonction de **répéteur**.

En contrepartie, le débit global est divisé par deux, puisque les bornes doivent partager le même canal pour communiquer entre elles et avec les stations. La borne recevant une trame de la part d'une station, doit, en effet, la retransmettre à l'autre borne et utilise pour cela la même fréquence.

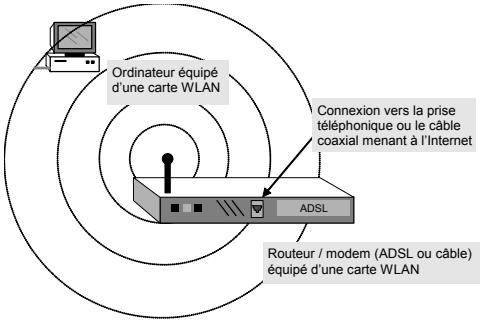


Le système de distribution est ici bâti sur un WLAN.

Le WLAN à la maison

La solution la plus simple et la moins chère repose sur une architecture ad-hoc, si les contraintes des distances et d'atténuation des signaux vous le permettent. Dans le cas contraire, l'installation d'une ou plusieurs bases sera nécessaire pour bâtir des architectures similaires à celles qui viennent d'être montrées.

Si cependant, vous disposez d'un abonnement à Internet *via* un accès ADSL ou le câble, la solution idéale repose sur un routeur modem sans fil, équipement regroupant les fonctions de point d'accès Wi-Fi, de modem ADSL (ou câble) et de routeur.

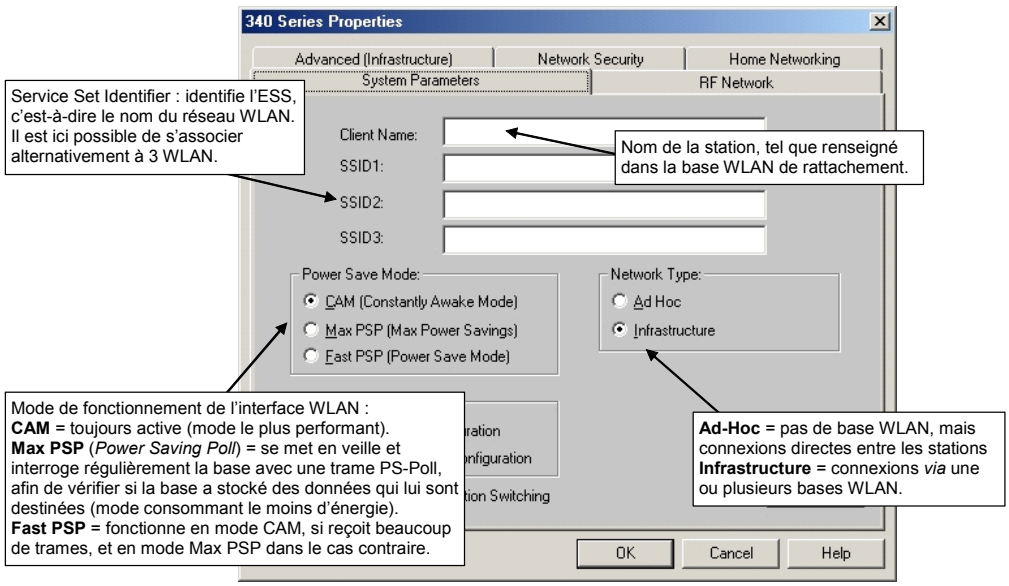


La fonction de routeur permet à tous vos ordinateurs de partager l'abonnement à Internet, qui est généralement limité, par votre fournisseur d'accès, à une seule adresse IP.

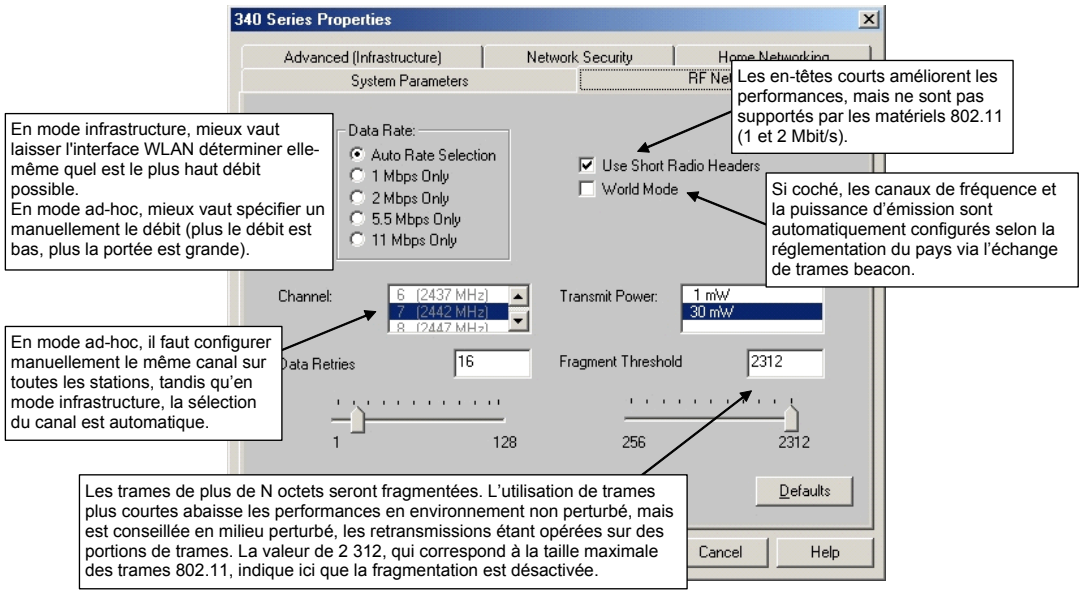
Certains équipements intègrent également un commutateur Ethernet équipé de 4 ou 8 ports 10/100 Mbit/s, vous permettant ainsi de connecter vos équipements historiques, comme les imprimantes, qui ne disposent pas encore d'interface WLAN.

La configuration des postes client

Les paramètres système



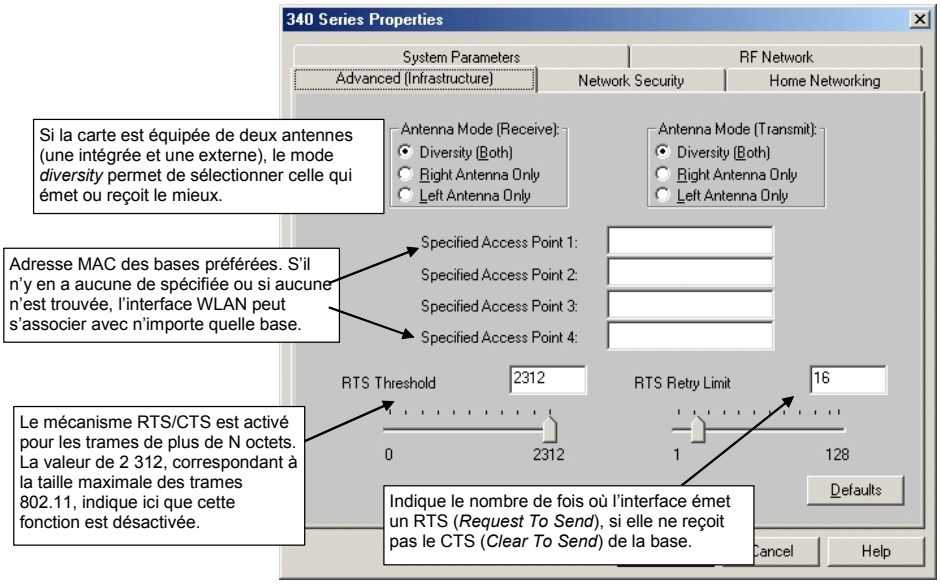
Les paramètres de communication



La technique DSSS employée par la norme 802.11b diffuse les signaux sur une largeur de bande de 22 MHz et impose un espacement de 25 MHz entre les canaux afin d'éviter toute interférence entre les signaux.

La France n'autorisant que 4 canaux, de 10 à 13, vous ne pouvez donc sélectionner qu'un seul canal pour votre réseau. En revanche, aux États-Unis et dans les pays européens respectant les directives de l'ETSI (*European Telecommunications Standards Institute*), il est possible de co-localiser 3 réseaux en sélectionnant les canaux 1, 6 et 11, respectant ainsi l'espacement de 25 MHz nécessaire pour éviter les interférences. La bande passante offerte aux utilisateurs est ainsi de 3 fois 11 Mbit/s.

Les paramètres propres au protocole 802.11b



Le mécanisme RTS/CTS induit un overhead propre aux trames de services échangées et aux temps d'attente avant les transmissions.

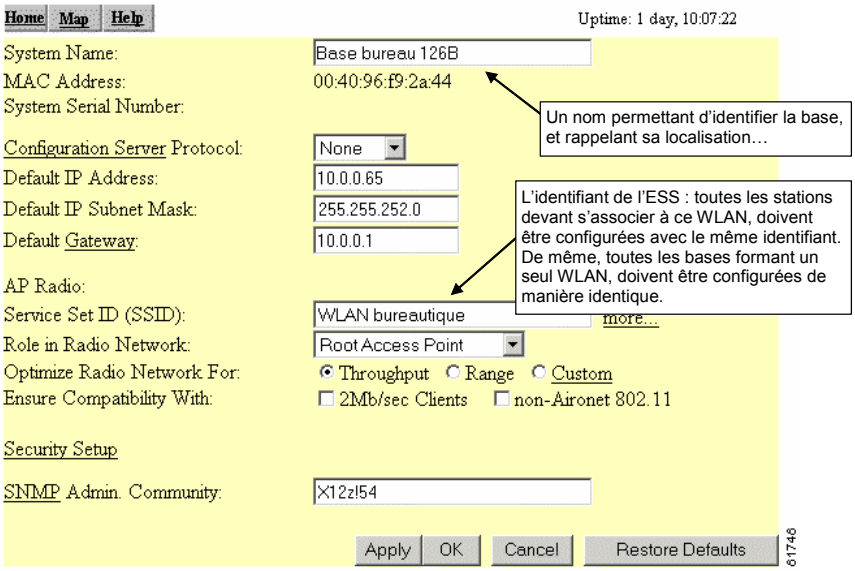
Son activation permet de limiter les collisions, et notamment de résoudre le problème du « nœud caché » : deux stations A et B sont situées de part et d'autre d'une borne, mais de par leur éloignement, elles sont hors de portée des signaux ; A ne peut ainsi pas recevoir les signaux émis par B et inversement.

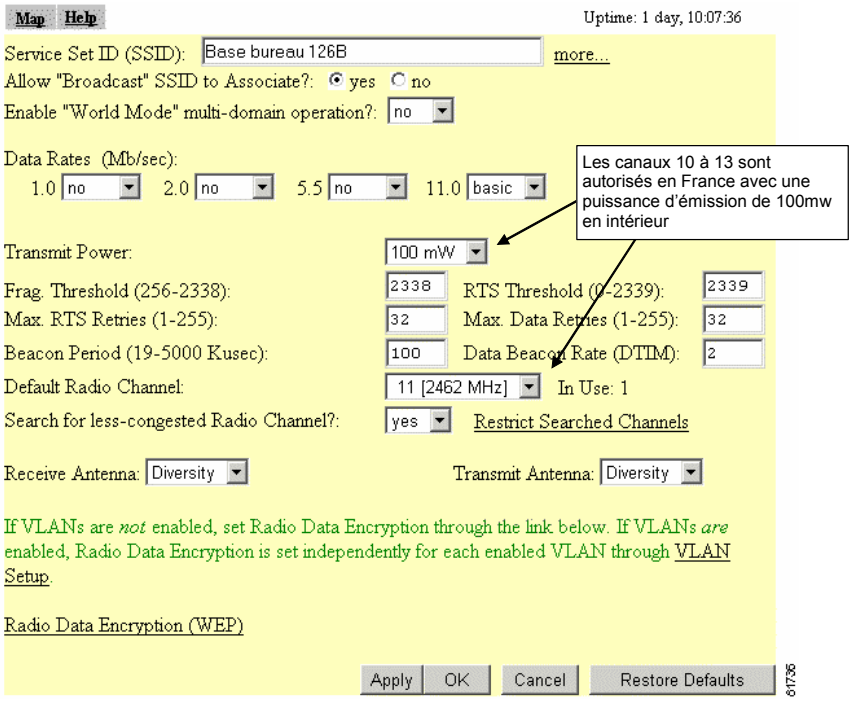
L'activation du mécanisme RTS/CTS est recommandée pour un WLAN supportant un fort trafic, du fait d'un nombre important d'utilisateurs ou de volumes de données transférés. Il est également recommandé si les signaux sont perturbés par des interférences ou par des surfaces métalliques favorisant les chemins multiples (voir description de ce phénomène à la fin de ce chapitre). Il peut, en revanche, être désactivé dans des configurations à faible trafic, ce qui permet d'augmenter les performances globales du WLAN.

Si vous constatez un nombre élevé de collisions et/ou de faibles débits, vous pouvez abaisser le seuil à partir duquel le RTS/CTS est activé.

La configuration des points d'accès

La configuration des bases est similaire à celles des stations. L'administrateur doit s'assurer que les paramètres sont identiques dans tous les composants de son WLAN.





Comment sécuriser son réseau ?

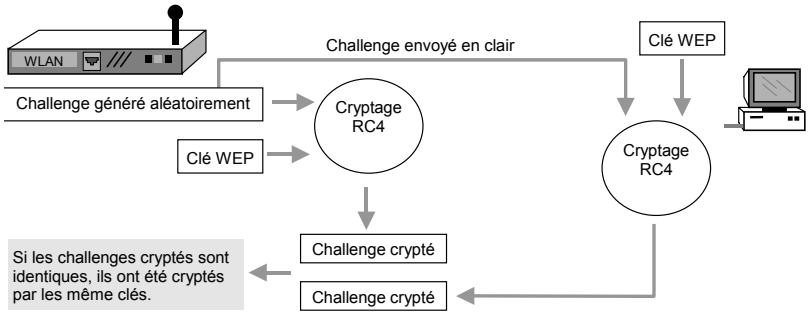
La sécurité devient une préoccupation majeure à partir du moment où les ondes radio peuvent être interceptées de manière complètement anonyme. Les standards 802.11a et b proposent trois mécanismes de bases pour renforcer (un peu) la sécurité : l'identification, les adresses MAC et le chiffrement.

La sécurité de base : WEP

La base peut identifier les stations autorisées à participer au WLAN en enregistrant leurs adresses MAC. Cette méthode présente le double inconvénient de devoir mettre à jour la liste à chaque installation d'une nouvelle carte sur une des quelconques stations de votre réseau, et d'être inefficace sur le plan de la sécurité dès lors qu'il est aisé de changer l'adresse MAC inscrite dans la PROM de la carte réseau.

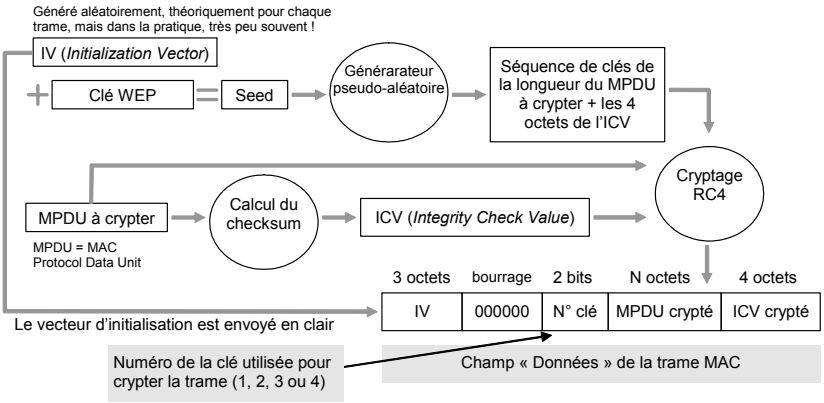
Le SSID (*Service Set Identifier*), programmé dans les bases et les stations, permet d'identifier le WLAN, appelé ESS (*Extended Service Set*) dans la terminologie 802.11. Il est donc tentant d'utiliser cet identifiant comme mot de passe. L'inconvénient est qu'il est statique et qu'il faut le changer régulièrement, et cela sur tous les composants du WLAN. En outre, le SSID est affiché en clair dans les outils de configuration.

Le protocole WEP (*Wired Equivalent Privacy*) offre un mécanisme d'authentification rudimentaire sur la base d'un mot de passe, appelé clé WEP, programmé dans tous les équipements, à l'instar du SSID. D'une part, la mise à jour de cette clé est fastidieuse, mais en plus sa large diffusion auprès des exploitants ou des utilisateurs augmente le risque de sa divulgation.



Le protocole WEP permet également de crypter les données transportées par la trame, hors en-tête. L'algorithme de chiffrement utilisé est RC4 (clé secrète de 40 bits en standard, jusqu'à 104 bits selon les constructeurs), qui ne présente pas la meilleure robustesse contre les attaques.

Deux clés sont générées : une clé de session pour les communications entre la base et la station, ainsi qu'une clé de groupe pour les flux multicasts.



On le voit, les quatre éléments proposés en standard ne permettent pas d'assurer une sécurité optimale de votre réseau sans fil. Des audits, réalisés en France et aux États-Unis, montrent qu'il est même possible de pénétrer sans difficulté un réseau WLAN, simplement en se promenant dans la rue avec son ordinateur équipé d'une carte 802.11.

WPA et 802.11i

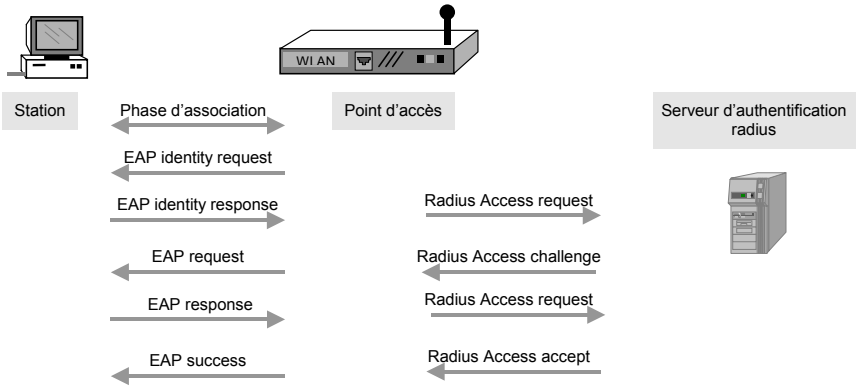
La solution repose sur le standard WPA et mieux encore, sur le standard 802.11i. WPA (*Wi-Fi Protected Access*) est un sous-ensemble de 802.11i, dont il reprend les composants déjà figés et ne requérant pas d'upgrade matériel :

- EAP (*Extensible Authentication Protocol*), normalisé 802.1x, pour l'authentification des équipements ;
- TKIP (*Temporal Key Integrity Protocol*) renforçant la sécurité du protocole WEP.

Le protocole EAP (RFC 2284) permet aux stations et aux bases de s'authentifier *via* un serveur d'authentification, soit intégré à la base, soit hébergé sur un serveur Radius.

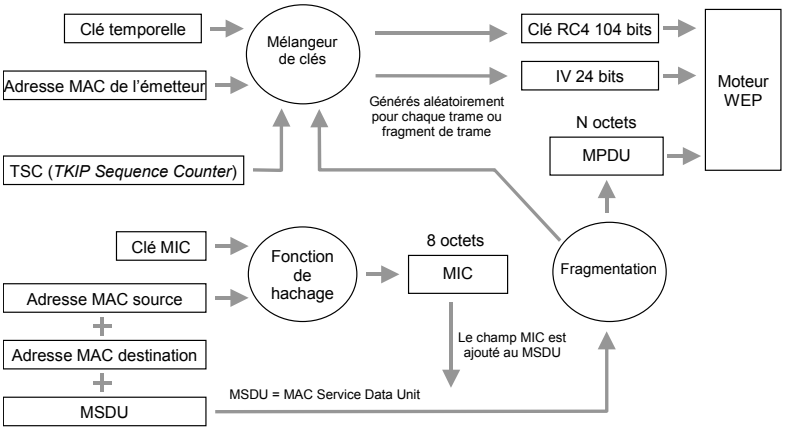
Avant que la station ne soit authentifiée, la base autorise celle-ci à ne communiquer qu'avec le serveur d'authentification. Le mécanisme d'authentification peut alors être de toute nature : un mot de passe, un challenge, un certificat, une token card ou encore une carte à puce.

Déclinaisons d'EAP	Mécanisme d'authentification
EAP-TLS	Certificat Windows XP (RFC 2716)
PEAP (*Protected EAP*)	Mot de passe unique (OTP – *One Time Password*)
EAP-MD5	Mot de passe (challenge avec algorithme de hachage MD5)
EAP-SIM	Carte à puce
LEAP (*Lightweight EAP*)	Mot de passe géré par la carte (propriétaire Cisco)

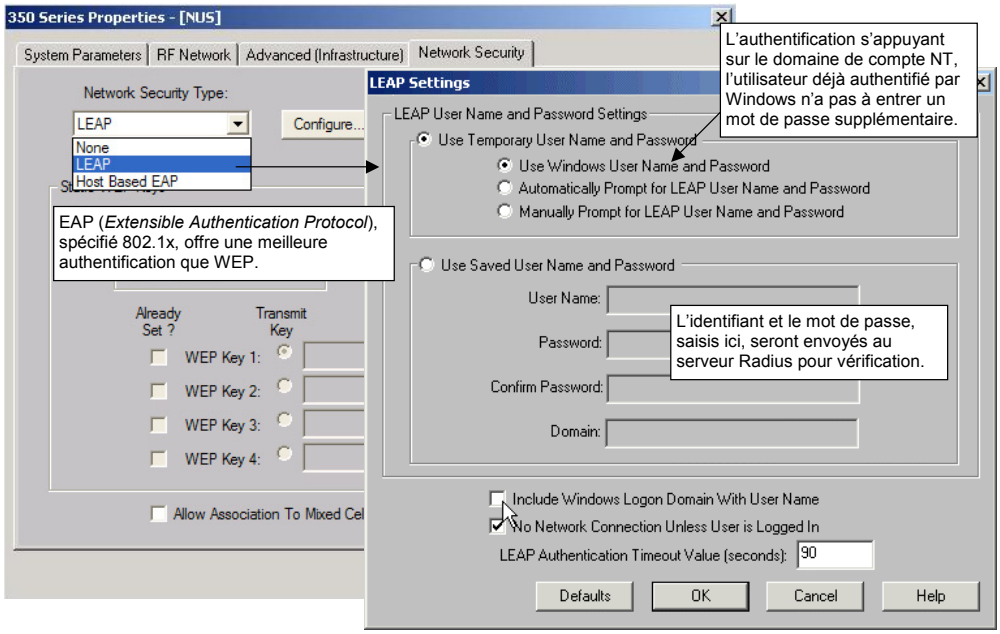


Le protocole TKIP est une amélioration de WEP en ce :

- qu'il impose un vecteur d'initialisation et une clé RC4 différents pour chaque trame ;
- qu'il permet de détecter toute modification illégale intervenant sur des trames, en calculant une clé de hachage irréversible transportée dans le champ MIC (*Message Integrity Code*) ;
- qu'il lance des contre-mesures si une tentative de modification ou d'attaque sur le mot de passe est détectée.



La norme 802.11i spécifie également l'utilisation de l'algorithme de chiffrement AES (*Advanced Encryption Standard*) plus robuste que RC4.



La boucle locale radio

La BLR, en abrégé, couvre le dernier kilomètre (*the last mile* en anglais), c'est-à-dire la portion de réseau entre le client (une maison) et le point d'accès le plus proche au réseau opérateur. Elle constitue de fait une solution intéressante pour que des opérateurs alternatifs, nouveaux venus sur le marché, puissent proposer des offres concurrentes aux boucles locales filaires proposées par un opérateur historique tel que France Télécom en France.

Les technologies proposées répondent à un besoin où émetteur et récepteur sont fixes : le point d'accès (une antenne sur un toit ou un pylône) et la maison d'un particulier.

Technologie	Bande de fréquences	Largeur de bande par cellule	Taille approximative de la cellule	Débit
MMDS (*Multichannel Multipoint Distribution Service*)	2,5 à 2,7 GHz (USA) 3,5 GHz (France)	~200 MHz	10 à 15 km	1 à 2 Mbit/s
LMDS (*Local Multipoint Distribution Service*)	28 à 38 GHz (USA) 26, 38 et 32 (France)	~500 MHz	5 à 8 km	155 Mbit/s à 500 Mbit/s
IEEE 802.16 IEEE 802.16a	10 à 66 GHz 2-11 GHz			2 à 155 Mbit/s

Ces technologies, notamment de par les longueurs d'onde utilisées, ne tolèrent aucun obstacle ou objet réfléchissant entre l'émetteur et le récepteur autour d'une zone en forme d'ellipse appelée zone de Fresnel. Elles sont également gênées par les objets réfléchissant les ondes, tels un plan d'eau, des surfaces métalliques ou, dans une moindre mesure, les gouttes de pluie : le récepteur est ainsi susceptible de recevoir plusieurs signaux décalés dans le temps, l'original et les réflexions de ce même signal. Ce phénomène est connu sous le nom de **chemins multiples**. D'autres objets, comme les arbres, absorbent les ondes et contribuent à les affaiblir : l'émetteur peut ainsi recevoir un signal trop faible pour être interprété.

Les dénominations MMDS et LMDS recouvrent une gamme d'application liée à une bande de fréquence autorisée, et reposent sur des technologies propriétaire.

Le développement de la boucle locale radio devrait venir du comité IEEE 802.16 qui définit les bases du WMAN (*Wireless Metropolitan Area Network*). À la différence des technologies MMDS et LMDS, ce standard spécifie les couches physique et MAC ainsi que l'interfaçage avec les couches supérieures (Ethernet, IP, ATM).

La couche physique repose sur plusieurs techniques de multiplexage variant selon les applications : temporel adaptatif appelé DAMA-TDMA (*Demand Assignment Multiple Access-Time Division Multiple Access*), OFDM (*OrthogonalFrequency Division Multiplexing*) ou OFDMA (*OFDM Access*).

Côté européen, le comité BRAN (*Broadband Radio Access Networks*) de l'ETSI travaille au projet HIPERMAN, similaire au 802.16.

Les micro-ondes point à point

Une autre application de la transmission sans fil dans la gamme des micro-ondes (2 à 40 GHz) est la connexion point à point (*PtP Microwave*). Toujours soumise aux aléas climatiques, au brouillard notamment, les débits peuvent néanmoins atteindre plusieurs centaines de Mbit/s. Étant donné la nature point à point de ce type de liaison, ce genre d'application ne nécessite bien souvent aucune licence d'exploitation.

Les technologies sous-jacentes, qui ne sont pas nouvelles puisqu'elles sont mises en œuvre depuis plus de dix ans notamment aux États-Unis, sont cette fois-ci directement utilisables par les entreprises pour leurs besoins propres. Des constructeurs proposent, en effet, des équipements dédiés permettant d'interconnecter des réseaux locaux, jusqu'à des débits pouvant atteindre 622 Mbit/s, entre deux immeubles séparés de quelques centaines de mètres à plusieurs kilomètres.

DEUXIÈME PARTIE

Les réseaux IP

5

Démarrer son réseau IP

Des architectures viennent d'être élaborées, des réseaux viennent d'être construits, des protocoles sont utilisés. Mais comment tout cela s'imbrique-t-il ? Comment IP, TCP et Ethernet fonctionnent-ils ensemble ?

Ce chapitre est l'occasion d'aller plus loin dans la connaissance de ces protocoles et de votre réseau, et de vous donner par la même occasion une vision plus globale des réseaux. Car, comprendre, c'est pouvoir construire des réseaux de plus en plus complexes comme le requièrent les applications multimédias d'aujourd'hui.

Jusqu'à présent, nous ne nous sommes préoccupés que du matériel mais, avec le spanning tree, introduit au chapitre 3, nous devons désormais nous préoccuper du paramétrage logiciel des équipements réseau.

Comprendre, c'est donc maîtriser le fonctionnement de son réseau.

Deuxième exemple, celui de l'adresse IP que nous avons utilisé sans en bien comprendre les tenants et aboutissants. Cet aspect logiciel doit maintenant être expliqué, car les choix que vous prenez lorsque vous commencez par construire un petit réseau peuvent ensuite peser bien des années plus tard, lorsque celui-ci a pris de l'ampleur.

Comprendre, c'est donc anticiper et faire les bons choix pour l'avenir.

Dans ce chapitre, vous apprendrez ainsi :

- à définir un plan d'adressage IP ;
- à comprendre l'interaction des protocoles Ethernet et IP ;
- le fonctionnement d'un réseau local ;
- le fonctionnement des protocoles IP, TCP et UDP.

Le plan d'adressage IP

À plusieurs reprises déjà, nous avons parlé d'adresses IP sans vraiment nous en préoccuper. Il est vrai que nous n'en avions pas réellement l'usage ; il suffisait simplement de saisir une adresse unique pour chaque station du réseau.

Mais notre réseau prend de l'ampleur et nous devons désormais organiser l'affectation des adresses. Les ISP ne procèdent pas autrement au sein de l'Internet.

Comme cela a été décrit au chapitre 3, une adresse IP s'écrit avec quatre numéros, compris entre 1 et 255, séparés par des points, par exemple 192.162.0.1. Une partie de cette adresse désigne un réseau, l'autre le numéro de station au sein de ce réseau. Jusqu'à présent, nous nous sommes arrangés pour configurer toutes nos stations dans le même réseau IP.

On peut se poser la question suivante : pourquoi faut-il des adresses IP alors qu'il existe déjà des adresses MAC ? D'abord, Ethernet est un réseau local, qui n'a donc qu'une portée géographique limitée. Ensuite, il existe des dizaines de réseaux de niveau 1 et 2 différents avec chacun un adressage physique qui lui est propre. Or, les PC, même situés sur des réseaux différents, doivent pouvoir communiquer ensemble. Il faut donc un protocole de niveau supérieur, dit de niveau 3 (couche réseau), qui permet de fédérer ces réseaux avec un adressage unique. On trouve ainsi IP sur Ethernet et PPP, mais aussi sur Token-Ring, ATM, etc.

POURQUOI UN PLAN D'ADRESSAGE ?

L'objectif premier du plan d'adressage est d'éviter la **duplication** accidentelle des adresses. Pour l'adressage MAC, un plan n'est pas utile car les adresses sont affectées aux cartes par les constructeurs. En revanche, l'affectation des adresses IP relève de votre responsabilité, ou de celle du NIC pour le réseau public Internet.

Le plan d'adressage permet également de **contrôler** le fonctionnement de votre réseau IP. En effet, l'affectation des adresses IP doit répondre à des règles précises sous peine d'aboutir à des dysfonctionnements (connexions impossibles, voire intermittentes, etc.).

En définitive, le plan d'adressage permet **d'organiser l'exploitation** de votre intranet.

IP permet aussi de partitionner les réseaux. En effet, de nombreux protocoles utilisent abondamment les broadcasts et multicasts, et il est préférable de limiter la diffusion de ces types de trames. Si votre intranet est connecté à l'Internet, il n'est pas envisageable de recevoir des trames multicast et broadcast émises par un employé de la société X.

De plus, l'interconnexion des sites coûte cher compte tenu des distances. Il est donc judicieux de limiter le trafic afin de ne pas surcharger inutilement les liaisons par des broadcasts.

La démarche

Tout d'abord, il est conseillé de retenir un adressage privé, c'est-à-dire complètement séparé de celui de l'Internet, ceci pour des questions de simplicité et de sécurité. Il est toujours possible d'opter pour un adressage public, mais l'obtention de telles adresses est très difficile, car il faut justifier de leur usage auprès des organismes de régulation de l'Internet.

L'ADRESSAGE IP V4 (RFC 791)

IP (*Internet Protocol*) définit **un réseau virtuel** reposant sur des réseaux physiques de différente nature (Ethernet et PPP, par exemple). Pour ce faire, IP utilise un **adressage logique** différent de l'adressage physique (MAC, PPP ou autre).

Une adresse IP est découpée en un numéro de réseau et un numéro de station au sein de ce réseau. Il existe trois **classes d'adresses** unicast en fonction de la taille du réseau (c'est-à-dire du nombre de stations par réseau). Pour différencier la partie réseau (*subnet*) de la partie station (*host*), IP utilise un **masque** dont tous les bits à 1 représentent la partie réseau.

Classe A — Masque naturel = 255.0.0.0

0	7 bits pour le n° de réseau de 1 à 127	24 bits pour le n° de station, de 1 à 16 777 214

126 réseaux
de 1.0.0.0 à 126.0.0.0

0.x.x.x	Réservé
127.x.x.x	Adresse de boucle locale (loopback)
x.255.255.255	Broadcast : toutes les stations sur le réseau x

Classe B — Masque naturel = 255.255.0.0

1	0	14 bits pour le n° de réseau, de 1 à 16 383	16 bits pour le n° de station, de 1 à 65 534

16 382 réseaux
de 128.1.0.0 à 191.254.0.0

128.0.x.x	Réservé
191.255.x.x	Réservé
x.x.255.255	Broadcast : toutes les stations sur le réseau x.x

Classe C — Masque naturel = 255.255.255.0

1	1	0	21 bits pour le n° de réseau, de 1 à 2 097 151	8 bits pour le n° de station de 1 à 254

2 097 150 réseaux
de 192.0.1.0 à 223.255.254.0

192.0.0.x	Réservé
223.255.255.x	Réservé
x.x.x.255	Broadcast : toutes les stations sur le réseau x.x.x

Deux valeurs sont réservées dans la partie station de l'adresse : 0 pour désigner le réseau lui-même et 255 (tous les bits à 1) pour désigner toutes les stations au sein de ce réseau (broadcast).

Il existe également une classe d'adresses multicast permettant de désigner des groupes de stations.

Classe D — Pas de masque

1	1	1	0	28 bits pour le n° de groupe, de 1 à 268 435 456

268 435 455 groupes
de 224.0.0.0 à 239.255.255.255

224.0.0.0	Réservé
224.0.0.1	Tous les groupes sur ce réseau local

Des n° sont déjà réservés (*well known group*)

La classe E (premiers à bits 11110) définit une classe d'adresses expérimentales. Elle n'est jamais utilisée.

L'adresse 255.255.255.255 désigne toutes les stations sur le réseau de l'émetteur du paquet (broadcast IP).

Il se peut donc que vous utilisiez des adresses déjà affectées sur l'Internet, mais cela n'a pas d'importance car votre intranet est isolé. Cela ne vous empêchera cependant pas de l'interconnecter avec l'Internet.

La seconde décision concerne le choix de la classe d'adresse IP. Ce choix dépend du nombre de stations présentes sur votre réseau. Si ce nombre dépasse 254, une classe B s'impose. Une classe A n'est pas utile, car une classe B offre 65 534 adresses de stations, ce qui est largement suffisant. De plus, une classe A est limitée à 126 réseaux IP, ce qui, pour les grands réseaux, peut être un handicap.

En résumé, notre choix s'est provisoirement porté sur un plan d'adressage privé de classe B, ce qui nous donne 16 382 réseaux possibles contenant chacun 65 534 stations. Aux sections suivantes, d'autres considérations viendront modifier ce choix.

Les principes de base

L'adressage IP est très souple et permet de faire tout ce que l'on veut. Afin d'éviter toute mauvaise surprise, il est conseillé de suivre les principes suivants :

- Règle 1 : un réseau IP ne doit pas chevaucher plusieurs sites.
- Règle 2 : il peut y avoir plusieurs réseaux IP sur un site.
- Règle 3 : s'il y a plusieurs réseaux IP sur un site, choisir des numéros contigus. Cela simplifiera le routage et permettra aux routeurs d'agréger les routes selon la méthode CIDR (*Classless Inter-Domain Routing*) expliquée au chapitre 12.
- Règle 4 : limiter le nombre de réseaux IP. Cela simplifiera les connexions à l'Internet.

Le protocole IP impose qu'une station se trouvant dans un réseau IP ne puisse pas communiquer directement avec une station se trouvant dans un autre réseau IP, même si elles sont connectées au même segment Ethernet. Les réseaux sont segmentés de manière logique ; en d'autres termes, ils sont partitionnés.

La solution repose sur l'utilisation d'un **routeur** dont le rôle est d'interconnecter les réseaux IP, quelle que soit leur localisation géographique.

On verra au chapitre 12 qu'il existe un moyen de lever cette contrainte imposée par IP.

De toute façon, l'utilisation d'un routeur s'impose dès que vous devez relier deux sites sur de longues distances. L'Internet comporte des dizaines de milliers de routeurs. Donc, autant prendre en compte cette contrainte dès le début de l'élaboration du plan d'adressage.

QU'EST-CE QU'UN ROUTEUR ?

Un routeur est un commutateur de niveau 3, c'est-à-dire qui commute les protocoles de la couche réseau, tels que IP. La commutation des paquets IP est plus complexe que celle des trames Ethernet. On emploiera donc plutôt le terme de **routage**.

Ce mécanisme consiste à analyser l'adresse de destination du paquet IP et à le transmettre sur le bon port (appelé **interface**). Il utilise pour cela des algorithmes de routage, tels que **OSPF** (*Open Shortest Path First*) qui permettent de calculer les meilleures routes en fonction des numéros de réseau IP.

Comme pour les PC, une interface routeur est associée à au moins un réseau IP.

Impact sur l'Internet

Tôt ou tard, l'interconnexion de votre réseau avec l'Internet sera nécessaire. Comment éviter que vos adresses internes entrent en conflit avec celles de l'Internet ?

La solution repose sur l'utilisation de la **translation d'adresses** (cf. chapitre 17). Cette technique permet de masquer votre plan d'adressage privé vis-à-vis des utilisateurs situés sur l'Internet.

Une solution complémentaire à la première repose sur le non-routage de certaines adresses. La RFC 1918 précise que certaines adresses ont été réservées pour l'adressage privé. Le respect par tous les ISP de cette RFC garantit que ces adresses ne seront jamais routées sur l'Internet.

Réseaux réservés (RFC 1918)	Espace d'adressage
10.0.0.0	1 réseau de classe A
De 172.16.0.0 à 172.31.0.0	16 réseaux de classe B
De 192.168.0.0 à 192.168.255.0	256 réseaux de classe C

On peut donc utiliser ces adresses pour notre réseau privé, sans que cela soit pour autant une obligation. L'essentiel d'une interconnexion avec l'Internet repose, en effet, sur la translation d'adresses.

Or, pour les grands réseaux, le nombre de réseaux IP à translater est source de complexité : s'il y a quarante sites mais un seul point de sortie vers l'Internet, le firewall devra prendre en compte quarante réseaux IP dans ses règles de translation d'adresses.

Afin de simplifier cette configuration, il faudrait donc pouvoir ne translater qu'un réseau IP au niveau du firewall (respect de la règle 4) tout en ayant autant de subnets IP que nécessaire pour notre intranet. La solution repose sur la création de sous-réseaux IP.

Les sous-réseaux IP

Le principe des sous-réseaux (*subnet*) consiste à étendre le nombre de bits désignant la partie réseau. Le nombre de stations par sous-réseau diminue donc d'autant.

Classe	Masque naturel	Nombre de bits affectés au numéro de réseau	Extension possible : nombre de bits affectés au sous-réseau
A	255.0.0.0	8	+ 1 à + 22 bits
B	255.255.0.0	16	+ 1 à + 14 bits
C	255.255.255.0	24	+ 1 à + 6 bits

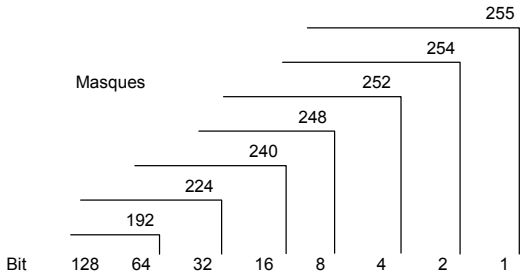
La partie station de l'adresse doit comporter au moins 2 bits afin que cette dernière soit valide.

	Numéro de station	Conclusion
1 bit	1 = broadcast 0 = ce réseau	**Interdit**. Il ne reste aucun bit pour désigner un sous-réseau ou une station.
2 bits	11 = broadcast 00 = ce réseau 01 = station n° 1 10 = station n° 2	**OK**. 2 est le nombre minimal de bits devant être réservés aux sous-réseaux et stations.

La notation décimale (octet par octet) est rendue difficile lorsque le sous-réseau ne porte pas sur un multiple de 8 bits. C'est pourquoi la notation « / [*nombre de bits affectés à la partie réseau*] » est plus souvent utilisée.

Numéro de réseau / nombre de bits réservés à la partie réseau	Masque	Commentaire
10.0.0.0 / 10 subnet de +2 bits	255.192.0.0	Permet de créer 4 sous-réseaux, de 10.0 à 10.3
10.0.0.0 / 16 subnet de +8 bits	255.255.0.0	On dit que la classe A est « subnettée » sur une classe B
194.50.0.0 / 19 subnet de +3 bits	255.255.224.0	Permet de créer 8 sous-réseaux
194.50.0.0 / 24 subnet de +8 bits	255.255.255.0	On dit que la classe B est « subnettée » sur une classe C

Figure 5-1.
Les masques de sous-réseaux.



Notre choix initial portait sur une classe B. Si nous voulons limiter le nombre de réseaux IP et conserver la même souplesse que la classe B, il faut donc retenir une classe A « subnettée » sur une classe B.

Cela nous offrirait 256 sous-réseaux. Si, dans le futur, ce chiffre était dépassé, on pourrait toujours ajouter un autre réseau de classe A (il ne ferait pas partie de la RFC 1918, mais ce n'est pas réellement important) et le « subnetter », ou ajouter une classe B à notre plan d'adressage. Notre but est simplement de limiter le nombre de réseaux IP.

Nous choisissons donc l'adresse de classe A, 10.0.0.0, issue de la RFC 1918. Étant donné le subnet choisi, notre masque sera donc : 255.255.0.0. Mais ce choix est encore provisoire.

Méthode d'affectation des réseaux LAN

Le plus simple est d'affecter les réseaux par site (respect de la règle 1). Au lieu d'affecter séquentiellement le numéro, on peut l'incrémenter de 4 ou 8, ce qui laisse la possibilité d'étendre le subnet affecté au site (respect de la règle 2). L'ajout d'un réseau sur un site se traduira donc par l'affectation du numéro de réseau suivant (respect de la règle 3).

Une première version de notre plan d'adressage serait donc la suivante :

Réseau	Site
10.0.0.0/16	Paris : 1 réseau de 65 534 stations
De 10.1.0.0/16 à 10.3.0.0/16	Non affecté (réservé aux extensions de Paris)
10.4.0.0/16	Toulouse : 1 réseau de 65 534 stations
De 10.5.0.0/16 à 10.7.0.0/16	Non affecté (réservé aux extensions de Toulouse)
Etc.	
De 10.248.0.0 à 10.255.0.0	Réseaux non affectés

L'incrément de 4 a été soigneusement choisi, de manière à obtenir des réseaux contigus. Ainsi, le site de Paris dispose de quatre réseaux : 10.0.0.0, 10.1.0.0, 10.2.0.0 et 10.3.0.0, avec chacun un masque à 255.255.0.0. Mais cette manière de découper les réseaux est quelque peu rigide, car la région de Paris peut comprendre à la fois des petits sites et des gros sites.

Une autre façon de voir les choses est de considérer le réseau 10.0.0.0 avec le masque 255.252.0.0 (soit 10.0.0.0/14), ce qui offre 262 142 adresses (65 536 × 4 –2) pour le subnet 10.0.0.0 affecté à Paris (de 10.0.0.0 à 10.3.255.255).

Figure 5-2.
Création d'un subnet.

Adresse	10.						0.		0.0
Masque	255.						252.		0.0
Bit		128	64	32	16	8	4	2	1
/14	8 bits						+ 6 bits		

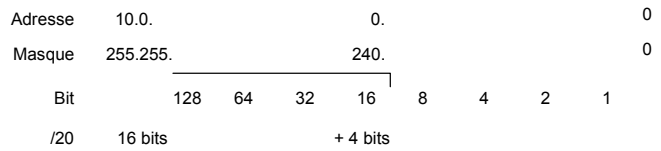
En définitive, notre plan d'adressage se présente en réalité sous la forme suivante :

Subnets du réseau 10.0.0.0/8	Site
10.0.0.0/14 255.252.0.0	Région parisienne
10.4.0.0/14 255.252.0.0	Région toulousaine
10.8.0.0/14 255.252.0.0	Strasbourg
Etc.	
De 10.248.0.0 à 10.255.252.0	Réseaux non affectés

Au sein de ce réseau, il est alors possible de créer d'autres subnets dont la taille varie en fonction de l'importance du site. En faisant varier la longueur du masque on crée ainsi des **subnets variables** (RFC 1219).

Par exemple, au sein de la plage d'adresses affectée à la région parisienne, on peut réserver le subnet suivant à un site de moyenne importance : 10.0.0.0/20 (masque égal à 255.255.240.0), soit 4 094 adresses (16 × 256 –2), de 10.0.0.1 à 10.15.255.254.

Figure 5-3.
Création d'un deuxième subnet.



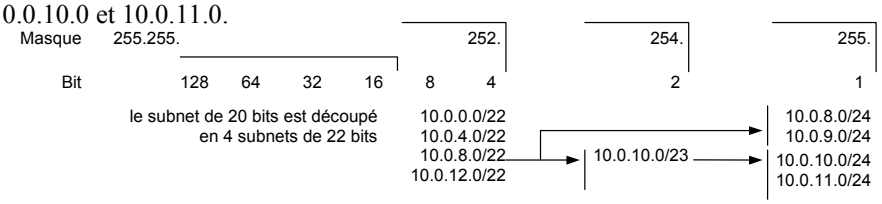
Au sein de ce site, il peut ensuite être nécessaire de créer des réseaux de différentes tailles, par exemple un réseau principal et de nombreux petits sous-réseaux dédiés connectés, par exemple, à un firewall.

Subnets du réseau 10.0.0.0/14	Fonction
10.0.0.0/22 255.255.252.0	Réseau principal (1 022 adresses)
10.0.4.0/22 255.255.252.0	Réservé à l'extension du réseau principal (*) ou à la création d'un deuxième réseau
10.0.8.0/24 255.255.255.0	Réseaux dédiés au firewall (254 adresses)

(*) Si le réseau principal est étendu, il suffit de changer le masque qui devient 255.255.248.0, ce qui donne le réseau 10.0.0.0/21.

Le réseau 10.0.10.0/23 peut également être découpé en deux subnets de classe C (masque de 24 bits) 10.0.10.0 et 10.0.11.0.

Figure 5-4
Extension des subnets.



Les subnets de classe C ainsi créés (10.0.8.0, 10.0.9.0, etc.) peuvent à leur tour être découpés en de très petits réseaux, juste assez grands pour connecter un routeur et quelques machines.

Subnets du réseau 10.0.8.0/24	Fonction du réseau dédié
10.0.8.0/27 255.255.255.224	Serveurs publics (30 adresses)
10.0.8.32/27 255.255.255.224	Réservé (30 adresses)
10.0.8.64/26 255.255.255.192	Accès distants (62 adresses)
10.0.8.128/28 255.255.255.240	Accès externes (14 adresses)
10.0.8.144/28 255.255.255.240	Réservé (14 adresses)
10.0.8.160/27 255.255.255.224	PABX (30 adresses)
10.0.8.192/26 255.255.255.192	Réservé (62 adresses)

Une autre manière d'appréhender la subtilité du subnetting qui vient d'être opéré est de considérer la grille de découpage suivante.

<table>
<tr><th></th><th>2 subnets
de 128 (– 2) adresses</th><th>4 subnets
de 64 (– 2) adresses</th><th>8 subnets
de 32 (– 2) adresses</th><th>16 subnets
de 16 (– 2) adresses</th></tr>
<tr><td rowspan="16">Plage d'adresses
au sein du subnet</td><td rowspan="8">0 – 127</td><td rowspan="4">0 – 63</td><td rowspan="2">0 – 31</td><td>0 – 15</td></tr>
<tr><td>16 – 31</td></tr>
<tr><td rowspan="2">32 – 63</td><td>32 – 47</td></tr>
<tr><td>48 – 63</td></tr>
<tr><td rowspan="4">64 – 127</td><td rowspan="2">64 – 95</td><td>64 – 79</td></tr>
<tr><td>80 – 95</td></tr>
<tr><td rowspan="2">96 – 127</td><td>96 – 111</td></tr>
<tr><td>112 – 127</td></tr>
<tr><td rowspan="8">128 – 255</td><td rowspan="4">128 – 192</td><td rowspan="2">128 – 159</td><td>128 – 143</td></tr>
<tr><td>144 – 159</td></tr>
<tr><td rowspan="2">160 – 191</td><td>160 – 175</td></tr>
<tr><td>176 – 191</td></tr>
<tr><td rowspan="4">192 – 255</td><td rowspan="2">192 – 223</td><td>192 – 207</td></tr>
<tr><td>208 – 223</td></tr>
<tr><td rowspan="2">224 – 240</td><td>224 – 239</td></tr>
<tr><td>240 – 255</td></tr>
<tr><td>Masque</td><td>/25</td><td>/26</td><td>/27</td><td>/28</td></tr>
</table>

Les plages réservées permettront d'étendre les plages déjà affectées si le nombre de stations devient plus important que prévu. Ainsi, le réseau « Serveurs publics » pourra être étendu en diminuant le masque de 1 bit, afin de donner le subnet 10.0.8.0/26 (255.255.255.192).

Il est à noter que la création d'un sous-réseau fait perdre chaque fois deux adresses.

La technique du *subnetting* permet de gérer la pénurie d'adresses publiques sur l'Internet. En effet, la création de réseaux IP taillés sur mesure évite le gaspillage d'adresses ; par exemple, le réseau 10.0.0.0/16 offre 65 534 adresses qui seront loin d'être toutes utilisées. Sur votre réseau privé, vous avez cependant plus de latitude. Mais attention aux évolutions qui peuvent être rapides, par exemple lors de la fusion de deux sociétés.

Méthode d'affectation des réseaux WAN

L'interconnexion des réseaux (abordée aux chapitres suivants) nécessite également des adresses, mais en moins grand nombre que pour les réseaux LAN.

Par exemple, sur une liaison point à point, seules deux adresses sont nécessaires, une pour chaque extrémité. Le subnetting sur 30 bits, qui offre deux adresses, permet de créer un réseau juste dimensionné pour ce besoin.

Nous pourrions utiliser une des plages de notre réseau 10, mais il est cependant plus intéressant d'utiliser un autre réseau IP, et cela pour plusieurs raisons :

- Les adresses des réseaux WAN ne sont pas diffusées sur l'ensemble du réseau ; elles ne sont connues qu'entre routeurs adjacents.
- Les adresses n'ont donc pas besoin d'être connues des réseaux utilisateurs.
- Utiliser une plage d'adresses distincte permet de mieux identifier les liaisons WAN.

Bien que cela ne soit pas une obligation, nous préférons donc utiliser une autre plage d'adresses de la RFC 1918. Une classe B suffira amplement.

Nous pouvons donc réserver une plage de notre réseau 172.16.0.0/16, que nous « subnetterons » comme suit :

Subnets de 172.16.0.0/16	Fonction
172.16.0.0/30 255.255.255.252	Liaison Paris-Toulouse
172.16.0.4/30 255.255.255.252	Liaison Paris-Strasbourg
etc.	En tout : 16 384 subnets de 2 adresses

Pour les interconnexions multipoints, il suffira de réduire le masque d'autant de bits que nécessaires pour les subnets considérés. En général, les réseaux multipoints WAN sont rares et comprennent peu d'adresses en comparaison des LAN.

Méthode d'affectation des stations au sein des réseaux

Chaque nœud IP doit posséder une adresse IP. Cela concerne les PC et les Macintosh, les serveurs (NT, Unix, etc.), les imprimantes, les routeurs, les concentrateurs et commutateurs administrables (pour les agents SNMP), etc.

Il est tentant de découper la plage d'adresses en autant de parties qu'il y a de types de matériels. Cela n'apporterait cependant rien ni sur un plan technique, ni sur un plan organisationnel.

L'expérience montre, de plus, qu'une telle pratique n'est pas pérenne : soit la taille de la plage que l'on avait réservée est insuffisante (davantage de PC que prévu, par exemple), soit, à la longue, personne ne respecte une marche à suivre qui est trop contraignante (par exemple, s'il faut installer un PC en urgence, on prend la première adresse disponible).

Il est, en revanche, intéressant de prévoir un découpage simple entre les équipements terminaux (PC, serveurs, imprimantes, etc.) et les équipements réseau (les routeurs, les agents SNMP des concentrateurs et des commutateurs, etc.). Cela permet de mieux contrôler les flux du réseau. Dans le cas d'un subnet de classe B, on peut se risquer à créer une troisième plage réservée aux serveurs.

Plage d'adresses	Affectation
De 0.1 à 24.255	Équipements réseau (routeurs, hubs, switches, etc.). 6 399 adresses (de 1 à 6 399)
De 25.0 à 49.255	Serveurs NT, Unix, etc. 6 400 adresses (de 6 400 à 12 799)
De 50.0 à 255.254	Postes de travail (PC, etc.) 52 735 adresses (de 12 800 à 65 534)

Dans le cas d'un subnet sur une classe C, le plus simple est de ne pas affecter de plage d'adresses par type d'équipement, car la probabilité de collision est encore plus forte qu'avec une classe B. L'affectation des adresses pour les équipements réseau et serveur peut commencer par le bas de la plage et s'incrémenter ensuite, tandis que celle pour les PC peut commencer par le haut de la plage et se décrémenter ensuite.

Plage d'adresses	Affectation
De .1 à .254	Équipements réseau (routeurs, hubs, switches, etc.) et serveurs NT, Unix, etc.
De .254 à .1	Stations de travail (PC, etc.)

On peut constater que le plan d'adressage doit prendre en compte de nombreux paramètres liés à des notions qui n'ont pas été introduites : routage, translation d'adresse, affectation dynamique, connexion à l'Internet et contrôle de flux. Les chapitres suivants vous permettront de juger de la pertinence ou non du plan d'adressage qui vous est proposé.

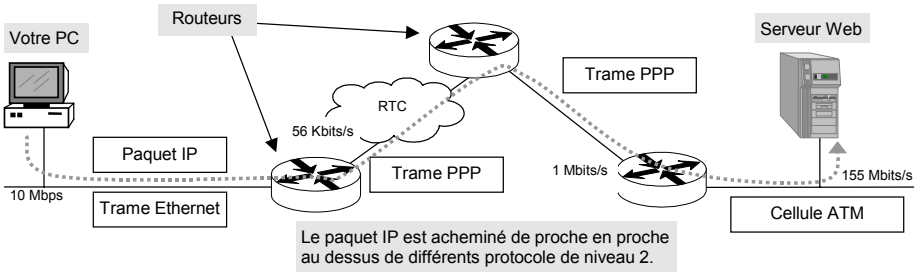
L'encapsulation des protocoles

Il existe différents supports de transmission (câbles en cuivre ou en fibre optique, faisceaux hertziens) et différents moyens d'accéder à ces supports (accès partagé par détection de collision, par jeton, par partage fixe de bande passante, etc.). Cela implique l'utilisation de nombreux protocoles de niveau 1 (couche physique) adaptés à chaque situation.

La couche liaison, telle que PPP, permet de masquer aux couches supérieures les particularités du niveau physique et ses contraintes. Mais il arrive qu'une norme spécifie les couches 1 et 2 : c'est le cas d'Ethernet et d'ATM (*Asynchronous Transfert Mode*).

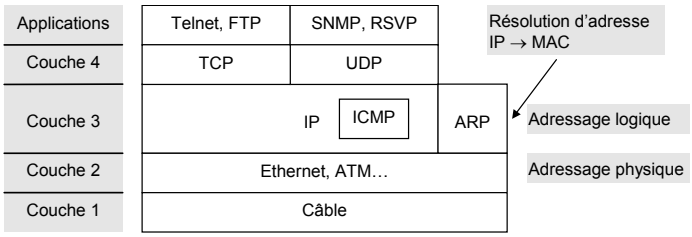
La couche de niveau 3 (couche réseau), telle que IP, peut donc utiliser différents réseaux en recourant aux services de PPP ou en s'adaptant directement sur une autre couche liaison.

Figure 5-5. *Paquet IP au-dessus de différents réseaux.*



On peut établir l'analogie suivante : le paquet IP est une voiture ; les pneus et les suspensions sont les protocoles de niveau 2 qui réalisent l'adaptation aux routes que sont les réseaux physiques. Vous roulez ainsi sur un chemin de terre (le RTC), puis sur une nationale (Ethernet) et enfin sur une autoroute (ATM), mais toujours avec la même voiture. Éventuellement, vous changez de pneus ou de suspensions, afin de vous adapter au terrain. De même, le paquet IP peut emprunter le RTC (avec une trame PPP), un réseau Ethernet (avec une trame Ethernet) ou un réseau ATM (avec une cellule ATM).

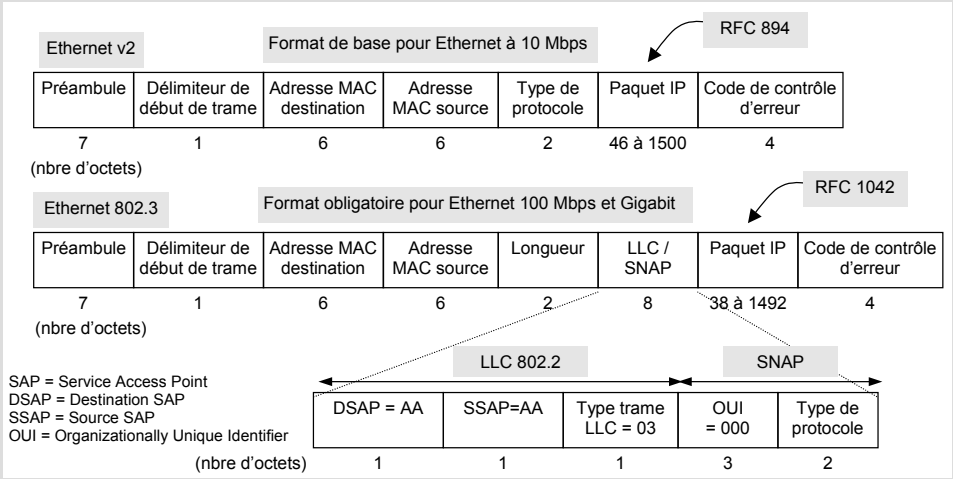
Figure 5-6. *Modèle en couches des protocoles Internet.*



L'ENCAPSULATION D'IP DANS ETHERNET (RFC 894 ET 1042)

Il existe deux formats de trames : **Ethernet v2**, également appelée DIX (du nom des constructeurs Digital, Intel et Xerox), et **Ethernet IEE 802.3**. Il y a donc deux façons d'envoyer un paquet IP sur Ethernet : directement dans une trame Ethernet v2 (RFC 894) ou dans une trame 802.3 *via* un en-tête LLC/SNAP (RFC 1042).

Si la valeur du champ " Type de protocole/Longueur " est supérieure à 1 500 (correspondant au nombre maximal d'octets pour le champ contenant le paquet IP), il s'agit d'une trame Ethernet v2 (le champ a alors la signification " Type de protocole "). Dans le cas contraire, il s'agit d'une trame Ethernet 802.3 (le champ a alors la signification " Longueur ").



Il existe différents types de **trames LLC** (IEEE 802.2) impliquant différents modes de fonctionnement. Pour IP, seule la trame de type *Unnumbered Information* (type 03) est utilisée (trame simple sans acquittement). Elle est également utilisée pour transporter IP dans ATM (*Classical IP*), Token-Ring et FDDI.

La **couche SNAP** (*Sub Network Access Protocol*) est nécessaire, car la trame LLC (*Logical Link Control*) ne contient pas de champ équivalent au champ " Type " de la trame Ethernet v2. Le SAP (*Service Access Point*) utilisé pour transporter SNAP est 170 (0xAA). On retrouve ce principe d'adaptation avec d'autres protocoles comme Frame-Relay ou ATM (voir chapitre 11).

Dans le dernier champ de l'en-tête SNAP, on retrouve enfin le **type de protocole** (*ethertype*) utilisé dont les valeurs sont identiques à celles du champ de même nom de la trame Ethernet v2 (**0x0800 pour IP**, 0806 pour ARP, etc.).

Il est à noter que le **MTU** (*Maximum Transfert Unit*), c'est-à-dire les données utiles transportées dans la trame Ethernet, est plus important avec Ethernet v2, l'encapsulation 802.3 faisant perdre 8 octets.

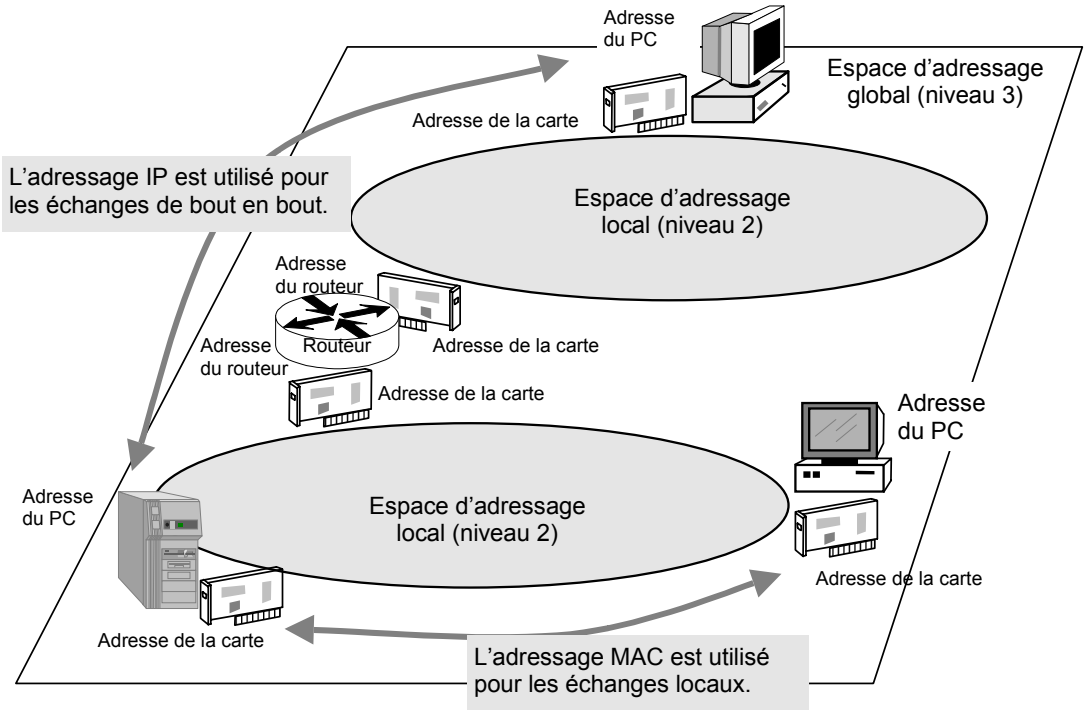
L'adressage

Toutes les couches réseau, de la couche physique à l'application en passant par les couches liaison, réseau et transport, utilisent des adresses afin d'identifier l'émetteur et le destinataire. Chaque couche utilise un système d'adressage spécifique qui répond à un besoin précis.

L'adressage de niveau 2 est géographiquement limité à un réseau local ou à une liaison point à point d'un réseau étendu.

L'adressage de la couche 3 permet d'identifier les stations à un niveau supérieur. Il assure la continuité entre des réseaux physiques qui utilisent différents systèmes d'adressage.

Figure 5-7.
Le rôle de l'adressage.

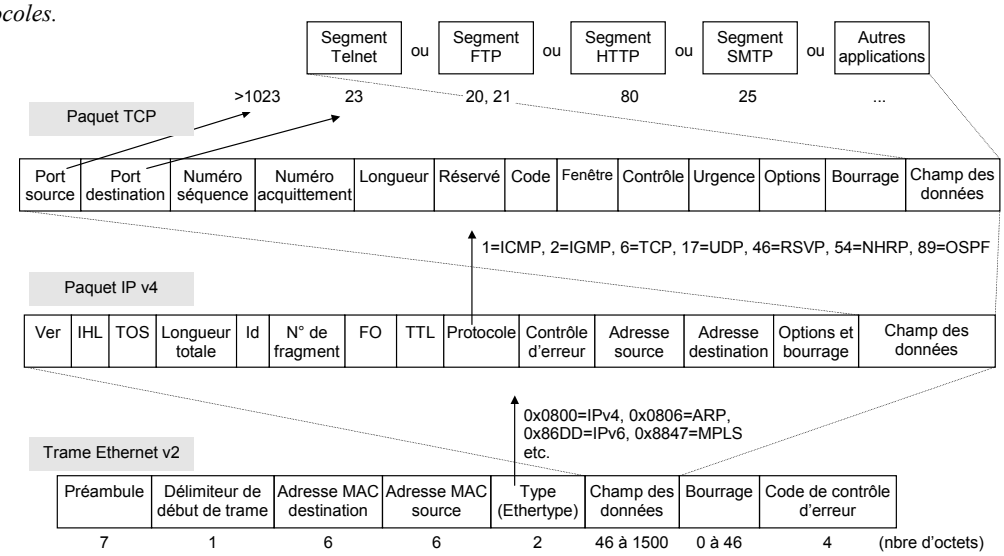


Enfin l'adressage de niveau 4 permet d'identifier les applications qu'utilisent les services de la couche transport.

Le multiplexage

Chaque couche réseau dispose d'un champ pour identifier le type de protocole encapsulé dans le champ de données. Ethernet identifie ainsi qu'il transporte un paquet IP, IP identifie qu'il transporte des données TCP, et TCP identifie l'application qui a rempli son champ de données.

Figure 5-8.
Le principe de l'encapsulation des protocoles.



Les champs « Type », « Protocole » et « Port » permettent à chaque couche de savoir à quelle couche supérieure remettre les données reçues. Le site www.iana.org/numbers.html recense ainsi toutes les valeurs affectées aux protocoles de la famille TCP/IP ou à ceux qui utilisent IP.

Il est ainsi possible d'envisager toutes les combinaisons d'encapsulation, telles que celle spécifiée par le protocole STUN (*Serial Tunneling*) qui permet de transporter dans un paquet IP une trame SDLC (*Synchronous Dala Link Control*) qui est un protocole de niveau 2 utilisé dans les réseaux SNA d'IBM. On pourra ainsi trouver l'encapsulation : « SDLC → TCP → IP → SNAP → LLC → Ethernet 802.3 ».

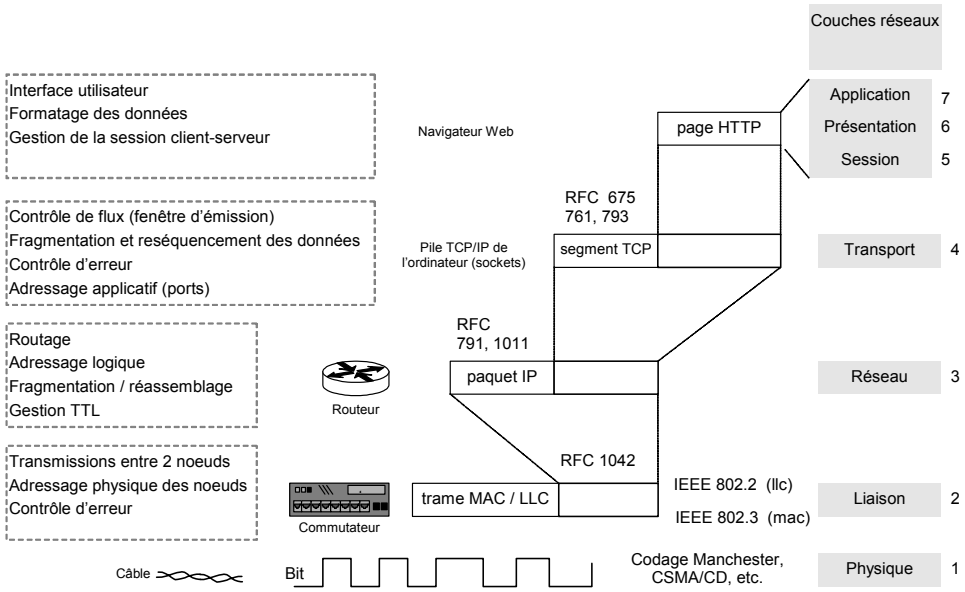
En théorie, tout protocole peut donc être encapsulé dans n'importe quel autre protocole. Dans la pratique, on utilise cette facilité pour répondre à une contrainte particulière, telle que le transport des flux SNA dans un réseau IP.

Dans le cas d'une navigation sur le web, l'empilement des protocoles est : « segment HTTP → TCP (port 80) → IP (protocole 6) », puis toutes sortes de réseaux de transport, tels que PPP, Frame-Relay, ATM, etc.

Sur votre réseau, l'encapsulation sera : « IP → SNAP (protocole 2048) → LLC (SAP 170) → Ethernet 802.3 ».

Par défaut, Windows 2000 encapsule les paquets IP dans les trames Ethernet v2. Il est possible d'activer l'encapsulation LLC/SNAP en ajoutant le paramètre « ArpUseEtherSNAP » de type REG_DWORD à la clé de registre « HKEY_local_machine\System\CurrentControlSet\Services\Tcpip\Parameters », et en fixant sa valeur à 1.

Figure 5-9. *Le rôle des couches réseau.*

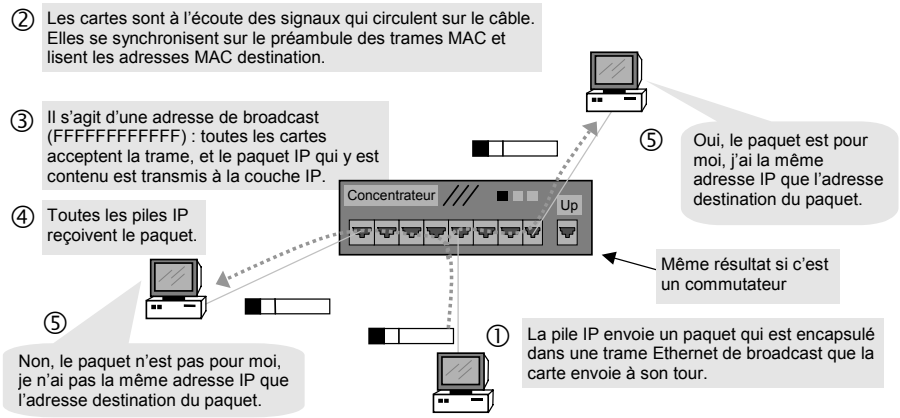


Ce modèle en couches simplifie la programmation des protocoles en leur assignant des rôles précis, et offre plus de souplesse par le jeu des encapsulations possibles.

L'échange de paquets IP

Une carte réseau ne se préoccupe que des adresses MAC pour envoyer et recevoir des données. En revanche, une application telle que Telnet ne connaît que l'adresse IP qui est purement logique : une pile IP recevant un paquet IP ne le prendra en compte que si l'adresse de destination du paquet correspond à l'adresse IP qui a été paramétrée dans le PC. Dans le cas contraire, il sera ignoré.

Figure 5-10. *Échange de paquets IP.*



Par ailleurs, l'adresse MAC de la station changera si la carte réseau est changée (en cas de panne, par exemple). De même, son adresse IP peut être modifiée à tout moment à l'aide des outils de configuration Windows (en cas de déménagement, par exemple).

L'exemple précédent montrait un paquet IP envoyé dans une trame de broadcast. Ce moyen d'opérer est pratique mais très consommateur de bande passante puisque la trame est propagée à travers tout le réseau. Sauf quand cela est nécessaire, un paquet IP est envoyé dans une trame unicast, c'est-à-dire directement au PC concerné. Mais comment connaître l'adresse MAC de la carte du PC destinataire alors que vous ne connaissez que l'adresse IP de sa pile IP ?

Cela est, par exemple, le cas lorsque vous lancez la commande suivante, qui permet de vous connecter à un serveur Unix :

```
Telnet 192.50.10.1
```

L'application Telnet va demander à la couche TCP d'ouvrir une connexion avec l'adresse IP indiquée, et va transmettre son paquet à la couche IP (avec l'adresse de destination indiquée). Cette dernière va encapsuler le segment TCP dans un paquet IP, puis l'envoyer à la carte. Mais la carte ne sait pas quoi faire d'une adresse IP ; elle ne sait gérer que des adresses MAC : une trame Ethernet ne contient qu'une adresse MAC qui permet aux autres cartes de la prendre ou non en compte.

La solution repose sur un mécanisme qui réalise la correspondance entre l'adresse MAC du PC destinataire et son adresse IP.

On pourrait utiliser une table de correspondance statique Adresse MAC ↔ Adresse IP. Mais cela serait fastidieux, car il faudrait relever les adresses MAC des stations ainsi que les adresse IP, et paramétrer la table sur tous les PC. Cela est inimaginable étant donné le nombre important de PC et les nombreux changements d'adresses qui interviennent. On perdrait en plus l'avantage de dissocier l'adresse physique de l'adresse logique. En outre, un PC peut être configuré avec plusieurs adresses IP.

LE POINT SUR IPV4 (RFC 791)

IP (*Internet Protocol*) est un protocole de niveau 3 (couche réseau) qui découpe les réseaux locaux en réseaux logiques indépendamment de leur implémentation physique. Ce protocole permet donc d'envoyer des données à travers **un réseau logique** reposant sur des réseaux physiques de différente nature (Ethernet et PPP, par exemple). Pour ce faire, IP utilise un **adressage logique** différent de l'adressage physique (MAC, PPP ou autre).

4 bits	4 bits	8 bits	16 bits
Version = 4	Longueur de l'entête	TOS (Type Of Service)	Longueur totale du paquet (en nombre d'octets, de 1 à 65 535)
Identification du fragment : permet de savoir à quel paquet IP appartient ce fragment			0 / D F / M / Position du fragment exprimée en multiples (de 0 à 8 191) de 8 octets
TTL (Time To Live) durée de vie du paquet		Protocole encapsulé dans le champ donnée	Contrôle d'erreur (checksum) sur l'entête du paquet IP uniquement
Adresse IP source (celle de la station qui a émis le paquet)			
Adresse IP destination (celle de la station à qui ce paquet doit être envoyé)			
Options (champ factultatif de longueur comprise entre 1 et 3 octets)			Bourrage

Cette couche se contente de **router** (c'est-à-dire acheminer) le paquet à travers un réseau IP : les paquets peuvent être perdus (pas de garantie d'acheminement), contenir des erreurs (sauf sur l'en-tête, qui est contrôlé) ou arriver dans le désordre. IP **fragmente** les paquets dont la taille excède celle des trames (le MTU, *Maximum Transfer Unit)*. Les fragments sont routés indépendamment les uns des autres comme autant de paquets, mais IP **assemble** dans le bon ordre les fragments d'un même paquet original. Le bit " M " positionné à " 0 " indique que ce paquet est le dernier fragment d'une série lorsque le bit " DF " est positionné à " 0 ".

Le bit " DF " positionné à " 1 " indique que l'émetteur ne souhaite pas fragmenter les paquets. Dans cas, si un routeur reçoit un paquet dépassant le MTU, il envoie à l'émetteur un message *icmp_destination unreachable* type 4 lui indiquant le MTU. Il est alors de la responsabilité de l'émetteur d'ajuster la taille de ces paquets ou d'autoriser la fragmentation.

Le champ **TTL** est décrémenté de 1 chaque fois que le paquet passe par un routeur. Si la valeur atteint zéro, le routeur détruit le paquet. Ce mécanisme évite aux paquets de rester trop longtemps sur le réseau, soit parce qu'ils tournent en boucle (suite à une erreur de routage), soit parce qu'ils traversent trop de routeurs. La valeur initiale du TTL est fixée par la station émettrice (de 32 à 128, en général).

D = délai d'acheminement court
T = débit élevé R = Grande fiabiltié

C = Option recopiée dans tous les fragments

Classe / Numéro
0 / 2 IP security Option
0 / 3 Routage lâche
0 / 7 Enregistrement des routes
0 / 9 Routage strict défini par la source
2 / 4 Horodatage des paquets

TOS	Priorité	D	T	R	0	0

Options	C	Classe	Numéro

Le champ **TOS** n'existe plus en tant que tel. Il a été redéfini par les RFC 2474 (DiffServ) et 3168 (ECN) qui sont détaillées au chapitre 14.

La résolution d'adresse

La seule solution est donc une résolution d'adresse automatique, c'est-à-dire un mécanisme permettant de trouver l'adresse MAC en connaissant uniquement l'adresse IP. C'est le rôle du protocole **ARP** (*Address Resolution Protocol*) lié à la couche IP. Ce protocole gère une table de correspondance dynamique Adresse MAC ↔ Adresse IP, appelée **cache ARP**. Vous pouvez en visualiser le contenu à l'aide de la commande Windows suivante :

```
arp -a
00:40:0b:4b:25:d2   190.50.1.253
```

Si vous n'avez pas communiqué récemment avec un autre PC, la table sera vide : les entrées sont, en effet, effacées au bout d'un certain temps. Sous Windows, une entrée ARP (adresse MAC / adresse IP) est supprimée au bout de deux minutes si le PC n'a pas dialogué avec la station cible (selon le mécanisme TTL, *Time To Live*). Dans tous les cas, l'entrée reste au maximum dix minutes en mémoire, puis elle est supprimée.

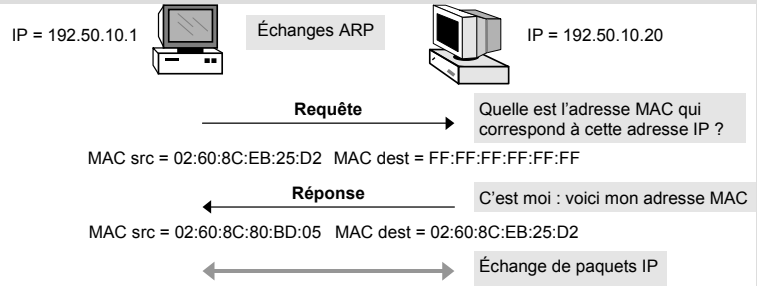
Si l'adresse IP recherchée n'est pas dans le cache, ARP va alors envoyer un paquet de requête encapsulé dans une trame Ethernet de broadcast. Cette dernière va donc être lue par toutes les cartes réseau.

LE POINT SUR ARP (RFC 826)

Pour obtenir l'adresse MAC d'une station ne connaissant que son adresse IP, la pile TCP/IP émet une requête ARP (Address Resolution Protocol) dans une trame Ethernet de broadcast dont le champ « Type » contient la valeur **0x0806**.

Chaque pile IP recevant un tel paquet compare alors son adresse avec celle figurant dans le champ « Adresse protocole destination ».

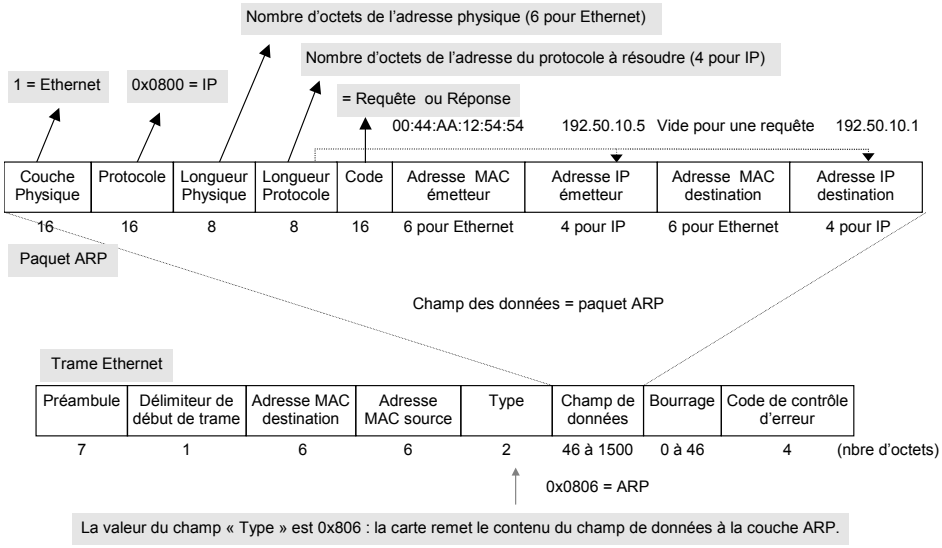
S'il y a correspondance, la couche ARP envoie un paquet de réponse en remplissant le champ « Adresse physique destination » avec l'adresse MAC de sa carte. Dans le cas contraire, le paquet est ignoré.



Donc, seule la station dont l'adresse IP correspond à celle demandée par la requête envoie en réponse un paquet contenant sa propre adresse MAC.

La résolution inverse, c'est-à-dire l'obtention de l'adresse IP à partir de l'adresse MAC, est réalisée par le protocole **RARP** (*Reverse ARP* – RFC 903).

Figure 5-11. *Format d'un paquet ARP dans une trame Ethernet.*



Seule la station configurée avec l'adresse IP recherchée va répondre en renvoyant son adresse MAC.

Une fois l'adresse résolue, le paquet IP peut être envoyé dans une trame MAC unicast dont l'adresse de destination est celle de la station cible.

L'échange de données entre applications

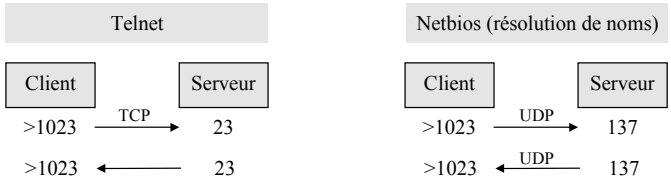
Une application utilise les services de la couche transport avec qui elle échange des données à travers une interface de programmation livrée avec la pile TCP/IP. Sous Unix, il s'agit des Sockets ; sous Windows de Winsock.

La couche transport est soit **TCP** (*Transport Control Protocol*) soit **UDP** (*User Datagram Protocol*), qui est une version allégée de TCP.

Le protocole TCP agit en **mode connecté**, ce qui implique que le client demande l'ouverture d'une connexion préalablement à tout échange. Par exemple, lorsque vous entrez la commande Windows « Telnet 192.50.10.1 », le programme client Telnet demande à TCP d'ouvrir une connexion à un serveur Telnet qui est en attente, c'est-à-dire à l'écoute du **port** TCP 23.

Inversement, UDP agit en mode **non connecté**, ce qui permet à deux machines d'échanger des données à tout moment, sans entrer dans une phase de connexion. Par exemple, lorsque vous voulez vous connecter à un serveur de fichiers Windows NT, votre PC émet une demande de résolution de nom à destination d'un serveur WINS qui est à l'écoute sur le **port** UDP 137.

Figure 5-12.
Utilisation des ports TCP et UDP.



La pile TCP/UDP du client choisit généralement un port source supérieur à 1023, et incrémente cette valeur à chaque nouvelle session ouverte simultanément à d'autres déjà actives.

LE POINT SUR UDP (RFC 768)

UDP (*User Datagram Protocol*) permet simplement à une application d'avoir accès au réseau IP. Ce protocole n'offre aucune garantie d'acheminement, aucun mécanisme de reprise sur erreur, ni de contrôle de flux, et ne vérifie pas la duplication des paquets. Tous ces contrôles doivent être opérés par les autres couches réseau. En revanche, les paquets remis à l'application le sont sans erreur.

16 bits	16 bits
Port UDP source	Port UDP destination
Nombre d'octets de l'entête et des données	Contrôle d'erreur : checksum portant sur l'entête UDP, une partie de l'entête IP et les données

Les champs **port source** et **port destination** servent à identifier une application (par exemple, 23 pour Telnet). Ces valeurs sont réservées et enregistrées par l'IANA (www.iana.org/assignments/port-numbers).

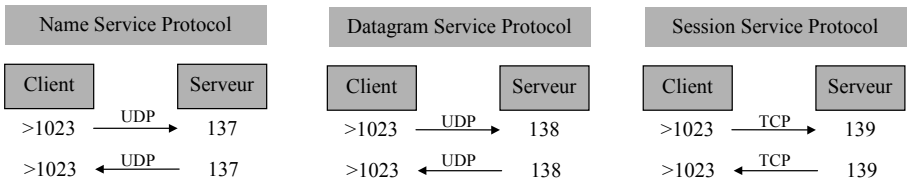
Les clients et les serveurs manipulent des noms (une machine Unix, un serveur de fichiers NT) et s'échangent des données à travers des ports qui leur sont réservés par l'IANA (voir en annexe « La gestion de l'Internet »). Sur l'Internet, le service DNS permet de convertir les noms en adresses IP (voir chapitre 7), tandis que, dans le monde Microsoft, on utilise encore le service WINS pour convertir des noms **Netbios** en adresses IP.

On peut considérer que ce protocole est situé au niveau de la couche 5 (couche session) : il permet, en effet, d'établir et de gérer des sessions entre applications. Dans le monde Internet, les applications comme Telnet, votre navigateur Web, FTP, etc.) gèrent elles-mêmes tous les mécanismes situés au-dessus de la couche transport, c'est-à-dire TCP et UDP.

À l'origine, Netbios circulait nativement dans des trames Ethernet, mais de nos jours, il est encapsulé dans IP (RFC 1001 et 1002).

Par exemple, le partage de fichiers et la messagerie Exchange utilisent le protocole Netbios sur le port TCP 139. Par ailleurs, les serveurs WINS s'échangent des données sur le port TCP 42.

Figure 5-13.
Netbios sur IP.



LE POINT SUR TCP (RFC 793)

TCP (*Transport Control Protocol*) est un protocole de niveau 4 (couche transport) qui permet à deux applications (un client et un serveur) d'échanger des données en leur masquant les mécanismes réseau. Les segments TCP sont transportés dans des paquets IP de type 6 (champ protocole = 6).

TCP offre un service de **bout en bout** (entre deux entités, quelle que soit leur localisation) en **mode connecté** (un client doit se connecter à une application serveur). Il utilise pour cela un adressage applicatif basé sur des **ports TCP**. Chaque application est identifiée par un numéro de port réservé (*well known port*). Généralement, le client choisit un numéro de port aléatoire supérieur à 1023 comme port source. Le serveur lui répond sur ce port.

4 bits	4 bits	8 bits						16 bits
Port TCP source								Port TCP destination
Numéro de séquence des segments émis (sens émetteur→récepteur) = N° du premier octet du champ « Donnée »								
Numéro d'acquittement des segments reçus (sens émetteur→récepteur) = ok j'ai bien reçu les N-1 octets, j'attends le Nème								
Longueur de l'entête	Réservé (6 bits)	URG	ACK	PSH	RST	SYN	FIN	Fenêtre d'émission permettant d'acquitter plusieurs segments en même temps
Contrôle d'erreur (checksum) portant sur l'entête TCP, une partie de l'entête IP et les données								Pointeur d'urgence = nombre d'octets urgents restant à envoyer
Options								Bourrage

Les bits de contrôle ont la signification suivante :

URG (*Urgent*)	Indique que les données doivent être remises sans délai à l'application.
ACK (*Acknowledge*)	Acquittement d'un ou de plusieurs segment selon le n° d'acquittement
PSH (*Push*)	Indique à la couche TCP d'envoyer et de remettre les données sans attendre le remplissage des tampons d'émission et de réception.
RST (*Reset*)	Ferme la connexion TCP suite à un problème.
SYN (*Synchhronize*)	Le récepteur doit synchroniser son n° d'acquittement avec le n° de séquence reçu
FIN	Demande de déconnexion.

Des options peuvent être négociées entre entités TCP (champ " Option "), par exemple la taille maximale des segments transportés.

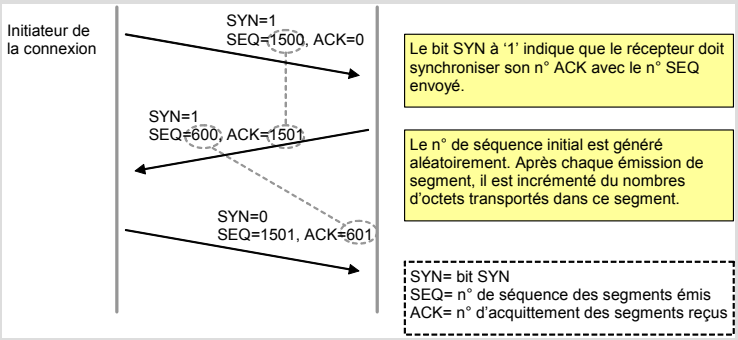
La couche TCP assure le **contrôle d'erreur** et le **séquencement** des segments (les segments sont remis dans le même ordre que lors de leur émission). La taille de la **fenêtre d'émission** indique au récepteur le nombre d'octets qu'il pourra à son tour envoyer sans attendre d'acquittement. Elle permet également de demander la retransmission à partir du premier segment en erreur (manquant ou erroné).

La couche TCP mesure le temps écoulé entre l'émission d'un segment et la réception de l'accusé de réception correspondant, et calcule ainsi une moyenne glissante du temps de réponse (*Round Trip Time*). Elle utilise l'algorithme de **Karn** pour déduire la valeur de ses temporisateurs. Ainsi, plus le temps de réponse est long, plus TCP attendra longtemps l'accusé de réception avant de retransmettre. De même, TCP estime le nombre de segments perdus : plus celui-ci augmente, plus la fenêtre d'émission est réduite. Ces mécanismes permettent à TCP de **contrôler le flux** de données en fonction de l'état du réseau (perte de segments et débits) et donc d'éviter une surcharge du réseau par un nombre croissant de retransmissions devenues inutiles.

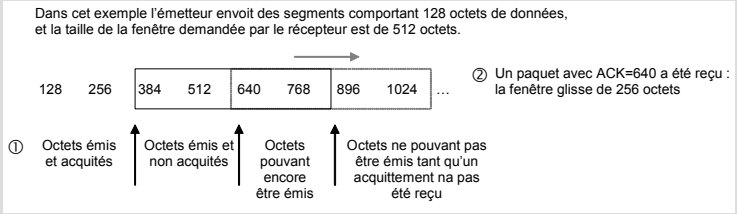
•••

LE POINT SUR TCP (SUITE)

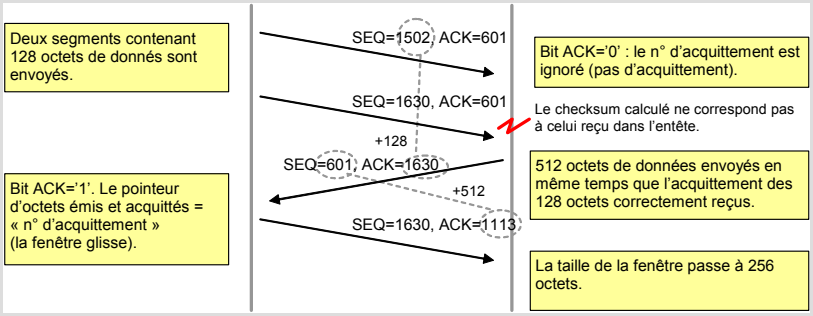
Une connexion TCP s'effectue par l'échange des trois segments suivants (*three ways handshake*) :



Pour gérer la fenêtre glissante, chaque extrémité gère deux fenêtres par connexion, une pour l'émission (cf. schéma ci-dessous) et une autre pour la réception.



L'acquittement TCP est cumulatif : un acquittement de numéro de séquence N signifie que les octets précédents ont bien été reçus.



Nombre d'applications se contentent d'envoyer peu d'octets, voire un seul octet par segment TCP, comme cela est le cas de Telnet (cf. chapitre 10). Sur un réseau non encombré et/ou à haut débit, on peut se permettre d'encapsuler un octet dans un segment TCP puis dans un paquet IP (*overhead* de 40 octets), de procéder à la même opération pour le caractère en écho, puis de générer deux acquittements, un pour le caractère envoyé et un autre pour le caractère en écho. Mais dans des conditions inverses, cet overhead n'est pas supportable. Ainsi, lorsque TCP constate que les temps de réponse sont longs, il active l'algorithme de **Nagle**, qui consiste à accumuler les octets pour les envoyer dans un même segment, plutôt que d'envoyer chaque caractère dans des segments individuels.

6

Administrer son réseau IP

Votre réseau d'entreprise est désormais opérationnel. Il s'étend sur plusieurs sites. Cependant, les problèmes vous guettent. En effet, plus le réseau est important, plus la probabilité qu'une panne survienne à un endroit ou à un autre est importante. C'est statistique.

En outre, plus un réseau est important, plus il est difficile à gérer. Il convient donc d'utiliser des outils qui simplifient sa gestion et diminuent donc le nombre potentiel de pannes.

Dans ce chapitre vous apprendrez :

- à utiliser les outils de base pour le diagnostic réseau ;
- à installer un serveur DHCP qui vous facilitera la vie.

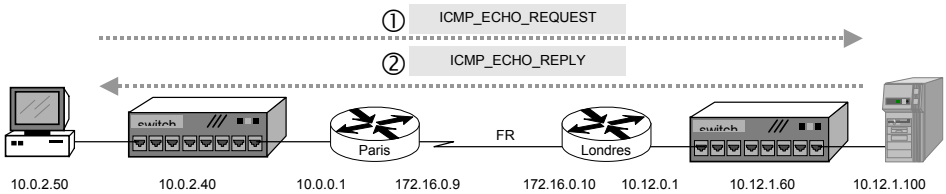
Les utilitaires de base

Deux programmes doivent faire partie du « kit de survie » de tout administrateur réseau : ping et Traceroute.

Le ping

La commande de base est le **ping** (*Packet Internet Groper*, ou *ping-pong*). Ce programme permet de savoir si une station IP est active et, plus généralement, si le réseau fonctionne correctement entre deux points, par exemple, entre votre PC à Paris et le serveur de Londres. Autre fonction intéressante, le programme donne également le temps de réponse mesuré.

Figure 6-1.
Le ping.



La commande ping s'appuie sur le protocole ICMP (*Internet Control Message Protocol*) qui fait partie de la couche IP. Ce protocole est utilisé pour toutes les opérations qui ont trait à la gestion du réseau IP, et ce, de façon transparente pour les utilisateurs.

Invite de commandes

```
C:\>ping 10.12.1.100

Pinging [10.12.1.100] avec 32 octets de données :

Réponse de 10.12.1.100 : octets=32 temps=80ms TTL=128
Réponse de 10.12.1.100 : octets=32 temps=90ms TTL=128
Réponse de 10.12.1.100 : octets=32 temps=80ms TTL=128
Réponse de 10.12.1.100 : octets=32 temps=90ms TTL=128

C:\>_
```

Le temps de réponse varie entre 80 ms et 90 ms. C'est le temps mis par le paquet ICMP_ECHO_REQUEST pour aller jusqu'au serveur + le temps mis par le paquet ICMP_EGHO_REPLY pour revenir jusqu'au PC.

Sous Windows NT, le TTL par défaut est de 128. Vous pouvez le modifier dans la base des registres Windows : HKEY_LOCAL_MACHINE\System\CurrentControlSet\Services\Tcpip\Parameters\DefaultTTL.

Le traceroute

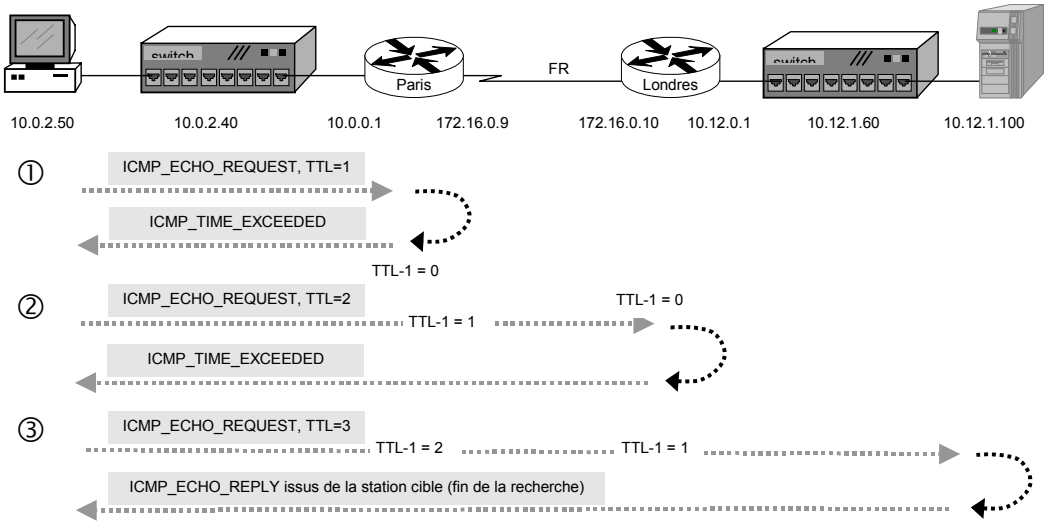
L'autre commande de base est le **traceroute**. Ce programme utilise des mécanismes propres à IP et ICMP pour afficher à l'écran la route empruntée par un paquet IP en plus du temps de réponse. Son principe de fonctionnement est le suivant :

- Le programme envoie un paquet ICMP_echo_request à destination de la machine cible avec un TTL (*Time To Live*) égal à 1.
- Le premier routeur reçoit ce paquet, décrémente le TTL de 1, et constate qu'il est égal à 0. Il détruit le paquet, et renvoie à l'émetteur un message ICMP_Time_exceeded.
- Le programme enregistre l'adresse IP du routeur qui a envoyé ce message ainsi que le temps écoulé depuis l'émission du paquet ICMP_Echo_request.
- Le programme continue de même en incrémentant le TTL de 1 à chaque paquet ICMP_Echo_request émis. Le paquet ira donc un saut plus loin que le précédent, et le routeur suivant répondra.

Le mécanisme du TTL (*Time To Live*) est expliqué dans l'encart « Le point sur IP v4 », au chapitre 5.

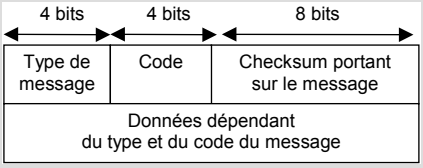
Certaines implémentations de traceroute utilisent un paquet UDP sur un port quelconque à la place d'un paquet ICMP_Echo_request. La RFC 1393 (statut expérimental) propose, quant à elle, un autre algorithme qui repose sur un message ICMP_Traceroute (type 30). Il est cependant rarement implémenté, aussi bien sur les stations que sur les routeurs.

Figure 6-2.
Le traceroute.



LE POINT SUR ICMP (RFC 792, 950 ET 1256)

ICMP (*Internet Control Message Protocol*) regroupe près de trente types de messages permettant aux nœuds d'un réseau d'échanger des informations de gestion relatives à IP.



Excepté pour les messages Echo et Timestamp, le paquet ICMP contient une copie partielle du paquet original ayant causé l'erreur (en-tête IP + 8 premiers octets des données).

Le tableau de la page suivante recense les types et codes existants à ce jour. Lorsque cela n'est pas précisé, les messages sont générés par les routeurs et les stations IP. Sauf indication contraire, les messages ICMP sont définis dans la RFC 792.

Type	Message	Code et description	
3	Destination Unreachable	0	*network unreachable* : le routeur ne connaît pas la route
		1	*host unreachable* : le routeur ne peut pas trouver la station
		2	*protocol unreachable* : le protocole demandé n'est pas actif.
		3	*port unreachable* : aucun programme ne répond sur ce port TCP ou UDP.
		4	*fragmentation needed and DF set* : indique à l'émetteur que le routeur doit fragmenter un paquet qui dépasse le MTU, alors que la fragmentation n'est pas souhaitée par ce même émetteur.
		5	*source route failed*
11	Time exceeded	0	Si un routeur reçoit un paquet avec un TTL à 0, il envoie ce message à l'émetteur.
		1	Si une station n'obtient pas tous les fragments d'un message dans le temps imparti, elle envoie ce message à l'émetteur.
12	Parameter Problem	0	Des paramètres incorrects ou inconsistants dans l'en-tête du paquet IP ont été détectés (un pointeur indique la position de l'erreur dans l'entête)
4	Source Quench	0	Le routeur, ou la station, est congestionné (ou régule le trafic selon un algorithme propre) et demande à l'émetteur de réduire son flux. Les paquets en excès peuvent être détruits.
5	Redirect	0	*for the Network* : le routeur a détecté une meilleure route et indique à la station quel routeur solliciter
		1	*for the Host* : idem pour une station
		2	*for the TOS and Network* : idem avec le champ TOS correspondant.
		3	*for the TOS and Host* : idem pour une station.
8	Echo request	0	Demande au récepteur de renvoyer un Echo reply (un identifiant et un numéro de séquence identifient le message).
0	Echo reply	0	Réponse à un Echo request. Les données contenues dans le message Echo request doivent être reportées dans ce message.

• • •

LE POINT SUR ICMP (SUITE)

Type	Message	Code et description	
13	Timestamp	0	Indique le nombre de millisecondes écoulé depuis 00h00 GMT lorsque le message Timestamp a été envoyé. Utilisé pour évaluer le temps de transit.
14	Timestamp reply	0	Indique le nombre de millisecondes écoulé depuis 00h00 GMT lorsque le message Timestamp a été reçu, ainsi que la valeur de ce nombre lorsque la réponse a été envoyée.
15	Information request	0	(obsolète) Permettait aux stations d'obtenir leur adresse IP. Mécanisme remplacé par les protocoles RARP, puis BOOTP et DHCP.
16	Information reply	0	Idem (obsolète).
17	Mask request	0	(RFC 950) Permet à une station d'obtenir le masque IP de son sous-réseau.
18	Mask reply	0	(RFC 950) Réponse du routeur à une demande de masque.
9	Router Advertisement	0	(RFC 1256) Émis périodiquement par un routeur pour indiquer l'adresse IP de son interface. Permet aux routeurs de découvrir leurs voisins, et aux stations de découvrir leur passerelle par défaut.
10	Router Solicitation	0	(RFC 1256) Lors de son initialisation, une station peut demander à un routeur de s'annoncer immédiatement. Les routeurs n'envoient, en principe, aucune sollicitation, mais attendent les annonces.

Invite de commandes

```
C:\>tracert 10.12.1.100

Trace l'itinéraire vers 10.12.1.100
avec un maximum de 30 tronçons :

  1    10 ms    10 ms    10 ms   10.0.0.1
  2    90 ms    80 ms    80 ms   172.16.0.10
  3    90 ms    90 ms    80 ms   10.12.1.100

Routage terminé.

C:\>
```

La version Windows de traceroute procède à trois essais à chaque saut et affiche trois temps de réponse.

Les adresses affichées sont celles des interfaces des routeurs qui ont émis le message ICMP_TIME_EXCEEDED et, en dernier, celle de la station cible qui, elle, renvoie un ICMP_ECHO_REPLY.

La commande ping sera plutôt utilisée pour savoir si un nœud IP est actif et joignable ainsi que le temps de réponse de bout en bout. Elle offre en outre davantage de possibilités de paramétrage (taille et nombre des paquets, enregistrement de la route, etc.). La commande traceroute permet, quant à elle, de savoir quelle route est empruntée (par exemple, le chemin principal ou celui de backup) et quelles parties du réseau engendrent les temps de réponse les plus longs.

De nombreux utilitaires de ce type sont disponibles sur les sites ftp.lip6.fr, tucows.club-internet.fr et www.winsite.com.

Observer ce qui se passe sur son réseau

Si votre réseau présente un dysfonctionnement et que, malgré toutes vos investigations, vous n'avez pas trouvé d'où provient le problème, il ne vous reste plus qu'à l'ausculter, c'est-à-dire observer les données qui y circulent.

Même lorsque le réseau semble bien fonctionner, il n'est pas inutile d'y jeter un coup d'œil, car bien souvent des erreurs (collision, paquets corrompus, flux non identifié, trafic censé ne pas être présent sur ce segment, etc.) se produisent. Ces erreurs ne sont alors pas perceptibles, mais peuvent le devenir sous certaines conditions, par exemple lorsque la charge réseau augmente. Une **maintenance préventive** permet donc d'éviter le pire.

L'**analyseur réseau** est l'outil tout indiqué pour ce type de situation. Il permet :

- de capturer toutes les trames qui circulent sur un segment Ethernet ;
- d'analyser le contenu de toutes les couches réseau, de la trame aux données applicatives en passant par le paquet IP ;
- de déterminer si des erreurs se produisent (collision, erreur de transmission, etc.) et en quelle proportion ;
- de connaître les temps de réponse précis (au millième de seconde près).

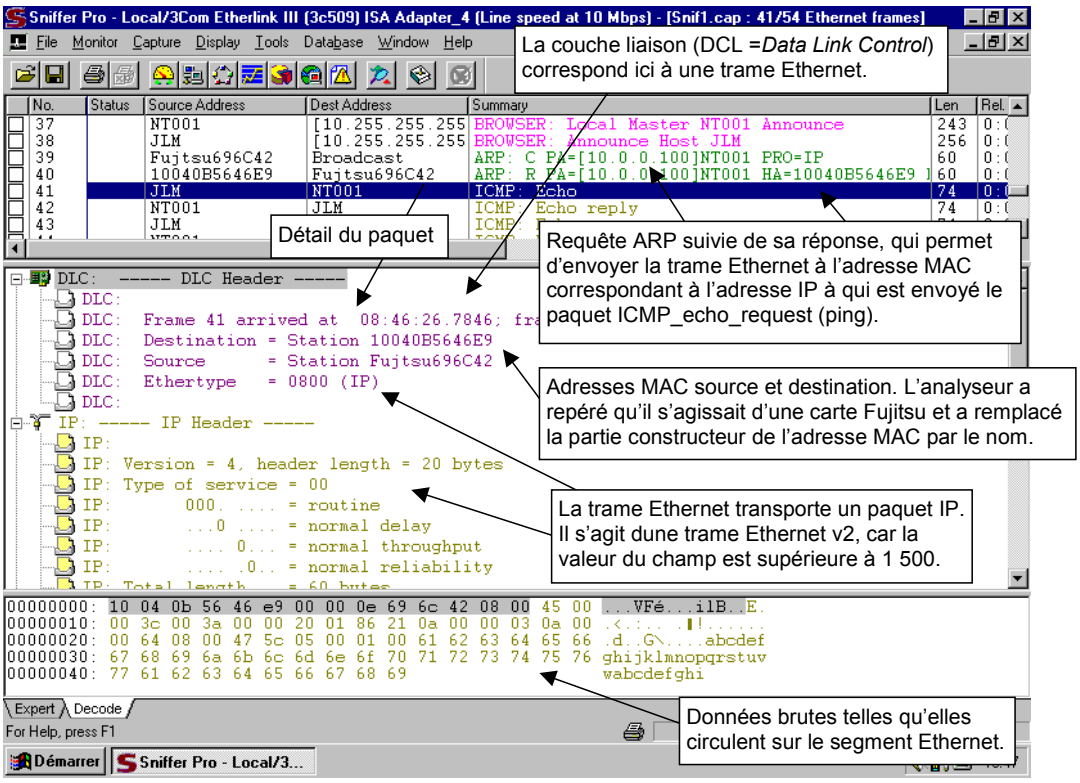
En positionnant des filtres, il est possible de suivre précisément les échanges entre deux stations, soit à partir de leurs adresses MAC, soit à partir de leurs adresses IP.

COMMENT UN ANALYSEUR RÉSEAU FONCTIONNE-T-IL ?

Un analyseur réseau est un logiciel capable de décoder idéalement tous les protocoles existants, du niveau 2 au niveau session. Il fonctionne de concert avec une carte réseau, de préférence haut de gamme, capable de capturer toutes les trames, même à pleine charge.

Le coût d'un tel produit dépend donc du nombre de protocoles reconnus et de la carte d'acquisition : Ethernet, Token-Ring ou ATM pour les LAN, et série synchrone pour les liaisons WAN en Frame Relay, ATM, etc.

La carte réseau doit fonctionner en mode **promiscus**. Dans un mode de fonctionnement normal, une carte ne prend en compte que les trames multicast et de broadcast, ainsi que celles dont l'adresse de destination MAC correspond à celle qui est programmée dans sa mémoire (PROM, Flash, etc.). En mode promiscus, la carte prend en compte toutes les trames. Toutes les cartes réseau ne supportent pas le mode promiscus.



L'analyseur réseau peut également être utilisé comme outil d'analyse pour :

- mesurer la volumétrie générée par une application entre un client et un serveur ;
- évaluer la part de responsabilité du réseau, du serveur et du client dans le temps de réponse global entre un client et un serveur ;
- surveiller la charge du réseau pendant 24 heures ou sur une semaine ;
- déterminer quelles stations génèrent le plus de trafic ;
- déterminer la répartition du trafic par protocole, par adresse IP, etc.

Enfin, l'analyseur réseau offre souvent des fonctions évoluées, telles que :

- une minuterie (*triger*) qui permet de déclencher et d'arrêter la capture sur réception d'une trame particulière (adresse, données, etc.) ;
- un générateur de trafic pour vérifier le comportement du réseau et des applications à pleine charge (par exemple si trop d'erreurs surviennent à partir d'une certaine charge, le câblage est sans doute en cause) ;
- la possibilité de rejouer un échange de trames préalablement capturées.

Piloter son réseau

Si votre réseau prend de l'ampleur — le nombre des équipements (routeurs, concentrateurs, commutateurs, etc.) augmente, et ces derniers sont répartis sur différents sites —, il devient de plus en plus nécessaire de centraliser la gestion des équipements.

Pour ce faire, la famille des protocoles TCP/IP propose le protocole SNMP (*Simple Network Management Protocol*) qui permet l'échange d'informations entre **une station d'administration** (le client) et des **agents** (les serveurs) implantés dans chaque équipement. On parle alors d'**agent SNMP** ; celui-ci se présente sous forme d'un petit programme qui répond aux requêtes SNMP émises par la station d'administration.

Quelle station d'administration ?

Une station d'administration, ou **plate-forme d'administration**, est constituée d'un ordinateur sous Windows NT/2000 ou sous Unix, ainsi que d'un logiciel, tel qu'OpenView de Hewlett-Packard, Tivoli d'IBM, Unicenter de Computer Associates, Spectrum de Cabletron, etc.

Ces logiciels haut de gamme (environ 15 000 €) sont en fait des boîtes à outils sur lesquelles s'installent des **modules dédiés** à chaque constructeur (CiscoView pour les équipements Cisco, Optivity pour ceux de Bay Networks, etc.). Ces modules peuvent, par ailleurs, fonctionner de manière autonome.

L'intérêt d'une telle plate-forme est de fédérer la gestion d'un parc d'équipements hétérogène autour d'une gestion centralisée des alarmes et d'une carte réseau sur laquelle s'affichent les équipements découverts dynamiquement.

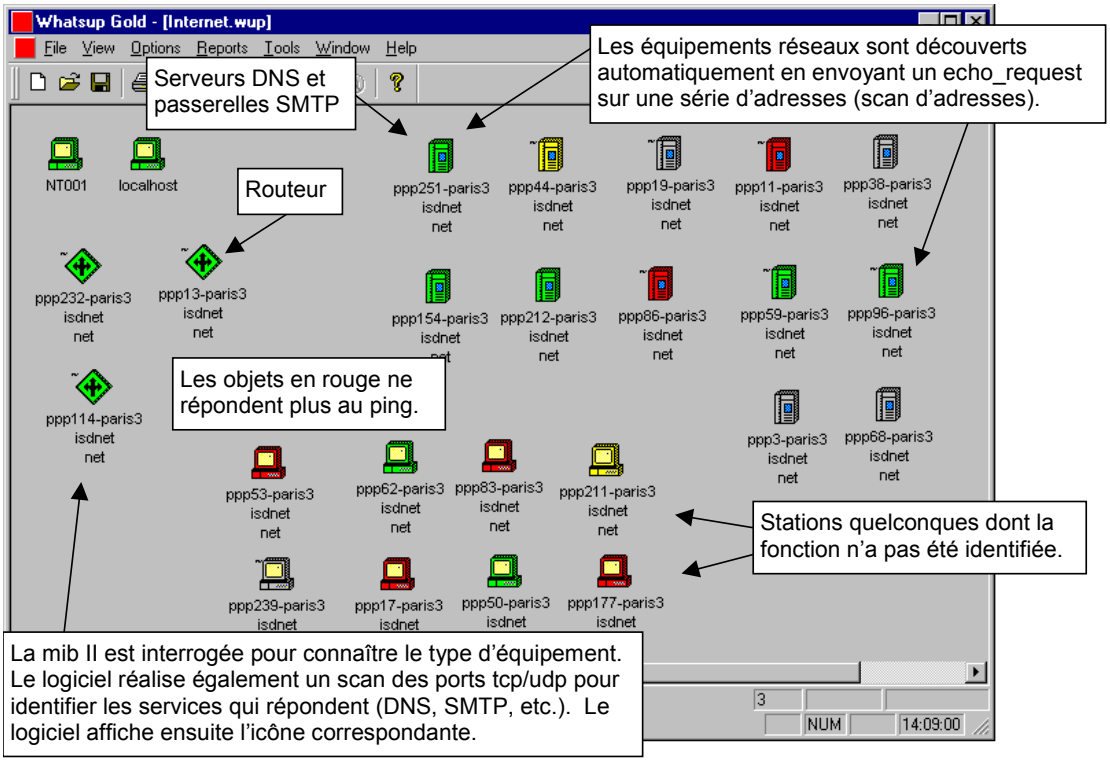
En fait, les plates-formes d'administration nécessitent un paramétrage très important dont le coût peut être plus élevé que celui du matériel et du logiciel réunis. Elles sont pour cela réservées aux grands réseaux, notamment chez les opérateurs.

Pour un réseau de plus petite taille, il est préférable d'utiliser les modules des constructeurs en mode autonome : si vous disposez d'un parc d'équipements homogène, nul besoin d'investir dans une « usine à gaz ». L'intérêt est de pouvoir visualiser graphiquement les équipements et de cliquer sur les cartes et ports que vous voulez configurer.

Certains administrateurs de grands réseaux se dispensent même de ce type de logiciel, préférant utiliser Telnet, TFTP et les fichiers de configuration en mode texte, le ping et le traceroute étant utilisés pour les dépannages quotidiens.

Pour quelle utilisation ?

Si vous désirez néanmoins visualiser et surveiller votre réseau de manière graphique, vous pouvez toujours utiliser des petits logiciels, tels que *Whatsup*. Le principe est identique à celui des plates-formes, mais avec un peu moins de fonctionnalités.



Cette carte a été obtenue en scannant les adresses IP d'un fournisseur d'accès à l'Internet (ISP). Le logiciel a ainsi trouvé des routeurs, des serveurs DNS, des passerelles de messagerie SMTP, ainsi qu'un certain nombre de stations non identifiées.

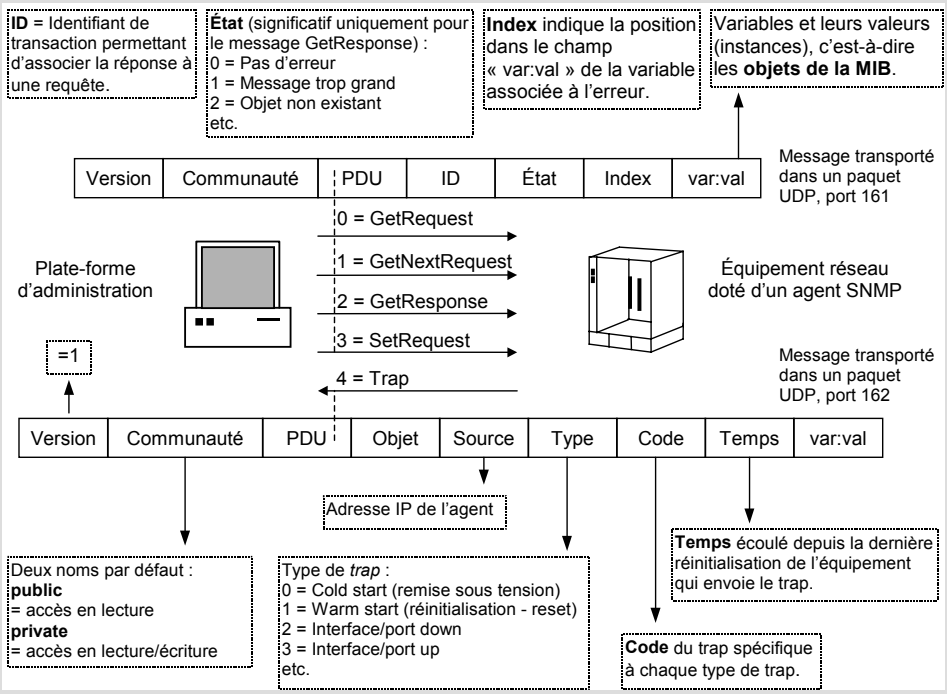
La première tâche est d'agencer les icônes qui apparaissent dans le désordre. L'administrateur peut ensuite dessiner un réseau et positionner les objets dessus, de manière à faire correspondre la carte à la réalité.

En sélectionnant une icône, il est alors possible d'opérer plusieurs actions sur l'équipement : ajouter des informations complémentaires, interroger son agent SNMP, surveiller des paramètres de cet agent, positionner des seuils d'alerte, etc.

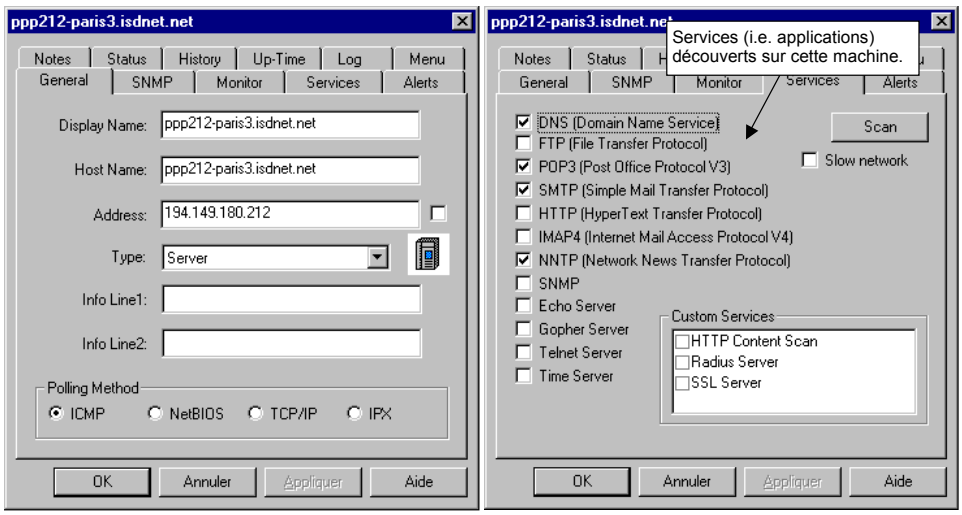
LE POINT SUR SNMP V1 (RFC 1157, 2571, 2572)

Le protocole SNMP (*Simple Network Management Protocol*) est utilisé pour piloter tous les équipements du réseau (routeurs, commutateurs, concentrateurs, serveurs, etc.) à partir d'une **station d'administration**.

Il est ainsi possible de **configurer** à distance les équipements (activation d'une interface, ajout d'une adresse IP, etc.) et de récupérer les paramètres actifs. Inversement, un équipement peut envoyer une alarme à la station d'administration *via* un **trap SNMP**.



Presque tous les équipements réseau intègrent un **agent SNMP**. Ce logiciel réalise l'interface entre les **requêtes SNMP** et la base de données **MIB** (*Management Information Base*) qui regroupe tous les paramètres de l'équipement.



Le problème de ce type de logiciel est qu'il faut sans cesse le mettre à jour, car le réseau ne cesse d'évoluer. L'autre problème tient à la gestion des alertes : le logiciel doit être très précisément paramétré pour ne générer que des alarmes réelles. Ces deux activités peuvent prendre beaucoup de temps à l'administrateur.

```
RFC1213-MIB DEFINITIONS ::= BEGIN
IMPORTS
      mgmt, NetworkAddress, IpAddress, Counter, Gauge, TimeTicks
FROM RFC1155-SMI
OBJECT-TYPE
FROM RFC-1212;
```

Importe ces groupes de la MIB SMI et tous les objets de la MIB-I.

Cette MIB définit le groupe MIB-II situé sous le groupe Management.

```
mib-2  OBJECT IDENTIFIER ::= { mgmt 1 }
       DisplayString : := OCTET STRING
       PhysAddress ::= OCTET STRING
-- groups in MIB-II
       system OBJECT IDENTIFIER ::= { mib-2 1 }
       interfaces    OBJECT IDENTIFIER ::= { mib-2 2 }
       at     OBJECT IDENTIFIER ::= { mib-2 3 }
       ip     OBJECT IDENTIFIER ::= { mib-2 4 }
       icmp   OBJECT IDENTIFIER ::= { mib-2 5 }
       tcp    OBJECT IDENTIFIER ::= { mib-2 6 }
       udp    OBJECT IDENTIFIER ::= { mib-2 7 }
       egp    OBJECT IDENTIFIER ::= { mib-2 8 }
```

Définition de l'objet **system**, identifiant n° 1 dans le groupe **mib-2**

Noms et identifiants dans la MIB 2.

```
        -- cmot       OBJECT IDENTIFIER ::= { mib-2 9 }
        transmission OBJECT IDENTIFIER ::= { mib-2 10}
        snmp   OBJECT IDENTIFIER ::= { mib-2 11 }

-- the System group
sysDescr OBJECT-TYPE
        SYNTAX  DisplayString (SIZE (0..255))
        ACCESS  read-only
        STATUS  mandatory
        DESCRIPTION
        "Par ex. : Cisco 761 Software Version c760-i…"
::= { system 1 }

sysObjectID OBJECT-TYPE
        SYNTAX  OBJECT IDENTIFIER
        ACCESS  read-only
        STATUS  mandatory
        DESCRIPTION
        "Par Ex. : Cisco2503"
::= { system 2 }
.......................................

ifNumber OBJECT-TYPE
        SYNTAX  INTEGER
        ACCESS  read-only
        STATUS  mandatory
        DESCRIPTION "Nombre d'interfaces, par exemple, 3"
::= { interfaces 1 }
.......................................
END
```

Description de l'objet **sysDescr**, identifiant n° 1 dans le groupe **system**.

Description de l'objet **sysObjectID**, identifiant n° 2 dans le groupe **system**.

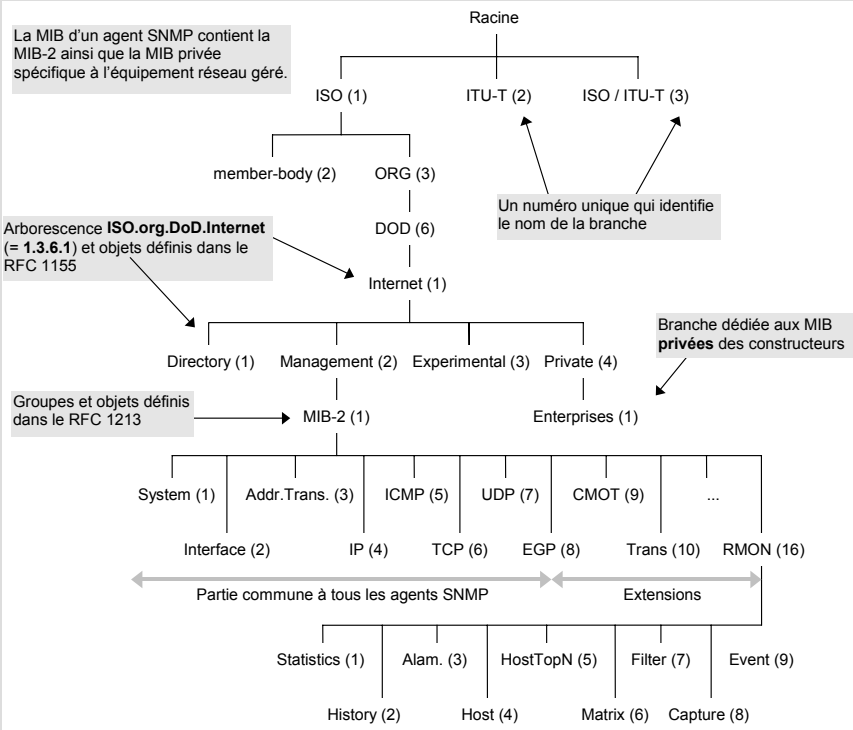
Description de l'objet **ifNumber**, identifiant n° 1 dans le groupe **interfaces**.

Pour que la station d'administration puisse interroger la MIB, il faut en premier lieu la **compiler** à partir du fichier en syntaxe ASN.1. Le fichier texte est alors intégré sous une autre forme (binaire généralement) dans le gestionnaire de MIB du logiciel d'administration.

Le moyen le plus simple de visualiser la MIB d'un équipement réseau est d'utiliser le module dédié, propre à chaque constructeur. La manipulation des variables est alors transparente, puisque le module affiche graphiquement l'équipement, par exemple, un commutateur. Il suffit alors de cliquer sur un port et de sélectionner les options qui vous sont proposées : activation/désactivation, vitesse (10, 100 ou *autosense*), nombre d'octets émis et reçus, etc.). Par ailleurs, l'interface graphique affiche les éléments dans différentes couleurs en fonction des alarmes (*trap*) qui lui sont remontées.

LE POINT SUR LA MIB (RFC 1212, 1213, 1155 ET 2863)

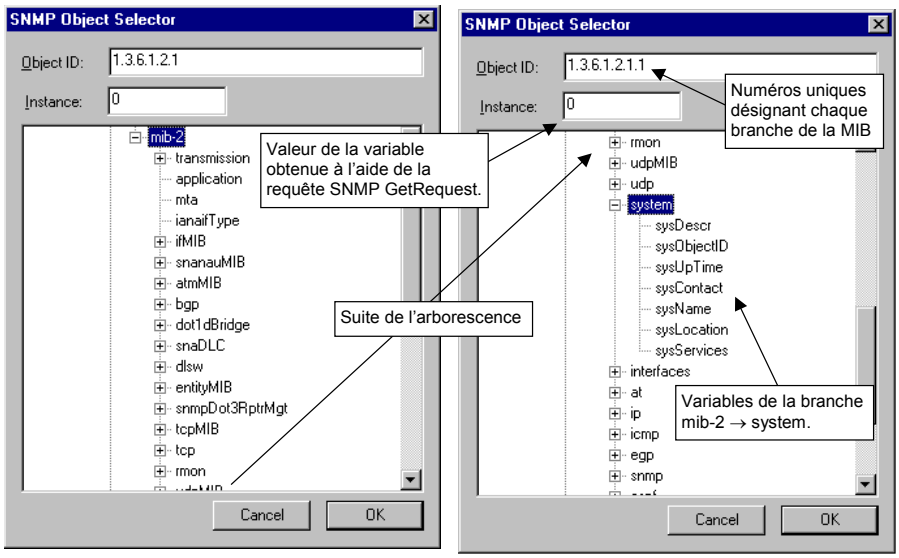
Les agents SNMP interagissent avec la **MIB** (*Management Information Base*) qui contient tous les paramètres de l'équipement réseau. Cette base de données se présente sous forme d'arborescence, normalisée ISO, dans laquelle une branche est réservée à l'Internet. Chaque objet de l'arborescence est identifié par un numéro.



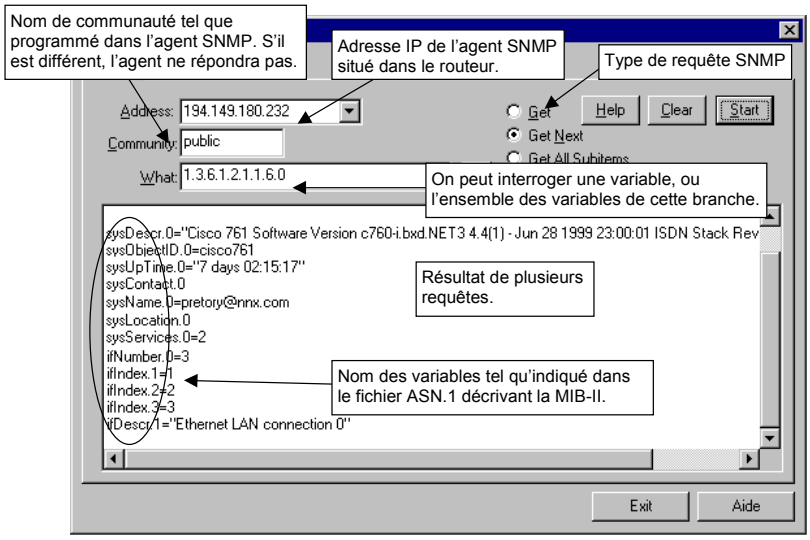
La structure de cette base de données est décrite dans la syntaxe **ASN.1** (*Abstract Syntax Notation 1*) normalisée ISO 8824. Un fichier MIB comporte deux parties (RFC 1212), la première décrivant les types et groupes d'objets (macro **Definitions**), la seconde décrivant les objets (série de macro **Object-Type**).

La macro " *Definitions* " contient deux **clauses** :
Import : importe des définitions d'autres fichiers MIB.
Object Identifier : définit un nouveau groupe (nom + identifiant).
La macro Object-Type contient trois clauses obligatoires, qui prennent les valeurs suivantes :
Syntax = integer | object identifier | octet string | networkaddress | ipaddress.
Access = read-only | read-write | write-only | not-accessible.
Status = mandatory | optional | obsolete | deprecated.
Et quatre autres facultatives, qui prennent les valeurs suivantes :
Description = " texte décrivant l'objet ".
Reference = référence à un autre objet.
Index = noms d'objets dans un objet structuré, par exemple, " ifIndex " pour l'objet *ifEntry* qui contient une liste d'interfaces.
Defval = valeur par défaut de l'objet, par exemple, " sysDescr " pour une syntaxe *Object Identifier*, ou " 1 " pour une syntaxe *Integer*.

La seconde solution est de parcourir la MIB à l'aide d'un ***browser* de MIB**. Cet outil permet de visualiser l'arborescence et d'interroger ou de modifier chaque variable.

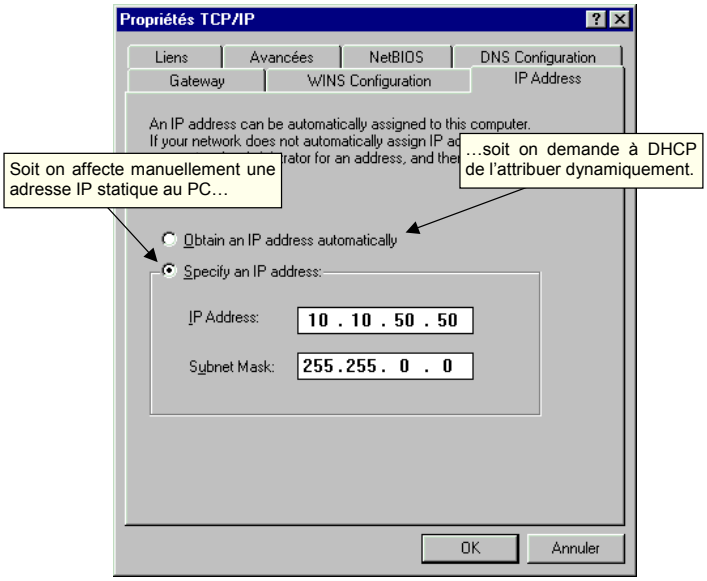


Grâce au browser de MIB, il est possible d'interroger une variable en particulier, puis de programmer des actions automatiques, comme une interrogation périodique du débit entrant et sortant d'une interface, de manière à suivre l'évolution de la charge d'un lien. Avec une série de valeurs, il sera par la suite possible de produire un graphique.



Configurer automatiquement ses PC

La configuration des PC peut s'avérer fastidieuse et être source d'erreurs : adresses dupliquées, masques incorrects, etc. Le protocole **DHCP** (*Dynamic Host Configuration Protocol*) permet d'automatiser ces tâches à l'aide d'un serveur qui héberge les configurations. La plupart des piles IP, dont celle de Microsoft, intègrent un client DHCP. L'unique configuration nécessaire sur les PC consiste à indiquer l'option « Obtenir l'adresse IP par un serveur DHCP » lors de la configuration de la carte réseau.



Par défaut, le PC envoie sa requête à tous les serveurs DHCP et sélectionne généralement le premier qui répond. Il est cependant possible de choisir le serveur DHCP en modifiant la clé de registre « HKEY_local_machine\System\CurrentControlSet\Services \VxD\DHCP\DhcpInfo00\DhcpIPAddress ».

Il s'agit de la même configuration que nous réalisons lorsque nous nous connectons à l'Internet. En effet, les fournisseurs d'accès Internet (les ISP) utilisent systématiquement un serveur DHCP pour attribuer les adresses IP aux PC qui se connectent à leur réseau *via* un modem. Il s'agit généralement d'une adresse publique prise dans le plan d'adressage affecté officiellement à l'opérateur par le RIPE (*Réseau IP Européen*).

Quelle utilisation de DHCP ?

Que vous disposiez de 10 ou 10 000 postes de travail, DHCP sera tout aussi simple à configurer et, dans tous les cas, il vous facilitera la vie.

Ce protocole permet, avant tout, d'affecter une adresse IP à une station pendant une durée limitée. À chaque initialisation — et lorsque la période de validité est expirée —, le PC demande une nouvelle adresse. Cela procure plusieurs avantages :

- Il n'y a plus de risque d'erreur lié à une configuration manuelle.
- Lorsqu'un PC est déplacé et qu'il change de réseau IP, il n'est plus nécessaire de modifier son adresse IP.
- En considérant que tous les PC ne se connectent pas en même temps, on peut utiliser un pool de 253 adresses (une classe C) pour connecter 500 PC, par exemple. Un ratio de un pour quinze est généralement utilisé par les ISP. Cela permet de pallier la pénurie d'adresses publiques.

On peut également affecter l'adresse de façon permanente, mais on perd alors tous les avantages énumérés précédemment.

Plus intéressant, l'utilisation de DHCP peut être étendue à la configuration de tous les paramètres réseau du PC (liés à la famille TCP/IP ou à d'autres protocoles), tels que le routeur par défaut, le masque IP ou encore le TTL par défaut. Cela procure de nouveaux avantages pour l'administrateur réseau :

- Tous les équipements réseau disposent des mêmes paramètres, ce qui assure une meilleure stabilité de fonctionnement de l'ensemble.
- Tout changement de configuration réseau est automatisé. Des opérations complexes, telles que la migration vers un nouveau plan d'adressage ou l'application d'un paramètre TCP permettant d'optimiser le réseau, sont rendues extrêmement simples et rapides.

Les matériels réseau (routeurs, agents SNMP, etc.) ainsi que les serveurs doivent disposer d'adresses fixes, car ils doivent être connus de tous. Ils peuvent faire appel à DHCP pour obtenir une adresse permanente que vous aurez préalablement réservée ou pour obtenir des paramètres de configuration IP.

LES OPTIONS DHCP (RFC 2132)

La RFC 2132 précise les principales options qui peuvent être affectées par un serveur DHCP. Parmi les plus importantes, on trouve (les numéros d'options sont indiqués entre parenthèses) :

- le masque de l'adresse IP (004) ;
- l'adresse IP des serveur DNS et le nom du domaine DNS dans lequel est situé la station (006 et 015) ;
- le nom de la station ;
- l'adresse IP des serveurs WINS et le type de nœud Netbios (044 et 046) ;
- des paramètres IP, TCP et ARP tels que le MTU (026), le TTL (023), la durée du cache ARP (035), etc. ;
- des routes statiques par défaut ainsi que l'adresse du routeur par défaut (033 et 003) ;
- les serveurs de messagerie SMTP et POP (069 et 070) ;
- divers serveurs par défaut tels que web (072), News (071), NTP (042), etc. ;
- des paramètres relatifs à DHCP (durée de validité de l'adresse, etc.) ;
- les types de messages DHCP (DISCOVER, REQUEST, RELEASE, etc.).

D'autres RFC peuvent décrire des options spécifiques (par exemple la RFC 2244 pour des paramètres Novell). La liste exhaustive des options officiellement reconnues est disponible sur http://www.iana.org.

Certains concentrateurs ou commutateurs peuvent télécharger leur système d'exploitation (un fichier exécutable appelé image de boot) ou un fichier de configuration en utilisant le sous-ensemble **BOOTP** également pris en charge par le serveur DHCP.

Avant de commencer, vous pourrez prendre en compte les recommandations suivantes :

- Installez au moins un serveur par site pour des questions de performance et de charge sur les liaisons WAN.
- Pour des questions de sécurité (surtout lorsque le nombre de PC est important), il est préférable de configurer deux serveurs par site, chacun gérant un pool d'adresses.
- La durée de validité des paramètres doit être limitée dans le temps dans le cas où les stations ne sont jamais éteintes (ce qui est le cas des serveurs, par exemple). Une durée de 12 ou 24 heures permet de couvrir une journée de travail et de diffuser de nouveaux paramètres assez rapidement. Pour les serveurs et équipements réseau, une durée plus longue peut être définie, mais l'application d'un nouveau paramètre prendra plus de temps. Vous pourrez de toute façon modifier la durée à tout moment.

Enfin, toutes les piles IP ne prennent pas l'ensemble des options possibles en charge. Il convient donc de vérifier que celle que vous utilisez accepte les options que vous voulez distribuer *via* DHCP.

Comment configurer un serveur DHCP ?

Pour installer le serveur sous Windows NT, il faut se rendre dans la configuration des services IP : « Démarrer→ Paramètres→ Panneau de Configuration→ Réseau→Services→Ajouter ».

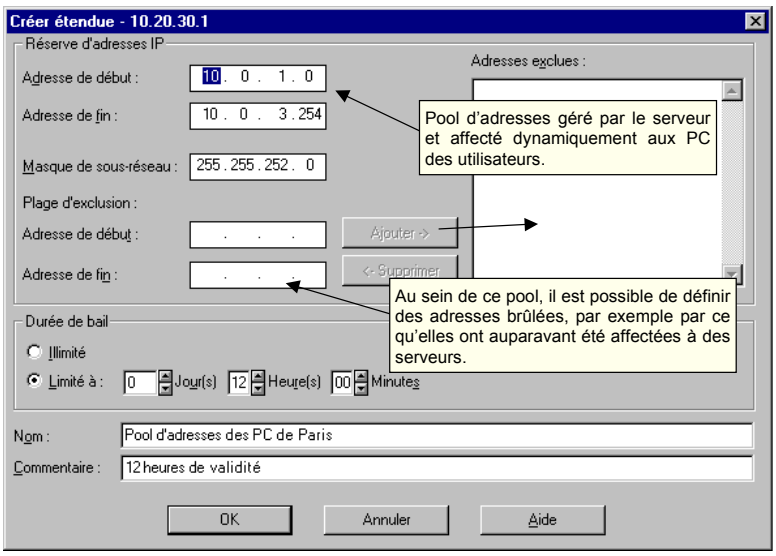
Pour configurer le serveur DHCP, cliquez sur « Démarrer→Programmes→Outils d'administration→ Gestionnaire DHCP ». Nous prenons l'exemple de Windows NT, mais le principe est le même sous Unix.

Définir les pools d'adresses

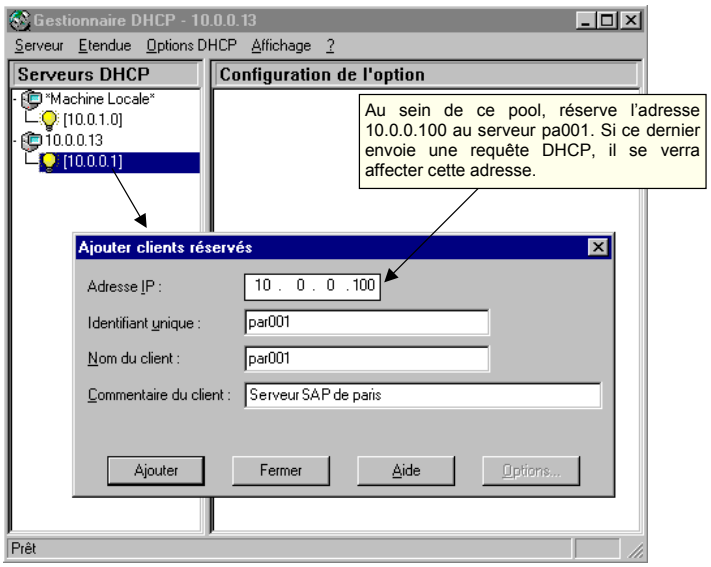
La première étape consiste à définir des pools d'adresses dans lesquels le serveur va piocher pour les affecter aux stations qui en feront la demande. Conformément à notre plan d'adressage, nous avons découpé notre espace d'adressage en trois parties.

Plage d'adresses	Affectation
De 0.1 à 0.64	Équipements réseau (routeurs, hubs, switches, etc.). Les adresses seront fixes, et seules les options seront distribuées (éventuellement des images de boot). Cela implique de référencer tous les équipements concernés (noms et adresses MAC notamment). Durée de validité des options = 1 semaine.
De 0.65 à 0.255	Serveurs NT, Unix, etc. Les adresses seront fixes, et seules les options seront distribuées. Les options peuvent être communes à tous les serveurs. Durée de validité des options = 1 semaine.
De 1.0 à 3.254	Postes de travail (PC, etc.). Adresses et options affectées dynamiquement. Durée de validité = 12 heures.

Rappelons qu'il est souvent souhaitable de limiter le découpage à deux tranches, une pour les équipements réseau et serveur, et une pour les postes de travail. Les deux premières parties peuvent donc être fusionnées.
À Paris, le serveur DHCP prendra ainsi en charge la plage d'adresses allant de 10.0.0.1 à 10.0.3.254, découpée en deux pools (appelés *scope* ou *étendue* chez Microsoft). Le *scope* dédié aux stations commencera à 10.0.1.0. En cas de saturation de la première tranche, il sera toujours possible de modifier la plage d'adresses affectée à ce *scope*.



La configuration du pool d'adresses pour les équipements réseau et les serveurs nécessite, en plus, de réserver les adresses *via* le menu « Etendue→Ajouter adresse réservée ». Avec les clients Windows, le seul moyen d'identifier les stations est d'utiliser l'adresse MAC de la carte réseau. La norme prévoit cependant que l'identifiant puisse être une chaîne de caractères quelconque, le nom de l'utilisateur ou du PC, par exemple.

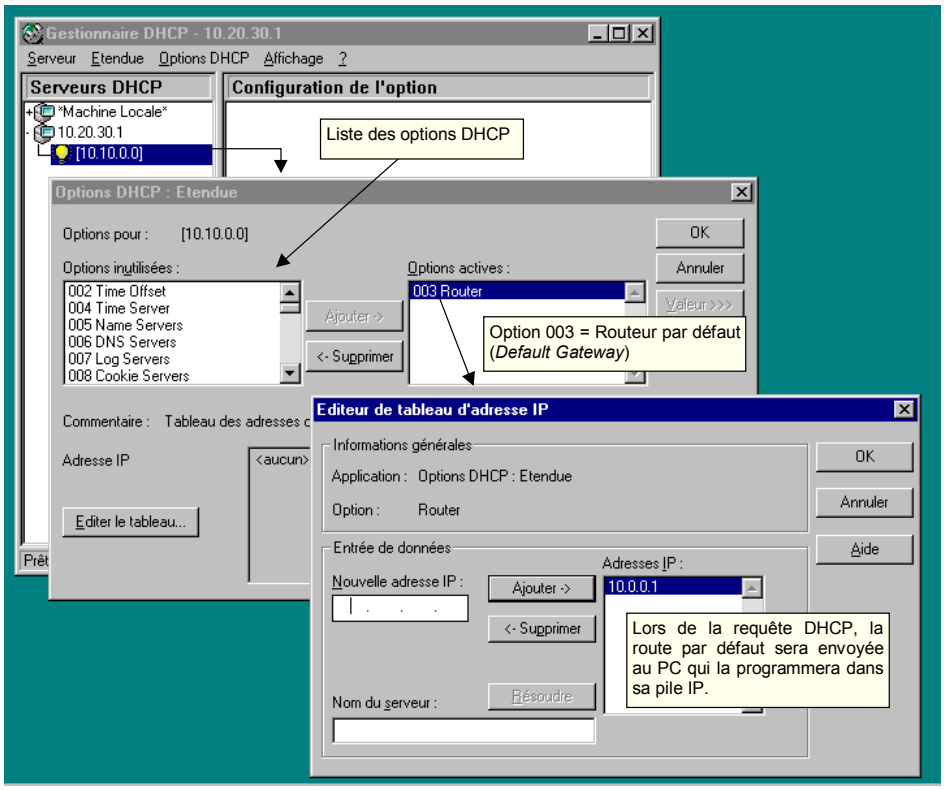


Cette opération peut être fastidieuse, car il faut relever les adresses MAC de tous les équipements concernés (qui sont cependant en nombre moins élevé que les PC). De plus, lorsqu'une carte réseau est changée, il faut mettre à jour la base de données DHCP (c'est une opération en principe peu fréquente, mais il faut y penser le jour où cela arrive).

Si vous ne voulez pas utiliser DHCP pour cette classe d'équipements, il suffit de ne pas ajouter de pool d'adresses. Le plan d'adressage facilite votre choix.

Définir les options à distribuer

Les options peuvent être définies à trois niveaux : soit globalement pour tous les pools d'adresses, soit pour chaque pool (*scope*), soit encore individuellement pour chaque client.
Les options communes à tous les nœuds du réseau — par exemple le TTL par défaut ou l'adresse d'un serveur NTP servant de référence à la mise à l'heure des horloges — peuvent être définies globalement dans le menu « Option DHCP → Global ».



Si vous voulez affecter comme passerelle par défaut l'adresse IP de la station elle-même (voir chapitre 9), il faut ajouter et positionner à « 1 » la clé de registre suivante au niveau du pool « HKEY_local_machine\System\ CurrentControlSet\ Services\ DHCPServer\ Subnets\ a.b.c.d\ SwitchedNetworkFlag », où a.b.c.d est l'adresse IP du pool.

LE POINT SUR DHCP (RFC 2131)

DHCP (*Dynamic Host Configuration Protocol*) permet à une station d'obtenir l'intégralité de ses paramètres IP (plus de 65 options recensées à ce jour), ce qui épargne à l'administrateur de devoir configurer manuellement chaque poste de travail. DHCP est une extension du protocole **BOOTP** ; il utilise le même format de paquet.

8 bits	8 bits	8 bits	8 bits
1 = Requête 2 = Réponse	Type d'adresse physique (1=MAC Eth)	Longueur de l'adresse physique	Saut incrémenté de 1 par les routeurs
XID : numéro identifiant de manière unique la transaction. La réponse du serveur doit contenir le même XID que la demande du client			
Nombre de secondes depuis que le client a initialisé sa demande	B	Indicateurs (non utilisés) B = bit de Broadcast	
ciaddr - Adresse IP du client si celui-ci la connaît Il peut la connaître s'il demande une prolongation de l'affectation de l'adresse.			
yiaddr - Adresse IP affectée par le serveur			
siaddr - Dans son message DHCP_OFFERT, le serveur DHCP indique ici son adresse IP que le client devra indiquer en retour dans son message DHCP_REQUEST			
giaddr - Adresse IP du routeur ayant relayé le message DHCP. Si cette adresse est non nulle, le serveur sait que la requête a traversé au moins un routeur.			
chaddr - Adresse physique du client (16 octets) Dans le cas d'Ethernet, il s'agit de l'adresse MAC			
sname - Nom optionnel du serveur (64 octets maximum terminés par un 0)			
file - Nom optionnel du fichier à télécharger (128 octets maximum terminés par un 0)			
Options - Liste des paramètres supplémentaires affectés au client (les options son listées dans le RFC 2132) ainsi que le type de message DHCP (DISCOVER, REQUEST, etc.).			

Si la pile IP (et notamment le module ARP) peut fonctionner sans adresse IP, le **bit de broadcast** peut être mis à 0 dans les requêtes DHCP, ce qui permet au serveur de renvoyer ses réponses dans des trames unicast (à l'adresse MAC indiquée par le client dans le champ *chaddr*). Dans le cas contraire, le bit de broadcast est positionné à 1 par le client, et le serveur répond dans des trames de broadcast MAC (FF-FF-FF-FF-FF-FF).

Cependant, si le champ ***giaddr*** est non nul, cela veut dire que la requête a transité par un routeur. Le serveur envoie alors le paquet DHCP à cette adresse IP (et donc à l'adresse MAC du routeur *via* ARP). Le port UDP de destination est alors 67 (celui du serveur) — au lieu de 68 qui désigne le client —, ce qui permet au routeur d'identifier les paquets DHCP à traiter (voir plus loin).

• • •

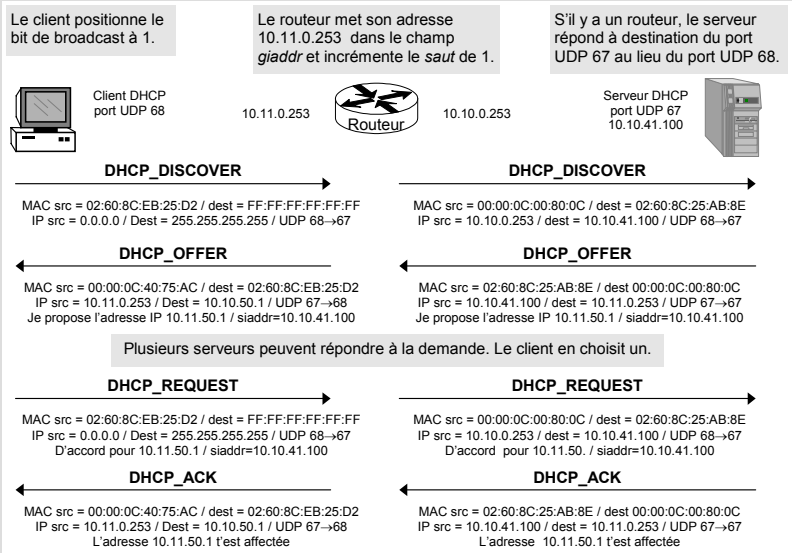
LE POINT SUR DHCP (SUITE)

Si le client possède déjà une adresse IP (champ *ciaddr* non nul), il peut demander des paramètres de configuration complémentaires (les **options DHCP**) en envoyant le message DHCP_INFORM. Le serveur envoie alors sa réponse à l'adresse IP indiquée (donc dans une trame MAC unicast).

Un client effectue sa requête en deux temps :

- Tout d'abord, il recherche un serveur DHCP, et attend les offres du ou des serveurs.
- Ensuite, il confirme sa demande auprès du serveur qu'il a choisi, et attend une réponse lui confirmant que l'adresse IP a bien été réservée.

Les options sont également négociées au cours de cet échange : le client indique celles déjà configurées dans sa pile IP, et les serveurs lui proposent les leurs.



Le client envoie toujours ses requêtes dans des trames de **broadcast MAC** pour plusieurs raisons :

- La requête initiale (DHCP_DISCOVER) permet de découvrir plusieurs serveurs (un serveur principal et un serveur de secours).
- Lors de la requête de confirmation (DHCP_REQUEST), la plupart des piles IP ne peuvent activer le module ARP sans adresse IP.
- Le client ne sait pas si le serveur est situé sur le même réseau ou s'il est séparé par un routeur. Dans ce dernier cas, si la trame est destinée à l'adresse MAC du serveur, elle ne traversera pas le routeur, sauf s'il fonctionne en mode proxy ARP (voir chapitre 9), ce que la station ne peut présupposer.

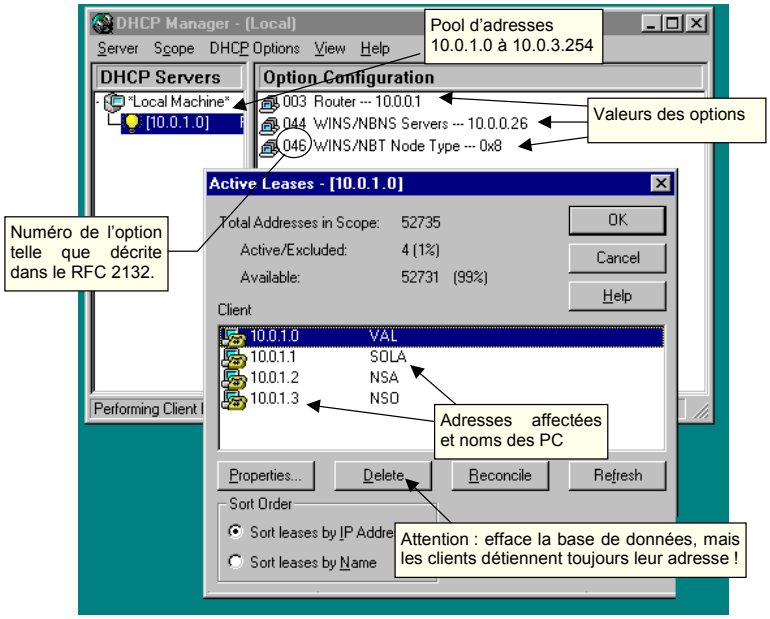
Configuré en relais DHCP/BOOTP, le routeur convertit les broadcast MAC/IP en adresses unicast à destination du serveur DHCP ou BOOTP.

À l'approche de l'expiration de la période de validité, le client demande à renouveler son bail auprès du serveur (DHCP_REQUEST) en indiquant son adresse IP dans le champ *ciaddr*. Le serveur peut alors proposer la même adresse, une nouvelle, ou encore accepter celle demandée par le client.

En principe, c'est au client qu'il incombe de vérifier que l'adresse allouée par le serveur n'est pas utilisée par une autre station (le serveur peut, en effet, être situé de l'autre côté d'un routeur). Le client génère à cet effet une requête ARP sur l'adresse qui vient de lui être allouée.

La même option peut être définie à plusieurs niveaux, mais les options individuelles ont priorité sur les options d'un pool, qui elles-mêmes ont priorité sur les options globales.

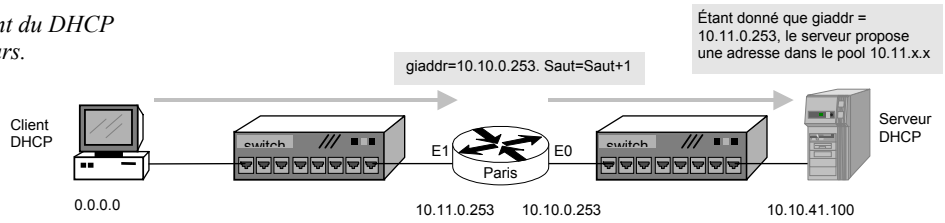
L'exemple suivant montre les options définies pour un pool, ainsi que les adresses déjà affectées à des clients.



Configurer les routeurs

Si les stations sont situées sur un réseau IP autre que le serveur DHCP, les requêtes DHCP doivent transiter par un routeur. Or, ce type d'équipement ne transmet jamais les trames de broadcast MAC, car il est justement conçu pour délimiter les domaines de broadcast (voir l'encart « Le point sur Ethernet » au chapitre 3). Il faut donc configurer explicitement les routeurs afin de pouvoir relayer les requêtes DHCP.

Figure 6-3.
Fonctionnement du DHCP avec les routeurs.



```
interface ethernet 1
ip helper-address 10.10.41.100
ip helper-address adresses ip d'autres serveurs DHCP
```

Avec la commande précédente, les trames de broadcast MAC dont le port UDP est égal à 67 seront transmises dans une trame unicast (*via* la résolution ARP) à destination du serveur DHCP ou BOOTP et, en retour, vers le client.

LE POINT SUR BOOTP (RFC 951 ET 1542)

BOOTP (*Bootstrap Protocol*) permet à un équipement réseau d'obtenir son adresse IP ainsi que le nom d'un fichier à télécharger *via* **TFTP** (*Trivial File Transfer Protocol*). Il peut s'agir d'un **fichier de configuration** ou d'un **exécutable** (appelé image de boot), tel qu'un système d'exploitation ou un micro-code. Ce protocole ne permet pas d'affecter dynamiquement les adresses, et nécessite donc de connaître les adresses MAC ou d'affecter un nom aux équipements qui émettent des requêtes.

DHCP a repris exactement les mêmes spécifications que BOOTP en étendant ses possibilités. Un serveur DHCP prend en charge les requêtes BOOTP (compatibilité ascendante), alors que l'inverse n'est pas possible.

Des équipements réseau qui ne disposent pas de mémoire flash, tels que des concentrateurs, des commutateurs ou des serveurs d'accès distants, utilisent BOOTP pour télécharger leur code exécutable.

Pour la petite histoire, la RFC 1542, relative aux extensions de BOOTP (l'équivalent des options DHCP), discute de l'utilité du bit de broadcast en rappelant le paradigme de la poule et de l'œuf : une station ne peut pas fonctionner sans adresse IP ; cependant, elle envoie et reçoit des paquets IP qui doivent justement lui permettre d'obtenir cette adresse IP ; mais elle ne peut pas traiter ces paquets puisqu'elle ne dispose pas d'adresse IP, etc.

Installer plusieurs serveurs

Si plusieurs serveurs sont utilisés (en partage de charge et en redondance), le pool d'adresses doit être découpé afin d'éviter les doubles affectations. La répartition peut se faire à parts égales, comme suit :

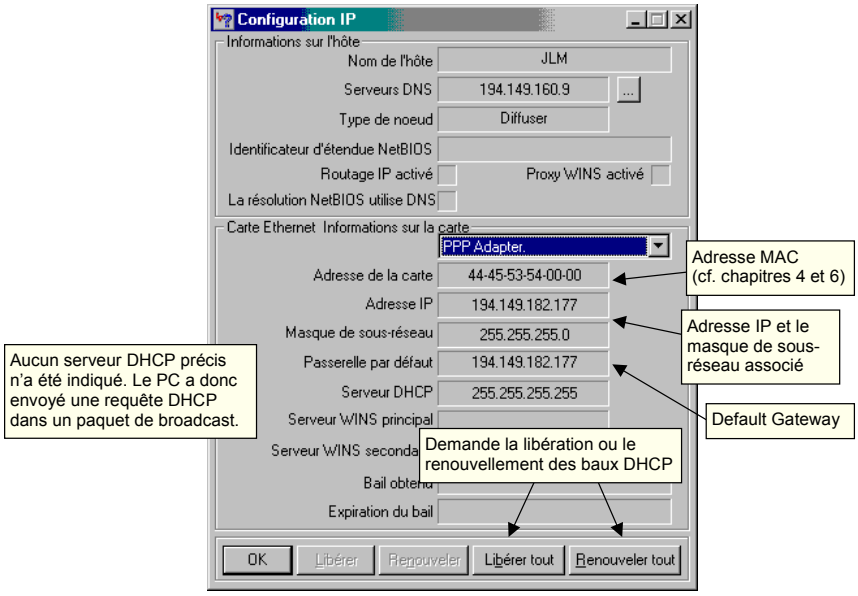
- le premier serveur affecte les adresses comprises entre 10.0.1.0 et 10.0.2.127 ;
- le second serveur affecte les adresses comprises entre 10.0.2.128 et 10.0.3.254.

Les adresses affectées de manière fixe doivent être réservées de manière identique sur chaque serveur.

Le serveur de Microsoft ne permet pas de mettre en place un réel partage de charge avec une redondance complète. Pour cela, d'autres serveurs DHCP plus perfectionnés existent sur le marché.

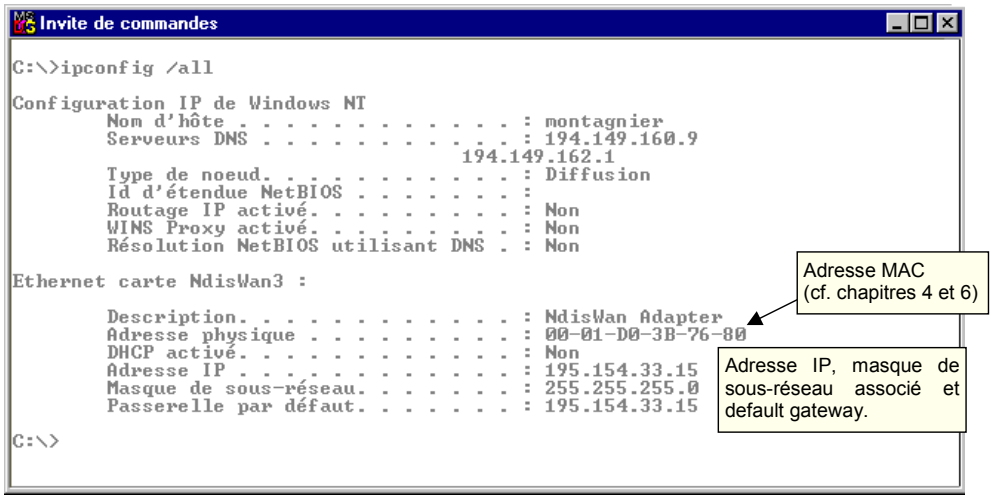
Vérifier la configuration de son PC

Plusieurs utilitaires permettent de vérifier le paramétrage TCP/IP de son PC. Sous Windows 9.x et Me, il s'agit de la commande **winipcfg**.



Sous Windows NT, la commande équivalente est **ipconfig** :

- **ipconfig /all** affiche tous les paramètres réseau.
- **ipconfig /release** envoie un message DHCP_RELEASE au serveur pour libérer l'adresse IP.
- **ipconfig /renew** envoie un DHCP_REQUEST pour demander le prolongement de la validité de l'adresse IP, ou un DHCP_DISCOVER si la station ne possède pas d'adresse.



Sous les deux environnements, la commande **route** permet de visualiser la table de routage de la pile IP.

Invite de commandes

```
C:\>route -p add 192.168.0.0 mask 255.255.255.0 10.10.0.253

C:\>route print

Itinéraires actifs :

  Adresse réseau          Masque réseau  Adresse passerelle  Interface  Métrique
          0.0.0.0              0.0.0.0        10.10.0.1     10.10.50.50       1
        127.0.0.0            255.0.0.0        127.0.0.1       127.0.0.1       1
      192.168.0.0      255.255.255.0       10.10.0.253     10.10.50.50       1
         10.10.0.0        255.255.0.0       10.10.50.50     10.10.50.50       1
     10.10.50.50  255.255.255.255          127.0.0.1       127.0.0.1       1
   10.10.255.255  255.255.255.255          10.10.0.1     10.10.50.50       1
       224.0.0.0        224.0.0.0          10.10.0.1     10.10.50.50       1
 255.255.255.255  255.255.255.255          10.10.0.1     10.10.50.50       1

C:\>_
```

Route permanente (sauvegardée dans la Registry).

Ajoute une route statique : pour envoyer un paquet vers le réseau 192.168.0.0, le PC le transmettra au routeur 10.0.0.253.

Route par défaut (defaut gateway)

Adresse de loopback utilisée pour vérifier que la pile IP fonctionne bien (par exemple avec un ping)

Groupe multicast par défaut (cf. chapitre 15)

Enfin, la commande **Netstat** permet de vérifier le bon fonctionnement de la couche TCP/IP. Elle permet, par exemple, de visualiser les connexions actives.

Invite de commandes

```
C:\>netstat

Connexions actives

  Proto  Adresse locale        Adresse extérieure      Etat
  TCP    montagnier:1025       localhost:1026          ETABLIE
  TCP    montagnier:1026       localhost:1025          ETABLIE
  TCP    montagnier:1328       ADSERVER58:80           ETABLIE
  TCP    montagnier:1641       cctvw02.cctec.com:80    TEMPS D'ATTENTE
  TCP    montagnier:1666       cio-sys.cisco.com:80    ETABLIE
  TCP    montagnier:1667       cio-sys.cisco.com:80    ETABLIE
  TCP    montagnier:1676       www.smartbooks.com:80   ATTENTE_FERMER
  TCP    montagnier:1677       www.smartbooks.com:80   ATTENTE_FERMER
  TCP    montagnier:1678       www.smartbooks.com:80   ATTENTE_FERMER
  TCP    montagnier:1679       www.smartbooks.com:80   ATTENTE_FERMER

C:\>
```

Noms obtenus par le DNS et correspondant à l'adresse IP cible.

Numéros de ports TCP source et destination.

7

La gestion des noms

Partie de la câblerie et de la basse filasse, nous abordons dans ce chapitre les couches applicatives, c'est-à-dire les services réseau. Pour ainsi dire, nous nous élevons dans les couches supérieures du réseau.

Car, de nos jours, l'administrateur réseau ne peut pas se contenter de fournir le transport des données. Il doit en faciliter l'accès à ses utilisateurs.

En outre, plus le nombre d'utilisateurs est élevé, plus il est important de simplifier les tâches administratives.

Il convient donc d'utiliser des outils qui simplifient la vie des utilisateurs et celle des exploitants. Le service de nom, ou DNS (*Domain Name System*), est le premier d'entre eux.

Dans ce chapitre vous apprendrez ainsi :

- à comprendre le fonctionnement du DNS ;
- à définir un plan de nommage ;
- à configurer les serveurs DNS ;
- à configurer les PC ;
- à interroger la base de données du DNS.

Présentation du DNS

Pour vos utilisateurs et vous-même, il serait bien plus pratique d'utiliser des noms de machines plutôt que des adresses, à l'instar de ce qui se fait sur l'Internet. De même que l'on accède à www.3com.com, il serait bien plus simple d'accéder à votre serveur au moyen du nom www.societe.fr plutôt que de son adresse IP.

La première solution repose sur l'utilisation des fichiers hosts (localisés dans /etc sous Unix, et dans \Windows sous Windows 9x). Ce fichier contient simplement la correspondance entre adresses IP et noms de machines :

```
10.0.0.1      par001
10.0.0.100    nt001
10.0.0.101    mail
```

L'inconvénient de cette méthode est qu'il faut configurer le fichier sur chaque poste de travail, et ce, à chaque changement d'adresse ou de nom. On pourrait imaginer une distribution automatique de ce fichier, mais cette opération serait très complexe et source d'erreur avec les PC. De plus, l'espace de nommage est « plat » (tous les noms sont au même niveau).

La seconde solution, de loin la meilleure, repose donc sur l'utilisation d'un système DNS (*Domain Name System*). C'est celle utilisée à grande échelle sur l'Internet (plusieurs millions de machines référencées début 2001) et qui convient également pour vos 10 ou 10 000 PC.

Les composants du DNS

Le DNS est utilisé sur l'Internet pour naviguer sur le Web et pour envoyer des messages. Il s'agit ici de recréer un DNS privé, c'est-à-dire réservé à nos utilisateurs.

Il faut pour cela définir :

- un espace de nommage hiérarchique découpé en domaines ;
- des serveurs gérant des bases de données ;
- des clients appelés *resolver* ;
- un protocole d'échange entre clients et serveurs d'une part, et entre serveurs d'autre part.

Tous ces composants sont décrits dans une série de RFC dont les premiers remontent à 1987.

Élaborer un plan de nommage

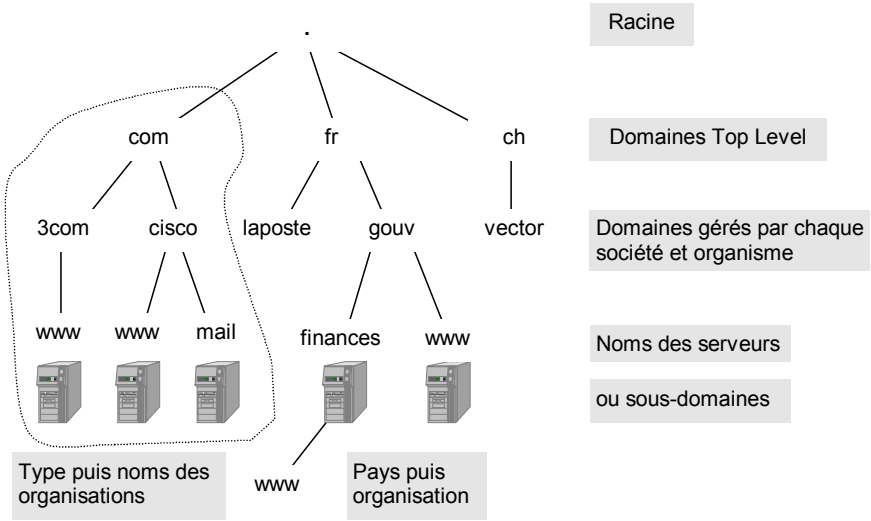
Le nommage DNS est organisé sous forme d'arbre, avec une racine et des domaines qui lui sont rattachés. Le plan de nommage consiste donc à définir cette arborescence et la manière d'affecter des noms aux objets (les feuilles de l'arbre).

Définir l'arborescence DNS

Dès le début de votre réflexion, vous serez confronté au dilemme classique : définir une arborescence qui reflète l'organisation de la société ou une arborescence qui reflète son implémentation géographique ?

Par expérience, l'une n'est pas meilleure que l'autre, car toutes deux sont soumises aux aléas des changements. L'approche organisationnelle est soumise au changement du nom de la société (suite à une décision stratégique, à un rachat, etc.) ou du service (suite à une réorganisation), tandis que l'approche géographique est soumise aux déménagements.

L'Internet a d'ailleurs retenu les deux approches.



Pour notre DNS privé, donc à usage purement interne, nous n'avons pas besoin de faire référence au pays et au nom de la société, mais plutôt à la ville, au nom du site et au nom des directions (ou, selon la terminologie propre à chaque société, des divisions, des *Business Units*, etc.).

Nous pouvons cependant retenir la même approche mixte en fonction du degré de centralisation de chaque département.

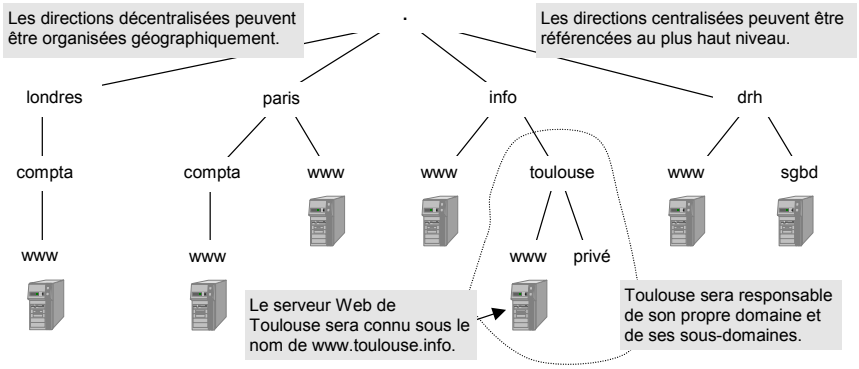
Représentation dans le DNS	Organisation centralisée	Organisation décentralisée
Racine	Par convention, le point ('.')	Par convention, le point ('.')
Domaine Top Level	Nom de la direction	Nom de la ville
Domaine (optionnel)	Nom de la ville	Nom de la direction

Il est conseillé de ne retenir que les invariants ou, pour être exact, les éléments qui sont susceptibles de changer le moins souvent. Dans notre cas, ce sont les directions (c'est-à-dire les services situés au sommet de la hiérarchie organisationnelle) et les villes principales où est implantée notre société.

L'espace de nommage est indépendant de la localisation géographique : tous les immeubles d'une même ville peuvent donc être référencés dans le même domaine DNS.

Par exemple, l'indication d'éléments pouvant changer fréquemment, tels que le nom du service au sein d'une direction ou le nom du site dans une ville, est source de complication, et entraîne un travail supplémentaire de reconfiguration. Plus vous ajouterez de niveaux au DNS, plus vous devrez effectuer de mises à jour. En revanche, si un service au sein d'une direction devait garder son autonomie, on pourrait envisager de lui déléguer la gestion de son sous-domaine.

Figure 7-1. *Définition d'un DNS privé.*



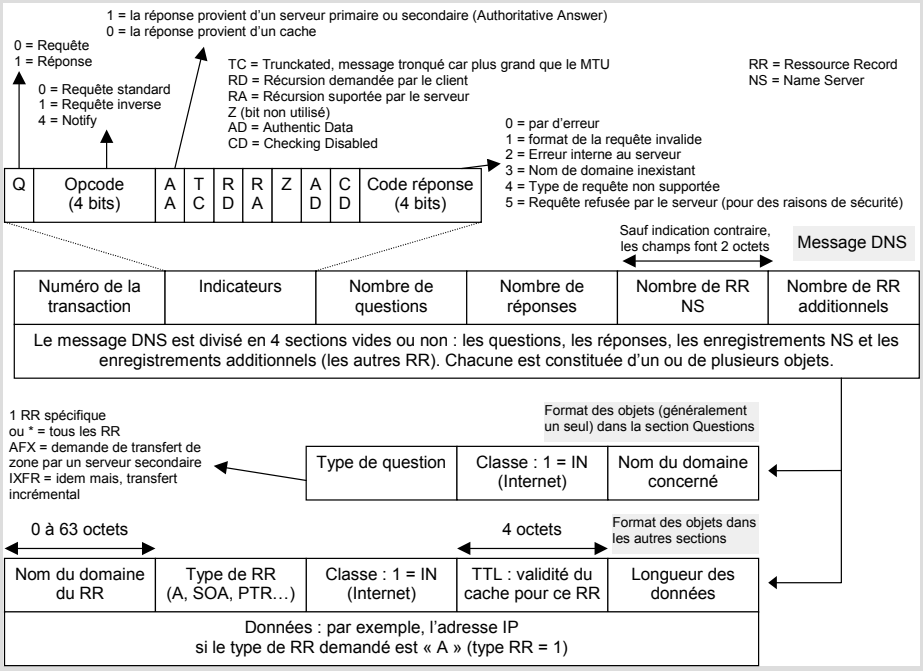
Le DNS offre donc beaucoup de souplesse :

- Une même machine (ayant une adresse IP) peut être référencée dans plusieurs domaines.
- Une même machine peut être référencée sous plusieurs noms principaux et/ou sous un nom principal associé à des alias.
- La gestion d'un domaine peut être déléguée à un nouveau service devenu autonome, ou être reprise de façon centrale.
- Une direction peut gérer son propre DNS (avec sa propre racine).
- Les machines peuvent être référencées dans plusieurs DNS.
- Un DNS peut comporter une centaine de niveaux.

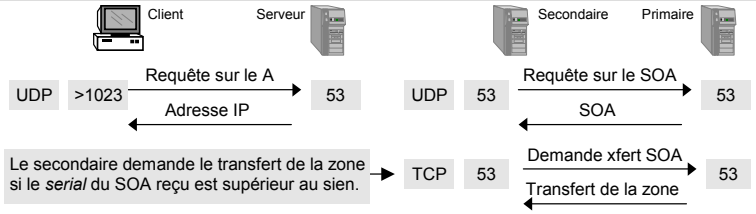
On le voit, le DNS permet toutes sortes de fantaisies. Il est donc de votre responsabilité d'en assurer la cohérence et la simplicité. Ainsi, les fonctionnalités présentées ci-dessus doivent-elles plutôt être utilisées pour faciliter les périodes de transition lors de changements d'organisation ou de déménagement ou, d'une manière générale, pour répondre à des situations exceptionnelles.

LE POINT SUR LE DNS (RFC 1034, 1035, 1995, 1996, 2181)

Le DNS (*Domain Name System*, quelquefois appelé *Domain Name Service*) définit une base de données découpée en **zones** (constituées d'un **domaine** et de ses sous-domaines) correspondant à une partie d'un **arbre hiérarchique**. Un serveur peut avoir autorité sur une ou plusieurs zones. S'il est **primaire**, il est maître de la zone. S'il est **secondaire**, il dispose d'une copie qu'il demande régulièrement au [serveur] primaire. S'il est **racine**, le serveur a autorité sur la racine de l'arbre. Les serveurs racines peuvent **déléguer leur autorité** aux serveurs gérant le premier niveau de domaines, appelés domaines **Top Level**, et ainsi de suite. Tous les serveurs font également office de **cache** pour les requêtes DNS émises par les clients appelés **resolver**.



Afin de diminuer la taille des messages DNS, les noms apparaissant plusieurs fois peuvent être remplacés par des pointeurs de deux octets indiquant la position de l'unique exemplaire du nom. La longueur maximale d'un nom d'objet est de 63 octets, et celle d'un nom de domaine complet (y compris celui de l'objet) est de 255 octets. Les messages DNS transitent dans des paquets UDP ou TCP selon les cas (port 53).



La RFC 1995 précise un mode de transfert incrémental (seules les modifications sont transférées). La RFC 1996 spécifie, quant à elle, un mécanisme permettant au primaire de notifier au secondaire que le SOA vient d'être modifié. Cela évite d'attendre la fin de la période indiquée dans le paramètre *refresh*.

Standardiser le nommage des objets

Il n'est pas nécessaire de mettre les postes de travail dans la base DNS, car ces derniers ne sont que des clients ; ils ne sont pas connus en tant que serveur. Seuls sont renseignés dans les bases DNS les équipements réseau (pour les exploitants) et les serveurs (pour les utilisateurs).

Il est conseillé d'adopter un codage différent et adapté à chaque type d'objet, c'est-à-dire à sa nature et à sa fonction. Par exemple, les serveurs ne seront pas nommés de la même manière que les routeurs. Le nom principal doit correspondre à un invariant, et les alias à la particularité du moment.

Si la sécurité est privilégiée et que le nom ne doit pas permettre d'identifier ces éléments, des noms neutres peuvent être retenus (noms de musiciens, de fleurs, de planètes, etc.).

Objet	Invariant	Variant
Serveur	Système d'exploitation (Unix, NT)	Fonction (Web, messagerie, base de donnée, etc.)
Routeur	Marque (Cisco, 3com, etc.) et fonction de routage	Localisation
Concentrateur	Marque (3com, etc.) et fonction de concentration	Localisation et, éventuellement, emploi de réseaux différents (Ethernet, ATM, etc.)
Commutateur	Marque (Cisco, 3com, etc.) et fonction de commutation	Localisation et, éventuellement, emploi de réseaux différents (Ethernet, ATM, etc.)
Serveur d'accès distant	Marque (Cisco, 3com, etc.) et fonction d'accès distant	Localisation

La localisation peut être considérée comme étant un variant de faible impact à partir du moment où l'équipement doit de toute façon être reconfiguré (changement d'adresse IP, par exemple).

Il faut privilégier un nommage simple des équipements auxquels on accède le plus fréquemment : les serveurs (auxquels accèdent les utilisateurs) et les routeurs et serveurs d'accès distants (auxquels accèdent les exploitants réseau). Le nom doit :

- Ne comporter que des caractères alphanumériques (minuscules et/ou majuscules) et des tirets (-). Ils constituent, en effet, le plus petit dénominateur commun dans le monde de l'informatique. Les autres caractères sont à proscrire, car certains systèmes ne les acceptent pas.
- Être court, afin d'être facile à mémoriser et rapide à saisir au clavier.
- Contenir une ou deux alternances de noms et de chiffres pour en améliorer la lisibilité.
- Contenir un tiret au maximum pour séparer deux champs alphabétiques ou numériques, afin d'en améliorer la lisibilité.

Équipement	Codage retenu	Nom principal (nom système)	Exemple
Serveur NT	Nom neutre (planètes)	mars pluton	mars.marseille. mars.paris.
Unix	Nom neutre (musiciens)	mozart schubert	mozart.toulouse.info.
Routeur	Ville sur trois lettres et numéro d'ordre sur trois chiffres	par001, par002 toul001	par001.info. par001.paris.
Serveurs d'accès distants	Préfixe svc, suivi du codage identique aux routeurs	ras-par001 ras-mar002	ras-mar002.info.

Un serveur peut être connu sous plusieurs noms :

- Un premier qui désigne sa nature et qui correspond au nom système défini lors de l'installation. Il ne change pas (sauf lors d'une réinstallation).
- Un ou plusieurs autres qui désignent sa ou ses fonctions et qui correspondent à une ou plusieurs applications (base de donnée, Web, etc.)

Pour les utilisateurs, le meilleur moyen d'identifier et de mémoriser le nom d'un serveur est de lui donner le nom de l'application dont ils se servent. Un **alias** correspondant à la fonction du serveur pourra donc être défini et correspondre à la fonction du serveur (il varie au cours du temps alors que le nom système ne change que si l'on réinstalle complètement la machine).

Fonction	Codage retenu	Alias	Exemple
Serveur Web principal	Convention universelle	www	www.paris. www.toulouse.info. www.drh.
Autres serveurs Web	Convention suivie d'un numé-ro d'ordre sur un chiffre	www1 www2	www.paris. www1.paris. www2 .paris.
Messagerie	Mail suivi d'un numéro d'ordre	mail mail1	mail.paris. mail1.info.
Autres applications	Nom de l'application	compta rivage	compta.paris. rivage.info.

Certains serveurs centraux, c'est-à-dire communs à l'ensemble de la société, pourront être situés juste sous la racine. On aura, par exemple, www pour le serveur servant de point d'entrée à toute la société : il contiendra les informations du jour et des liens vers les autres serveurs Web. On pourra aussi y placer les serveurs qui réalisent l'interconnexion des messageries.

Les concentrateurs et les commutateurs ne sont généralement pas accessibles au moyen d'une connexion Telnet, mais *via* SNMP et une station d'administration. Le codage des noms peut donc être plus complexe. Il peut refléter leur localisation géographique et, éven-

tuellement, leur fonction, afin de faciliter leur identification. Dans notre cas, nous avons choisi le codage suivant : [type]-[ville][étage][immeuble].

Champ	Signification	Code
[type]	Type de matériels, sur une lettre	**H** = Hub, **S** = switch, **A** = ATM, **F** = Frame Relay, **I** = FDDI
[ville]	Abréviation de la ville, sur trois lettres	**par** = Paris, **tou** = Toulouse **tur** = Tour, etc.
[immeuble]	Lettre majuscule identifiant un immeuble dans la ville	Le code dépend de la ville **D** = Descartes, **M** = Montparnasse, etc.

Ce qui donne, par exemple, h-par05a, s-tou05D, etc.

Afin de faciliter leur lecture et leur traitement dans des bases de données, il est préférable que chaque champ ait une longueur fixe.

Configurer les serveurs DNS

Le DNS est géré par des serveurs qui assurent plusieurs fonctions :

- la gestion d'une **zone**, c'est-à-dire d'un domaine et de ses sous-domaines : on dit que le serveur a **autorité** sur la zone ;
- l'échange des bases de données au sein d'une zone : un serveur est désigné **primaire** pour une zone et distribue la base de données aux serveurs **secondaires** ;
- le **relais** des requêtes DNS d'un client sur un nom situé dans une autre zone ;
- le **cache** des requêtes clients de manière à limiter les requêtes ;
- enfin, un serveur doit être désigné **racine** de l'arbre DNS.

Tout serveur DNS est serveur cache et peut être à la fois ou seulement :

- **primaire** pour un ou plusieurs domaines ;
- **secondaire** pour un ou plusieurs autres domaines ;
- **racine**.

Comme pour le nommage, le DNS offre de nombreuses possibilités et peu de contraintes :

- Il faut au moins un serveur racine.
- Il faut un et un seul serveur primaire par domaine.
- Toutes les modifications de la base de données d'une zone doivent être réalisées sur le serveur qui a autorité (le serveur primaire).
- Il peut y avoir aucun ou autant de serveurs secondaires par domaine.
- Les fonctions primaires et secondaires sont exclusives pour un domaine donné.
- Mais un serveur peut être à la fois primaire pour un domaine et secondaire pour un autre.

- Un serveur DNS peut ne servir que de cache (c'est-à-dire n'être ni racine, ni primaire, ni secondaire).
- L'administrateur peut déléguer la gestion d'un domaine faisant partie de la zone d'autorité à un autre serveur qui devient alors primaire pour le domaine.
- Les clients et les serveurs DNS peuvent être situés n'importe où sur le réseau : sur des lieux géographiques différents et sur des réseaux IP différents.
- Il peut y avoir plusieurs DNS distincts.
- Les postes de travail et les serveurs peuvent être référencés dans n'importe quel domaine.

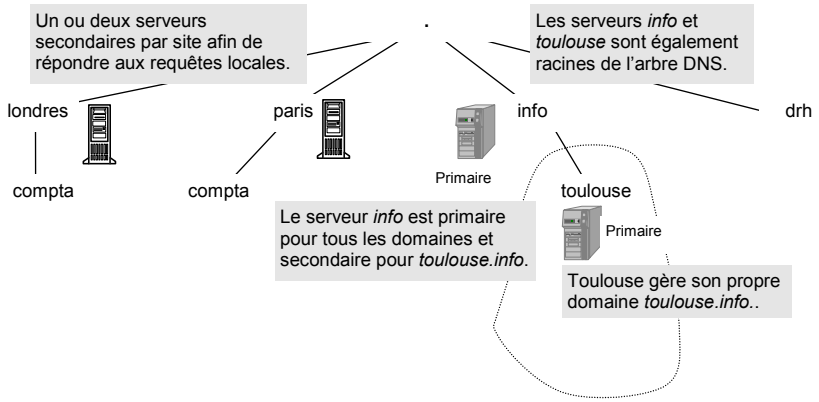
Le DNS permet de faire tout et n'importe quoi. Il convient donc de respecter quelques règles afin de conserver une certaine cohérence au sein de votre société. Ainsi, vous devez prévoir :

- au moins un serveur DNS par site afin de ne pas surcharger les liaisons WAN ;
- au moins un serveur secondaire en partage de charge et en secours ;
- un arbre DNS pour toute la société, même si chacun gère son propre domaine.

Les bonnes pratiques du DNS sont exposées dans la RFC 1912.

Selon ces principes, nous avons décidé de créer l'architecture présentée ci-après.

Figure 7-2.
Les serveurs DNS.



Chaque serveur DNS gère une partie de la base de données DNS, celle représentant la zone sur laquelle il a autorité. Le serveur de Toulouse ne contiendra ainsi que les objets situés dans la zone « toulouse.info ».

Sous Windows NT, la manipulation de la base de données est réalisée à l'aide du Gestionnaire DNS, situé dans le menu « Démarrer → Programmes → Outils d'administration ». Celui-ci sauvegarde les données dans la base des registres Windows (clé \HKEY_LOCAL_MACHINE\System\CurrentControlSet\Services \Dns) et maintient une copie sous forme de fichiers ASCII dans le répertoire \Winnt\System32\Dns. Ces fichiers respectent le format des serveurs DNS sous Unix, tels que ceux produits par la version appelée BIND et développée à Berkeley.

LA BASE DE DONNÉES DU DNS

La base de données consiste en des fichiers ASCII contenant des **enregistrements** appelés **RR** (*Resource Record*) dont le format générique est le suivant :

nom [ttl] [IN] RR paramètres

nom Désigne le nom du domaine. S'il est omis, c'est le même que celui du SOA.

ttl Time To Live. Indique pendant combien de secondes le *resolver,* qui aura fait une requête, pourra conserver cet enregistrement dans son cache. Le TTL est défini au niveau du SOA (et donc pour tous les RR situés dedans), mais peut, optionnellement, être défini pour chaque RR. Le délai de validité du cache est le maximum des valeurs *ttl* et *minimum*.

IN Désigne la classe d'adresse, dans notre cas l'Internet (on trouve aussi CH pour les adresses Chaos).

RR Nom de la ressource (SOA, NS, A, LNAME, NX, PTR, HINFO, etc.).

Les principaux enregistrements sont SOA, NS, A, CNAME, MX, PTR, HINFO et AAA.

nom [ttl] [IN] **SOA** name e-mail serial refresh retry expire minimum

(*Start Of Authority*). Indique le nom de la zone pour laquelle le serveur a autorité (secondaire ou primaire), ainsi que les paramètres de mise à jour de la base de données.

name Désigne le serveur primaire de la zone.

e-mail Désigne l'adresse e-mail de la personne responsable de la zone.

serial Désigne le numéro de version du SOA.

refresh Indique le nombre de secondes au bout duquel le serveur secondaire doit redemander le SOA au serveur primaire.

retry Indique le nombre de secondes qui s'écoule entre deux tentatives de téléchargement du SOA par le serveur secondaire.

expire Indique le nombre de secondes au bout duquel le SOA ne sera plus valable après le délai indiqué par *refresh*. Au-delà de ce délai, le serveur secondaire ne doit plus répondre aux requêtes concernant ce SOA.

minimum Indique la valeur minimale du TTL des RR de cette zone.

[nom] [ttl] IN **NS** nom_du serveur_DNS
(*Name Server*) Indique le serveur (primaire ou secondaire) qui a autorité sur la zone.

[nom] [ttl] IN **A** adresse_IPv4
(*Address*) Indique l'adresse IPv4 qui correspond au nom demandé dans une requête.

alias [ttl] IN **CNAME** nom_principal_de_l'objet
(*Canonical Name*) Indique l'alias d'un objet.

nom [ttl] IN **MX** priorité nom_du_MTA
(*Mail eXchanger*) Indique le nom du MTA acheminant les messages à destination du domaine. 65535 = priorité basse, 1 = priorité haute.

inverse [ttl] IN **PTR** nom_machine
(Pointer) Indique le nom qui correspond à l'adresse IP demandée dans une requête.
inverse Désigne l'adresse IP inverse, suivie de in-addr.arpa.

[nom] [ttl] IN **HINFO** matériel système
(*Host Information*) Champ d'information concernant l'objet (type de machine et d'OS).

[nom] [ttl] IN **AAAA** adresse_IPv6
Indique l'adresse IPv6 qui correspond au nom demandé (RFC 1886).

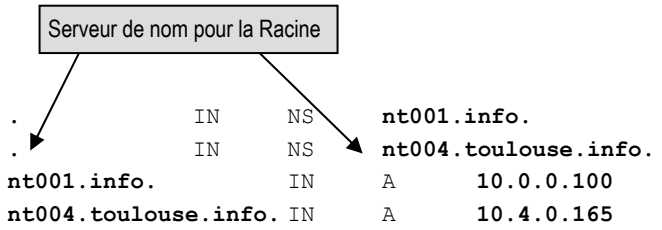
• • •

LA BASE DE DONNÉES DU DNS (SUITE)

Les noms de serveurs indiqués dans les RR ne doivent pas être des alias.
Il existe d'autres enregistrements peu ou pas utilisés ou encore à l'état expérimental : WKS, TXT (RFC 1035), AFSDB, RP, X25, ISDN, RT (RFC 1183), NSAP (RFC 1706), GPOS (RFC 1712), SRV (RFC 2052) et KX (RFC 2230).

Configurer le fichier cache

Tous les serveurs participant à votre DNS doivent connaître les serveurs racines. Ces informations résident dans le **fichier cache** (attention, celui-ci n'a rien à voir avec le serveur appelé cache). Sous Windows NT, ce fichier est situé dans le répertoire \Winnt\System32\Dns\Cache.dns. Il doit être édité manuellement sur chaque serveur DNS de votre société :



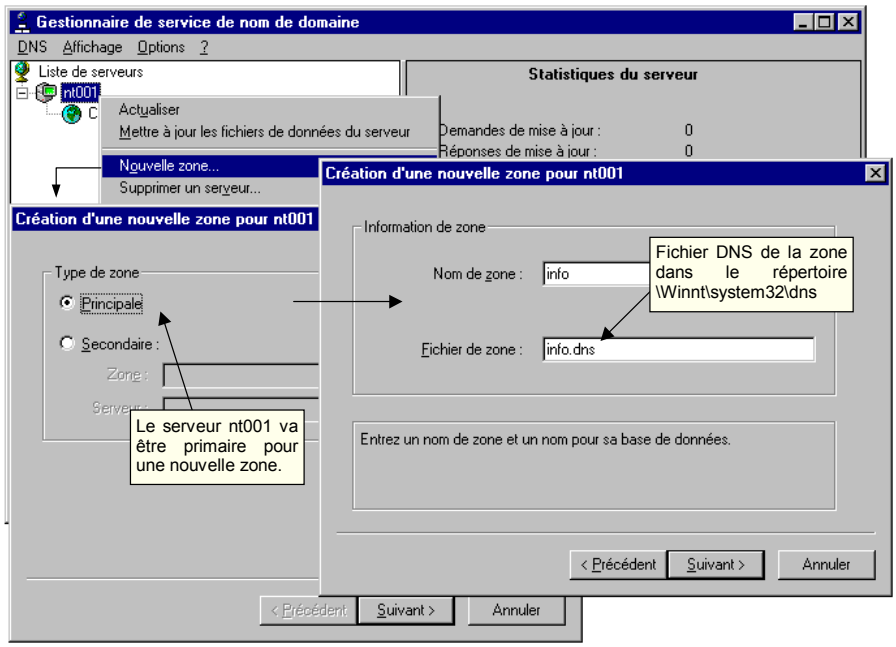
Les instructions NS (*Name Server*) et A (*Address*) indiquent les adresses IP des serveurs racines. Il peut y en avoir autant que nécessaire, afin d'assurer la redondance et le partage de charge.

Si vous voulez construire un DNS intranet directement rattaché à l'Internet, il faut y indiquer les serveurs racines officiels. (Vous pouvez vous procurer la dernière version sur le site ftp://ftp.rs.internic.net/domain/named.root.) Vous obtenez alors le fichier cache suivant (extrait) :

```
;
; formerly NS.INTERNIC.NET
;
.                       3600000  IN  NS    A.ROOT-SERVERS.NET.
A.ROOT-SERVERS.NET.     3600000      A     198.41.0.4
;
; formerly NS1.ISI.EDU
;
.                       3600000      NS    B.ROOT-SERVERS.NET.
B.ROOT-SERVERS.NET.     3600000      A     128.9.0.107
.........
```

Configurer un serveur primaire

Dans le DNS, il est nécessaire de configurer un serveur primaire unique pour chaque zone. Un serveur primaire pour une zone est primaire pour le domaine situé en tête de cette zone, ainsi que pour tous les sous-domaines.



Cela a pour effet de produire le fichier C:\Winnt\System32\Dns\info.dns :

```
;
;  Database file info.dns for info zone.
;      Zone version:  1
;

@                       IN   SOA nt001.info. administrateur.info. (
                                 1            ; serial number
                                 3600         ; refresh
                                 600          ; retry
                                 86400        ; expire
                                 3600       ) ; minimum TTL

@                            NS  nt001
nt001                        A   10.0.0.100
nt004.toulouse               A   10.4.0.165
```

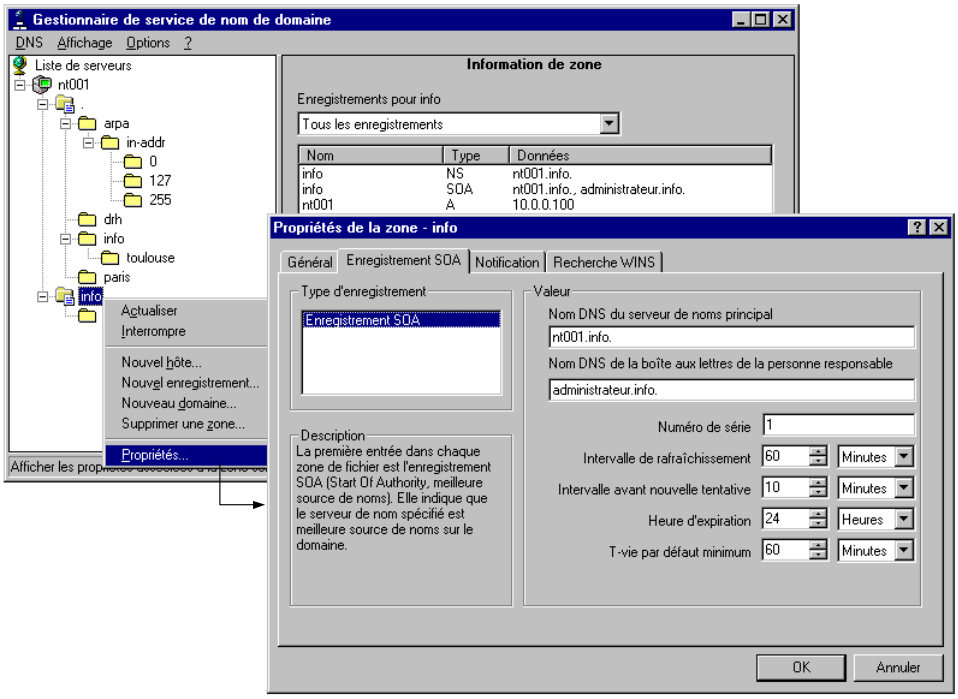
Description de la zone info.

Les parenthèses doivent être utilisées si les paramètres tiennent sur plusieurs lignes.

Référence relative à la zone en cours, ici “ info. ”.

Serveur primaire pour la zone en cours.

Que ce soit sous NT ou sous Unix, il est possible d'éditer manuellement ce fichier pour modifier les paramètres de la zone. Avec le gestionnaire DNS de Windows, on obtient l'écran suivant.



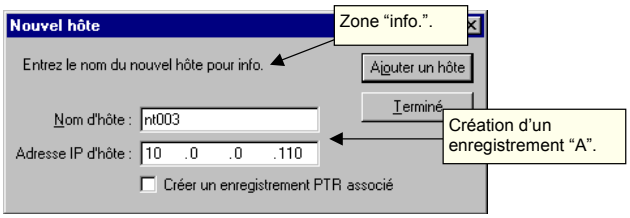
Les paramètres de la zone (SOA, *Start Of Authority*) sont définis au niveau du serveur primaire pour être ensuite utilisés par les serveurs secondaires. Le champ « Temps par défaut minimum » pourra également être utilisé par les resolvers pour déterminer la durée de conservation de l'enregistrement dans leur cache.

Si vous voulez limiter les échanges entre serveurs, mieux vaux positionner le paramètre « refresh » à une valeur assez élevée, généralement 12 à 24 heures, sauf durant les périodes de changement. En prévision de telles périodes, vous pouvez réduire cette valeur à 1 heure (voire moins selon vos besoins).

Vous pouvez également activer un mécanisme de notification automatique qui permet au serveur secondaire d'être averti immédiatement de tout changement. À réception de ce signal, le serveur secondaire déclenchera alors un transfert de zone. Il suffit pour cela d'ajouter, dans l'onglet « Notification », les adresses IP des serveurs secondaires.

Activer la résolution de nom

Généralement, l'administrateur commence par remplir la base de données d'enregistrements « **A** ». Ce type d'enregistrement permet aux utilisateurs d'obtenir l'adresse IP correspondant au nom, qui leur est plus familier. C'est toute l'utilité du DNS.



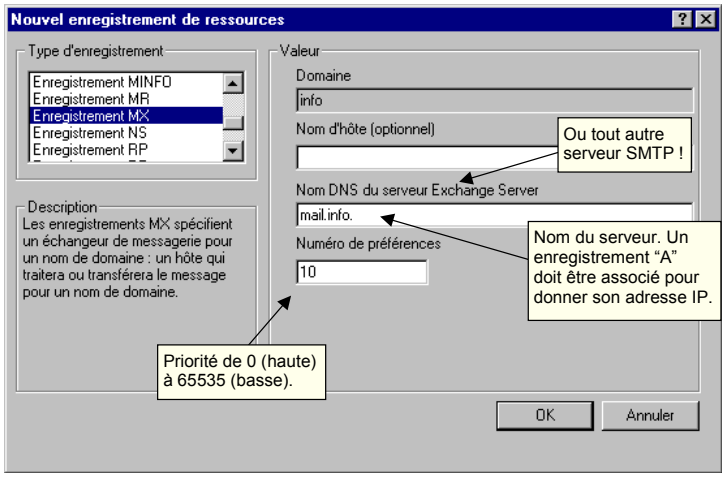
Cela a pour effet de créer l'enregistrement suivant :

```
nt003  IN     A      10.0.0.110
```

Désormais, le serveur DNS enverra au client l'adresse IP indiquée ci-dessus, à toute requête concernant ce nom.

Activer le routage de la messagerie

Un autre grand utilisateur du DNS est la messagerie SMTP. Chaque MTA (*Message Transfer Agent*) s'appuie en effet sur les enregistrements « **MX** » pour router les messages vers le prochain MTA.



En définitive, deux enregistrements doivent être créés :

```
info.  IN     MX     10     mail.info.
@      IN     MX     10     mail
mail   IN     A      10.0.0.100
```

Ces deux notations sont équivalentes.

Ainsi, tous les messages à destination de xxx@info seront routés vers le serveur « mail » situé dans le domaine « info ».

Avec le mécanisme des priorités, il est possible de définir des serveurs de secours au cas où le MTA principal serait surchargé ou en panne. Il suffit pour cela de créer un autre enregistrement « MX » de priorité plus basse que celui créé précédemment :

```
info.  IN    MX    20    mail2.info.
mail2  IN    A     10.0.0.101
```

Enfin, il est également possible de définir un serveur « poubelle » recueillant tous les messages à destination de domaines inconnus :

```
*      IN    MX    100   mail9
```

Du bon usage des alias

L'utilisation d'un alias permet de dissocier la fonction (Web, messagerie) de la nature des serveurs (nt001, unix002) ou de leur localisation dans un domaine ou un autre :

```
www    IN    CNAME  nt003.info.
pluton IN    CNAME  ux001.paris.compta.
```

Grâce à l'utilisation des alias, le changement des noms pluton et www suite à un déménagement sera plus facile à opérer que celui de nt003 qui est référencé par des enregistrements « A ».

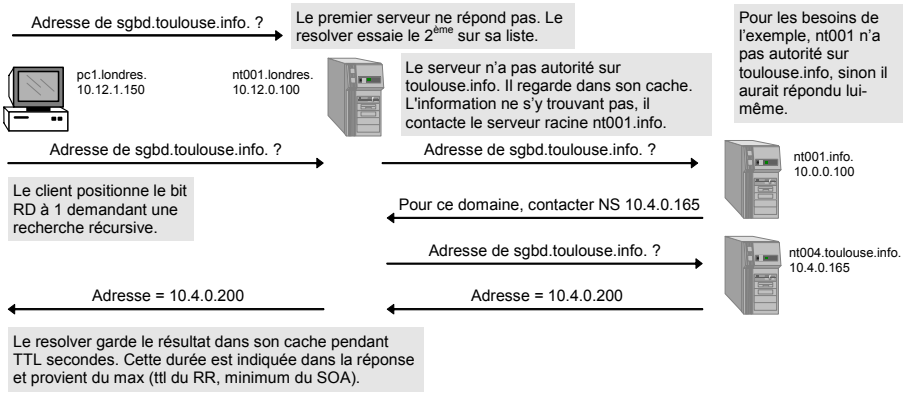
Par exemple, on doit déplacer l'application intranet sur une nouvelle machine plus puissante. Il suffit de modifier le CNAME comme suit :

```
www    IN    CNAME  nt004.toulouse.info.
```

Configurer un serveur racine

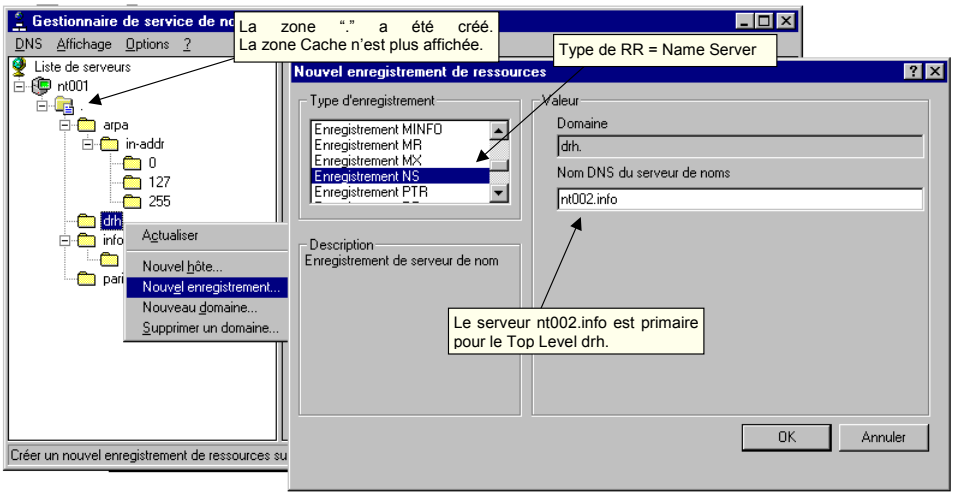
Les serveurs racines ont tout simplement autorité pour la zone « **.** », c'est-à-dire sur l'ensemble de l'arbre DNS. Leur rôle est d'indiquer aux serveurs ne pouvant répondre aux requêtes de leurs clients, l'adresse du serveur pouvant les aider.

Figure 7-3. *Recherche DNS récursive.*



La recherche peut également être effectuée de manière itérative. Dans ce cas, le serveur DNS indique au client le nom du serveur DNS qui peut l'aider, et il revient au client d'effectuer lui-même la recherche.

Dans notre cas, nous commençons par déclarer un serveur primaire pour la zone « **.** » (ne pas choisir le nom de fichier « . », mais, par exemple, le nom « racine.dns »). Puis, nous déclarons les serveurs primaires des domaines Top Level.



Cette action produit le fichier suivant :

```
.       IN      SOA     nt001.info.  jlm.nt001.info. (
                2       ; serial : 2ème version
                1800    ; refresh : toutes les 30 min.
                300     ; retry : toutes les 5 minutes
                604800  ; expire : 48 heures après
                86400)  ; minimum : 24h dans le cache
                NS      nt001.info
                NS      nt004.toulouse.info

nt001.info              A       10.0.0.100
nt004.toulouse.info     A       10.4.0.165

paris                   NS      nt001.info
info                    NS      nt001.info
drh                     NS      nt002.info

nt002.info.             A       10.0.0.124
```

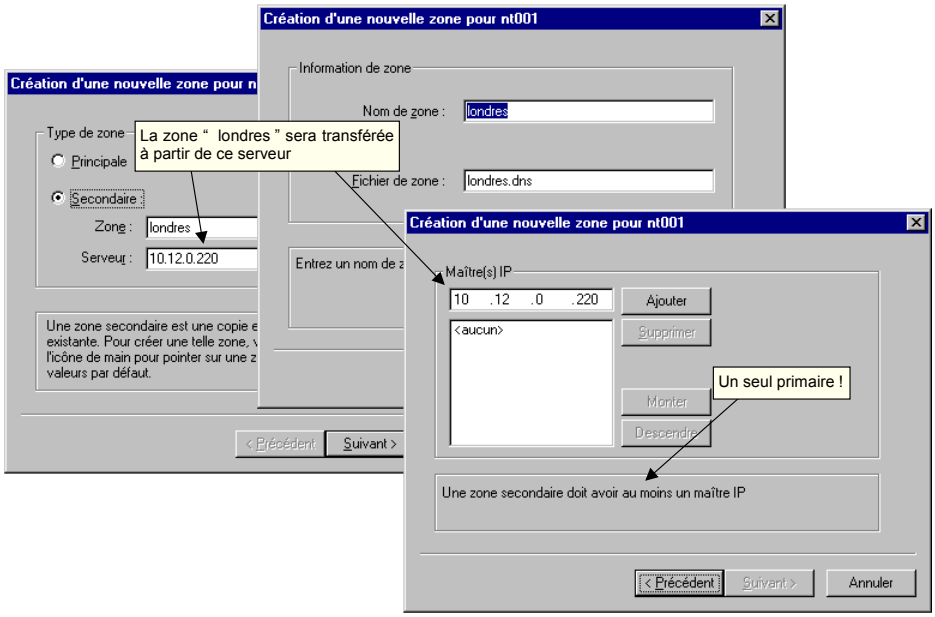
Liste des serveurs ayant autorité sur la racine.

Liste des serveurs ayant autorité sur les Top Level paris, info et drh.

Configurer un serveur secondaire

Un serveur secondaire pour une zone est un serveur qui a autorité sur ladite zone, mais qui ne peut pas modifier la base de données correspondante. Pour cela, il demande auprès du serveur primaire **le transfert de la zone**, ce qui lui permet ensuite de répondre aux requêtes de la même manière qu'un serveur primaire.

Sur ce type de serveur, une zone est déclarée secondaire en indiquant l'adresse IP du serveur primaire à partir duquel transférer la zone.



Il est à noter que le volume de données représenté par un transfert de zone est peu important, d'autant plus si les serveurs supportent une mise à jour incrémentale (RFC 1995).

Configurer un serveur cache

Tous les serveurs DNS conservent en mémoire cache les réponses des requêtes précédentes. Le serveur cache joue le même rôle, mais présente la particularité de n'avoir aucune autorité sur une quelconque zone : le serveur n'est ni primaire ni secondaire.

La configuration d'un serveur cache se résume donc à renseigner uniquement le fichier « cache.dns » qui liste les serveurs racines.

La durée de validité des enregistrements dans le cache est déterminée par le maximum des deux valeurs suivantes : le paramètre « *minimum TTL* » du SOA et le TTL de chaque enregistrement si celui-ci est spécifié.

Déléguer l'autorité à un autre serveur

Pour des questions d'organisation, il est maintenant opportun de déléguer la gestion de la zone « toulouse.info » aux exploitants de ce site. Pour cela, la procédure est la suivante :

1. Configurer le serveur de Toulouse en secondaire de « toulouse.info », afin de transférer cette zone.
2. Configurer ensuite le serveur de Paris en secondaire pour la zone « toulouse.info ».
3. Basculer enfin le serveur de Toulouse en primaire pour la zone « toulouse.info ».
4. Modifier les enregistrements NS sur les différents serveurs concernés, de manière à pointer sur le nouveau serveur primaire.
5. Si l'ancien serveur primaire ne doit pas faire office de serveur secondaire du domaine « toulouse.info », supprimer toute référence à cette zone.

Sous le gestionnaire DNS de Windows, il faut, en plus, créer un domaine « toulouse.info », de manière à extraire les données du fichier « info.dns » dans un autre fichier, « toulouse.info.dns », par exemple.

Dans le serveur primaire de « info », la zone déléguée apparaîtra alors comme suit :

```
;
;  Delegated sub-zone:  toulouse.info.
;
toulouse              NS    nt001
;  End delegation
```

Les domaines de résolution inverse

Un domaine spécifique, appelé domaine de résolution inverse et noté **in-addr.arpa** pour IPv4 et **ip6.int** pour IPv6, est utilisé pour trouver un nom à partir d'une adresse IP. Un domaine correspondant au réseau 10.4.0.0 sera ainsi noté 4.10.in-addr.arpa. (notation inverse de l'adresse IP) et l'adresse 2001:xxx:C7D9 sera notée 9.D.7.C.xxx.1.0.0.2.ip6.int.

Les domaines de résolution inverse se manipulent de manière identique aux domaines de noms. Chacun des 4 nombres de l'adresse IP est considéré comme étant un domaine qui peut être délégué. Un serveur peut ainsi être primaire pour le domaine 10.in-addr.arpa, et déléguer son autorité pour 15.10.in-addr.arpa.

La résolution inverse permet :

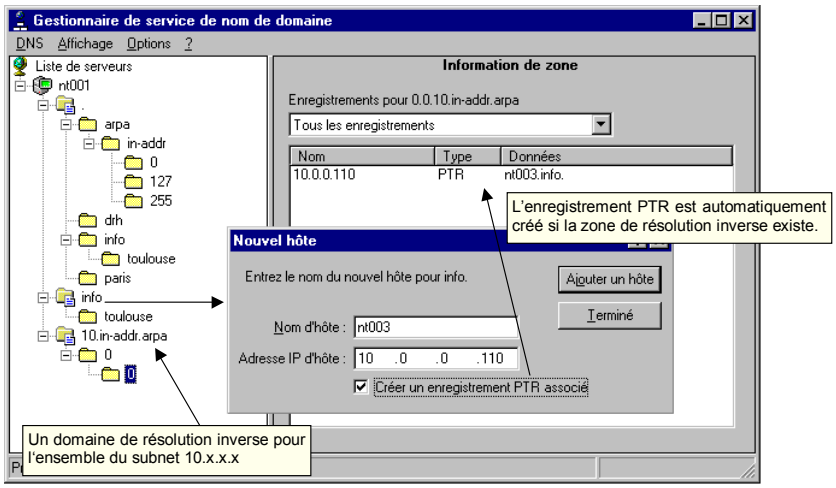
- De découvrir les routeurs d'un sous-réseau IP. Le client reçoit le nom des routeurs et peut ensuite lancer une requête pour en connaître les adresses IP. Pour des questions de sécurité, il est donc déconseillé d'utiliser cette facilité.
- À un serveur de s'assurer qu'un client qui se présente avec une adresse IP appartient bien à un domaine autorisé et/ou est bien celui qu'il prétend être. Certains serveurs FTP de l'Internet nécessitent que les clients soient référencés dans la base DNS sous forme d'enregistrement PTR.

- De référencer les noms de réseaux permettant à une station d'administration d'afficher les noms au lieu des adresses IP.
- Au programme **nslookup** (voir plus loin) de fonctionner correctement. En effet, le serveur à partir duquel la commande est lancée doit être référencé.

Une arborescence similaire doit donc être créée. Et, là encore, le DNS permet toutes sortes de fantaisies.

Par exemple, il n'y a pas obligatoirement de correspondance entre domaines de noms et domaines de résolution inverse. Les serveurs primaires pour un réseau peuvent ainsi ne pas l'être pour les domaines qui le contiennent. La création automatique de PTR est alors rendue difficile.

Il est donc conseillé de limiter le nombre de ce type de zones à une par subnet IP principal, le réseau 10 dans notre cas.



Le fichier d'initialisation

Le serveur DNS de Microsoft peut fonctionner sans l'interface graphique, à la manière des premiers serveurs DNS sous Unix. Le système fonctionne alors uniquement à partir des fichiers ASCII situés dans le répertoire \Winnt\System32\Dns.

Pour passer en mode texte, il faut supprimer la variable *EnableRegistryBoot* de la base des registres Windows au niveau de la clé « HKEY_LOCAL_MACHINE\System\ CurrentControlSet\Services\Dns\ Parameters ».

Un nouveau fichier doit alors être renseigné : il s'agit du fichier d'initialisation du DNS tel qu'utilisé sous Unix, appelé **boot** :

```
;
; Fichier Boot de nt001.info
;
cache        .              cache.dns
primary      info           info.dns
secondary    toulouse.info 10.4.0.165      toulouse.info.dns
primary      10.in-addr.arpa        10.in-addr.arpa.dns
```

Nom du fichier cache indiquant les serveurs racines.

Adresse IP du serveur primaire

Précise les zones et noms des fichiers de chaque SOA pour lequel le serveur est primaire ou secondaire.

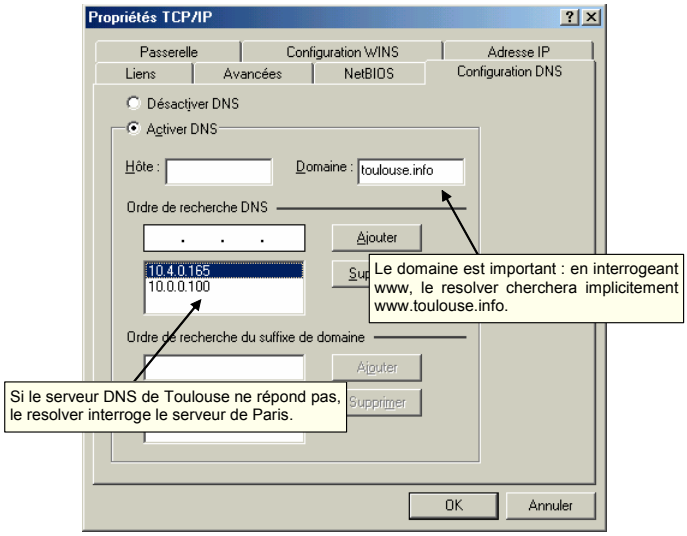
Configurer les clients DNS

Dans la terminologie DNS, un client est appelé **resolver**. Ce petit morceau de programme est directement sollicité par toutes les applications fonctionnant en réseau, telles que le navigateur Web. Il se contente d'interroger des serveurs DNS.

Sa configuration consiste simplement à indiquer le domaine dans lequel il se trouve, ainsi que la liste des adresses IP des serveurs DNS susceptibles de répondre à ses requêtes.

Une fois de plus, le DNS offre beaucoup de souplesse :

- Le client interroge le premier serveur de la liste puis, s'il ne répond pas, le second et ainsi de suite.
- Les serveurs peuvent être de type cache, primaire ou secondaire.
- Le client peut être situé dans un domaine différent des serveurs qu'il interroge.



L'utilisateur peut formuler une demande sur un **nom complet**, tel que « www.info. » (un nom terminé par un point). Aucun suffixe ne sera alors ajouté.

S'il lance une recherche sur un **nom relatif**, par exemple www (un nom qui n'est pas terminé par un point), le resolver le recherchera par défaut dans le même domaine que celui configuré. S'il reçoit une réponse négative, il ajoutera le suffixe « .info », puis un autre suffixe en fonction de la configuration.

Vérifier le fonctionnement du DNS

L'utilitaire de base pour tout administrateur DNS est le **nslookup** qui s'exécute à partir d'une fenêtre DOS de Windows NT :

Se positionne en mode debug afin d'afficher le détail des échanges.

Une recherche récursive est demandée.

Nom et adresse du serveur DNS de rattachement.

```
>set d2
>set recurse
>www.3com.com.
> Serveur:  nt001.info.
Address:  10.0.0.100

------------
SendRequest(), len 30

    HEADER:
opcode = QUERY, id = 5, rcode = NOERROR
header flags:  query, want recursion
questions = 1,  answers = 0,  authority records = 0,  additional = 0

    QUESTIONS:
       www.3com.com, type = A, class = IN

------------
------------

Got answer (301 bytes):

    HEADER:
opcode = QUERY, id = 5, rcode = NOERROR
header flags:  response, want recursion, recursion avail.
questions = 1,  answers = 1,  authority records = 6,  additional = 6
```

On demande l'adresse IP du serveur www situé dans le domaine 3com.com.

La question porte sur un enregistrement " A ".

Dans sa réponse, le serveur DNS indique qu'il supporte la recherche récursive.

Nombre de RR par section

```
    QUESTIONS:
       www.3com.com, type = A, class = IN
    ANSWERS:
    ->  www.3com.com
       type = A, class = IN, dlen = 4
       ttl = 3645 (1 hour 45 secs)
    AUTHORITY RECORDS:
    ->  3COM.COM
       type = NS, class = IN, dlen = 9
       nameserver = FOUR11.3COM.COM
       ttl = 38819 (10 hours 46 mins 59 secs)
...... suivent 5 autres RR ......
    ADDITIONAL RECORDS:
    ->  FOUR11.3COM.COM
       type = A, class = IN, dlen = 4
       internet address = 129.213.128.98
       ttl = 92825 (1 day 1 hour 47 mins 5 secs)
...... suivent 5 autres RR ......
```

La réponse est structurée en 4 sections : question, réponse, enregistrements NS et les autres enregistrements.

Par défaut, l'enregistrement « A » est demandé, mais il est possible de demander d'autres types d'enregistrements, tels que le « SOA » :

```
>set d2
>set recurse
>set querytype=soa
>www.3com.com.
> Serveur:  ns1.club-internet.fr
Address:  194.117.200.10


------------
SendRequest(), len 26
    HEADER:
opcode = QUERY, id = 5, rcode = NOERROR
header flags:  query, want recursion
questions = 1,  answers = 0,  authority records = 0,  additional = 0

    QUESTIONS:
       3com.com, type = SOA, class = IN
------------
------------
```

```
Got answer (320 bytes):

    HEADER:
opcode = QUERY, id = 5, rcode = NOERROR
header flags:  response, want recursion, recursion avail.
questions = 1,  answers = 1,  authority records = 6,  additional = 6

    QUESTIONS:
       3com.com, type = SOA, class = IN
    ANSWERS:
    ->  3com.com
       type = SOA, class = IN, dlen = 42
       ttl = 68436 (19 hours 36 secs)
       primary name server = four11.3com.com
       responsible mail addr = hostmaster.3com.com
       serial  = 1998090100
       refresh = 10800 (3 hours)
       retry   = 3600 (1 hour)
       expire  = 604800 (7 days)
       default TTL = 10800 (3 hours)
    AUTHORITY RECORDS:
    ->  3com.com
       type = NS, class = IN, dlen = 2
       nameserver = four11.3com.com
       ttl = 38922 (10 hours 48 mins 42 secs)
...... etc. ......
```

On retrouve les paramètres du SOA tels que l'administrateur de 3com les a définis.

8

En route vers la sixième dimension d'IP

À peine digérés les protocoles du monde IP, voilà qu'un nouveau apparaît. Fini donc IP, dont on apprend qu'il était le quatrième du nom, et bienvenue dans le monde IPv6. Le mystère demeure quant à ce qui est advenu des autres versions.

Avec IPv6, il faut presque tout réapprendre, à commencer par le format des adresses. Heureusement, la plupart des protocoles existants sont soit adaptés (BGP, DNS, RSVP…) soit reprennent les principes de leur équivalent en IPv4 (ICMP, OSPF…). Nous pouvons donc réutiliser la plupart des concepts que nous avons appris.

Dans ce chapitre, vous découvrirez ainsi :

- l'adressage ipv6 ;
- les principes de fonctionnement du protocole ;
- les moyens de migrer de IPv4 vers IPv6.

Genèse d'IPv6

Les besoins

Début 2002, il ne restait que 36 % des 4,3 milliards d'adresses publiques IPv4 possibles. En particulier, l'Asie, zone économique à forte croissance, ne disposait que de 20 % de l'espace d'adressage.

En outre, de nouvelles applications d'Internet doivent voir le jour, comme les PDA (*Personal Digital Assistants*), les téléphones portables de troisième génération (les fameux UMTS) ou encore la domotique (eh oui, certains réfrigérateurs sont déjà connectés à Internet !) ; même la console de jeu X-Box de Microsoft supporte déjà IPv6.

Besoins réels ou non, applications du XXIe siècle ou de science-fiction, la pression sur les adresses se fait de plus en plus forte, ne serait-ce que par le biais des connexions à Internet que nous connaissons déjà aujourd'hui. Ainsi en Asie, la Chine migre vers IPv6 et le Japon investit massivement dans un programme de migration.

Et même si la pression vient du marché grand public, des constructeurs tels que Microsoft et Cisco, en quête de nouveaux débouchés dans les entreprises, souhaitent également accélérer le mouvement.

Les solutions

Face à la pénurie d'adresses qui se profilait en 1990, deux types de solutions ont vu le jour, l'une consistant à ajouter des fonctions autour d'IPv4 (avec la translation d'adresses notamment) et l'autre consistant à développer un nouveau protocole, appelé IPv6.

Il est vrai que la translation d'adresses complexifie la gestion des réseaux et atteindra ses limites. Cependant, les réseaux IPv6 sont toujours en phase pilote, et l'on s'achemine plutôt vers une migration lente par la mise en place de passerelles aux frontières des réseaux publics et privés.

Néanmoins, les concepteurs d'IPv6 ont vu les choses en grand : un espace d'adressage immense de 3,4 10^{38} adresses et l'intégration de nombreuses fonctionnalités, qui ne sont que des rajouts dans IPv4 :

- la gestion de la QoS DiffServ ;
- la gestion des flux avec les labels ;
- la gestion de la sécurité avec l'authentification et le chiffrement des données IPsec ;
- l'affectation automatique des adresses ;
- le support de la mobilité, qui permet de conserver son adresse partout dans le monde (fonction de *roaming* utilisée par les téléphones portables).

Mais pour ne pas retomber dans les limitations d'IPv4, les concepteurs ont sans doute répété les erreurs d'ATM : un protocole à tout faire et qui remet en cause l'existant. Car, avec IPv6, il faut apprendre à manipuler un nouveau système d'adressage et un nouvel en-tête ; réécrire de nouveaux protocoles (ICMP, IS-IS, SNMP…) et adapter les existants (OSPF,

BGP…) ; puis installer de nouveaux logiciels dans les routeurs et les ordinateurs (couches TCP/IP, browser, Telnet, FTP, ping, etc.).

Si la lenteur de la migration est rattrapée par des évolutions d'IPv4, alors IPv6 suivra le chemin d'ATM, c'est-à-dire qu'il restera cantonné dans les réseaux des opérateurs.

Mais enfin, tôt ou tard, de près ou de loin, vous serez confronté à IPv6 ; donc autant s'informer dès maintenant.

Comparaison avec IPv4

Afin de donner quelques points de repère à ceux qui connaissent désormais IPv4, voici quelques points de comparaison.

Tout d'abord les adresses : 128 bits au lieu de 32 et plusieurs formats à la place de la notion de classe. On retrouve certes une partie réseau et une partie station, fixée à 64 bits chacune, mais le découpage et le format de chaque partie sont différents. La première utilise des bits variables pour désigner les subnets, permettant ainsi d'agréger les routes selon le principe du CIDR (*Classeless Inter-Domain Routing*), alors que la partie station correspond à l'adresse Ethernet MAC au format 64 bits.

Le système d'adressage est donc plus compliqué, mais les adresses peuvent être configurées automatiquement avec ou sans DHCP, ce qui simplifie grandement la tâche des administrateurs. Toute la complexité de l'adressage est, en effet, reportée sur les routeurs.

Ceux qui connaissent le protocole IPX et Novell ne seront pas dépaysés.

Ensuite, une série de procédures IPv4 est remplacée. Ainsi :

- le protocole ARP est remplacé par une procédure simple reposant sur deux messages ICMPv6 « *Neighbor Solicitation* » et « *Neighbor Advertisement* » ;
- la découverte des routeurs, optionnelle avec IPv4 et désormais obligatoire, repose sur deux messages ICMPv6 « *Router Solicitation* » et « *Router Advertisement* » ;
- les adresses de broadcasts sont remplacées par des adresses multicasts permettant de désigner tous les nœuds ou tous les routeurs ;
- le protocole IGMP est remplacé par une procédure similaire, la MLD (*Multicast Listener Discovery*) reposant sur trois messages ICMPv6.

Enfin, les couches TCP et UDP doivent être mises à jour car le contrôle d'erreur porte sur une partie de l'en-tête IP qui n'est pas la même entre les deux versions.

Remarque sur la terminologie

Dans le jargon d'IP, un **nœud** désigne une machine TCP/IP connectée au réseau. Parmi celles-ci, on distingue les **routeurs** et les **stations** (*hosts*). Un **lien** est le réseau sur lequel est connecté le nœud, soit une liaison point à point (lien série PPP, PVC ATM ou Frame Relay), soit un réseau local tel qu'Ethernet. Dans ce dernier cas, le lien désigne le domaine de broadcast MAC délimité par des routeurs.

L'ADRESSAGE IPv6 (RFC 2373, 2374, 2450)

Une adresse IPv6 est codée sur 128 bits et s'écrit sous la forme de 8 chiffres hexadécimaux représentant chacun 16 bits séparés par " : ", comme par exemple **2001:0000:0000:A7CF:02D0:B7FF:FE26:CB2A**.

Plusieurs zéros consécutifs peuvent être abrégés par deux signes " :: " consécutifs, et ce, une seule fois dans l'adresse, comme par exemple, **2001::A7CF:02D0:B7FF:FE26:CB2A**.

Pour les adresses compatibles IPv4, les 32 derniers bits peuvent conserver l'écriture standard, comme par exemple, **0:0:0:0:0:0:176.16.10.3** qui s'écrit également **::176.16.10.3**.

Le masque indique le nombre de bits codant la partie réseau de l'adresse (subnet), les autres bits codant un éventuel sous-réseau et la partie station (*host*), comme par exemple : **2001:0660:0100:0:800:1:2:3/41**, équivalent à 2001:660:100:0:800::/41, qui précise que les 41 premiers bits codent le réseau. Le masque ne peut jamais dépasser 64 bits, les 64 derniers bits étant réservés à l'identifiant d'interface.

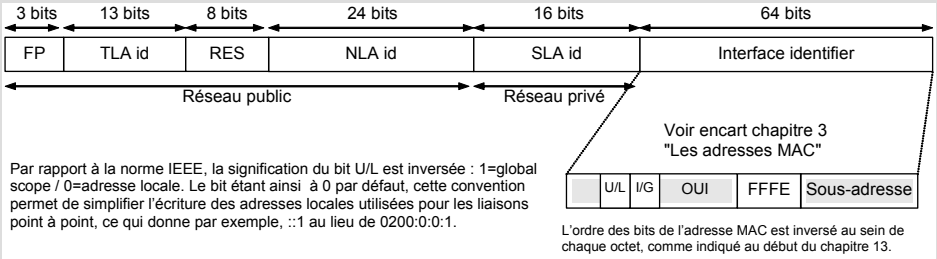
Des **formats de préfixe** (i.e. les premiers bits de l'adresse) spécifient un format d'adresse.

Les adresses commençant par…	sont affectées aux…
0000 001 (préfixe 02/7)	Adresses NSAP
0000 010 (préfixe 04/7)	Adresses IPX
001 (préfixe 2/3)	Adresses unicasts globales (agrégeables selon les principes CIDR)
1111 1110 10 (préfixe FE8/64)	Adresses unicasts locales de lien (non routables en dehors du lien)
1111 1110 11 (préfixe FEC/48)	Adresses unicasts locales de site (non routables en dehors du site)
1111 1111 (préfixe FF/16)	Adresses multicasts
::x.x.x.x (préfixe ::/96)	Nœuds IPv4 d'adresse x.x.x.x supportant les adresses IPv6
::FFFF:x.x.x.x (préfixe ::FFFF/96)	Représentation interne d'un nœud IPv4 d'adresse x.x.x.x

Par ailleurs, les adresses ::1 et :: sont réservées :

- 0:0:0:0:0:0:0:1 pour l'adresse de **loopback** qui désigne le nœud lui-même ;
- 0:0:0:0:0:0:0:0 pour l'adresse **non spécifiée** qui indique l'absence d'adresse sur l'interface.

La RFC 2374 structure les **adresses unicasts globales** (i.e. publiques) de façon **hiérarchique** :



Par rapport à la norme IEEE, la signification du bit U/L est inversée : 1=global scope / 0=adresse locale. Le bit étant ainsi à 0 par défaut, cette convention permet de simplifier l'écriture des adresses locales utilisées pour les liaisons point à point, ce qui donne par exemple, ::1 au lieu de 0200:0:0:1.

L'ordre des bits de l'adresse MAC est inversé au sein de chaque octet, comme indiqué au début du chapitre 13.

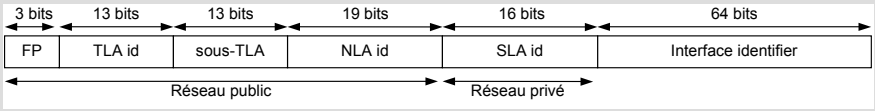
Champ	Description
FP	*Format Prefix* (=001) : identifie le type d'adresse global unicast
TLA	*Top Level Agregation identifier* : identifie un organisme gérant un réseau public tel que le IANA pour l'Internet (cf. RFC 2450).
RES	Réservé pour étendre les champs TLA et NLA (notamment aux délégations régionales de l'IANA (cf. RFC 2450).
NLA	*Next-Level Aggregation identifier* : identifie un opérateur réseau tel un ISP (cf. RFC 2450).
SLA	*Site-Level Aggregation identifier* : identifie 65 535 sous-réseaux au sein d'une organisation (équivalent aux subnets IPv4).
Interface identifier	Adresse de l'interface au format IEEE EUI-64. Reprend l'adresse MAC IEEE 802.3 en ajoutant FFFE avant la sous-adresse.

- - -

L'ADRESSAGE IPV6 (SUITE)

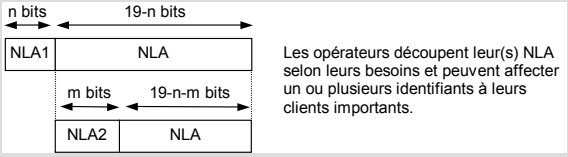
Ainsi, la RFC 2450 affecte à l'Internet le **TLA 0001** provenant du **format de préfixe 2**, ce qui correspond au réseau **2001/16**. Le IANA, qui détient ce préfixe, délègue ensuite des sous-TLA de 13 bits aux autorités régionales telles que l'APNIC (2001:0200::/29), l'ARIN (2001:0400::/29), ou le RIPE en Europe (2001:0600::/29). Ces organismes distribuent ensuite ces adresses aux opérateurs tels les ISP (*Internet Service Provider*), qui les allouent ensuite à leurs clients, ces derniers disposant alors des bits du champ SLA pour leur réseau privé (voir l'annexe « La gestion de l'Internet » à ce sujet).

Le format de l'adresse unicast globale retenu pour l'Internet est le suivant :

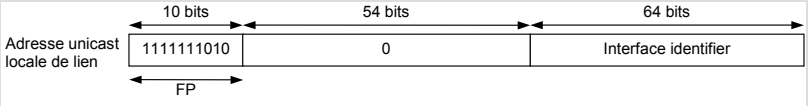


Le champ sous-TLA correspond aux 8 bits du champ RES, plus 5 bits prélevés sur le champ NLA.

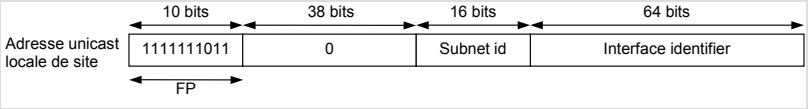
Le champ NLA peut être structuré de manière à permettre l'agrégation de routes en respectant, par exemple, la topologie du réseau de l'opérateur.



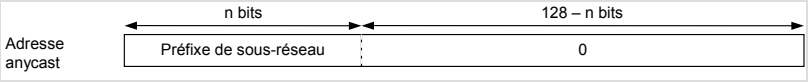
Une **adresse locale de lien** n'est pas routée au-delà du domaine de broadcast délimité par des routeurs. Ce type d'adresse est utilisé pour la communication des nœuds voisins, (découverte de routeurs, résolution d'adresse, etc.), et pour l'adressage des liaisons point à point (lien série ou tunnel).



Une **adresse locale de site** est utilisée dans un espace d'adressage privé sans possibilité de routage sur le réseau public.



Une adresse unicast peut être affectée à plusieurs interfaces et à des routeurs différents : on parle alors d'adresse ***anycast***. Sur les 128 bits, le préfixe de sous-réseau est variable, la partie interface étant égale à 0. Étant donné qu'il n'y a aucun moyen de différencier une adresse anycast d'une adresse unicast, elle doit être explicitement déclarée et traitée comme telle dans les routeurs.



L'affectation des préfixes peut être consultée sur le site www.iana.org/assignments/ipv6-address-space.

Configuration des nœuds

À l'initialisation d'un nœud, IPv6 affecte à chaque interface une **adresse locale de lien** (préfixe FE80::/64).

Si un routeur est présent sur le lien de l'interface, celle-ci se voit affectée d'une ou de plusieurs **adresses globales** (préfixe 02/3) ou **locales de site** (préfixe FEC/48), en fait une par préfixe annoncé par le ou les routeurs.

L'**identifiant d'interface** (adresse MAC au format 64 bits dans le cas d'Ethernet) est ajouté à chacun de ces préfixes.

De plus, chaque nœud IPv6 se met à l'écoute des adresses multicasts suivantes :

Préfixe	Signification
FF01::1	Toutes les interfaces du nœud
FF02::1	Tous les nœuds du lien
FF02 ::1:FFxx :xxxx	Tous les nœuds se terminant par xx:xxx (adresse de sollicitation de nœud)

Sous Windows 2000 ou XP, nous pouvons visualiser la configuration résultante.

```
ipv6 if 4
Interface 4 (site 1): Local Area Connection
uses Neighbor Discovery
uses Router Discovery
  link-layer address: 00-d0-b7-26-cb-2a
  preferred link-local fe80::02d0:b7ff:fe26:cb2a, life infinite

  multicast interface-local ff01::1, 1 refs, not reportable
  multicast link-local ff02::1, 1 refs, not reportable
  multicast link-local ff02::1:ff26:cb2a, 1 refs, last reporter

link MTU 1500 (true link MTU 1500)
current hop limit 128
reachable time 44500ms (base 30000ms)
retransmission interval 1000ms
DAD transmits 1
```

Adresse MAC

Adresse de lien

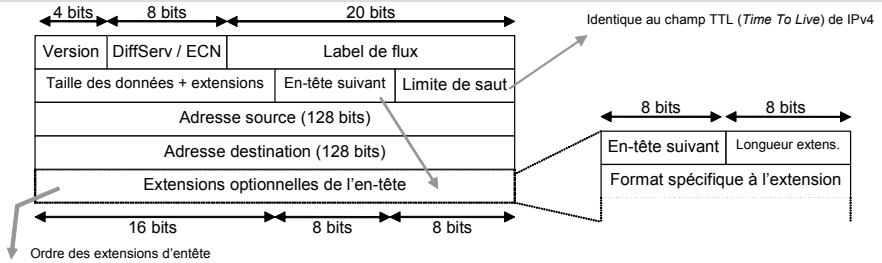
Adresse de sollicitation de nœud

Paramètres par défaut

En l'absence de routeur, une station ne peut générer qu'une adresse locale de lien.

LE POINT SUR IPV6 (RFC 2460)

Le protocole repose sur un **en-tête fixe de 40 octets**, éventuellement suivi **d'extensions** de tailles variables, comportant des options utilisées occasionnellement.



valeur	Champ « En-tête suivant »	Description
59	No next header	Aucune donnée ne suit l'en-tête IPv6
0	Hop-by-hop options	Options examinées par chaque nœud : Bourrage, Jumbogram (RFC 2675), et Alerte routeur (RFC 2711) qui indique au routeur de traiter le message multicast même s'il n'est pas membre de ce groupe.
60	Destination options	Options examinées par le destinataire et les routeurs listés dans l'option "Routing"
43	Routing	Liste des routeurs à traverser : impose un chemin de routage lâche (*loose routing*)
44	Fragment	Identifie le numéro de fragment du paquet IP, si celui-ci a été fragmenté parce qu'excédant le MTU
51	Authentication	Cf. RFC 2402
50	Encapsulation security payload	Cf. RFC 2406
60	Destination options	Options uniquement examinées par le destinataire
…	(1)	Le paquet de la couche supérieure suit immédiatement l'en-tête IPv6. Il n'y a pas d'en-tête optionnel.
Ex: 58	= ICMPv6	Cf. RFC 2463

(1) Fonctionnellement Identique au champ "Protocol" de l'en-tête IPv4. Voir www.iana.org/assignments/protocol-numbers

Excepté l'extension *hop-by-hop*, les routeurs intermédiaires n'ont pas besoin d'analyser les extensions, accélérant par là le traitement du paquet IP.

En-tête IPv6 Next header = 6	Paquet TCP	
En-tête IPv6 Next header = 43	En-tête Routing Next header = 6	Paquet TCP

La couche IP a pour rôle de **router** les paquets individuellement et de traiter les flots de paquets identifiés par des **labels de flux**. Les paquets peuvent être perdus ou arriver dans le désordre. À la différence de son prédécesseur, IPV6 ne procède à **aucun contrôle d'erreur**, cette fonction faisant double emploi avec celle réalisée par la couche liaison tel Ethernet ou Frame Relay.

Les stations IP **fragmentent** également les paquets pour qu'ils puissent tenir dans les MTU des différentes couches liaison traversées. À l'arrivée, les paquets sont **réassemblés** (à la différence d'IPv4, les routeurs ne procèdent à aucune fragmentation, car le calcul du MTU ne le nécessite pas). Le champ " **hop limit** " est décrémenté d'une unité chaque fois que la paquet traverse le routeur. Arrivé à une valeur de 0, le paquet est détruit au cas où cela serait le résultat d'une boucle dans le réseau.

Le site www.ietf.org/html.charters/ipv6-charter.html permet de suivre les évolutions du protocole dont certains aspects, tels les labels de flux, ne sont pas encore figés.

Sur les routeurs et selon le plan d'adressage retenu, il faut configurer les préfixes d'adresses globales et locales de site.

```
ipv6 unicast-routing
int e0
ipv6 address 2001:0660:0100:A7CF::/64 eui-64
```

Préfixe d'adresse unicast globale

Indique qu'il faut générer un identifiant d'interface à partir de l'adresse MAC au format 64 bits.

Notre routeur diffuse alors le préfixe (i.e. les subnets) qu'il supporte sur le réseau local de son interface Ethernet « e0 ».

En plus des adresses précédemment mentionnées, le routeur est à l'écoute des adresses anycasts et multicasts suivantes :

Préfixe	Signification
FF01::2	Toutes les interfaces du routeur
FF02::2	Tous les routeurs du lien
FF05::2	Tous les routeurs du site

```
show ipv6 int e0
Ethernet0 is up, line protocol is up
  IPv6 is enabled, link-local address is FE80::0A00:58FF:FE5A:C7D9
  Global unicast address(es):
    2001:0660:0100:A7CF:0A00:58FF:FE5A:C7D9, subnet is
                                          2001:0660:0100:A7CF/64
  Joined group address(es):
    FF02::1
    FF02::2
    FF02::1:5A:C7D9
  MTU is 1500 bytes
  ICMP error messages limited to one every 500 milliseconds
  ND reachable time is 30000 milliseconds
  ND advertised reachable time is 0 milliseconds
  ND advertised retransmit interval is 0 milliseconds
  ND router advertisements are sent every 200 seconds
  ND router advertisements live for 1800 seconds
  Hosts use stateless autoconfig for addresses.
```

Adresse unicast locale de lien

Adresse unicast globale

Groupes multicasts par défaut

Adresse de sollicitation de nœud

La station récupère le préfixe et construit l'adresse en y ajoutant son adresse MAC au format 64 bits. Nous pouvons alors la visualiser sur notre station Windows XP.

```
ipv6 if 4
Interface 4 (site 1): Local Area Connection
...
  preferred address 2001:0660:0100:a7cf:02d0:b7ff:fe26:cb2a,
                                                  infinite/infinite
...
```

Adresse unicast globale

Validité permanente

```
FP      |   TLA    | s-TLA  | NLA  | SLA | Interface = adresse MAC
unicast     IANA     Renater  IdF           *Intel:    :n° NIC
2           001    :  0660  : 0100 : A7CF: 02D0:B7FF:FE26:CB2A
```

Le RIPE disposant du sous-TLA 0600/29 a alloué 0660/35 à Renater (voir www.ripe.net/ripencc/mem-services/registration/ipv6/ipv6allocs.html) qui gère les 29 bits d'espace d'adressage restant (13 bits de NLA structurés en 63 plaques régionales de 127 campus et laboratoires et 16 bits de SLA, soit 65 535 réseaux). Par exemple, le NLA 0100/41 correspond à la région Ile-de-France et 0180/41, à Grenoble.

Le champ « Interface » a été construit à partir de l'adresse MAC 00-D0-B7-26-CB-2A (D0-B7 indiquant que le fabriquant de la carte réseau est Intel), la valeur du bit U/L (indiqué par une *) étant inversée.

La couche IP créé également une pseudo-interface de **loopback** comme, par exemple, sur notre PC :

```
ipv6 if 1
Interface 1 (site 0): Loopback Pseudo-Interface
does not use Neighbor Discovery
link-level address:
  preferred address ::1, infinite/infinite
 link MTU 1500 (true link MTU 1500)
 current hop limit 1
 reachable time 0ms (base 0ms)
 retransmission interval 0ms
 DAD transmits 0
```

Adresse réservée à l'interface de loopback.

Sa portée reste locale au nœud.

L'ADRESSAGE MULTICAST (RFC 2373 ET 2375)

Une adresse multicast, identifiée par le préfixe FF/16, permet de gérer des groupes de diffusion.

8 bits	4 bits	4 bits	80 bits	32 bits
11111111	000T	Scope	0	Identifiant de groupe

Le bit T à "1" indique que le groupe est temporaire.
Le champ "Scope" définit la portée du groupe multicast.

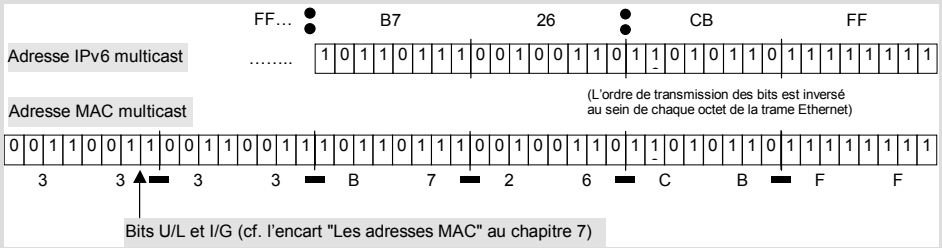
Scope	Portée du groupe	Préfixe
1	Nœud local	FF01
2	Lien local	FF02
5	Site local	FF05
8	Organisation locale	FF08
E	Portée globale	FF0E

Un certain nombre de groupes permanents est utilisé pour échanger des messages de service ICMP.

Préfixe	Groupe permanent
FF01:1	Toutes les interfaces du nœud
FF01:2	Toutes les interfaces du routeur
FF02::1	Tous les nœuds du lien
FF02::2	Tous les routeurs du lien
FF02::D	Tous les routeurs PIM du lien
FF05::2	Tous les routeurs du site
FF05::1:3	Tous les serveurs DHCP du site

Le site www.iana.org/assignments/ipv6-multicast-addresses tient à jour la liste des groupes réservés.

Sur un réseau Ethernet, les paquets multicasts IP sont envoyés dans des trames multicasts dont l'adresse MAC de destination commence par " 33-33 ", les 32 derniers bits de cette adresse provenant des 32 derniers bits de l'adresse IP.



À titre de comparaison, vous pouvez consulter le schéma au début du chapitre 15, qui montre le même principe appliqué à IPv4.

L'**adresse de sollicitation de nœud** a pour but de minimiser le nombre de nœuds devant traiter la trame multicast diffusée sur le réseau local. Son préfixe est **FF02::1:FF**xx:xxxx/104, " xx:xxx " désignant les 24 derniers bits d'une adresse IP.

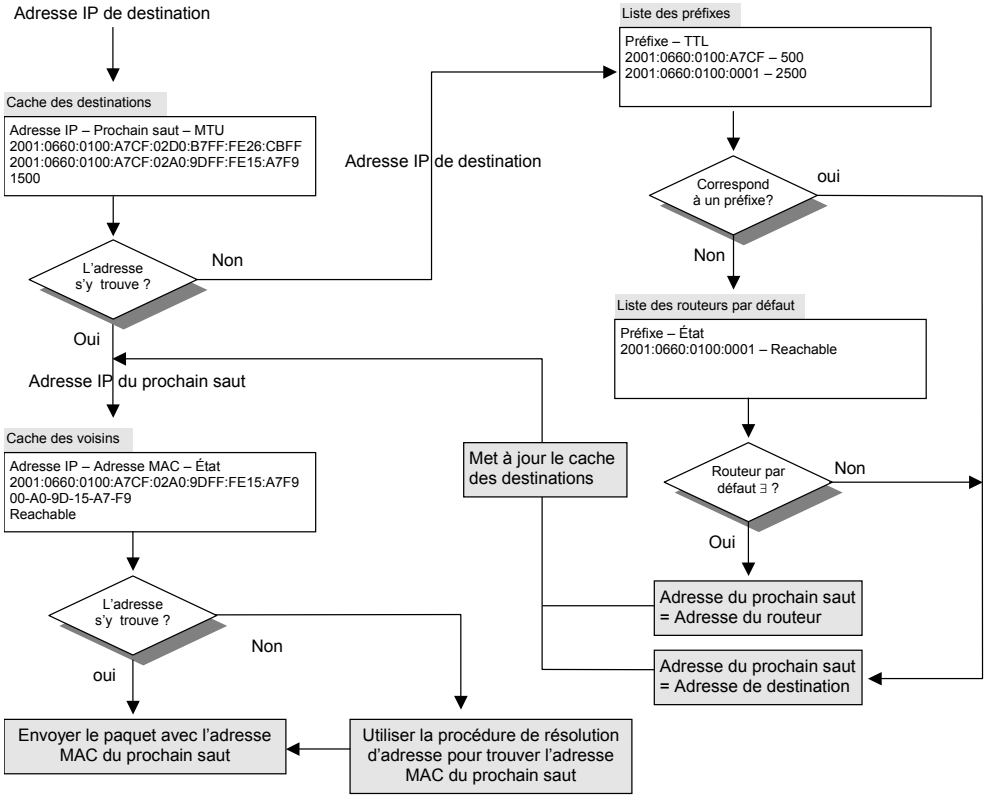
Ainsi, le nœud d'adresse de lien FE80::02D0:B7FF:FE26:CB2A écoutera également les adresses FE02::1:FF26:CB2A. Et seuls les nœuds dont l'adresse MAC se termine par 26-CB-2A prendront en compte la trame multicast 33-33-FF-26-CB-2A, les autres ne perdant pas de temps à l'analyser.

Fonctionnement standard des nœuds

Chaque nœud IPv6 conserve en mémoire les informations suivantes :

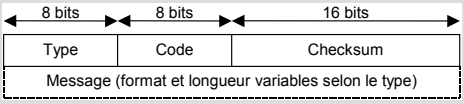
- **Le cache des voisins**, équivalent au cache ARP d'IPv4, indique la liste des adresses IPv6 avec les adresses de niveau 2 associées, par exemple les adresses MAC Ethernet.
- **Le cache des destinations** contient les adresses IP auxquelles le nœud a récemment envoyé un paquet, avec pour chacune d'entre elles, l'adresse du prochain saut ainsi que le MTU du chemin.
- **La liste des préfixes**, provenant des messages d'annonce des routeurs (messages ICMP 134), qui contient les préfixes directement joignables (i.e. situés sur le même lien).
- **La liste des routeurs** qui ont envoyé une annonce (message ICMP 134) et parmi lesquels l'un est susceptible d'être élu routeur par défaut pour ce nœud.

L'algorithme d'envoi des paquets permet de décrire le format et l'utilisation de ces tables.



LE POINT SUR ICMPv6 (RFC 2463 ET 3122)

ICMP (*Internet Control Message Protocol*) regroupe tous les messages de services nécessaires aux communications entre nœuds. ICMP correspond, en effet, à une série de procédures très diverses allant de la découverte des routeurs voisins à la notification d'erreurs.



Type	Message	Code	Description
1	Destination unreachable	0	Le routeur ne connaît pas la route
		1	La communication est interdite par un pare-feu
		2	L'adresse destination dépasse le domaine de routage de l'adresse source
		3	La destination n'est pas joignable
		4	Aucun programme ne répond sur le port TCP ou UDP
2	Packet too big	0	La taille du paquet dépasse le MTU de la couche liaison
3	Time exceeded	0	Le paquet a été reçu avec une limite de saut à 0
		1	Tous les fragments d'un paquet n'ont pas été reçus dans le temps imparti
4	Parameter problem	0	Erreur détectée dans l'en-tête ou une de ses extensions
		1	Un code d'extension d'en-tête inconnu a été trouvé
		2	Une option inconnue a été trouvé
128	Echo request	0	Demande au récepteur de renvoyer un Echo reply (un identifiant et un numéro de séquence identifient le message)
129	Echo reply	0	Réponse à un Echo request. Les données contenues dans le message Echo request doivent être reportées dans ce message.
130	Multicast listener query	0	Demande de rapport envoyée par le routeur
131	Multicast listener report	0	Rapport de participation au groupe envoyé par le membre
132	Multicast listener done	0	Le membre indique au routeur qu'il quitte le groupe
133	Router solicitation	0	Émis par une station pour découvrir les routeurs sur son lien
134	Router advertisement	0	Envoyé par un routeur périodiquement ou en réponse à un message de sollicitation pour indiquer les paramètres du lien
135	Neighbor solicitation	0	Demande de résolution de l'adresse IPv6 en adresse MAC
136	Neighbor advertisement	0	Réponse à une demande de résolution d'adresse, envoyée spontanément en cas de changement d'adresse de niveau 2
137	Redirect	0	Le routeur indique au nœud qu'un autre routeur connaît une meilleure route
141	Inverse neighbor discovery solicitation	0	Demande de résolution de l'adresse MAC en adresse IPv6
142	Inverse neighbor discovery solicitation	0	Réponse à une demande de résolution d'adresse inverse

La découverte des nœuds voisins (RFC 2461 et 3122)

Les voisins sont les nœuds situés sur le même domaine de broadcast Ethernet ou sur le même lien série (PPP, Frame Relay ou ATM).

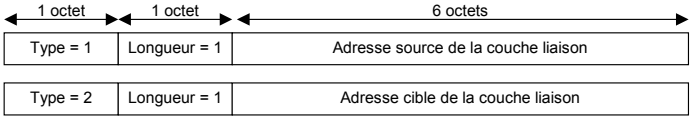
Derrière le terme ND (*Neighbor Discovery*), se cache une série de procédures qui sont :

- utilisées par les stations pour la **découverte des routeurs**, des **préfixes réseaux** et des **paramètres** utilisés (MTU, nombre maximal de sauts,…) ;
- utilisées par les routeurs pour la fonction de **redirection** vers une meilleure route ;
- utilisées par les nœuds en général pour la **résolution d'adresses**, la vérification de **l'état des voisins** et la **recherche du prochain saut** pour aller vers un réseau donné.

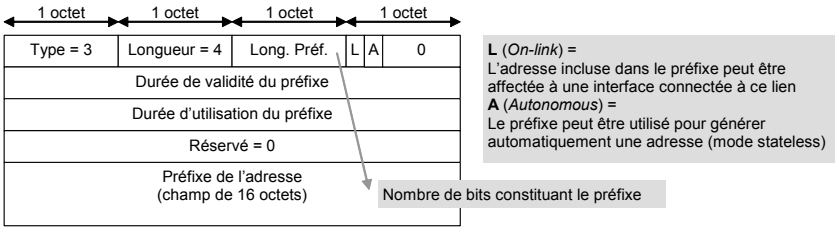
Toutes ces procédures reposent sur 5 messages ICMP, contenant un en-tête de message ND et éventuellement des options.

Les options utilisées dans les messages de découverte

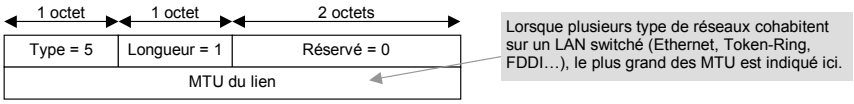
Options « Adresse de lien source » et « Adresse de lien destination » utilisées dans divers messages.



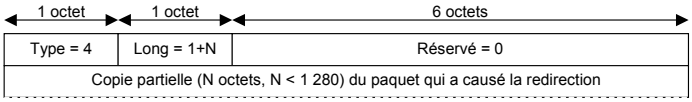
Option « Information de préfixe » utilisée dans les messages d'annonce de routeurs.



Option « MTU » utilisée dans les messages d'annonce de routeurs.



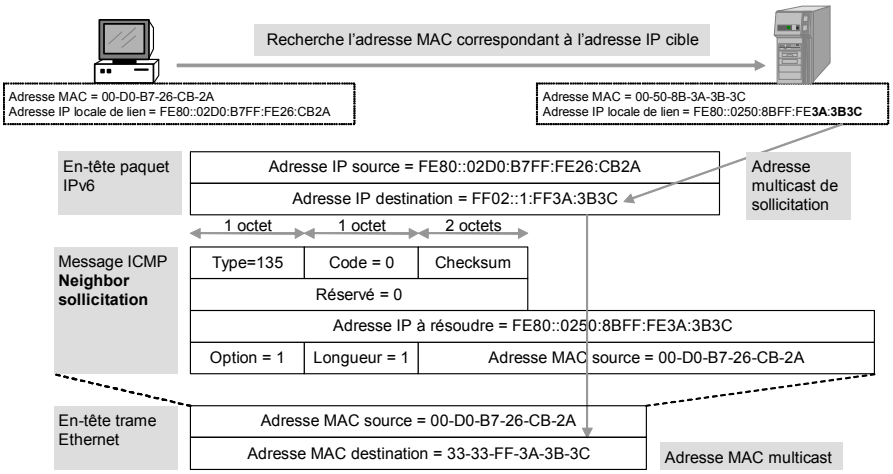
Option « Redirection » utilisée par les routeurs.



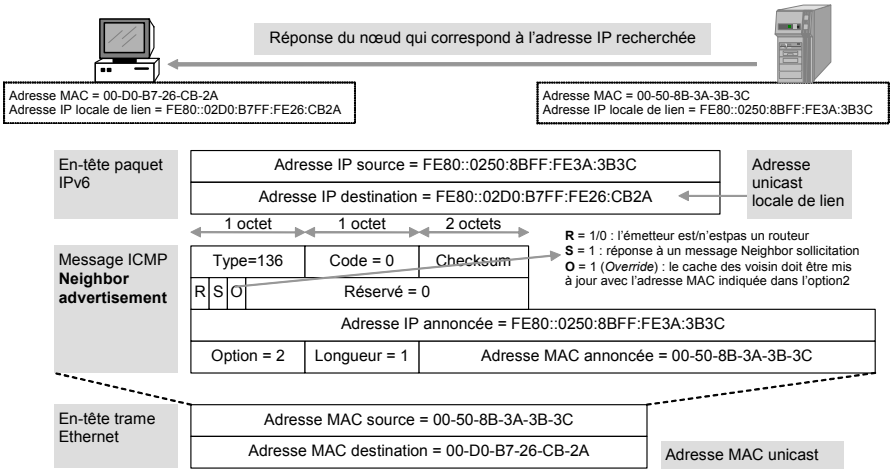
Le champ « Longueur » indique le nombre de blocs de 8 octets, données et en-tête compris.

La résolution d'adresse

La procédure est relativement simple. Un nœud désirant connaître l'adresse MAC à partir de son adresse IP, émet une requête multicast dans le groupe de sollicitation de nœuds correspondant à cette même adresse IP.



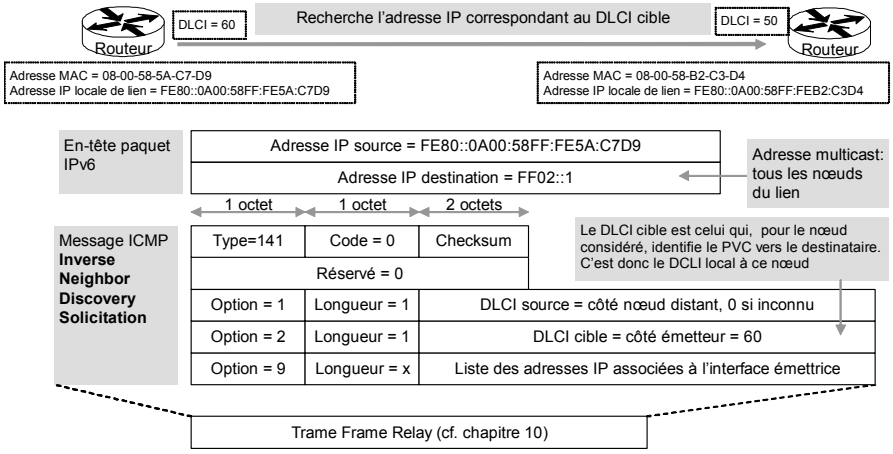
Seuls les nœuds dont l'adresse MAC se termine par les 24 derniers bits de l'adresse MAC de la trame multicast vérifient si l'adresse IP demandée correspond à la leur, et seul le nœud qui répond par l'affirmative renvoie une annonce à l'émetteur.



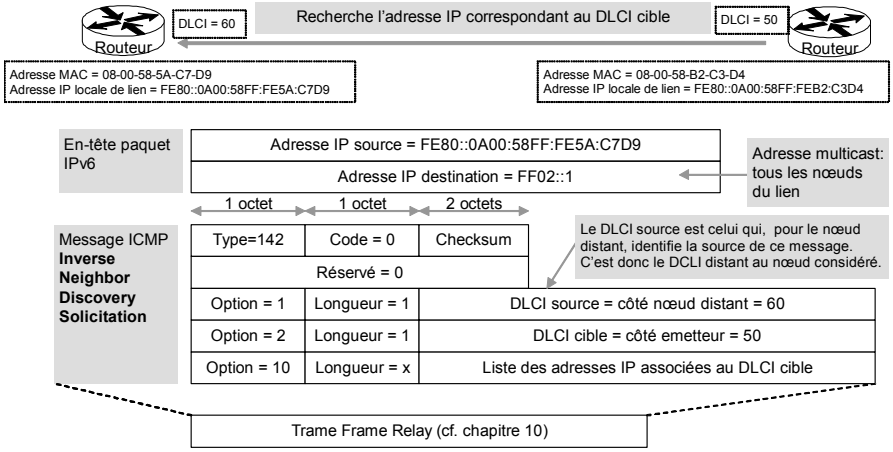
Les informations ainsi collectées permettent de mettre à jour le cache des voisins.

La résolution inverse d'adresse

Il est des cas, comme par exemple sur un lien Frame Relay, où connaissant l'adresse de niveau 2 du routeur d'en face (en l'occurrence le DLCI), il serait intéressant d'en connaître automatiquement l'adresse IP. La RFC 3122 définit la procédure dite de résolution inverse.



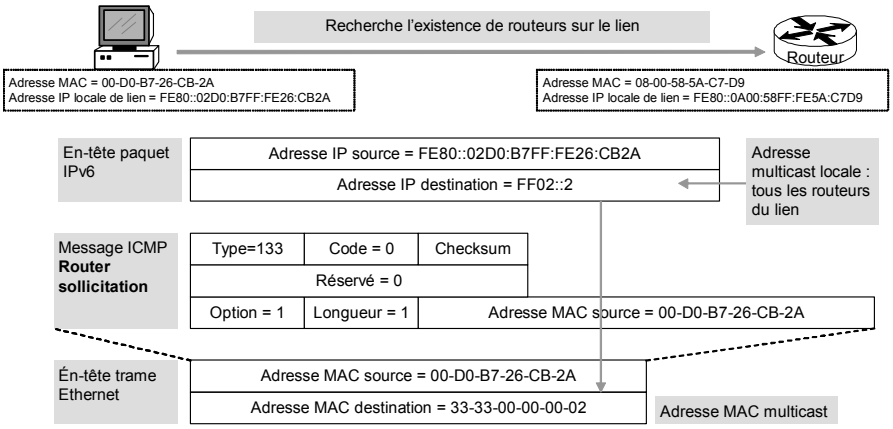
Si présente, l'option 9 indique la liste des adresses IP affectées à l'interface par laquelle la requête est envoyée.



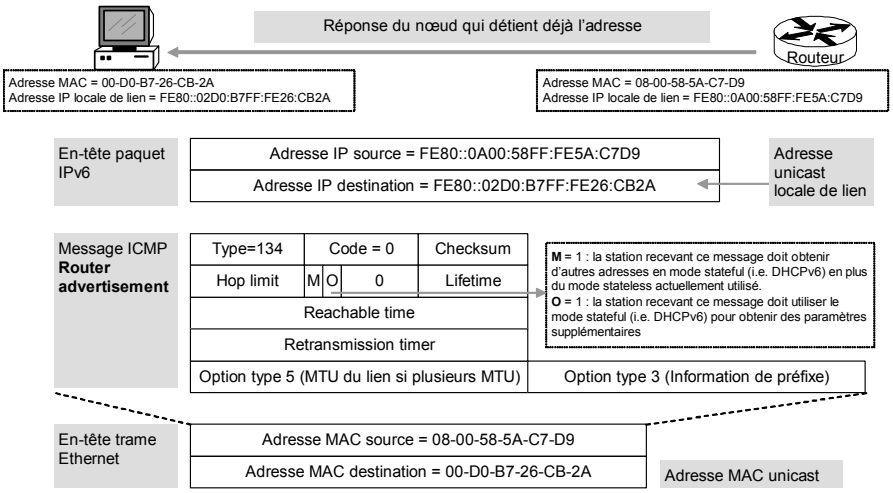
Dans le paquet de réponse, le récepteur met son DLCI dans l'option « Adresse de lien destination » qu'il obtient de la couche niveau 2 de la trame Frame Relay qu'il reçoit. Il y ajoute les adresses IP associées.

La découverte des routeurs

À son initialisation, tout nœud IPv6 recherche l'existence de routeurs sur son ou ses liens afin de déterminer s'il faut récupérer des préfixes et des paramètres.



Les routeurs envoient périodiquement des annonces et répondent également aux sollicitations comme celle montrée ci-dessus.



Dans ce message, le routeur communique à la station une série de paramètres précisés dans le tableau ci-après.

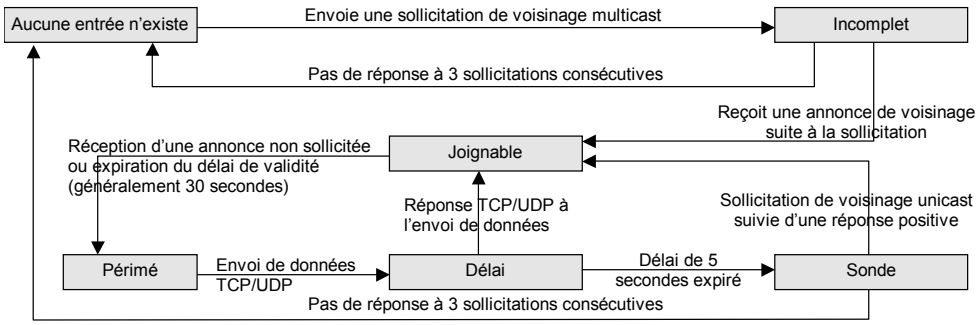
Paramètre	Signification
Hop limit	Nombre limite de sauts : valeur que la station doit utiliser dans les paquets qu'elle émet.
Lifetime	Durée, en ms, pendant laquelle le routeur peut être considéré comme routeur par défaut. Une durée de 0 ms indique que le routeur ne peut pas être considéré comme routeur par défaut.
Reachable time	Durée, en ms, pendant laquelle un nœud voisin peut être considéré comme joignable après avoir reçu une confirmation.
Retransmission timer	Délai, en ms, entre les envois de message de sollicitation de voisins.
MTU	Dans le cas de réseaux commutés avec différents supports (Ethernet v2 et 802.3, FDDI, Token-Ring…).
Information de préfixe	Préfixe annoncé par le routeur.

Sauf indication contraire, une valeur à 0 indique que le routeur ne spécifie pas de valeur.

La vérification de l'état des voisins

Une fois les voisins découverts, un nœud vérifie périodiquement leur état et en garde une trace dans le cache des voisins.

État	Signification
Incomplete	La résolution d'adresse est en cours.
Reachable	Le nœud a été récemment confirmé comme joignable.
Stale	Le nœud n'est apparemment pas joignable, mais son état ne sera pas vérifié avant que des données doivent lui être envoyées.
Delay	Le nœud n'est apparemment pas joignable alors que des données lui ont été envoyées. Attend un certain délai avant de lancer une vérification.
Probe	Le nœud n'est apparemment pas joignable, et une vérification est en cours.

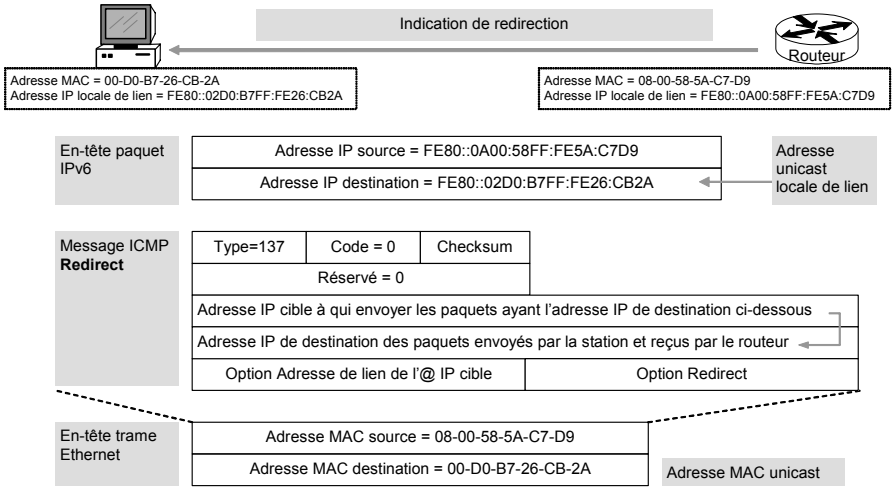


Si le voisin est le routeur par défaut et vient à ne plus être joignable, le nœud recherche un autre routeur dans sa liste de routeurs.

La redirection

Un routeur a la possibilité d'informer une station qu'il existe un routeur voisin disposant d'une meilleure route vers la destination spécifiée dans les paquets émis par la station.

De la même manière, il peut informer une station que son destinataire est directement joignable sans passer par un quelconque routeur. Ce cas de figure se produit lorsque plusieurs subnets IP cohabitent sur le même domaine de broadcast Ethernet.



À la réception de ce message, la station envoie les paquets à l'adresse MAC qui lui est indiquée.

L'auto-configuration des adresses (RFC 2462)

Lors de son initialisation, une interface IPv6 s'attribue une adresse de lien sur la base du préfixe FE80::/64 et de son adresse MAC dans le cas d'Ethernet. Des adresses supplémentaires sont créées si des routeurs lui envoient des annonces de préfixe.

Cette procédure d'affectation automatique des adresses est connue sous le nom de **stateless** par opposition au **mode stateful** qui fait intervenir un serveur DHCP.

L'affectation automatique des adresses

Le nœud fait appel à la procédure de détection des adresses dupliquées. En cas de détection positive, l'interface ne peut être activée. Dans le cas contraire, l'adresse est enregistrée.

Dans le cas d'une station, la procédure se poursuit par la recherche des routeurs voisins, et en cas de succès, la station sauvegarde les paramètres « Hop limit », « Reachable time », « Retransmission timer », et le MTU.

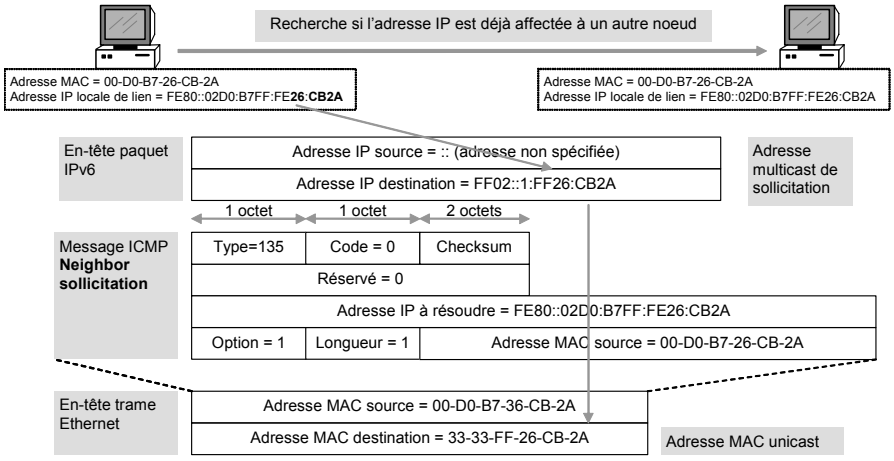
Pour chaque préfixe indiqué dans l'annonce de routeur, la procédure d'auto-configuration de la station se poursuit comme suit :

- Si l'indicateur « **On-Link** » de l'option d'information de préfixe est positionné à « 1 », le préfixe est ajouté à la liste des préfixes.
- Si l'indicateur « **Autonomous** » de l'option d'information de préfixe est positionné à « 1 », le préfixe et l'adresse MAC sont concaténés pour former une adresse, dont l'unicité est vérifiée à l'aide de la procédure de détection des adresses dupliquées.
- Si l'indicateur « **Managed address** » est positionné à « 1 » dans le message d'annonce du routeur, un protocole tel que DHCP doit être utilisé pour obtenir d'autres adresses.
- Si l'indicateur « **Other stateful** » est positionné à « 1 » dans le message d'annonce du routeur, un protocole tel que DHCP doit être utilisé pour obtenir d'autres paramètres.

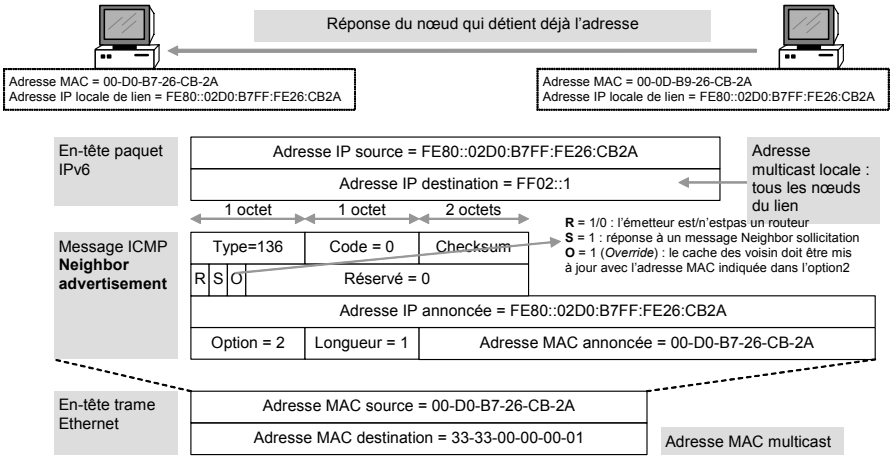
Les adresses IPv6 sont donc allouées dynamiquement, une de lien locale, et d'autres, si des routeurs voisins sont détectés.

La détection d'adresse dupliquée

Il s'agit du cas hautement improbable d'une duplication d'adresse, sans doute suite à une mauvaise configuration lorsque, par exemple, l'administrateur a voulu fixer l'adresse IP indépendamment de l'adresse MAC de la carte réseau.



Si aucune réponse n'est reçue, l'adresse est supposée être libre et est affectée à l'interface.



La découverte du MTU (RFC 1981)

Le MTU (*Maximum Transmission Unit*) est la taille maximale d'un paquet IP non fragmenté pouvant être transporté par la couche liaison. Étant donné qu'un paquet IP peut traverser des réseaux de nature très différente, le MTU résultant doit correspondre au plus petit MTU trouvé. Par exemple, le MTU d'une trame Ethernet v2 est de 1 500 octets, celui d'une trame Ethernet 802.3, de 1 492 octets et celui d'une trame Frame Relay, de 4 096 octets.

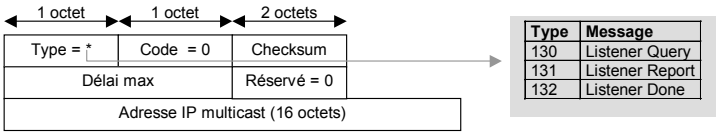
La procédure utilisée est la suivante :

1. Initialement, le MTU correspond à celui de la couche liaison de l'interface par laquelle le premier paquet est envoyé.
2. Si un routeur constate que la taille du paquet est supérieure au MTU de la couche liaison de l'interface vers laquelle il doit transmettre le paquet, il renvoie à l'émetteur un message ICMP « *Packet too big* » qui contient le MTU correspondant à la couche liaison en cause.
3. L'émetteur ajuste son MTU et renvoie le paquet en le fragmentant éventuellement. Il procède ainsi de suite, tant qu'un routeur situé sur la route d'acheminement du paquet lui envoie le même type d'erreur.
4. Toutes les cinq ou dix minutes, l'émetteur essaye d'augmenter son MTU, afin de détecter un éventuel changement dans la topologie du réseau. En cas d'échec, il reçoit les mêmes messages d'erreur.

En aucun cas, IPv6 n'accepte un MTU inférieur à 1 280 octets.

La gestion des groupes de diffusion (RFC 2710 et 2711)

La procédure MLD (*Multicast Listener Discovery*) permet d'enregistrer les stations appartenant à un groupe multicast. Elle est identique à IGMP (voir le début du chapitre 15) employé avec IPv4, mais utilise des messages ICMP équivalents.



Les messages MLD sont constitués d'un en-tête IP, d'une extension « Hop-by-hop » comportant l'option « *Router Alert* », et d'un des trois messages ICMP ci-dessus.

Les messages « *Listener Query* » sont envoyés par les routeurs pour rechercher les membres d'un groupe (recherche spécifique à une adresse multicast) ou les membres de tous les groupes (recherche générale).

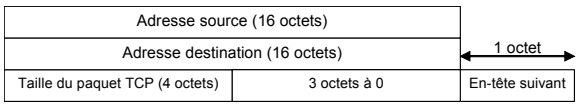
Les membres des groupes répondent à cette sollicitation ou se manifestent d'eux-mêmes en envoyant un message « *Listener Report* ». Dès que l'un d'eux quitte le groupe, il envoie un message « *Listener Done* ».

L'option « *Router Alert* », décrite par la RFC 2711, indique au routeur de traiter le message MLD, même s'il n'est pas à l'écoute de l'adresse multicast parce que n'ayant aucun membre pour ce groupe.

Migrer de la v4 à la v6

Théoriquement, les applications ne doivent pas être réécrites, car elles font référence aux couches TCP et UDP. Or, ces couches doivent être modifiées afin que le calcul du checksum porte sur une partie de l'en-tête IPv6 (voir la fin du chapitre 5).

Avec IPv6, TCP doit calculer le checksum sur les données et la partie d'en-tête suivante, appelée pseudo en-tête :



Dans le cas de Windows, Microsoft a développé une nouvelle couche TCP adaptée à IPv6 et distincte de celle reposant sur IPv4. En conséquence, les applications doivent être réécrites pour supporter la nouvelle interface Winsock conforme à la RFC 2553.

Par exemple, les primitives *gethostbyname* et *gethostbyaddr* doivent respectivement être remplacées par *getaddrinfo* et *getnameinfo*.

Pour plus d'informations, le lecteur pourra se reporter aux RFC suivants :

- RFC 2529 – « Transmission of IPv6 over IPv4 Domains without Explicit Tunnels » ;
- RFC 2893 – « Transition mechanisms » ;
- RFC 3056 – « Connection of IPv6 Domains via IPv4 Clouds (6to4) »;
- RFC 3142 – « An IPv6-to-IPv4 Transport Relay Translator ».

TROISIÈME PARTIE

Les réseaux étendus

9

Mettre en place sa première interconnexion

Comme leur nom l'indique, les réseaux locaux sont géographiquement restreints à un immeuble, voire à un campus. Votre société se développant, de nouveaux sites sont créés et les réseaux locaux se multiplient.

L'enjeu est désormais de connecter ces réseaux entre eux de sorte que tous les utilisateurs accèdent aux mêmes applications quelle que soit leur localisation.

C'est le rôle des réseaux WAN (*Wide Area Network*), c'est-à-dire des réseaux étendus. On parle également d'interconnexion de réseaux, de réseaux longue distance ou de réseaux intersite.

Dans ce chapitre, vous apprendrez :

- à choisir entre plusieurs solutions techniques et économiques ;
- à interconnecter deux sites ;
- à mettre en place une ligne spécialisée et une ligne de secours RNIS ;
- à configurer un routeur.

Le contexte

Notre réseau local est opérationnel, et les utilisateurs sont satisfaits. En plus du site parisien, il faut maintenant offrir le même type de service pour un nouveau site situé à Orléans. Qu'à cela ne tienne, il suffit d'appliquer les recettes qui ont déjà fait le succès de notre premier réseau.

Mais les utilisateurs de chaque site doivent communiquer entre eux : messagerie, transferts de fichiers et connexions aux serveurs Web sont demandés. Il faut donc interconnecter les réseaux locaux de ces deux sites. Le problème est qu'ils sont distants de plus de 100 km.

On pourrait utiliser les mêmes équipements Ethernet, des commutateurs par exemple. Mais la norme impose une longueur maximale aux liaisons Ethernet, au mieux quelques kilomètres en fibre optique. Ces contraintes proviennent de l'affaiblissement du signal d'une part, et du délai de propagation des trames, d'autre part. En effet, plus les distances sont grandes, plus le signal est affaibli et plus le délai de propagation des trames est élevé. Si ce dernier était plus élevé que celui imposé par la norme, les stations devraient attendre plus longtemps avant de pouvoir transmettre une trame, ce qui diminuerait considérablement le débit du réseau (moins de trames circuleraient en un laps de temps donné puisqu'il faudrait attendre plus longtemps avant de transmettre). À l'avenir, cette contrainte disparaîtra mais, pour le moment, nous devons encore en tenir compte.

Il faut donc employer d'autres techniques plus adaptées à ces contraintes et définir une architecture des réseaux étendus, appelés **WAN** (*Wide Area Network*) par opposition aux réseaux locaux, LAN (*Local Area Network*).

Les choix de base

Quel support de transmission ?

Le RTC (réseau téléphonique commuté) que nous avons utilisé pour nous connecter à Internet est un réseau étendu.

Quelques supports de transmission utilisés pour les réseaux étendus		Débit
RTC	Réseau téléphonique analogique utilisé pour transporter des données.	De 19,2 à 56,6 Kbit/s
RNIS (réseau numérique à intégration de service)	Réseau téléphonique numérique utilisé pour transporter des données. Très utilisé en interconnexion de LAN.	De 64 à 128 Kbit/s (plus rare : N x 64 Kbit/s)
LS (ligne spécialisée)	Liaison numérique en point à point entre deux sites. Très utilisée en interconnexion de LAN.	De 64 Kbit/s à 2 Mbit/s ou 34 Mbit/s

Quelques supports de transmission utilisés pour les réseaux étendus		Débit
DSL (*Digital Subscriber Line*)	Liaisons numériques en point à point entre deux sites. De plus en plus utilisées pour les accès à Internet.	De 64 Kbit/s à 6 Mbit/s
SDH (*Synchronous Data Hierarchy*)	Liaisons en fibre optique à haut débit utilisées par les opérateurs.	De 51 Mbit/s à plusieurs Gbit/s
ATM (*Asynchronous Transfert Mode*)	Liaisons en fibre optique à haut débit (emprunte également des supports SDH). ATM est également utilisé pour les LAN.	34, 155, 622 Mbit/s et plus
Frame Relay	Liaisons numériques à commutation de trames. Très utilisées en interconnexion voix/données.	De 64 Kbit/s à 8 Mbit/s (voire à 34 Mbit/s)

Ces supports de transmission nécessitent des modems adaptés dénommés CSU (*Channel Service Unit*) ou DCE (*Data Circuit-Terminating Equipment*), par opposition aux équipements DSU (*Data Service Unit*) ou DTE (*Data Terminal Equipment*) qui s'y connectent. Par exemple, le modem est un DCE, et le PC un DTE. Nous nous situons ici au niveau physique.

Quel protocole de niveau 2 ?

Les concepteurs auraient pu utiliser les mêmes trames Ethernet, et ainsi simplifier le problème, d'autant que la possibilité d'adressage est immense (2^{48} adresses). Ce pourrait être le cas, mais il aurait fallu élaborer un mécanisme spécifique pour ne pas propager dans tout le réseau les trames de broadcast et multicast. Cela aurait entraîné d'autres complications inextricables puisqu'il faut quand même les propager dans un certain périmètre (résolution d'adresses, etc.). De plus, une trame Ethernet comprend 18 octets, ce qui était considéré comme un overhead important lorsque le débit des réseaux étendus était limité (ce qui est d'ailleurs toujours le cas avec notre modem RTC). En définitive, les protocoles LAN ne sont pas adaptés aux réseaux étendus. Ce constat sera de moins en moins valable dans le futur.

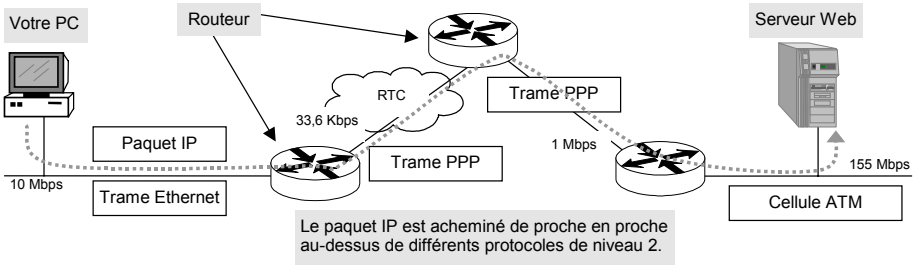
Quelques protocoles de niveau 2 utilisés pour les réseaux étendus	
PPP (*Point to Point Protocol*)	Utilisé sur les supports RTC, RNIS, LS et ADSL.
Frame Relay	Protocole point à point et multipoint voix et données. Utilisé sur des supports LS.
ATM (*Asynchronous Transfert Mode*)	Comme pour Ethernet, la norme définit les couches physique et liaison. ATM véhicule voix et données.

Les deux tableaux précédents font apparaître ATM comme étant le protocole universel : il fonctionne sur les LAN et les WAN et supporte la voix et les données. Mais, bien que très utilisé par les opérateurs sur leur réseau WAN, il est très peu utilisé en LAN à cause de son coût.

Quel équipement réseau ?

L'utilisation de multiples protocoles de niveau 2 pour transporter les paquets IP pose un nouveau problème en comparaison des architectures de réseaux locaux qui n'utilisaient qu'Ethernet. Celui-ci avait été entraperçu lorsque l'encapsulation des paquets au-dessus de plusieurs réseaux avait été étudiée au chapitre 5.

Figure 9-1. *L'acheminement d'un paquet IP.*



Comment, en effet, assurer la continuité d'adressage et de commutation au-dessus de protocoles aussi différents ? Réponse : le paquet IP est le seul lien commun. Il faut donc disposer d'équipements spécifiques qui permettent de :

- gérer les différents supports de transmission LAN et WAN ;
- traiter les paquets IP, c'est-à-dire utiliser les protocoles de niveau 3.

L'équipement qui répond à ces besoins est le **routeur**, c'est-à-dire un commutateur de niveau 3 (par opposition aux commutateurs de niveau 2, tels ceux trouvés sur Ethernet).

Quel opérateur ?

De plus, notre interconnexion de réseaux doit passer dans des zones du domaine public dont nous n'avons pas la maîtrise. Enfin, même en admettant que nous obtenions toutes les autorisations administratives nécessaires, la pose de câbles entre les deux sites reviendrait très chère.

La seule solution est de faire appel aux services d'opérateurs comme France Télécom, Cegetel, Colt, etc. Dans ce domaine, le marché offre un nombre impressionnant de solutions combinant techniques et niveaux de service.

Niveau de prestation	Description technique	Service fourni
1. Support de transmission (couche physique)	LS, xDSL, ATM (connexions point à point) et RNIS (multipoint)	Supervision de la ligne (option : garantie de temps de réparation)
2. Réseau fédérateur (couches physique et liaison)	Accès *via* LS et RNIS au backbone Frame-Relay, ATM, etc., de l'opérateur	Réseau fourni et exploité par l'opérateur + support client avec engagements de résultats
3. Interconnexion de réseaux locaux (couches 1, 2 et 3)	Support de transmission + réseau fédérateur + routeur	Réseau étendu de bout en bout fourni et exploité par l'opérateur + support client, avec engagement de résultat

Les coûts associés à ces services sont de différentes natures :

- frais uniques de mise en service ;
- frais mensuels fixes en fonction du débit des lignes et de la qualité du service ;
- et, de moins en moins, frais mensuels variables en fonction de la consommation.

De quoi avons-nous besoin ?

Pour notre première interconnexion, nous allons restreindre notre choix au plus simple et au moins cher.

D'une liaison entre les deux sites

Les sites de Paris et d'Orléans étant distants de 112 km à vol d'oiseau, il faut obligatoirement passer par un opérateur. Mais quel support de transmission utiliser et à quel débit, et quel type de service demander ?

À quel débit ?

Le dimensionnement des liaisons est un exercice délicat et important, car il va influer sur les temps de réponse du réseau et donc sur la satisfaction des utilisateurs. À l'inverse des LAN, les débits des réseaux étendus sont limités à cause des coûts qu'ils entraînent.

Le débit dépend de trois facteurs : le type de trafic, le volume de données généré et les temps de réponse requis.

Application	Type de trafic	Volume généré	Temps de réponse requis
Sauvegarde	Transfert de gros et très gros fichiers	Élevé	Faible à moyen
Partage de fichiers avec des serveurs Windows NT	Transfert de petits et gros fichiers	Élevé	Moyen
Messagerie	Transfert de petits et gros fichiers	Moyen à élevé	Moyen à élevé
Base de données client-serveur	Transactionnel (transfert de petits et moyens fichiers)	Faible à élevé	Moyen
Connexion aux serveurs Web	Transactionnel (transfert de petits et moyens fichiers)	Faible	Moyen
Connexion Telnet sur un serveur Unix	Conversationnel (écho distant)	Faible	Élevé (< 300 ms)

Les transferts de fichiers perturbent fortement les flux conversationnels (Telnet) et, dans une moindre mesure, les petits flux transactionnels (serveurs Web). En conséquence, les sauvegardes et les transferts de gros fichiers doivent, de préférence, être effectués la nuit, ce qui présente le double avantage de ne pas gêner le travail des utilisateurs de jour et de répartir l'utilisation du réseau sur 24 heures.

Dans un cas simple comme le nôtre (messagerie, transfert de fichiers et connexion Web), un débit de 64 à 128 Kbit/s devrait être suffisant. Compte tenu du nombre d'utilisateurs (plusieurs centaines sur chaque site), un débit de 128 Kbit/s est plus sécurisant. Nous limiterons les risques en souscrivant un contrat d'une durée minimale d'un an, ce qui nous permettra de changer de débit facilement.

Lorsque notre réseau intersite aura une plus grande ampleur, nous devrons nous livrer à un calcul plus précis, comme nous le verrons au chapitre 10.

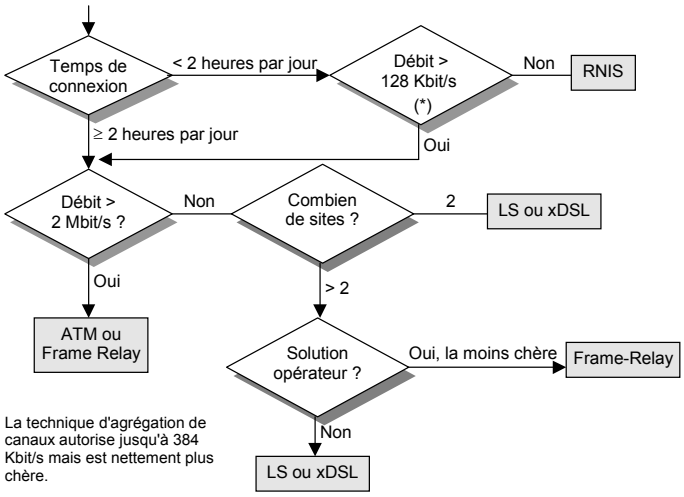
Avec quel support de transmission ?

La manière de calculer les frais mensuels varie d'un support de transmission à l'autre et, parfois, d'un opérateur à l'autre.

Support de transmission	Mode de facturation	Frais mensuels
RNIS	Durée de la communication et distance	Variables
LS / xDSL	Débit et distance	Fixes
Frame Relay	Débit garanti, débit du port et type d'équipement	Fixes
ATM	Débit et qualité de service	Fixes

Le premier critère à prendre en compte est donc le temps d'utilisation de la liaison, puis son débit.

Figure 9-2.
Choix d'un support pour une interconnexion de réseaux locaux.



NOTE Le seuil de deux heures par jour a été calculé pour une distance de 100 km.

Dans notre cas, la liaison devrait être utilisée toute la journée, quasiment en continu, compte tenu du nombre d'utilisateurs (quelques centaines) et des applications utilisées (flux client-serveur, connexions à des machines Unix, etc.). La LS point à point est donc appropriée.

Avec quel service opérateur ?

En ce qui concerne le type de service, la question est de savoir si vous voulez « faire » ou « faire faire ». En d'autres termes, voulez-vous réaliser vous-même l'interconnexion (solution privée), ou préférez-vous confier le projet à un spécialiste (solution opérateur) ?

Le premier critère à prendre en compte est celui de l'implication en termes de ressources humaines.

Tâches \ Qui fait ?	Solution privée	Service opérateur niveau 3
Étude technique (choix des supports, dimensionnement, etc.)	Vous	Vous, assisté de l'opérateur
Fourniture de la liaison	Opérateur	Opérateur
Achat des routeurs	Vous, auprès d'un distributeur réseau	Opérateur
Installation et configuration des routeurs	Vous ou le distributeur réseau	Opérateur
Exploitation du routeur (supervision, modification des paramètres, intervention lors des pannes, etc.)	Vous ou le distributeur, avec un contrat d'assistance	Opérateur
Résolution des pannes matérielles	Distributeur, avec un contrat de maintenance	Opérateur

La solution opérateur nécessite moins de ressources internes (vous êtes notamment moins sollicité), surtout avec le service de niveau 3. La solution privée revient à choisir le service de niveau 1 et implique donc de gérer un projet avec au moins deux interlocuteurs (le distributeur et l'opérateur).

Le second critère à prendre en compte est d'ordre financier.

Tâches \ Quel coût ?	Solution privée	Service opérateur niveau 3
Étude technique (choix des supports, dimensionnement, etc.)	Dépend de l'importance du projet	Coût identique
Mise en œuvre du support de transmission	Coût standard (mise en service + frais mensuels)	Coût moins élevé, car la LS d'accès est locale
Achat des routeurs	Coût d'investissement	Frais mensuels d'exploitation
Installation et configuration des routeurs	Coût d'investissement	Frais de mise en service
Exploitation	Contrat d'assistance sur site	Coût plus élevé, mais meilleure qualité de service
Résolution des pannes matérielles	Contrat de maintenance annuel : 8 à 12 % du coût d'achat	Moins cher

La solution opérateur devrait être la moins chère à long terme (2 à 5 ans) et apporter le moins de soucis. Cela n'est cependant pas toujours le cas, surtout pour de petits réseaux, qui plus est, limités à la France. Les économies les plus importantes sont, en effet, réalisées à l'international.

Si vous disposez d'un gros réseau, vous pouvez opter pour la solution intermédiaire qui consiste à retenir le service opérateur de niveau 2. Cette solution permet de conserver la maîtrise du routage IP et de modifier plus facilement la configuration de vos routeurs. Rien ne vous empêchera par la suite d'opter pour le service de niveau 3 (tous les opérateurs proposent des routeurs Cisco qui peuvent donc être repris).

Dans notre cas (deux sites à interconnecter), la meilleure solution est une LS avec une garantie de temps de rétablissement. Mais il serait intéressant de comparer les coûts avec un service de niveau 3 (LS plus routeur). Les résultats de cette comparaison peuvent varier en fonction de nombreux critères (le moment où elle est réalisée – car les prix évoluent vite –, les distances entre sites, le débit, etc.).

Admettons cependant que la solution privée soit la moins chère ou alors regardons faire l'opérateur qui fait ce que nous aurions dû faire.

De routeurs

Le routeur est le seul équipement permettant d'interconnecter deux sites sur de longues distances. Le réseau Internet n'est d'ailleurs constitué que de routeurs utilisant des liaisons spécialisées, ATM et Frame Relay.

Le plus simple est d'opter pour une configuration fixe comprenant une interface Ethernet et une interface WAN et ne supportant que le protocole IP. Le coût de cette configuration de base oscille entre 1 000 et 1 300 € HT. Il nous faut deux routeurs, un par site.

Quelques extensions matérielles et logicielles sont proposées.

Fonctionnalité	Description	Intérêt	Coût HT
Multiprotocole	Support des protocoles IP + IPX + Decnet + SNA, etc.	Si existant à supporter	± 1 500 € (nécessite davantage de mémoire)
Fonction pont (bridge)	Commutation de niveau 2	Pour les anciens protocoles non routables	± 150 € (ou 0 € car livrée souvent en standard)
Interface RNIS	Connexion au support de transmission RNIS	Peut être utilisée pour le secours ou le débordement	± 450 €

Dans notre cas, l'option RNIS est intéressante, car elle offre une solution de secours en cas de panne de la LS. Elle permet ainsi d'assurer la continuité de service avec la même qualité, le débit offert étant de 64 et 128 Kbit/s, soit exactement celui de notre LS.

L'autre intérêt du RNIS concerne le débordement : lorsque la LS est chargée à 80 ou 100 %, le routeur peut activer l'interface RNIS, offrant ainsi un débit supplémentaire de 64 Kbit/s, voire 128 Kbit/s. Cette fonction pourra être utilisée si le débit nécessaire a été sous-évalué.

QU'EST-CE QU'UN PONT ?

Le **pont** (bridge) est l'ancêtre du commutateur de niveau 2. Il a été le premier équipement utilisé pour **interconnecter** des segments Ethernet partagés, soit localement, soit *via* des réseaux étendus. Compte tenu de la technologie et des prix d'alors, il était équipé de deux à quatre interfaces LAN et WAN, et offrait une faible puissance comparée aux commutateurs d'aujourd'hui.

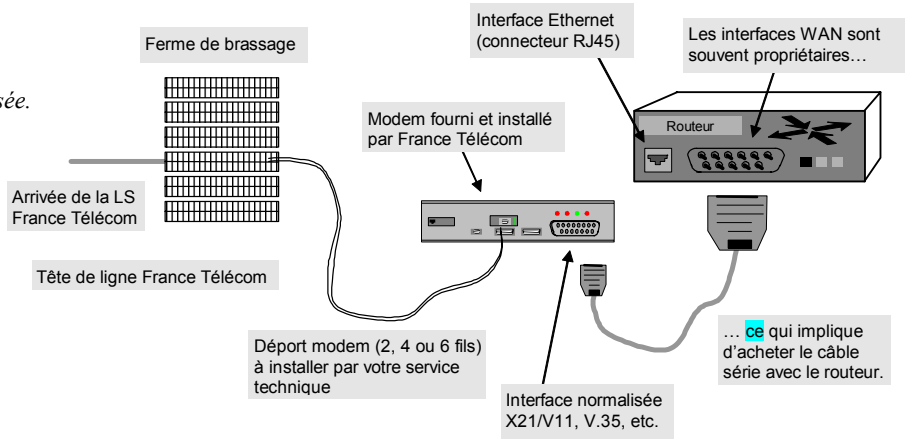
Les routeurs continuent de supporter cette fonctionnalité afin d'assurer la compatibilité avec l'existant.

De câbles

Encore des câbles ! Il faut bien sûr un cordon de brassage pour connecter le routeur à un concentrateur ou à un commutateur Ethernet. L'interface Ethernet du routeur étant strictement équivalente à une carte réseau d'un PC, un cordon identique (droit RJ45/RJ45) sera utilisé.

Il faut également des câbles spécifiques côté réseau de l'opérateur, en l'occurrence côté ligne spécialisée. Pour notre connexion à Internet, nous avions utilisé un câble série entre le PC et le modem, câble fourni avec le modem. Le principe est identique pour les routeurs : lorsque l'opérateur (France Télécom, par exemple) met en service la LS, il installe dans vos locaux un modem auquel vous connectez le routeur *via* un câble série adapté.

Figure 9-3.
Connexion du routeur à la ligne spécialisée.



La ligne spécialisée arrive dans un local technique réservé à France Télécom. La longueur des câbles série étant limitée, il faut déporter le modem au plus près du routeur qui se trouve dans un LTE ou une salle informatique. Cette partie de l'installation est privée (elle se déroule dans les locaux de votre société) et doit donc être réalisée par les services techniques de l'immeuble (souvent, les services généraux ou les téléphonistes). France Télécom se contente d'amener la LS dans son local, d'attendre la fin de l'installation du déport modem, puis d'installer le modem et de tester la LS avant de la mettre en service.

Les câbles série sont de différentes natures. Pour la connexion Internet, nous avions utilisé un câble V.24 (connexion série RS-232). Pour notre LS, l'opérateur France Télécom nous propose V.35 ou X21/V11.

Spécifications des interfaces					
Appellation usuelle	**Mécanique**	**Électrique**	**Fonctionnelle**	**Longueur du câble**	**Débit en Kbit/s**
V.24 (RS 232)	ISO 2110 25 broches	V.28	V.24	12 m	De 2,4 à 19,2
V.35 (RS 232)	ISO 2693 34 broches	V.11/V.10	V.24	15 m 10 m	48, 56, 64 128, 256
V.36	ISO 4902 37 broches	V.11/V.10	V.24	15 m 10 m	48,56, 64 128, 256
X21/V.11	ISO 4903 15 broches	V.11	X.24	100 m 50 m	De 64 à 1 024 1 920
G703	ETSI 300.166	G703	G703	300 m	2 048
G703/G704	ETSI 300.167 9 broches	G703	G704	300 m	De 256 à 1 984

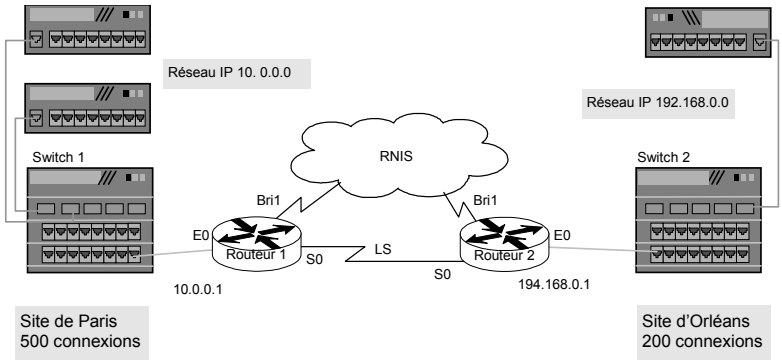
Malgré la normalisation, des variantes peuvent cependant exister. Les plus connues sont celles relatives à la norme V.35 qui accepte une version américaine et une version française en fonction du format de la prise. Les différences résident dans le diamètre des broches (plus grand sur le modèle américain) et dans le type de fixation (par vis dans le premier cas, par clip dans le second).

Comment faire fonctionner tout cela ?

Définir l'architecture

Les routeurs vont s'insérer dans un réseau existant. Il convient donc de définir l'architecture et de réfléchir au paramétrage des routeurs.

Figure 9-4.
Principe de l'architecture réseau.



Suivant les indications de notre plan d'adressage (voir chapitre 5), nous avons choisi pour notre site parisien une classe A subnettée sur 22 bits. En revanche, le site d'Orléans correspond à une nouvelle entité rachetée par la société, et une classe C avait été choisie par notre prédécesseur. Nous n'avons pas le temps de changer cette adresse et la conservons en l'état.

La question encore en suspens est de savoir si les interfaces WAN doivent également disposer d'une adresse IP et, dans l'affirmative, de quelle classe d'adresse.

Connecter un PC au routeur

Pour une configuration fixe, aucune préparation matérielle spécifique n'est nécessaire (pas de carte à installer, rien à démonter).

En revanche, le routeur est livré avec un logiciel non configuré (comme les PC). Vous pouvez le configurer dans votre bureau sans qu'il soit connecté au réseau, ou l'installer et le configurer sur place.

Les premiers paramétrages d'un routeur nécessitent de s'y connecter directement pour saisir des commandes. Un routeur est à l'image d'un PC : il contient un système d'exploitation (propriétaire) et une interface utilisateur (généralement en mode texte sans fenêtre ni souris).

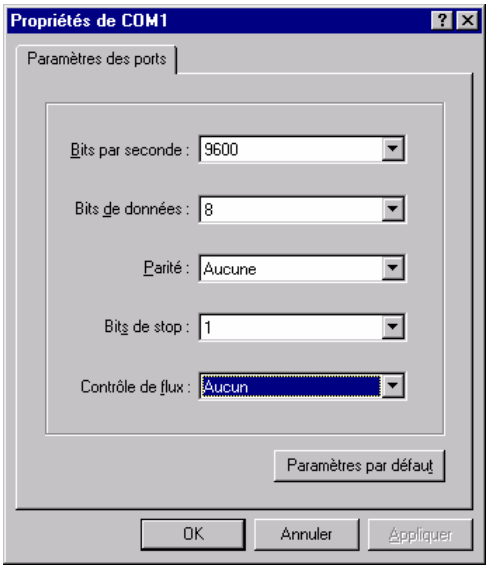
Vous avez donc besoin d'un câble série (RS-232 avec un connecteur V.24) pour raccorder le port console du routeur au port série d'un PC.

Le type de port série des routeurs est variable. Vous pouvez trouver un connecteur DB9 comme sur votre PC ou, plus rarement, un connecteur DB25 (l'équivalent d'un DB9 avec 25 broches) ou encore, de plus en plus souvent, une prise RJ45 femelle. Il faut donc trouver le bon câble (croisé !) qui dispose d'un connecteur DB9 femelle côté PC (DTE) et d'un connecteur adéquat mâle, côté routeur. De plus en plus souvent, ce cordon est livré avec le routeur.

Côté PC, vous devez utiliser le logiciel Hyperterminal en cliquant sur « Programme→Accessoires→Hyperterminal ». Ce logiciel est un émulateur VT (*Virtual Terminal*) qui va vous permettre de dialoguer en mode texte avec le routeur.

Saisissez alors le nom du profil que vous pourrez réutiliser par la suite, par exemple « routeur », puis cliquez sur « OK ». Choisissez ensuite le port série auquel vous avez connecté le câble reliant le PC au routeur : généralement COM1 ou COM2 (écran ci-contre).

Puis, saisissez les paramètres de la liaison série. Les valeurs dépendent du routeur employé ; elles sont généralement indiquées dans la documentation. Nos routeurs étant des routeurs Cisco, les valeurs sont les suivantes (deuxième écran ci-contre) :



Cliquez sur « OK », puis appuyez plusieurs fois sur la touche « Entrée » ou, si cela n'est déjà fait, allumez le routeur.

Vous devez vous retrouver avec un écran de bienvenue et, un $$prompt$$ vous invitant à saisir des commandes, ou bien vous obtenez un menu dans lequel vous vous déplacez à l'aide des touches fléchées du clavier.

Vous êtes maintenant prêt à configurer le routeur.

Configurer le routeur

Selon les constructeurs, la connexion réseau d'un routeur — l'équivalent d'une carte réseau d'un PC — s'appelle « port » ou « interface ». On parlera alors indifféremment de ports LAN ou WAN, d'interface série, etc.

Affecter les adresses IP

La première chose à faire est d'affecter les adresses IP à chaque interface. Côté LAN, on se réfère au plan d'adressage. La commande suivante réalise l'opération et active le logiciel IP dans le routeur.

```
int e 0
ip address 10.0.0.1 255.255.252.0
```

Abrégé de " Interface ethernet 0 "

Par contre, faut-il en affecter une aux interfaces série ? Si tel est le cas, il faut dédier un réseau IP pour seulement deux adresses. Même pour une classe C, 252 adresses sont gaspillées. Ceci étant, nous pouvons utiliser autant d'adresses que souhaitées puisque nous avons pris le parti d'un plan d'adressage privé. Cependant, il est conseillé de limiter le nombre de réseaux afin d'en simplifier la configuration.

Deux solutions sont conseillées : aucune adresse IP (pour les petits réseaux), ou un réseau IP dédié aux interfaces WAN en le subnettant (pour les petits et grands réseaux).

Aucune adresse IP	Adresse IP subnettée
Point à point uniquement avec HDLC ; PPP et Frame Relay en point à point.	Point à point et multipoint (Frame Relay, X.25, SMDS).
L'interface ne répond jamais au ping (pour savoir si elle est active) ; SNMP fonctionne.	L'interface répond au ping : cela permet de savoir rapidement si elle est active.
Autres fonctions propres au constructeur non supportées.	---

La commande suivante permet d'activer IP sur l'interface série sans lui affecter une adresse. Si des paquets sont générés par cette interface, l'adresse IP source sera celle de l'interface Ethernet.

```
int s 0
ip unnumbered e 0
```

Abrégé de " Interface serial 0 "

Par défaut, le routeur utilise un protocole de niveau 2, en général HDLC ou PPP. La commande suivante permet de forcer l'utilisation de PPP, ce dernier étant plus approprié aux liaisons point à point et à IP.

```
int s 0
encapsulation ppp
```

Activer le routage

Le routage des paquets IP s'appuie sur la partie réseau des adresses de destination des paquets. Il faut donc indiquer au routeur parisien que les paquets à destination d'Orléans doivent être envoyés sur l'interface série.

Dans notre cas, le plus simple est de configurer une route statique. Puisque notre interface série ne possède pas d'adresse IP, on indique explicitement son nom.

```
# Sur le routeur de Paris
ip route 192.168.0.0 255.255.255.0 s0
```

LE POINT SUR PPP (RFC 1661 ET 1662)

PPP (*Point-to-Point Protocol*) est un protocole utilisé sur des liaisons point à point **synchrones** (une LS) et **asynchrones** (avec un modem RTC, par exemple). Il est utilisé pour des débits variant entre 19,2 Kbit/s et 2 Mbit/s. Il est composé de trois éléments :

- des trames pour transporter les protocoles de niveau 3 ;
- de LCP (*Link Control Protocol*) pour établir, configurer et tester la liaison ;
- de NCP (*Network Control Protocol*) pour établir et configurer différents protocoles de niveau 3 (tels que IP, IPX ou Decnet) qui peuvent être multiplexés sur une seule liaison.

1 octet	1 octet	1 octet	2 octets	Longueur variable	2 octets
Fanion	Adresse	Contrôle	Protocole	Données	FCS

Le champ " Fanion " est un délimiteur de trame de valeur binaire " 01111110 ". Le champ " Protocole " indique le type de protocole de niveau 3 présent dans le champ de données (0021 pour IP, c021 pour LCP, c023 pour LQM, c223 pour CHAP, 8281 pour MPLS-CP, etc.). Le champ " FCS " (*Frame Check Sequence*) est un code de détection d'erreur.

Dans sa version asynchrone, chaque octet de la trame est transmis avec un bit start et un bit stop, mais sans bit de parité, puisque le champ FCS est utilisé pour détecter les erreurs.

Le protocole **LCP** permet d'établir, de configurer, de surveiller et de terminer les liaisons point à point. Il permet, tout d'abord, de négocier des options, telles que :

- le MTU (*Maximum Transmission Unit*) dont la valeur par défaut est celle d'Ethernet (1 500 octets) ;
- le choix du protocole d'authentification (CHAP, PAP ou aucun) ;
- le choix du protocole de contrôle de qualité (LQR ou aucun) ;
- la réduction du champ " Protocole " à un octet ;
- la suppression des champs " Adresse " et " Contrôle " lorsqu'ils ne sont pas utilisés.

LCP peut activer un protocole d'**authentification**, tel que **CHAP** (*Challenge Handshake Authentication Protocol*). Un mot de passe chiffré est échangé entre les deux nœuds (routeur ou PC).

LCP peut activer la procédure LQM (Link Quality Monitor — RFC 1989) qui coupe la liaison lorsque la qualité de service calculée tombe en dessous d'un seuil prédéfini. La qualité en émission est calculée en comparant le nombre total de paquets et d'octets transmis avec ceux reçus par la station distante. De même, la qualité en réception est calculée en comparant le nombre total de paquets et d'octets reçus avec ceux émis par la station distante.

Le protocole **NCP** se décline en autant de versions que de protocoles réseau supportés. On trouve ainsi IPCP (*IP Control Protocol*), DCP (*Decnet Phase IV Control Protocol*), MPLS-CP (*MPLS Control Protocol*), etc. Les trames échangées sont du même type que celles utilisées par LCP.

IPCP (RFC 1332) permet de négocier des options spécifiques, comme l'affectation des adresses IP ou la compression des en-têtes TCP/IP. Pour éviter la fragmentation des paquets TCP, la longueur maximale des données peut être négociée pour aller au-delà des 1 500 octets, et correspondre ainsi au MTU du protocole TCP.

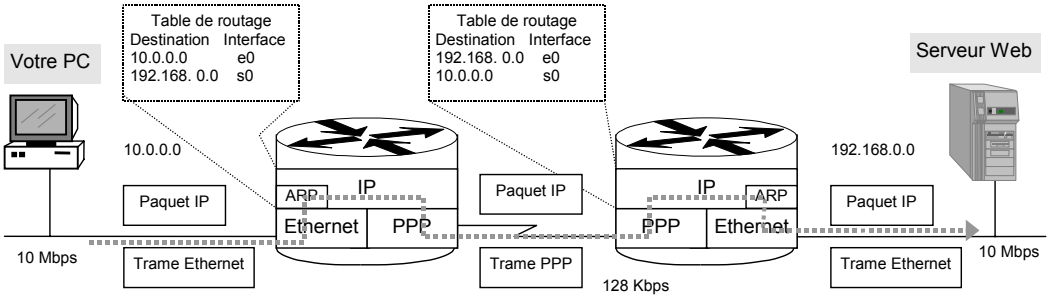
PPP se contente ensuite de véhiculer les paquets IP en offrant simplement la détection d'erreur.

La même configuration doit être appliquée au routeur d'Orléans.

```
int e0
ip address 192.168.0.1 255.255.255.0
int s0
ip unnumbered e 0
encapsulation ppp
^Z
ip route 10.0.0.0 255.255.252.0 s0
```

Grâce à ces commandes, les routeurs sont capables de router correctement les paquets.

Figure 9-5.
Principe du routage des paquets IP.



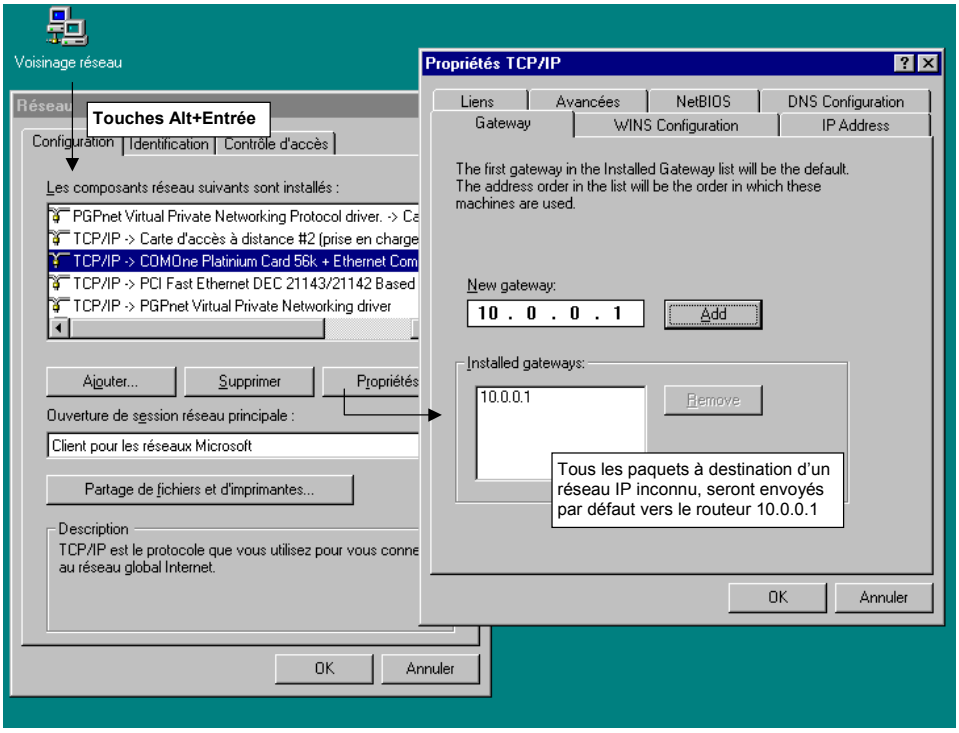
Puisqu'il n'existe qu'un seul chemin (la ligne série), il est possible de déclarer une route par défaut. Cette commande permet au routeur d'envoyer sur son interface série tous les paquets dont il ne connaît pas l'adresse de destination :

```
# Remplace la commande ip route 192.168.0.0 255.255.255.0 s0
ip route default s0
```

Configurer les postes de travail

La question est maintenant de savoir comment les PC vont pouvoir envoyer les paquets IP sur l'autre site. Réponse : selon le même principe que celui utilisé par les routeurs.

Il suffit, en effet, d'ajouter une route par défaut (*default gateway*).



Si une route par défaut existe déjà vers un autre routeur, il est toujours possible d'ajouter une route statique, comme suit :

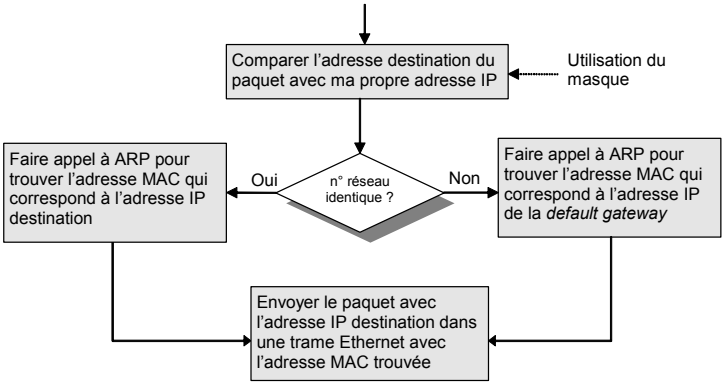
```
route add -p 192.168.0.0 mask 255.255.255.0 10.0.0.1
```

Pour aller vers ce réseau, je passe par ce routeur

Mais, comment la résolution d'adresse fonctionne-t-elle, étant donné que l'adresse MAC du serveur situé à Orléans nous est inconnue et que les trames de broadcast ARP ne peuvent pas passer par le routeur ?

On pourrait activer la fonction « Pont du routeur » et ne laisser passer que les broadcasts ARP, mais on perdrait alors l'avantage de la segmentation : le réseau de Paris recevrait des broadcasts MAC dont il n'a pas l'usage, et inversement.

En réalité, la solution retenue par la pile IP est la suivante :



En définitive, le PC recherchera l'adresse MAC du routeur et l'associera à l'adresse IP de destination. Les trames Ethernet seront ainsi envoyées directement au routeur. Ce dernier – grâce à l'adresse de destination IP – comprendra que le paquet ne lui est pas destiné. Mais, au lieu d'ignorer le paquet comme le font les PC (voir chapitre 5), il consultera sa table de routage et retransmettra le paquet vers la bonne interface (Ethernet ou autre, ici la série).

Le routeur d'Orléans reçoit le paquet. L'adresse IP de destination correspondant à un réseau auquel il est directement connecté, celui-ci fait appel à ARP pour obtenir l'adresse MAC qui correspond à cette adresse IP. Le routeur envoie alors le paquet IP dans une trame Ethernet dont l'adresse de destination est celle du serveur.

Il est également possible d'indiquer l'**adresse IP de la station** elle-même **comme passerelle par défaut**. Ce faisant, la pile IP considère que tous les subnets lui sont directement accessibles. Par conséquent, si ceux-ci sont tous situés sur le même segment Ethernet, la station pourra directement accéder à toutes les autres stations. Si les subnets se trouvent sur d'autres segments interconnectés par des routeurs, le routeur local doit être configuré en mode **Proxy ARP**, ce qui est le cas par défaut de nos routeurs Cisco.

LE POINT SUR PROXY ARP (RFC 1027)

Le protocole ARP (Address Resolution Protocol) permet à une station de connaître l'adresse Ethernet MAC d'une autre station en ne connaissant que son adresse IP. La trame de broadcast envoyée par ARP n'est diffusée que localement à un segment : elle est, en effet, bloquée par les routeurs, comme cela est la règle dans les réseaux locaux.

En mode proxy ARP, un routeur qui reçoit une requête ARP concernant une adresse IP dont il connaît le réseau (car l'adresse est présente dans sa table de routage) s'assure qu'il dispose de la meilleure route, puis répond à la requête en y mettant sa propre adresse MAC. En définitive, **il se substitue à la station cible** qui ne peut pas recevoir une telle requête puisqu'elle se trouve sur un autre segment Ethernet.

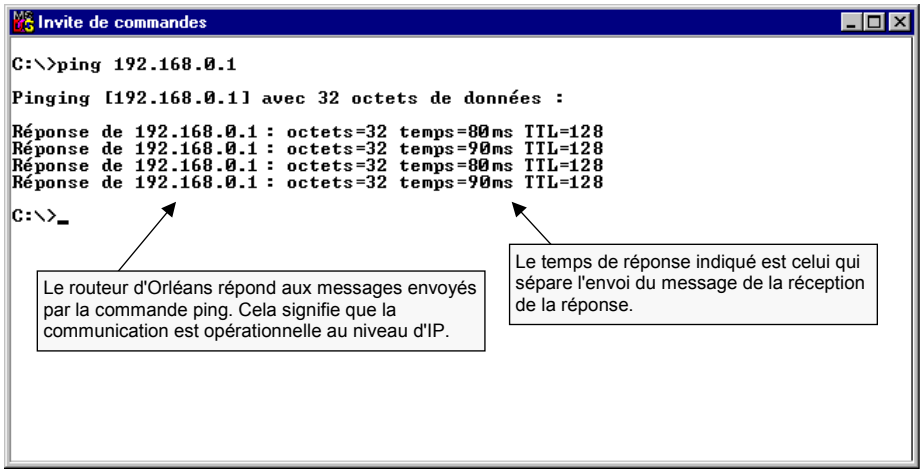
Ces deux mécanismes utilisés conjointement permettent, par exemple, de migrer d'un réseau de routeurs vers un réseau de commutateurs, ou d'un VLAN à plusieurs subnets IP vers un VLAN à un seul subnet. Si, de plus, vous utilisez un serveur DHCP (voir chapitre 6), ce dernier doit également être configuré en conséquence.

Voilà, vous venez de réaliser votre première interconnexion de réseaux.

Tester le réseau

Pour tester le bon fonctionnement de votre réseau, vous pouvez utiliser la commande **ping** (voir chapitre 6). Ce programme est disponible sur les routeurs et les PC Windows, les serveurs Unix, etc.

Il se contente d'envoyer des paquets à une adresse cible qui lui répond par un paquet en retour. De plus, le ping mesure le temps de réponse aller-retour, ce qui est une bonne indication sur les performances du réseau. Toute pile IP se doit de répondre à un ping.



Réduire l'overhead

Une fonctionnalité très souvent utilisée pour les liaisons bas débit (64 à 128 Kbit/s), est la compression des en-têtes TCP/IP (RFC 1144). Le principe repose sur le constat que les en-têtes varient peu d'un paquet à l'autre. Seuls les octets modifiés par rapport au précédent paquet sont donc transmis. La taille de l'en-tête est ainsi réduite de 40 (20 pour TCP et 20 pour IP) à 10 octets en moyenne. Cette fonctionnalité est d'autant plus efficace qu'il y a de nombreux petits paquets à traiter (connexions Telnet, par exemple).

```
interface serial 0
ip tcp header-compression
```

Mettre en place une liaison de secours

Notre liaison est en place : tout fonctionne. Mais que se passera-t-il si elle tombe en panne ? Plus de réseau. La solution consiste à mettre en place une liaison de secours. Là encore, la question du choix du support de transmission et du débit se pose.

Quels sont les choix ?

La première solution consiste à doubler la liaison principale : une seconde LS prend le relais de la première. En fonctionnement normal, les deux liaisons peuvent être utilisées en partage de charge.

L'autre solution consiste à utiliser Numéris, le réseau téléphonique numérique de France Télécom, généralement appelé RNIS (réseau numérique à intégration de service). Pour un coût nettement moins élevé qu'une LS, la liaison RNIS n'est activée que si la liaison principale est coupée.

Secours *via*…	Avantages	Inconvénients
RNIS	Permet le débordement. Facturé essentiellement à l'utilisation	Limité à 128 Kbit/s en standard et à 384 Kbit/s avec l'agrégation de canaux.
LS	Convient à une utilisation en partage de charge.	Facturée même si elle n'est pas utilisée.

La mise en place de deux LS, par exemple 2 × 64 Kbit/s à la place d'une seule liaison à 128 Kbit/s, permet de répartir la charge sur les deux liaisons. Cette solution nécessite cependant des mécanismes plus complexes que ceux utilisés jusqu'à présent : ils reposent sur des protocoles de routages qui sont le plus souvent propriétaires. En outre, deux LS à 64 Kbit/s coûtent plus cher qu'une seule à 128 Kbit/s.

Une liaison RNIS présente l'avantage de n'être activée qu'en cas de nécessité. Si elle est inutilisée, seul l'abonnement de base doit être payé : de 30 à 45 € contre 455 à 610 € par mois pour une LS 64 Kbit/s courte distance.

Globalement, le secours RNIS revient donc nettement moins cher que le doublement d'une LS. Pour toutes ces raisons, cette solution est donc plus adaptée à notre besoin.

Solutions alternatives

Des techniques d'**agrégation de canaux B** permettent d'utiliser *n* canaux B à 64 Kbit/s de manière à n'en faire qu'une seule liaison logique à *n* × 64 Kbit/s. Il est ainsi possible d'agréger les deux canaux B d'un accès de base (offrant un débit global de 128 Kbit/s) ou d'agréger les canaux B de plusieurs accès de base (généralement jusqu'à trois, autorisant un débit global de 384 Kbit/s). Cette technique requiert l'utilisation d'équipements spécifiques qui peuvent être onéreux.

Attention au coût des communications : l'agrégation de deux canaux B offre un débit de 128 Kbit/s mais équivaut à deux communications téléphoniques simultanées. Cette technique est donc deux fois plus coûteuse qu'une simple connexion à 64 Kbit/s.

Le choix du débit de la liaison de secours dépend de l'importance accordée à l'interconnexion :

- Si un **mode dégradé** est acceptable, une liaison à 128 ou 256 Kbit/s pourra être secourue à 64 Kbit/s.
- Inversement, si vous voulez conserver les performances en mode secours, il faut utiliser une seconde liaison qui aille au-delà des limitations de débit du RNIS.

LE POINT SUR LE RNIS (ITU SÉRIE I)

Le RNIS (réseau numérique à intégration de service — **ISDN**, en anglais) désigne le réseau téléphonique numérique, par opposition au RTC (réseau téléphonique commuté — **PSTN**, en anglais) qui en est la version analogique. En France, le RNIS est commercialisé par France Télécom sous le nom de Numéris.

Il existe deux types d'abonnement RNIS :

- L'**accès de base T0** qui offre 2 canaux à 64 Kbit/s et un canal de signalisation à 16 Kbit/s.
- L'**accès primaire T2** qui revient à fournir un système à 30 canaux de 64 Kbit/s chacun et un canal de 64 Kbit/s pour la signalisation (le premier canal, appelé verrouillage de trame, est dédié à la synchronisation du faisceau).

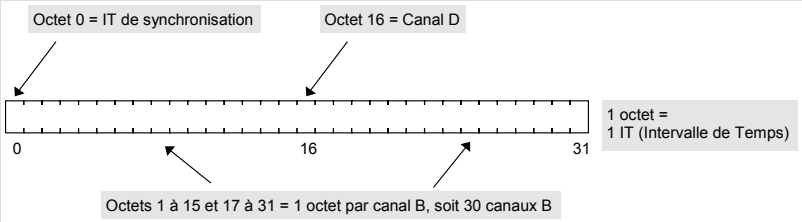
Un canal RNIS permet de transporter une liaison téléphonique, une liaison télécopie ou encore une liaison de données. Cette unité de base est appelée **canal B** (Base), tandis que le **canal de signalisation** est appelé canal D (Données).

L'accès de base est l'équivalent d'un réseau local, appelé **bus S0**, qui est géré par la TNR (terminaison numérique de réseau). Cinq terminaux au maximum (téléphone, télécopieur, routeur ou tout autre équipement pourvus d'une interface de base — BRI, *Basic Rate Interface*) peuvent se partager le bus dont l'accès est géré selon l'algorithme **CSMA-CR** (*Carriage Sense Multiple Access - Contention Resolution*).

Le canal D véhicule le protocole de signalisation de niveau 3 (couche réseau) **Q.931** qui permet d'établir et de gérer les communications entre deux terminaux RNIS (numérotation, identification de l'appelant, négociation des paramètres, etc.).

Les messages Q.931 sont transportés dans des trames de niveau 2 (couche liaison) gérées par le protocole **HDLC** (*High Data Link Control*) qui fournit des mécanismes d'acquittement de trames, de contrôle de flux et de reprise sur erreur pour le canal D. Le RNIS utilise une version adaptée, appelée LAP-D (*Link Access Procedure - D channel*).

La trame de l'accès primaire est constituée de 32 canaux de 64 Kbit/s : 30 canaux B, un canal D et un canal de synchronisation (30B+D, en abrégé). Le débit global de l'accès primaire est de 2 Mbit/s.

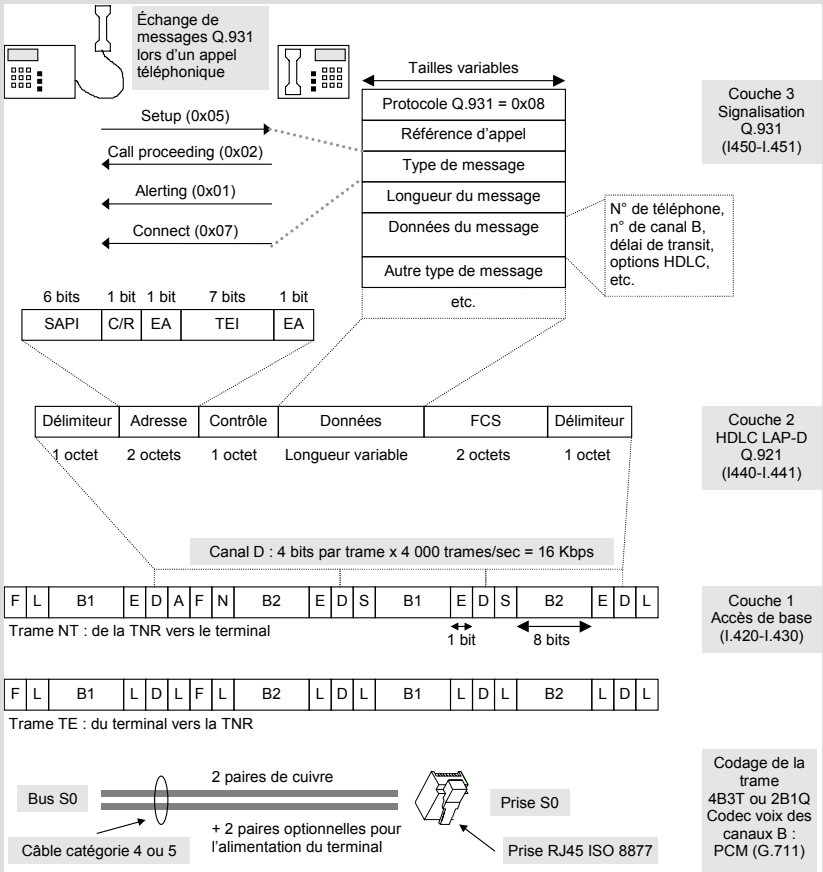


La trame de l'accès de base comprend 2 canaux B à 64 Kbit/s, 1 canal D à 16 Kbit/s et quelques bits de gestion (2B+D, en abrégé). Le débit global est de 192 Kbit/s. D'une longueur de 48 bits, elle est émise en 250 microsecondes (soit 4 000 trames par seconde).

• • •

LE POINT SUR LE RNIS (SUITE)

L'accès primaire et l'accès de base utilisent les mêmes protocoles de signalisation de niveaux 2 et 3.



Délimiteur	est de valeur binaire « 01111110 ». Afin d'éviter que les données lui soient identiques, un mécanisme d'**insertion de « 0 »** est utilisé à l'émission. Le principe consiste à insérer un « 0 » dès que 5 valeurs « 1 » consécutives ont été transmises (cet algorithme n'est bien sûr pas appliqué aux délimiteurs). À la réception, le 6e bit suivant 5 bits ayant une valeur de « 1 » est analysé. S'il est à « 0 », il est ignoré et, dans le cas contraire, il est considéré comme un délimiteur de trame.
Adresse	Comporte l'identifiant **SAPI** (*Service Access Point Identifier*) permettant de distinguer les applications utilisant les canaux B (téléphonie, télécopie, etc.) et l'identifiant **TEI** (*Terminal End-point Identifier*) qui permet d'identifier le terminal sur le bus. Le numéro 127 est réservé à la diffusion (*broadcast*).
Contrôle	indique le type de la trame (*Unnumbered*, *Information* ou *Supervision*), et contient un numéro de séquence. Le protocole utilise 20 formats de trames.
Données	contient les données utiles (260 octets au maximum).
FCS	(*Frame Check Sequence*) est un code de détection d'erreur de type CRC (*Cyclic Redundancy Check*).

• • •

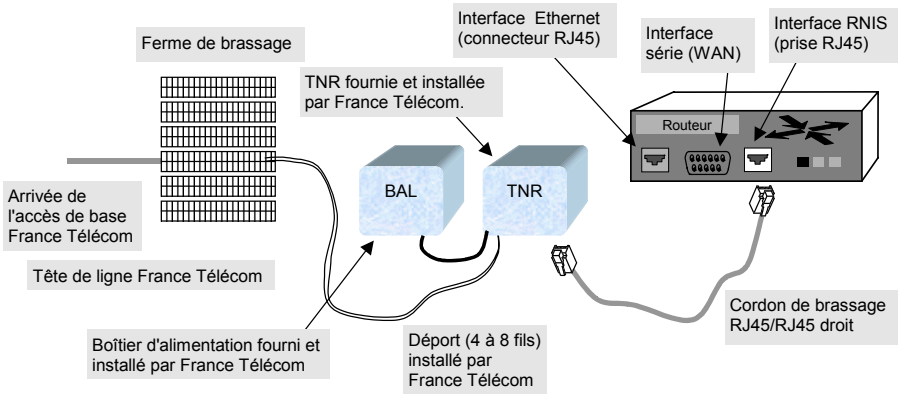
LE POINT SUR LE RNIS (FIN)

F	est un bit de synchronisation.
F_a	est un bit de synchronisation auxiliaire pour une utilisation future.
L	est le bit d'équilibrage qui sert à rétablir la composante énergétique du courant transportant le train de bits. L'objectif est d'obtenir une composante énergétique nulle afin de consommer le minimum d'énergie.
E	bit d'écho utilisé par la méthode d'accès au bus CSMA/CR (*Carrier Sense Multiple Access - Contention Resolution*).
A	bit utilisé pour l'activation du terminal.
S	bit non utilisé.
B1	est un champ de 8 bits qui véhicule les données du premier canal B.
B2	est un champ de 8 bits qui véhicule les données du deuxième canal B.
D	est le champ qui véhicule les données de signalisation du canal D.

Installation d'un accès de base T0

Comme pour notre LS, l'opérateur (France Télécom en France) installe la liaison RNIS dans la partie de vos locaux techniques qui lui est réservée. L'accès de base T0 se termine chez l'utilisateur par un petit boîtier appelé **TNR** (terminaison numérique de réseau). Celui-ci est, la plupart du temps, accompagné d'un coffret d'alimentation appelé BAL.

Figure 9-6. *Connexion du routeur au RNIS.*



Sur notre routeur d'Orléans, l'interface RNIS (appelée *BRI* pour *Basic Rate Interface*) doit être configurée de manière à ne s'activer que si la liaison principale (l'interface série) est coupée :

```
isdn switch-type vn3
interface Ethernet 0
 ip address 192.168.0.1 255.255.255.0
^Z
```

Interface RNIS de France Télécom (Version Numéris 3)

```
interface serial 0
 ip unnumbered e0
 encapsulation ppp
 backup delay 2 15

 backup interface bri 0
^Z
interface bri 0
 ip unnumbered e0
 encapsulation ppp
 dialer string 0144758899
 dialer-group 1
^Z
ip route 10.0.0.0 255.255.252.0 s0
dialer-list 1 protocol ip permit
```

Active l'interface RNIS 2 secondes après la chute de l'interface principale. Attend 15 secondes avant de couper la liaison RNIS lorsque l'interface principale est de nouveau opérationnelle.

L'interface de secours est la bri 0 qui appelle le numéro de téléphone de l'abonnement RNIS situé à Paris.

Définition d'une route statique.

En mode secours, un seul canal B est activé. Notre routeur est capable de gérer deux connexions simultanées, une sur chaque canal B, mais pas d'agréger les canaux.

Sécurisation de la liaison

Le problème avec le numéro de téléphone est qu'il peut être appelé par n'importe qui. Il faut donc appliquer quelques règles de base. La première est, bien sûr, d'inscrire votre numéro sur liste rouge. Le service d'identification de l'appelant pourra aussi être utile.

Au niveau du routeur, il est possible de réaliser une authentification à l'aide du protocole CHAP (*Challenge Handshake Authentication Protocol*) utilisé conjointement avec PPP. Le principe repose sur le partage d'une clé secrète par les deux routeurs. Lors de l'appel, le routeur appelé envoie un « challenge » au routeur appelant. La bonne réponse est liée à la clé secrète (le mot de passe).

```
hostname orleans

username paris password 7 14041BBEAB04

isdn switch-type vn3

interface Ethernet 0
 ip address 192.168.0.1 255.255.255.0
^Z
interface serial 0
```

Nom du routeur distant qui correspond au " *hostname paris* ".

Clé secrète CHAP (mot de passe)

Le chiffre 7 indique que le mot de passe n'est pas affiché en clair. Attention cependant : il faut saisir le mot de passe en clair lors de la configuration.

```
 ip unnumbered e0
 encapsulation ppp
 backup delay 2 15
 backup interface bri 0
^Z
interface bri 0
 ip unnumbered e0
 encapsulation ppp
 dialer in-band
 dialer map ip 10.0.0.0 name paris broadcast 0144758899
 dialer-group 1
^Z
ip route 10.0.0.0 255.255.252.0 s0
dialer-list 1 protocol ip permit

ip route 10.0.0.0 255.255.252.0 s0
```

Pour joindre le réseau 10.0.0.0, il faut appeler le routeur distant “ paris ”.

Numéro de téléphone RNIS du site parisien

Seul le protocole IP est permis.

Gestion du débordement

La même liaison RNIS peut également être utilisée pour absorber du trafic en surplus sur l’interface série :

```
int bri 0
backup load 80 10
```

Cette commande active l’interface RNIS si la ligne principale atteint une charge égale à 80 %, et la déconnecte lorsque la charge globale (principale + secours) redescend à 10 %.

Le cas de la panne d’un routeur n’est pas traité ici. Pour offrir cette sécurité supplémentaire, d’autres mécanismes doivent être activés (voir VRRP et HSRP au chapitre 12).

10

Architecture des réseaux étendus

Après avoir réalisé notre première interconnexion, nous pouvons passer à l'étape suivante qui consiste à créer un réseau WAN complet.

À l'inverse d'un LAN qui est par essence privé, un WAN nécessite d'emprunter des réseaux publics ou opérateur qui agissent sous licence octroyée par l'État.

À l'inverse de votre LAN sur lequel vous pouvez réaliser des excès de vitesse gratuitement jusqu'au Gigabit, la vitesse est limitée à quelques dizaines de Mbit/s sur les réseaux WAN. Et plus vous allez vite et loin, plus c'est cher.

Il faut donc choisir avec discernement la technologie à utiliser, et déterminer au mieux, en fonction de vos besoins et de votre budget, la route à emprunter (route départementale, voie rapide, autoroute de l'information avec péage, bretelle d'accès, etc.).

Dans ce chapitre, vous apprendrez ainsi :

- à choisir un niveau de service opérateur ;
- à connaître les technologies d'accès xDSL ;
- à connaître les réseaux haut débit SDH et WDM ;
- à dimensionner votre réseau WAN.

Les solutions disponibles sur le marché

Après notre première interconnexion, voilà qu'il faut interconnecter quatre nouveaux sites : Toulouse, Marseille, Strasbourg et Londres.

Quels sont alors les moyens mis à notre disposition ?

Les liaisons point à point

L'élément de base d'un réseau étendu est la **liaison** détenue et gérée par un **opérateur** : France Télécom et Cegetel en France, Deutsche Telekom en Allemagne, ATT et Worldcom (pour ne citer qu'eux) aux États-Unis, etc.

Une liaison peut être **filaire** (cuivre ou fibre optique) ou **hertzienne** (satellites ou émetteurs terrestres). Les liaisons filaires sont obligatoirement **point à point**, tandis que les liaisons hertziennes peuvent être point à point ou **multipoint**.

Les opérateurs louent leurs liaisons entre eux. Ils bâtissent leurs propres réseaux basés sur des liaisons qui leur appartiennent et sur d'autres qu'ils louent là où ils ne peuvent pas en construire.

La construction de liaisons filaires ou hertziennes requiert, en effet, des autorisations administratives (**licences**) pour pouvoir poser les câbles à travers des territoires appartenant à chaque État, lancer les satellites, obtenir les fréquences hertziennes et les exploiter commercialement. En France, l'ART (autorité de régulation des télécommunications) est responsable de l'attribution de ces licences aux opérateurs. Depuis peu, la libéralisation du commerce mondial incite à supprimer les derniers monopoles.

RÉSEAUX LOCAUX, ÉTENDUS ET INTERSITE

Le terme LAN (*Local Area Network*) désigne les **réseaux locaux**, dont le principal représentant est Ethernet. Le terme WAN (*Wide Area Network*) désigne les **réseaux étendus** dont les représentants les plus répandus sont les LS, Frame Relay et ATM.

On désignera par **réseau intersite** un réseau d'interconnexion de réseaux locaux reposant sur un réseau étendu et des **routeurs** qui réalisent l'interface entre les LAN et le WAN.

Au niveau international, l'organisme qui contrôle les activités télécoms (autorisations, projets internationaux, normes, etc.) est l'**ITU** (*International Telecommunication Union*) qui est affilié à l'ONU. Tous les opérateurs nationaux sont membres de l'ITU.

La plupart des liaisons internationales sont le fruit d'une coopération entre les opérateurs qui utilisent les bandes passantes proportionnellement à leur participation financière. Certains opérateurs possèdent en propre leurs liaisons internationales, et les louent à d'autres opérateurs.

Les réseaux opérateur

Les opérateurs utilisent des réseaux à haut débit pour transporter une variété de flux (voix, lignes spécialisées, Frame Relay, ATM, etc.) à des débits divers (de 64 Kbit/s à plusieurs Mbit/s). Leurs clients se raccordent à leur réseau *via* **une boucle locale**, qui généralement appartient et est gérée par un opérateur national historique (France Télécom, Deutsche Telekom, etc.) ou un opérateur local, nouveau venu, comme COLT ou Cegetel.

Le réseau d'un opérateur repose sur des fibres optiques qui souvent ne lui appartiennent pas. Il achète de la bande passante en gros à un autre opérateur, appelé pour cela « carrier de carrier ». Aux extrémités des fibres optiques, dans ses locaux techniques (POP et centre de télécommunications), l'opérateur connecte tous ses équipements SDH, ATM, Frame Relay, routeurs, modems, etc.

À partir des liaisons qui lui appartiennent en propre ou qu'il loue, un opérateur crée un ou plusieurs réseaux interconnectant ses propres sites. Ces liaisons se terminent par des multiplexeurs et des commutateurs (Frame Relay, ATM, SDH ou propriétaires). L'opérateur propose ensuite à ses clients de partager son réseau en leur revendant de la bande passante et en leur proposant un service d'exploitation.

Le but de l'opérateur est alors de partager son réseau par le plus grand nombre de clients possible, tout en séparant ces derniers. Le partitionnement logique du réseau est réalisé grâce aux **circuits virtuels**, qui permettent de constituer autant de **VPN** (*Virtual Private Network*) qu'il y a de clients. Le VPN est donc la base de l'offre de service des opérateurs réseau.

QU'EST-CE QU'UN CIRCUIT VIRTUEL ?

Les circuits virtuels sont des **liaisons logiques**, généralement point à point, se partageant un même support physique. Les données appartenant à un même CV ne peuvent être échangées avec celles des autres CV. Grâce aux CV, un réseau physique peut être **partitionné** en plusieurs réseaux logiques indépendants et étanches les uns des autres, créant ainsi ce qu'on appelle des **VPN** (*Virtual Private Network*). Pour l'utilisateur, le VPN se comporte comme s'il était seul sur le réseau alors que ce dernier est partagé.

Une entreprise désirant interconnecter ses sites, ou un particulier désirant communiquer à distance (téléphoner, se connecter à Internet ou à son intranet, etc.) devra faire appel aux services d'un opérateur, ou, si cela est économiquement plus intéressant, creuser ses propres tranchées pour y poser ses propres câbles afin de construire une infrastructure privative sur le domaine public, comme le font les opérateurs.

Le premier des réseaux opérateur est historiquement le **RTC** (ou réseau téléphonique commuté) qui permet aux entreprises et aux particuliers de communiquer par téléphone moyennant un abonnement et une facturation à la durée.

Le second type qui est apparu est la **LS** (ligne spécialisée, encore appelée ligne louée), qui permet aux entreprises d'interconnecter leurs sites moyennant un abonnement mensuel dont le coût dépend du débit utilisé et de la distance.

Aujourd'hui, les réseaux sont très variés : boucles optiques SDH, réseaux optiques Ethernet natifs, réseaux IP-MPLS sur fibre optique WDM, etc.

L'accès aux réseaux des opérateurs

Les sites et les liaisons qui constituent le réseau de l'opérateur peuvent être plus ou moins nombreux, et la **zone de couverture** de ce réseau peut être plus ou moins étendue.

Chaque client doit se raccorder au réseau de l'opérateur *via* des **liaisons d'accès** (filaires le plus souvent, ou hertziennes pour le téléphone mobile, par exemple) entre son site et les points d'accès appelés **POP** (*Point Of Presence*) de l'opérateur. La liaison d'accès est également appelée desserte locale ou, plus spécifiquement, **boucle locale**.

L'intérêt d'une telle solution est que le POP soit le plus près possible du site client afin que la liaison de raccordement, dont le prix dépend du débit et de la distance, coûte le moins cher possible. Le RTC est le réseau opérateur qui dispose du plus important nombre de POP : il s'agit d'un commutateur installé au coin de la rue, dans un immeuble. Mais, pour les autres réseaux, il n'en est pas de même. Par exemple, vous avez sans doute consulté les zones de couverture d'Itinéris, de Bouygues Télécom, de SFR, etc.

L'opérateur installe toujours un équipement d'extrémité, appelé **CPE** (*Customer Premises Equipment*), dans les locaux du client qui permet de gérer la liaison d'accès. Il s'agit, par exemple, d'un modem, d'un commutateur, etc., selon le niveau de service fourni.

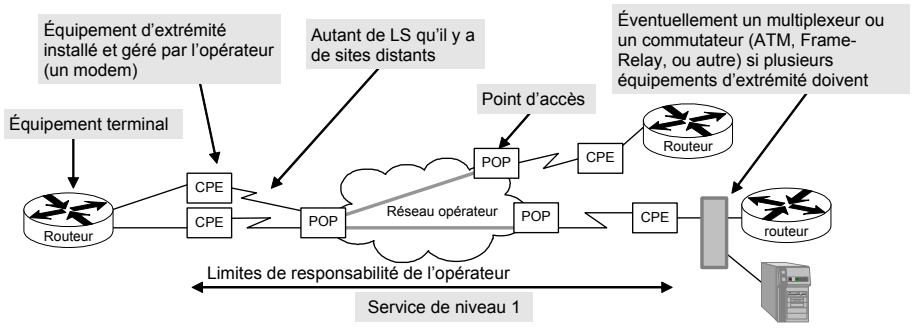
Les services proposés par les opérateurs

L'opérateur propose à ses clients d'utiliser son réseau en leur revendant de la bande passante ainsi que différents niveaux de service d'exploitation. À la diversité des technologies s'ajoute le maquis des services proposés, l'habillage commercial en quelque sorte. Bien qu'il en existe de nombreuses variantes, on trouve principalement trois types de service.

Niveau de prestation	Description technique	Service fourni	Exemples
1. Fourniture du support de transmission brut (couche physique)	Liaisons (le plus souvent point à point) pour créer un réseau privé	Support client de la liaison	Lignes spécialisées, liaisons VSAT
2. Réseau fédérateur (couches physique et liaison)	Création d'un réseau privé (le plus souvent multipoint) au-dessus de l'infrastructure opérateur	Réseau fourni et exploité par l'opérateur + support client	Réseau ATM, Frame Relay ; accès à Internet
3. Service à valeur ajoutée (couches physique, liaison et réseau)	Support de transmission + réseau fédérateur + équipements terminaux (routeur, téléphone, etc.)	Réseau de bout en bout (jusque dans le site du client) fourni et exploité par l'opérateur	RTC, RNIS, interconnexion de LAN, accès à Internet

Le service de niveau 1 permet aux clients de disposer du support de transmission qui leur est fourni : ils peuvent installer des multiplexeurs, des commutateurs ATM ou Frame Relay, des routeurs, etc., pour transmettre des données, de la voix ou de la vidéo.

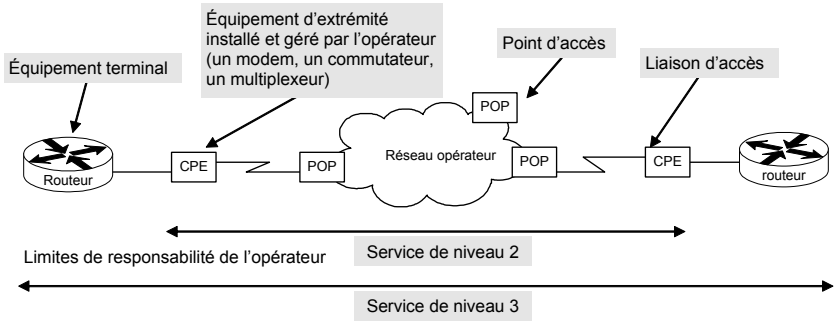
Figure 10-1. *Service opérateur de niveau 1 – lignes spécialisées.*



Le service de niveau 2 consiste à profiter de l'infrastructure réseau de l'opérateur, qui est vu comme un nuage (un réseau multipoint), sur lequel les sites du client sont raccordés *via* des liaisons locales et les POP.

Le service de niveau 3 est un service complet : l'opérateur propose un service clé en main *via* un réseau quelconque, avec les moyens et les technologies souhaités. Le client ne voit que le service : téléphone, interconnexion de réseaux locaux, etc.

Figure 10-2. *Services opérateur de niveaux 2 et 3.*



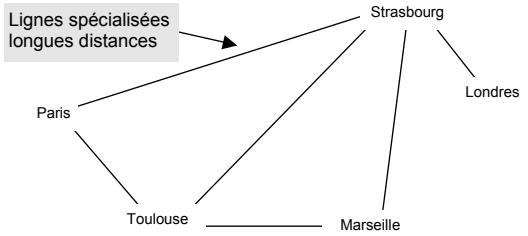
Les choix du client

Pour créer notre réseau étendu, la première question concerne le niveau de service que nous allons demander à l'opérateur. Du point de vue client, nous avons le choix entre deux types de solutions :

- **Solution privée** reposant sur le service de niveau 1. L'entreprise construit son réseau étendu.
- **Solution opérateur** reposant sur les services de niveaux 2 et 3. L'entreprise confie tout ou partie de son réseau étendu à un opérateur.

La première solution consiste à suivre le modèle de notre première interconnexion, afin de réaliser notre propre réseau multipoint reposant sur des lignes spécialisées point à point.

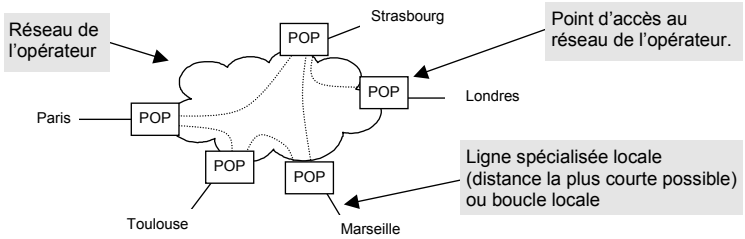
Figure 10-3.
Réseau privé reposant sur des LS.



Le réseau peut être en étoile sur un site principal, distribué selon l'importance relative de chaque site, ou plus ou moins maillé afin de répartir la charge et d'assurer le secours des liaisons.

La seconde solution repose sur l'utilisation d'un réseau opérateur et sur des LS ou, de préférence, des boucles locales raccordant les sites du client aux POP les plus proches.

Figure 10-4.
Réseau privé virtuel basé sur un réseau opérateur.



L'opérateur installe les liaisons d'accès locales (ou les commande auprès de l'opérateur local s'il n'a pas de licence) entre les sites du client et ses points d'accès.

Pour compliquer le choix des clients, il existe plusieurs types d'offres :

- Les **VPN de niveau 2** reposant sur **Frame Relay** et **ATM** (voir chapitre 11). L'opérateur garantit au client la bande passante demandée ainsi que le niveau de service souhaité, avec des engagements de résultats.
- Les **VPN de niveau 3**, reposant sur IP et **MPLS** (voir chapitre 12). De la même manière, l'opérateur s'engage sur une qualité de service équivalente, voire supérieure, à celle pouvant être offerte sur un VPN de niveau 2.
- Les **VPN IPsec**, également de niveau 3, reposant sur le protocole de cryptage IPsec (voir chapitre 17). Dans ce cas, il n'y a pas séparation logique, mais création d'un tunnel crypté qui garantit la confidentialité des informations vis-à-vis des autres utilisateurs. Si le réseau IP est celui de l'opérateur, celui-ci peut s'engager sur une qualité de service, généralement de niveau inférieur aux deux précédents VPN. Si le réseau IP est l'Internet, l'opérateur ne s'engage quasiment sur aucune qualité de service.

Dans tous les cas, le client doit signer un contrat de service, appelé **SLA** (*Service Level Agreement*), avec l'opérateur. Ce dernier s'engage à fournir une qualité de service (temps de réponse, débit garanti, taux de disponibilité, etc.) ainsi qu'à payer des pénalités au client en cas de non-respect de ses engagements.

Le réseau opérateur peut s'engager sur différents services :

- Taux de disponibilité. L'opérateur garantit que son réseau sera disponible à 99,9 % du temps (c'est une valeur courante).
- Bande passante. L'opérateur garantit que le client disposera du débit demandé (512 Kbit/s entre deux sites, par exemple) pendant 100 % du temps.
- Temps de transit. L'opérateur garantit le temps mis par un paquet pour aller d'un site à l'autre.

Ces garanties peuvent être valables de POP à POP (c'est-à-dire dans le réseau fédérateur de l'opérateur appelé backbone) ou de bout en bout (c'est-à-dire entre les sites du client), liaisons et équipements d'accès compris. C'est à vous de négocier les engagements en fonction du coût.

L'opérateur exploite son réseau à l'aide d'un ou plusieurs centres de supervision, de centres de support client (*Help Desk*) et d'équipes projet qui lui permettent d'offrir un service de guichet unique (interlocuteur unique) :

- L'opérateur se charge de toutes les commandes de LS d'accès auprès des opérateurs locaux, de l'installation des équipements, de la configuration de son réseau pour accueillir le VPN du client ainsi que de la gestion du projet.
- Quels que soient le site et le pays, le client peut appeler le centre de supervision pour lui signaler un problème, et inversement.
- Il n'y a qu'une facture, et le client choisit le mode de facturation qu'il souhaite : centralisée, répartie par site, etc.
- L'opérateur fournit des statistiques sur le VPN du client et les preuves que ses engagements ont été respectés. Cela n'empêche pas le client de mettre en place son propre système de contrôle.

Critère	Solution privée	Solution opérateur
Investissements	Importants (frais de mise en service et achat des équipements)	Faibles (frais de mise en service) ; pas d'achat d'équipement
Coûts de fonctionnement	Assez faibles en national, élevés à l'international	Assez élevés (coût des LS plus faible et forfait opérateur)
Exploitation	Assurée par la société (vous)	Assurée par l'opérateur (supervision des LS, exploitation, etc.)
Distances des LS	Longues entre les sites	Locales entre les sites et les POP de l'opérateur

Une entreprise a toujours la possibilité de passer d'une solution privée à une solution opérateur, autrement dit d'**externaliser** son réseau (ce que l'on appelle l'**outsourcing**).

L'opération inverse (la **réversibilité**) n'est pas dans l'air du temps et est toujours une opération délicate. La réversibilité peut aussi être appliquée pour changer d'opérateur en cas d'insatisfaction.

Les lignes spécialisées

Une LS (ligne spécialisée ou ligne louée) est un terme générique pour désigner une liaison point à point entre deux sites.

Pour l'utilisateur, une LS consiste en deux ou quatre fils (plus rarement six) de cuivre reliant deux de ses sites. Pour l'opérateur, ce n'est qu'une liaison d'accès à son réseau : la liaison cuivre qui part du site du client aboutit à un multiplexeur qui concentre d'autres LS (issues d'autres clients) sur des liaisons haut débit en fibre optique.

Les LS reposent sur deux technologies :

- T1 (États-Unis, Canada et Japon) et E1 (reste du monde), qui datent des années 60.
- HDSL (*High bit rate Digital Subscriber Line*), qui date des années 80.

Dénomination	Technologie	Codage en ligne	Distance
E1	MIC (*modulation par impulsions codées*) 32 canaux de 64 Kbit/s	Bipolar AMI (*Alternate Mark inversion*)	De 1 à 2 km sans répéteur 2 paires
T1	PCM (*Pulse Code Modulation*) 24 canaux de 64 Kbit/s	Codage bipolar AMI	De 1 à 2 km sans répéteur 2 paires
HDSL	DSL Trames transmises en 6 ms	2B1Q (2 bits / 1 signal quaternaire)	De 3,7 à 7,9 km sans répéteur 1, 2 ou 3 paires

Le débit offert par une E1/T1 est un multiple de 64 Kbit/s. Cette unité correspond à un canal dans un lien E1 à 2,048 Mbit/s (32 canaux à 64 Kbit/s) ou T1 (*Trunk-carrier, level 1 multiplexing*) à 1,544 Mbit/s (24 canaux à 64 Kbit/s + 8 Kbit/s de signalisation).

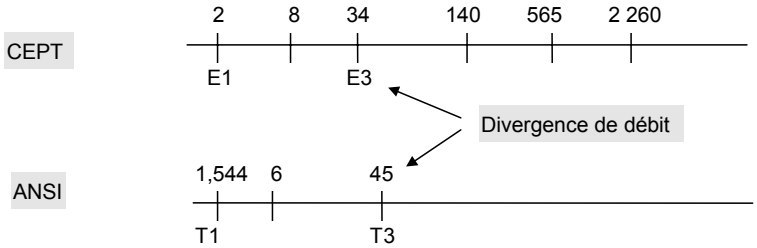
Les ondes sonores de la voix sont converties en signaux numériques à un débit de 64 Kbit/s, tout comme le mouvement est converti en 24 images/seconde par une caméra. Ce débit est lié aux limitations technologiques des années 60, et demeure encore l'unité de référence. Les progrès font que, aujourd'hui, on peut se contenter de 8 Kbit/s, voire moins.

L'unité de base est donc constituée d'un **canal** de 64 Kbit/s (dénommé DS0, *Digital Signaling 0*). Comme il est plus pratique et plus économique de transporter plusieurs canaux en même temps, ces derniers sont multiplexés au sein d'une liaison composite (*trunk*). Et, comme dans toutes les couches réseau (nous sommes ici au niveau physique), un protocole est nécessaire pour gérer ces canaux (début et fin du canal, synchronisation des horloges, etc.). Un canal est donc dédié à ce protocole souvent appelé canal de **signalisation**.

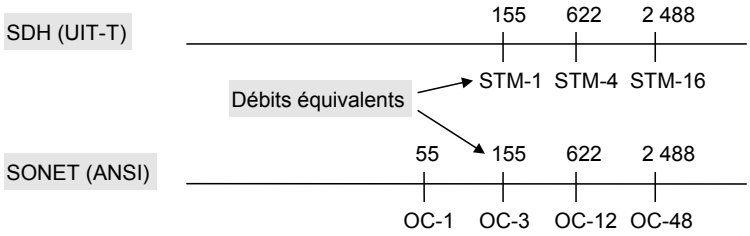
Liaison	Canaux de données utiles	Canaux de signalisation
RNIS (réseau téléphonique numérique)	2 canaux à 64 Kbit/s	1 canal à 16 Kbit/s
E1 en Europe : 2 048 Kbit/s	30 canaux à 64 Kbit/s	2 canaux à 64 Kbit/s
T1 aux États-Unis : 1 544 Kbit/s	24 canaux à 64 Kbit/s	1 canal à 8 Kbit/s

Ces échanges point à point sont réalisés entre deux multiplexeurs qui dialoguent *via* le canal de signalisation (des bits prélevés sur le débit global).

Un multiplexeur prend ainsi *n* canaux à 64 Kbit/s en entrée, et génère un signal de 2 048 Mbit/s en sortie pour une E1. Un autre type de multiplexeur prend 4 canaux à 2 Mbit/s en entrée, et génère un signal à 8 Mbit/s en sortie, et ainsi de suite. On définit ainsi une hiérarchie de débits **plésiochrones** (plusieurs horloges, une pour chaque type de multiplexeur) jusqu'à 34 Mbit/s en Europe et 45 Mbit/s aux États-Unis.



Aujourd'hui, cette cascade de multiplexage a été remplacée par des multiplexeurs permettant d'extraire directement la bande passante souhaitée. On définit ainsi une hiérarchie de débits **synchrones** (une seule horloge pour transporter plusieurs débits). La structure des trames et le protocole associé sont **SONET** (*Synchronous Optical NETwork*) aux États-Unis, et **SDH** (*Synchronous Digital Hierarchy*) en Europe. Ces réseaux forment le cœur des infrastructures réseau haut débit des opérateurs (de 51,84 Mbit/s à plus de 9 Gbit/s).



Équipements installés chez l'utilisateur

L'opérateur installe dans les locaux du client un **CSU/DSU** (*Channel Service Unit/Data Service Unit*) qui génère et reçoit les signaux sur la LS. Le CSU/DSU est l'équivalent numérique du modem : alors que ce dernier convertit les signaux analogiques en signaux numériques et inversement, le CSU/DSU convertit les signaux numériques des interfaces locales

(V.35, X21/V11, etc.) en signaux numériques adaptés aux longues distances, xDSL par exemple.

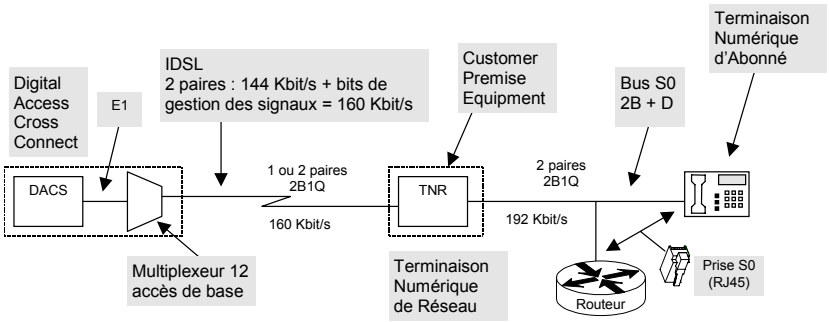
Le **CPE** (*Customer Premises Equipment*) est le terme générique désignant un équipement de raccordement installé chez le client. Il s'agit d'un CSU/DSU (équipement de base), d'un commutateur ATM ou Frame Relay, d'un FRAD ou d'un routeur. Ces derniers peuvent ou non intégrer un CSU/DSU.

Les technologies DSL

Les technologies PCM/MIC utilisées par les liaisons T1 et E1 depuis les années 60 sont aujourd'hui dépassées. Les lignes spécialisées offrent toujours une interface E1 ou T1 à leurs clients, mais le PCM a fait place aux technologies DSL (*Digital Subscriber Line*). Celles-ci utilisent des codages plus performants et des processeurs spécialisés dans le traitement du signal, les DSP (*Digital Signaling Processing*).

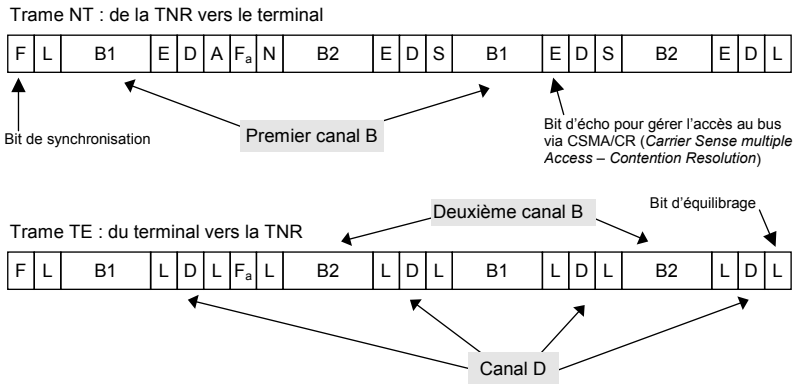
La première de ces technologies est **IDSL** (*Integrated Service Digital Network DSL*), couramment appelée RNIS (réseau numérique à intégration de service) et commercialisée sous le nom de Numéris par France Télécom. Le RNIS est la base du réseau téléphonique numérique ; il a été la première technologie numérique à être proposée aux particuliers. L'accès de base T0 offre deux canaux B de 64 Kbit/s, chacun pouvant être utilisé séparément (deux communications) ou agrégé pour offrir un débit de 128 Kbit/s. Un canal D de 16 Kbit/s est réservé à la signalisation (numérotation, réveil du terminal, etc.). Des bits supplémentaires permettent de gérer les signaux transmis sur le câble.

Figure 10-5.
L'accès de base IDSL.



Côté client, le RNIS se présente sous la forme d'un bus, appelé bus S0, sur lequel plusieurs équipements peuvent être connectés (téléphones, télécopieurs, routeurs, etc.).

Figure 10-6.
Les trames IDSL (accès de base S0).



La deuxième technologie qui est apparue est **HDSL** (*High bit rate DSL*) sur laquelle reposent quasiment toutes les liaisons d'accès (LS notamment). La raison est qu'une liaison E1/T1 nécessite de coûteux répéteurs tous les 1 à 2 km, alors que la portée du HDSL est d'au moins 3,7 km (jusqu'à 7,9 km sans répéteur). La réduction du nombre de répéteurs permet ainsi de réduire les coûts des LS de 30 à 50 % par rapport aux E1/T1.

La vraie révolution des technologies xDSL est que cette baisse de prix permet aux opérateurs de proposer des liaisons xDSL aux particuliers pour leurs connexions téléphonique et Internet.

Figure 10-7.
Accès E1 via HDSL.

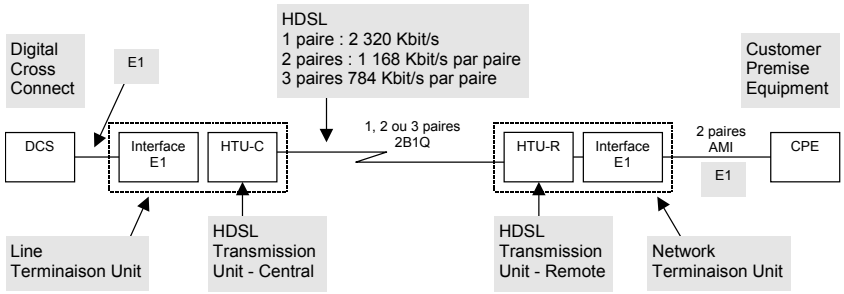
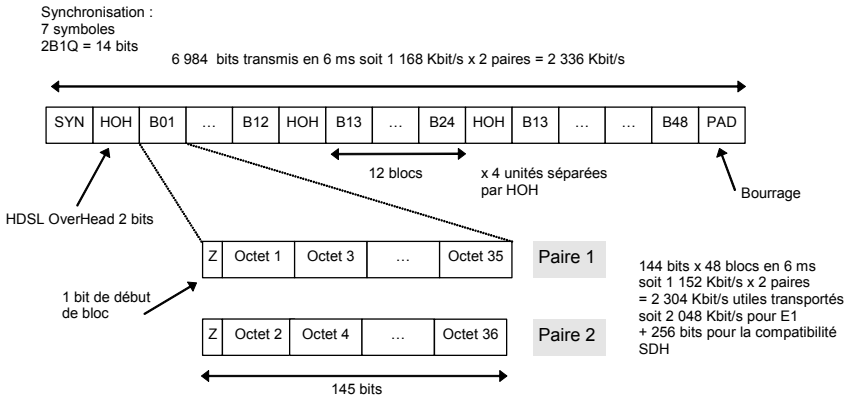


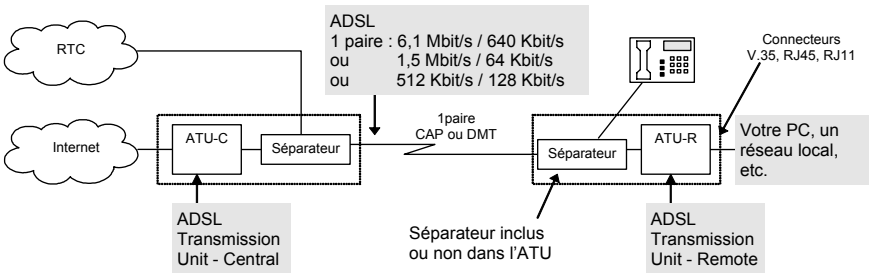
Figure 10-8.
Trames HDSL sur 2 paires.



S-HDSL (*Single pair HDSL*) et **SDSL** (*Symetric DSL*) sont des versions simplifiées de HDSL.

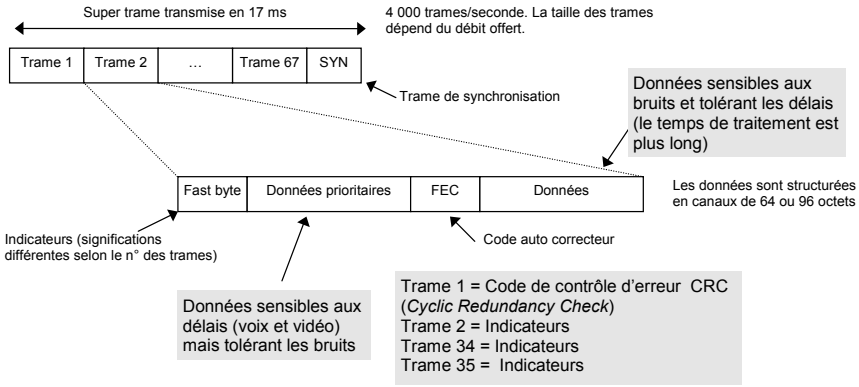
La norme DSL la plus avancée est **ADSL** (*Asymetric DSL*). Sa particularité est de présenter des débits différents selon le sens de la transmission : le débit à destination du client est plus important que celui offert à ce dernier en émission. Ce type d'accès est associé à un séparateur offrant un accès analogique pour les téléphones classiques en plus de l'accès numérique. Cela en fait une utilisation appropriée pour les particuliers.

Figure 10-9. *Accès ADSL pour les particuliers.*



Il est à noter que les TNR, HTU et ATU qui viennent d'être évoqués aux figures précédentes sont des modems numériques (CSU/DSU).

Figure 10-10. *Format d'une trame ADSL.*



RADSL (*Rate Adaptative DSL*) est une version d'ADSL qui permet d'augmenter ou de diminuer le débit selon ce que permet la qualité de la ligne. C'est en fait une évolution naturelle d'ADSL.

ADSL Lite (*G.Lite*), anciennement **CDSL** (*Customer DSL*)**,** est l'équivalent d'ADSL sans le séparateur (*splitter*). L'objectif est de diminuer les coûts en simplifiant l'installation de l'abonné (suppression du séparateur) ainsi que l'exploitation pour l'opérateur. Les fonction-

nalités sont les mêmes que celles fournies par ADSL/RADSL (connexion Internet + téléphone).

La dernière technologie en date est le **VDSL** (*Very high speed DSL*) qui utilise les paires torsadées en cuivre mais surtout la fibre optique. L'objectif premier de VDSL est de transporter des cellules ATM (*Asynchronous Transfer Mode*). La norme repose sur un multiplexage temporel (TDM, *Time Division Multiplexing*) aux débits normalisés STM (*Synchronous Transfer Mode*), ceux utilisés par SDH (*Synchronous Digital Hierarchy*).

Technologie	Débit descendant / Montant en Kbit/s	Distance maximale sans répéteur	Nombre de paires	Codage
HDSL – G.991.1 (*High bit rate DSL*)	1 544 / 1 544 2 048 / 2 048	3,7 km (24 AWG) 2,7 km (26 AWG)	2 1 à 3	**2B1Q**
ADSL – G.992.1 (*Asymetric DSL*)	1 544 / 512 6 144 / 640	5,5 km (24 AWG) 3,7 km (24 AWG)	1 1	**DMT**
RADSL (*Rate Adaptative DSL*)	1 544 / 512 6 144 / 640	4,6 km (26 AWG) 2,7 km (26 AWG)	1 1	**DMT**
ADSL lite – G.992.2 (*G.lite*)	1 544 / 512	5,5 km (24 AWG)	1	**DMT**
VDSL (*Very high speed DSL*)	13 000 / 1 500 52 000 / 2 300	1,3 km (24 AWG) 0,3 km (24 AWG)	1 ou FO	**DWMT / SLC**
S-HDSL – G.991.2 (*Symetric HDSL*)	768 / 768	3,7 km	1	**2B1Q**
SDSL (*Single pair HDSL*)	1 544 / 1 544 2 048 / 2 048	3 km 3 km	1 1	**2B1Q / CAP**
IDSL (*ISDN DSL*)	160 / 160		1 ou 2	**2B1Q**

Le débit descendant correspond aux flux allant du réseau (par exemple l'Internet) vers le client (vous), tandis que le débit montant correspond au flux allant du client vers le réseau.

Au sein d'une même norme, les débits et distances varient en fonction du diamètre des fils utilisés (norme AWG, *American Wire Gauge*).

Diamètre des fils	Distance maximale
22 AWG = 0,65 mm	7,9 km (26 000 pieds)
24 AWG = 0,5 mm	5,5 km (18 000 pieds)
26 AWG = 0,4 mm	3,7 km (12 000 pieds)

Les débits peuvent, par ailleurs, être augmentés si les distances sont raccourcies.

Codage utilisé par xDSL	Brève description
DMT (*Discrete Multi-Tone*)	Repose sur les transformations de Fourrier pour gérer et démoduler 256 sous-canaux (sous-porteuses).
DWMT (*Discrete Wawelet Multi-Tone*)	Repose sur une fonction mathématique, les ondelettes, plus performante que les transformations de Fourrier.
SLC (*Simple Line Code*)	Codage en bande de base à 4 niveaux.
QAM – 16 (Quadrature Amplitude Modulation)	2 amplitudes et 12 changements de phase permettent d'obtenir 16 signaux différents représentant 4 bits de données.
CAP (Carrierless Amplitude Phase)	Analogue à QAM, mais sans générer de porteuse.
PAM (*Pulse Amplitude Modulation*)	2B1Q est un exemple de code PAM à 4 niveaux : 2 bits de données sont codés en un signal quaternaire (4 niveaux électriques).

Les applications du xDSL

Suite à la libéralisation du marché des télécoms, l'enjeu commercial qui aiguise le plus les appétits est celui de la **boucle locale**, c'est-à-dire la liaison d'accès aux réseaux des opérateurs. HDSL couvre les besoins des entreprises en matière de réseau étendu, tandis que les besoins des particuliers (téléphone, télévision, accès à Internet – c'est-à-dire les services résidentiels) sont couverts par ADSL/RADSL/CDSL (G.lite).

Technologie	Applications / Marché visé
IDSL (accès de base RNIS, 2B+D)	Pour les entreprises Téléphonie numérique, accès à Internet et à l'intranet Interconnexion des réseaux locaux : accès principal si les temps d'utilisation sont faibles, secours d'une liaison principale et débordement en cas de surcharge de la ligne principale
HDSL	Accès primaire RNIS (2 Mbit/s) Liaison d'accès E1/T1 de 64 Kbit/s à 2 Mbit/s
ADSL, RADSL, CDSL	Services résidentiels : téléphone, télévision, vidéo à la demande, connexion à Internet
VDSL	Réseaux haut débit ATM sur cuivre et surtout sur fibre optique

Concernant nos réseaux locaux, PPP fonctionne directement au-dessus des LS (donc de HDSL) et de RNIS (IDSL) grâce aux routeurs qui prennent en compte les interfaces E1/T1 et S0.

Quant à ADSL, quatre modes d'accès aux canaux de données sont possibles :

- Mode **synchrone**. Les canaux sont accessibles sous forme d'un train de bits aux débits STM.
- Mode **adaptation de paquet**. Permet à plusieurs applications d'utiliser les canaux pour transporter leurs données selon leurs propres formats.
- Mode **paquet de bout en bout**. Les paquets IP sont envoyés directement dans les trames (sans se soucier de l'affection des canaux), et la commutation est effectuée au niveau d'ADSL sur la base des adresses IP contenues dans les paquets.
- Mode **ATM**. Permet de transporter les cellules ATM qui contiennent les trames PPP (qui transportent IP, et ainsi de suite).

Actuellement, aucune tendance n'est perceptible, car le marché est naissant et les équipements peu nombreux. En France, France Télécom promeut le mode ATM.

Les boucles SDH

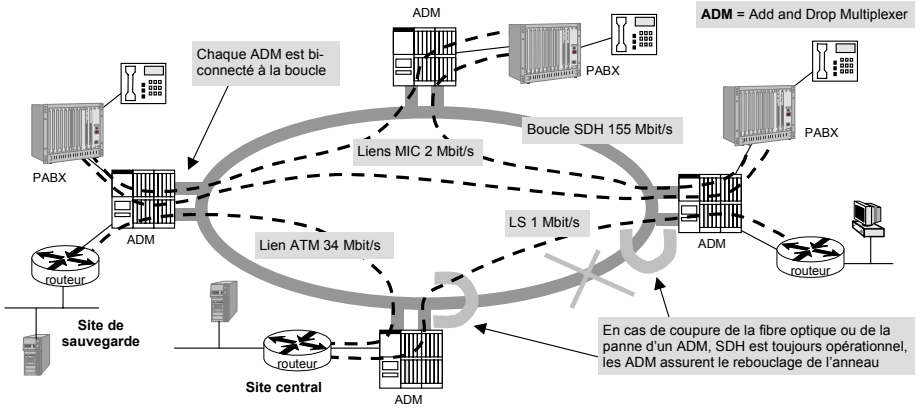
Dès l'origine des télécommunications, le premier objectif a été de partager une même liaison physique entre plusieurs flux comme on vient de le voir avec xDSL. Ce principe, appelé multiplexage, consiste à découper en tranches de temps simples (multiplexage temporel), puis à structurer ces tranches de temps en des formats de plus en plus évolués pour créer des trames.

La LS et la boucle locale ne concernent qu'un utilisateur, mais déjà sur cette portion de réseau, différents flux sont multiplexés : téléphonie et Internet si l'on prend l'exemple d'une boucle ADSL. Si l'on avance au cœur des réseaux des opérateurs, les tuyaux deviennent de plus en plus gros et généralistes, c'est-à-dire qu'ils doivent véhiculer des flux aussi divers que de la téléphonie et des données (lignes spécialisées, ATM, Frame Relay, ADSL, etc.), et à des débits très variés, allant de 64 Kbit/s à plusieurs centaines de Mbit/s.

La technique de multiplexage aujourd'hui utilisée par les opérateurs pour transporter tous ces flux à tous ces débits est SDH (*Synchronous Digital Hierarchy*), et son équivalent américain SONET (*Synchronous Optical Network*), sur de la fibre optique.

Et qui dit fibre optique, dit distances importantes, de l'ordre de plusieurs dizaines de kilomètres. SDH permet donc de constituer des réseaux à haut débit à l'échelle d'une agglomération. Il permet ainsi de véhiculer des lignes téléphoniques (canaux MIC, niveau 1), de créer des boucles locales (LS, ATM et Frame Relay, niveau 2), et de constituer des MAN (réseaux à haut débit). Les opérateurs mettent à profit cette technologie pour raccorder leurs grands clients à leur POP (points de présence).

Des boucles à 155 Mbit/s, 622 Mbit/s ou à 2,5 Gbit/s sont commercialisées auprès des grandes entreprises, leur permettant de réaliser ces mêmes fonctions pour leurs besoins propres, comme l'interconnexion à haut débit (2, 34, 45, 155 ou 622 Mbit/s) de leurs sites répartis autour d'une grande ville.



Dans cet exemple, la boucle SDH véhicule des liens MIC à 2 Mbit/s entre les PABX, des LS à 1 Mbit/s entre les routeurs, et une liaison ATM à 34 Mbit/s entre les deux principaux sites.

Sonet	SDH	Débit en Mbit/s
OC-1	--	51,84
OC-3	STM-1 (*)	155,52
OC-9	STM-3	466,56
OC-12	STM-4 (*)	622,08
OC-18	STM-6	933,12
OC-24	STM-8	1 244,66
OC-36	STM-12	1 866,24
OC-48	STM-16 (*)	2 488,32
OC-96	STM-32	4 976,64
OC-192	STM-64	9 953,28

(*) Débits standardisés et commercialisés par les opérateurs.

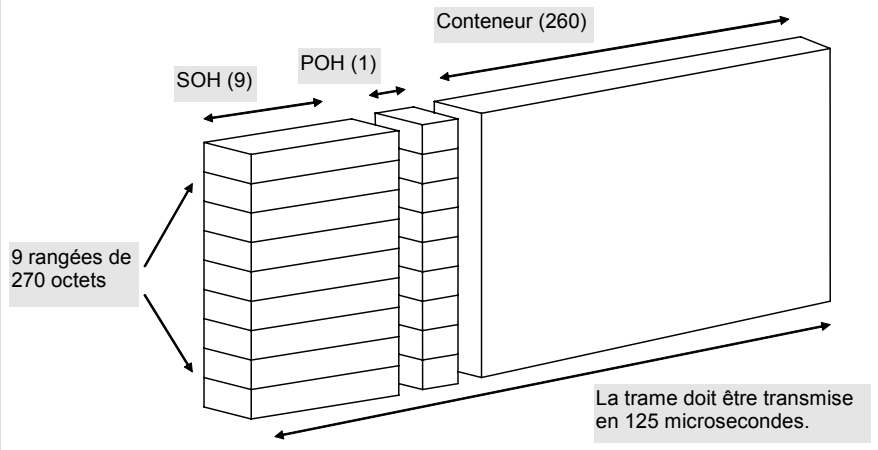
LE POINT SUR SDH (ITU-T G.709)

Une trame SDH se compose d'un en-tête appelé section de surdébit (SOH pour *Section OverHead*) de 81 octets, qui contient toutes les informations relatives à la structure de la trame ainsi que les données relatives à la signalisation Q.93B (norme G.771) et à l'administration. Au sein de la trame, les données utiles sont placées dans des conteneurs virtuels (norme G.709) adaptés à leur débit. Ces données peuvent être des cellules ATM, ou correspondre aux débits des réseaux plésiochrones (normes G.703 et G.704), tel celui à 2,048 Mbps.

Un conteneur est composé des données utiles et d'un en-tête de conduit (POH pour *Path OverHead*) de 9 octets qui assure notamment les fonctions de contrôle de la parité et d'identification du type des données transportées.

L'en-tête SOH comporte un pointeur permettant au débit des données transportées de varier dans le temps. Les données utiles peuvent ainsi « flotter » à l'intérieur du conduit qui leur est alloué. Cette technique permet d'adapter le débit lorsque les horloges de deux nœuds sont différentes.

À titre d'exemple, la section de surdébit occupe 5,184 Mbit/s sur une liaison STM-1, laissant 150,336 Mbps de bande passante pour les données utiles. Des débits supérieurs sont formés par entrelacement de trames STM-1. Dans le cas des trames STM-4 et STM-16, 4 ou 16 flux de cellules ATM à 150 Mbit/s sont transportés dans un conteneur de type 4.



Le SOH est constitué de 9 rangées de 9 octets, le POH de 9 rangées de 1 octet et les données utiles de 9 rangées de 260 octets. La transmission s'effectue de gauche à droite en partant de la rangée du haut. Cette technique d'entrelacement permet de limiter les conséquences d'une erreur de transmission, le raisonnement étant qu'il vaut mieux perdre 1 octet tous les N octets que N octets successifs.

Une trame STM-3 est composée de trois trames STM-1 entrelacées. L'en-tête est constitué de 3 x 9 octets (toujours sur 9 rangées), puis de 3 x 261 octets. Toutes les trames STM-n sont transportées en 125 ms. À titre de comparaison, la trame Sonet est composée de 9 rangées de 90 octets transmises en 125 ms.

Le multiplexage WDM

Le multiplexage proposé par SDH est cependant un peu rigide de par la taille incrémentale des conteneurs ; on passe, par exemple, d'un débit de 622 Mbit/s (STM-4) à 2,5 Gbit/s (STM-16).

Les nouvelles techniques de multiplexage sur fibre optique monomode s'appellent **WDM** (*Wavelength Division Multiplexing*) et **DWMD** (*Dense WDM*) dont le principe consiste à générer plusieurs faisceaux lumineux dans la même fibre, jusqu'à 256 actuellement. Chaque faisceau optique, indifféremment appelé canal, longueur d'onde ou lambda, transporte un flux de données de nature et **de débit quelconque**, alors que SDH multiplexe les données en entrée puis les transmet sur une seule longueur d'onde. Cependant, WDM ne prend que des signaux optiques en entrée, alors que SDH multiplexe des signaux d'entrée optique et non optique.

WDM peut ainsi transporter en même temps des signaux optiques OC-3, ATM, Gigabit Ethernet et même SDH, puisque ce dernier génère un signal optique. WDM permet donc de multiplier aisément les capacités des fibres optiques déjà existantes, sans remettre en question l'existant.

Comparaison	WDM	SDH / SONET
Topologie	Bus point à point ou boucle	Boucle
Signaux en entrée	Optiques	Optiques et électriques
Faisceaux optiques	Plusieurs centaines d'ondes	Une onde
Régénération du signal	Amplificateur optique	Régénérateur opto-électrique
Débit transporté	Débit quelconque par longueur d'onde en fonction du débit entrant	Débits par palier : STM-1, STM-3… STM-64.
Transport d'un lien Gigabit Ethernet	Une longueur d'onde de 1 Gbit/s	Trame STM-16 (2,48 Gbit/s) : 59,8 % de bande passante inutilisée
Distances maximales (dépend du débit)	Plusieurs centaines de km à quelques milliers	Plusieurs dizaines de km à quelques centaines

On le voit, toutes ces techniques de multiplexage peuvent s'imbriquer : ATM dans SDH ou dans WDM, SDH dans WDM, bien que nous soyons toujours au niveau physique, c'est-à-dire au niveau du multiplexage des signaux générés sur un câble.

Dimensionner les liaisons

Quelle que soit la solution retenue, le choix du débit est important, car il influe directement sur le coût des liaisons. Celui-ci est d'ailleurs d'autant plus élevé que la distance entre les deux sites est grande. Par exemple, une simple liaison 64 Kbit/s entre la France et la Chine coûte plusieurs milliers d'euros par mois. S'il s'agit d'une liaison locale vers un POP de

l'opérateur, le coût est moindre, mais ce dernier vous facturera de toute façon en fonction du débit.

Il ne faut donc pas surévaluer le débit par rapport à vos besoins, afin d'éviter de payer un surcoût inutile. Il ne faut pas non plus le sous-évaluer, car les utilisateurs exigent des temps de réponse corrects. La conception d'un **réseau intersite** (réseau d'interconnexion de réseaux locaux) résulte donc d'un compromis coûts/performances.

La démarche proposée pour dimensionner les liens repose sur trois étapes :

- identifier les flux générés par les applications ;
- estimer la volumétrie, soit à partir de statistiques existantes (facturation, traces relevées sur les équipements, etc.), soit à partir d'hypothèses ;
- déterminer une formule permettant de calculer le débit nécessaire.

Identifier les flux

Nos utilisateurs sont répartis sur six sites (le siège — Paris — et les directions régionales : Orléans, Toulouse, Marseille, Strasbourg et Londres). Ils veulent utiliser les mêmes applications et accéder aux mêmes données.

Le but de cette phase est de caractériser les flux de chaque application (type, périodicité) et d'identifier les acteurs qui émettent et reçoivent les données.

Étape 1 — Identifier les flux

Application	De ↔ vers	Objet	Type de flux	Système	Périodicité
Base de données	Directions régionales ↔ siège	Comptabilité, logistique	Client-serveur	Unix	Mise à jour : TLN* Consultation : TLJ*
Datawarehouse	Directions régionales ↔ siège	Activité commerciale	Transfert de gros fichiers	Unix	Mise à jour : TLN Consultation : TLJ
Messagerie	Intrasite et intersite		Transfert de fichiers (Word, Excel, applications métier)	Exchange SMTP	Toutes les 10 minutes entre MTA (un MTA par site)
Télécopie *via* la messagerie	Tous les sites	Échanges externes	Messagerie	Exchange	TLJ
Intranet	Tous les sites	Informations, DRH, accès aux bases de données	Transactionnel Client-serveur	Unix, NT	TLJ
Connexions Internet	Directions régionales ↔ Siège	Consultation Web messagerie	Transactionnel, transfert de fichiers	Accès Internet situé au siège	TLJ

*TLJ = tous les jours *TLN = toutes les nuits

Cette vision synthétique est une étape vers la traduction du langage utilisateur en langage informatique. C'est aussi un bon moyen de décrire les flux circulant au sein de la société (*workflow*) afin de bâtir le réseau qui lui soit le mieux adapté.

Les flux recensés peuvent être classés en trois catégories :

- les flux conversationnels ;
- les flux transactionnels ;
- les flux de type transfert de fichiers.

Il faut ajouter à cela les applications client-serveur qui peuvent, selon les cas, s'apparenter à la deuxième ou à la troisième catégorie.

Quel type de flux ?	Quelles caractéristiques ?	Quelles applications ?
Conversationnel	Trames courtes, fréquence soutenue	Connexions Telnet
Transactionnel	Trames moyennes (trafic montant) et longues par rafales (trafic descendant)	Serveurs intranet Connexions aux sites centraux (*via* des passerelles)
Transfert de fichiers	Trames longues, trafic soutenu	Serveurs bureautiques (FTP ou moniteur spécialisé sur TCP/ IP)
Client-serveur	Dépend de la position de la base de données et du module client	Requêtes SQL sous Unix, Windows NT, etc.

Ces flux doivent cohabiter au sein d'un même réseau intersite et être transportés simultanément sur une même liaison.

Les flux de type conversationnel

Les applications conversationnelles sont les plus courantes dans les mondes Unix et TCP/IP. Le protocole utilisé est Telnet. Le principe repose sur l'envoi d'un caractère avec écho distant. Une session étant établie entre un poste de travail et une machine, tout caractère frappé sur le clavier est envoyé à la machine, traité par cette dernière, et enfin renvoyé tel quel pour affichage, éventuellement avec d'autres attributs. Chaque caractère peut en effet déclencher une action comme l'affichage d'une fenêtre.

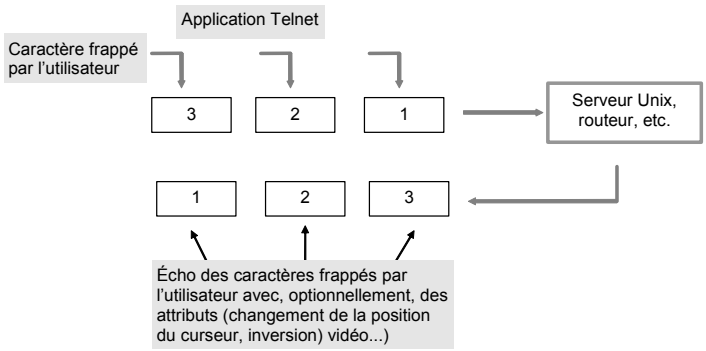


Figure 10-11. *Types de flux générés par des applications conversationnelles.*

Le type de flux qui en résulte est par conséquent irrégulier, car il dépend de l'activité de l'utilisateur et est composé de trames courtes.

Le temps de réponse est donc primordial pour ce type d'application. Il se doit d'être le plus régulier possible, le principe étant qu'un utilisateur s'habitue à un temps de réponse, même mauvais, pourvu qu'il soit régulier. Un maximum de 300 à 500 ms est généralement toléré. Quand plusieurs caractères sont saisis à la suite, ce temps est généralement réduit du fait de leur encapsulation dans le même paquet TCP (algorithme de Nagle — RFC 896).

```
! Activation de l'algorithme nagle sur un routeur Cisco
! Utile pour les connexions Telnet
! A desactiver pour X-Window
service nagle
```

Les flux de type transactionnel

Le mode transactionnel est le fonctionnement le plus courant pour les applications critiques sur les systèmes centraux. La technique consiste à envoyer un écran de saisie vers un terminal, à réaliser localement les modifications, puis à renvoyer les données modifiées vers le site central. Ces opérations sont contrôlées par un logiciel appelé moniteur transactionnel (CICS sous IBM et Tuxedo sous Unix).

Les flux générés entre serveurs intranet et navigateurs peuvent être assimilés au mode transactionnel, bien que le volume des pages HTML soit beaucoup plus important.

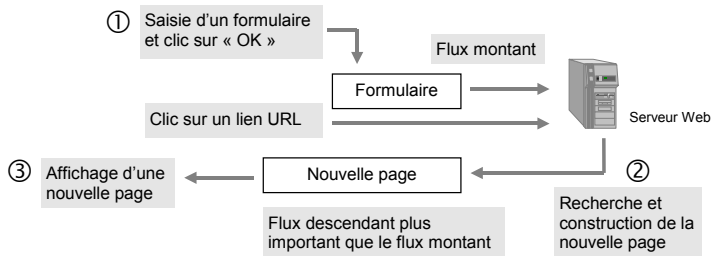


Figure 10-12.
Types de flux générés par des applications intranet.

Les flux générés sont caractérisés par un trafic descendant (serveur intranet vers navigateur) plus important que le trafic montant (les données du formulaire ou un clic sur une URL). La ligne est rarement mobilisée (2 à 4 transactions par minute) tandis que le transfert d'une page (de 4 à 50 Ko, voire plus) nécessite la presque totalité de la bande passante pendant quelques secondes. Le débit instantané requis est donc une donnée importante dans le calcul de la bande passante requise par ce type de flux.

Les flux de type transfert de fichiers

Ces flux sont caractérisés par des échanges soutenus et des trames longues. Leurs occurrences peuvent être prévisibles, dans la mesure où la majorité des transferts de fichiers est souvent associée à des traitements automatiques qui ont lieu en dehors des heures ouvrées, ce qui est le cas pour la sauvegarde ou la synchronisation de bases de données.

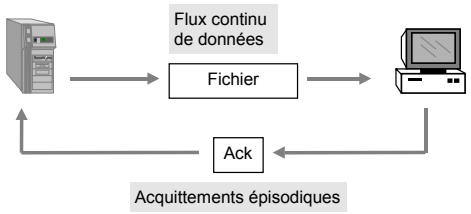


Figure 10-13.
Types de flux générés par des applications transactionnelles.

Pendant les heures ouvrées, ce type de flux peut dégrader les temps de réponse des flux transactionnels et surtout des flux conversationnels. Cette interférence peut être contrôlée par des mécanismes de priorité positionnés sur les équipements d'interconnexion tels que les routeurs.

Les flux client-serveur

Le concept client-serveur se décline en réalité sur plusieurs modèles :

- La base de données et la logique applicative sont situées sur le serveur. Le poste client soumet une requête puis attend les résultats qui, seuls, transitent par le réseau.
- Le serveur héberge la base de données ; la logique applicative réside sur le poste client. La puissance de traitement est donc reportée sur les postes client. Les échanges sur le réseau sont aussi fréquents que les manipulations de la base.
- La logique applicative et les données sont réparties entre le serveur et le client. La procédure d'interrogation consiste à extraire tout ou partie de la base de données centrale, puis à opérer des traitements spécifiques sur la base de données locale. Le réseau n'est sollicité que lors des extractions depuis la base de données. La synchronisation des bases peut intervenir en dehors des heures ouvrées.

Modèle	Application	Flux réseau
1. Base de données et application sur le serveur	Comptabilité, gestion commerciale, logistique, gestion des ressources humaines	Type transactionnel
2. Base de données sur le serveur et applications sur le client	Gestion commerciale Serveur Intranet	Type transactionnel (volumétrie faible) ou transfert de fichiers (volumétrie élevée)
3. Base de données et applications réparties entre le serveur et le client	Datawarehouse Architecture 3 tiers	Transfert de fichiers Transactionnel

L'architecture 3 tiers (client, serveur applicatif, serveur de base de données) génère des flux de type transactionnel, côté client, et de type transfert de fichiers, côté base de données.

Estimer la volumétrie

Les flux doivent ensuite être quantifiés, à partir de données existantes, ou sur la base d'hypothèses. Si l'on part d'un réseau existant, soit pour l'optimiser, soit pour le faire évoluer, on peut s'appuyer sur des statistiques indiquant les volumes échangés entre deux sites. Ces données peuvent être issues de facturations détaillées ou d'une phase d'audit consistant en une campagne de mesure sur le terrain.

La volumétrie est calculée différemment selon le type de flux. Souvent, elle doit être extrapolée à partir d'informations partielles. Ce travail doit donc être réalisé indépendamment pour chaque application susceptible d'être véhiculée par le réseau intersite. Les résultats doivent ensuite être consolidés sous forme de matrice de flux présentant les volumes échangés entre chaque site. L'échelle de temps généralement utilisée est une journée de travail ; cette périodicité permet en effet de lisser les variations.

La volumétrie globale pour un site est généralement issue d'une volumétrie unitaire estimée pour un utilisateur et calculée selon la formule suivante :

$$Vj = Vu \times U$$

Vj est le volume journalier à calculer pour un site.
Vu est le volume journalier estimé pour un utilisateur.
U est le nombre d'utilisateurs pour un site donné.

Les sections suivantes décrivent les manières d'estimer les volumétries quand l'existant est peu ou pas connu.

Étape 2 — Estimer la volumétrie

Applications	Exemples d'estimation de la volumétrie
Messagerie SMTP ou Exchange	10 messages par utilisateur et par jour x 100 Ko Synchronisation des annuaires
Transfert de fichiers FTP	*N* % des utilisateurs (ou des applications batch) = *X* Ko par jour
Transactionnelles sites centraux	100 à 200 écrans de 2 ou 4 Ko par utilisateur et par jour
Transactionnelles intranet	20 à 50 écrans de 4 à 50 Ko par utilisateur et par jour
Conversationnelles Telnet	Dépend des applications (faire des tests avec un analyseur réseau) ; un écran = 2 à 4 Ko
Services réseau (vidéotex, télécopie, etc.)	3 sessions vidéotex par jour et par utilisateur Un fax de 10 Ko par jour
Administration du réseau	0 à 10 sessions Telnet sur chaque routeur Configuration SNMP de 1 Ko par équipement et par jour 1 à 50 trap SNMP de 1Ko par jour Sondes RMON : 1 à 10 transferts de fichiers par jour (plusieurs centaines de Ko)

Volumétrie liée à la messagerie

Les volumes de données générés par une messagerie bureautique peuvent être modélisés sur la base des hypothèses suivantes :

- Environ 10 messages par jour et par utilisateur à destination d'un autre site (20 % des messages sont à destination d'un site extérieur, 80 % restent locaux).
- Environ 100 Ko par message. Cette valeur dépend beaucoup de l'utilisation qui est faite de la messagerie au sein de la société. Plus celle-ci est utilisée, plus les messages ont tendance à être importants (pièces jointes).
- La taille de l'annuaire est basée sur 100 octets par utilisateur.
- Synchronisation hebdomadaire (voire toutes les nuits) de l'annuaire : transfert depuis les sites distants vers le siège (si la gestion est décentralisée), consolidation de l'annuaire, puis transfert depuis le siège vers les sites distants.

Les messageries bureautiques transportent les messages sous forme de copies de fichiers entre les serveurs bureautiques. La périodicité des échanges dépend du paramétrage ; elle est généralement comprise entre 5 et 15 minutes. Ces transferts de fichiers occupent donc régulièrement la bande passante des liens.

Volumétrie liée aux transferts de fichiers

La volumétrie liée aux transferts de fichiers dépend des applications présentes au sein de la société. Son évaluation repose donc sur une analyse précise de l'existant et/ou des besoins. Elle peut être modélisée sous la forme N % des utilisateurs réalisant l'équivalent d'un transfert de X Ko par jour à destination d'un site distant.

Volumétrie liée aux applications transactionnelles site central

Dans la plupart des cas, on peut estimer qu'un utilisateur échange 100 à 200 écrans de 2 Ko à 4 Ko par jour avec le site central. Cette évaluation est bien sûr éminemment variable selon le contexte à considérer. La taille des écrans varie, par exemple, en fonction des applications, et la fréquence des échanges en fonction du type de travail de l'utilisateur (saisie intensive, consultation, etc.). Il convient donc d'estimer la volumétrie moyenne à partir de tests.

Volumétrie liée aux applications transactionnelles intranet

Même remarque que pour les applications transactionnelles, sauf que la taille des pages varie entre 4 Ko et 50 Ko, une page pouvant contenir des images GIF (fixes ou animées).

En prenant en compte les fichiers GIF, JPG et HTML, la moyenne constatée est de 4 Ko. Si l'on se réfère aux transferts de fichiers réalisés à partir d'Internet (documents .pdf, .txt ou .doc), la moyenne constatée est de 100 Ko. La moyenne peut atteindre plusieurs Mo si le téléchargement des exécutables (.exe) est autorisé.

Vous pouvez vérifier les valeurs propres à votre contexte en visualisant le contenu du cache de votre navigateur (recherchez un répertoire appelé « cache » situé dans le répertoire d'installation du navigateur).

Volumétrie liée à d'autres services

Différents services peuvent emprunter le réseau intersite, notamment en provenance de sites rattachés dans le cas où les passerelles de communication sont centralisées. Les hypothèses de travail qui peuvent être retenues sont les suivantes (il ne s'agit ici que d'indications, la volumétrie réelle étant liée à la nature des travaux réalisés par les utilisateurs) :

- Service de télécopie. Chaque utilisateur expédie en moyenne un fax de 10 Ko par jour.
- Service d'accès au vidéotex. Un quart des utilisateurs effectue 3 connexions vidéotex de 2 minutes par jour. Chaque connexion génère un flux de 5 Ko.
- Etc.

L'utilisation des applications multimédias pose d'autres problèmes, qui sont abordés au chapitre 13.

Rassembler toutes les données

Ayant ces abaques en tête, nous pouvons maintenant calculer les volumes pour notre cas. La première chose à faire est d'établir la matrice des flux présentant les types de flux et le nombre d'utilisateurs qui les génèrent.

Étape 3 — Matrice des flux

Depuis \ vers	Toulouse	Paris	Strasbourg	Etc.
Toulouse	---	400 intranet 40 Telnet	100 Telnet	
Paris		---		
Strasbourg	50 intranet		---	
Etc.				---

La volumétrie doit être calculée entre chaque site et dans les deux sens. Les liaisons étant de type *full duplex*, il convient de prendre la valeur la plus haute, ce qui permet de calculer le débit instantané nécessaire.

Étape 4 — Matrice volumétrique

De → vers	P → T	T → P	T → S	S → T	etc.
Flux transactionnels intranet					
Nombre d'utilisateurs	400	400	50	50	…
Volumétrie unitaire (en Ko)	50	0,5	50	0,5	…
Pages par jour et par utilisateur	10	10	10	10	…
Quantité / utilisateur / jour (en Ko)	500	5	500	5	…
Total journalier en Mo	200	2	25	0,25	…
Flux conversationnels Telnet					
Nombre d'utilisateurs	40	40	100	100	…
Volumétrie unitaire (en Ko)	4	0,5	0,5	4	…
Écrans par jour et par utilisateur	100	100	50	50	…
Quantité / utilisateur / jour (en Ko)	400	50	25	200	…
Total journalier en Mo	16	2	2,5	20	…
Etc.	…	…	…	…	…
Volume total en Mo	960	320	240	190	

Selon le sens de la connexion client-serveur, les flux montants (depuis le client vers le serveur) et descendants (depuis le serveur vers le client) apparaîtront dans la première colonne ou la deuxième. Au final, seul le maximum des deux flux doit être pris en compte. C'est lui qui déterminera la bande passante maximale requise.

Calculer les débits

Pour dimensionner une liaison, il convient d'estimer les besoins en termes de débit instantané. La formule de calcul généralement admise est la suivante :

$$Bp = Vj \times Th \times Ov \times \frac{1}{Tu} \times \frac{1}{3600} \times (8 \times 1{,}024)$$

La signification des paramètres est la suivante :

Bp est la bande passante instantanée calculée pour une liaison exprimée en Kbit/s.

Vj est le volume journalier, estimé en Ko. Cette valeur représente la somme des flux devant circuler sur le lien considéré (le maximum pris entre les flux montants et descendants).

Th est un coefficient permettant de calculer le trafic ramené à l'heure chargée. On considère généralement que le trafic journalier est concentré sur une heure chargée. Cette hypothèse part du constat que, sur 8 heures de travail, les utilisateurs sont le plus actifs sur deux périodes de pointe, entre 10 h et 11 h, et entre 15 h et 16 h. Les valeurs généralement admises sont comprises entre 20 % et 30 % du trafic journalier concentré sur une heure.

Ov est l'overhead généré par les protocoles de transport (TCP, IP, PPP). Ce coefficient est généralement affecté d'une valeur de 20 %. Il tient compte des en-têtes et des paquets de service (acquittements, etc.).

Tu est le taux maximal d'utilisation de la bande passante du lien. Cette correction permet de prendre en compte le fait que l'on utilise rarement 100 % du débit nominal d'un lien. Ce taux est généralement fixé à 80 % de la bande passante, ce qui donne un surdimensionnement du lien de l'ordre de 25 %. Pour des liaisons à haut débit, ce taux peut atteindre 90 %.

Le rapport 1/3600 permet de ramener la volumétrie sur une heure en secondes, tandis que le rapport 8*1,024 permet de convertir les kilo-octets en kilobits (1 octet = 8 bits, 1 Ko = 1 024 octets et 1 000 bits = 1 kilobit).

Si l'on prend les valeurs standard pour ces paramètres, la formule devient :

$$Bp = Vj \times 0{,}30 \times 1{,}2 \times \frac{1}{0{,}8} \times \frac{1}{3600} \times (8 \times 1{,}024)$$

soit, par exemple, une bande passante de 1 Mbit/s pour un volume journalier estimé à 1 Go.

Si la liaison doit servir de secours pour *n* autres liaisons de débit D_N sans que les performances ne soient dégradées, la bande passante du lien doit être augmentée de la somme de ces débits D_N. Dans notre cas, nous choisirons un mode dégradé, afin de limiter les coûts.

Liaison	Volume en Ko Vj	30 % à l'heure chargée Th	Overhead protocole Ov	Taux d'occupation du lien Tu	Débit du lien en Kbit/s
P ↔ T	960 000	0,3	1,2	0,8	984
T ↔ S	240 000	0,3	1,2	0,8	246
Etc.	...	...	...	...	...

Le débit du lien doit être arrondi à la valeur supérieure des débits proposés par les opérateurs, soit, dans notre cas, 1 Mbit/s entre Paris et Toulouse et 256 Kbit/s entre Toulouse et Strasbourg.

Tenir compte des temps de réponse

Pour des applications client-serveur reposant sur des extractions de données et assimilables à des transferts de fichiers, le critère performance se pose en termes de délai maximal de transfert des données. Il convient donc de calculer les débits nécessaires en fonction des délais acceptables et des volumes estimés :

$$Bp = \frac{Vo}{Tps}$$

Bp est la bande passante nécessaire.

Vo est le volume moyen (converti en kilobits) des données extraites suite à une requête.

Tps est le temps de réponse souhaité.

Cette démarche est à combiner avec une étude de coût, car il faut trouver un compromis avec la performance. Il faut donc recourir à une simulation des temps de réponse obtenus en fonction des débits des liens, et éventuellement les mesurer pour différentes tailles de requêtes.

Par exemple, le tableau suivant compare les temps de transfert de données de différentes tailles en fonction du débit de la ligne.

Volume en Ko	64 Kbit/s	128 Kbit/s
10	1,25 s	0,62 s
20	2,50 s	1,25 s
30	3,75 s	1,87 s
50	6,25 s	3,12 s
Coût mensuel HT	**1 200 €**	**2 200 €**

Il vous appartient alors de mettre en balance le coût et les performances souhaitées.

11

Bâtir un réseau de transport

Le réseau WAN qui interconnecte les réseaux LAN utilise les services d'un **réseau de transport**. Celui-ci véhicule également de la voix, de la vidéo, etc.

Le réseau de transport correspond aux couches physique (niveau 1) et logique (niveau 2). Sur le LAN nous avions Ethernet, sur le WAN, nous allons avoir ATM et Frame Relay.

Ces technologies répondent à des contraintes plus larges que celles d'une interconnexion de LAN. Notre réseau intersite n'est donc qu'un utilisateur parmi d'autres, tels un réseau de PABX pour la téléphonie ou des connexions entre salles de visioconférence.

Dans ce chapitre, vous apprendrez ainsi :

- à interconnecter des réseaux locaux *via* Frame Relay et ATM ;
- à configurer les circuits virtuels ;
- à gérer la qualité de service ;
- à connaître la signalisation et l'adressage.

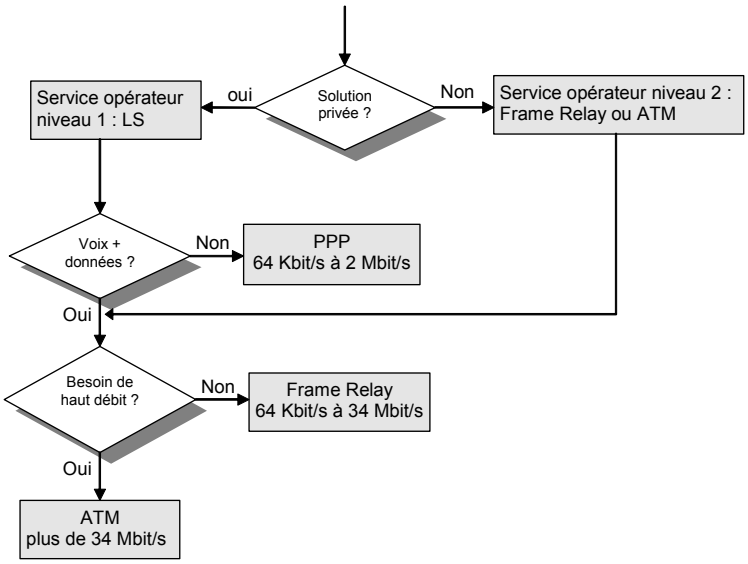
LS, Frame Relay ou ATM ?

En tant qu'utilisateur, seul le service compte, quelle que soit la technologie employée. Le choix de l'une des trois solutions qui s'offrent à nous se fera en fonction de nos besoins, de l'offre du marché et du coût, que ce soit dans le cadre d'une solution privée ou opérateur.

Comme nous l'avons vu au chapitre précédent, il est possible de choisir entre différentes technologies et différents niveaux de service.

Pour une interconnexion de réseaux locaux, la LS est la solution idéale. À partir d'un certain nombre de sites (une dizaine, voire moins à l'international), la solution opérateur reposant sur un réseau Frame Relay ou ATM devient plus rentable.

La technologie Frame Relay est actuellement la plus répandue, mais les opérateurs investissent dans ATM. En 1997, leurs réseaux reposaient pour 30 % sur une infrastructure Frame Relay et pour 45 % sur ATM. En 1998, la proportion était de 22 % pour Frame Relay et 61 % pour ATM.



Pour l'utilisateur, un accès Frame Relay revient cependant moins cher. Les opérateurs réservent donc ATM pour des accès à haut débit (34 Mbit/s et plus), et limitent les accès Frame Relay à 34 Mbit/s, voire à 8 Mbit/s.

Techniquement, ATM a tout pour s'imposer, mais l'histoire de l'informatique nous a montré que cela n'est pas un gage de pérennité. Rappelons-le, les protocoles qui se sont imposés sont les plus simples, les moins chers et surtout les mieux adaptés aux besoins des utilisateurs.

Par exemple, les principaux inconvénients d'ATM sont :

- un overhead très important (plus de 10 %) ;
- un manque de maturité dans la normalisation et les offres.

En revanche, le point fort d'ATM réside dans sa capacité multiservice : il permet de créer des réseaux locaux et étendus (LAN et WAN) et transporte les flux multimédias. Malheureusement, c'est justement sur ces points qu'ATM manque de maturité : les normes ne sont pas stabilisées, ou non satisfaisantes, et aucune offre ne permet de transporter simultanément les flux voix, données et vidéo.

En réalité, ATM est actuellement utilisé en tant que :

- réseau fédérateur WAN pour les réseaux fédérateurs (backbone) des opérateurs ;
- réseau fédérateur LAN par les entreprises, en association avec les réseaux virtuels LANE (*LAN Emulation*).

Mais, même sur ces créneaux, il est concurrencé par le Gigabit Ethernet.

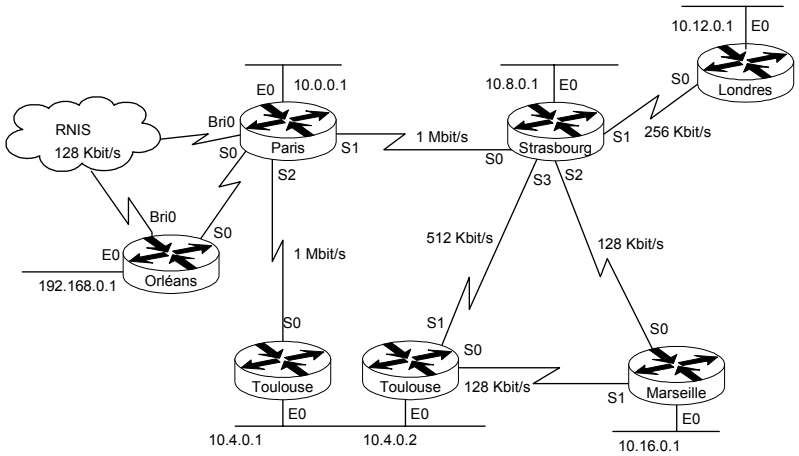
Critère	Frame Relay	ATM
Débit	De 64 Kbit/s à 45 Mbit/s	À partir de 34 Mbit/s
Qualité de service	Gestion des congestions, débit garanti	Gestion des congestions, débit garanti, trafic synchrone, priorités
Réseau	WAN	LAN et WAN
Application	Voix et données (vidéo *via* IP)	Voix, données et vidéo
DSU	FRAD	DXI ou ATM
Overhead pour un MTU de 1 500 octets	0,5 %	10,4 % au minimum
Adressage	Local (DLCI)	Local (VPI/VCI)
Normes	ITU Q.922 / Q.933	ITU I.361 à I.363 / Q.931 et ATM Forum
Transport d'IP	RFC 2427 (NLPID/SNAP)	RFC 1483 (LLC/SNAP) RFC 2225 (Classical IP)

À côté de ces deux protocoles, nous retrouvons nos LS (lignes spécialisées) qui sont dans tous les cas la base d'un réseau de transport Frame Relay ou ATM. Nous pouvons même décider de bâtir notre réseau uniquement sur des LS, comme nous l'avons fait pour notre première interconnexion.

Mettre en place un réseau de LS

La mise en place d'un réseau de LS implique de commander vous-même les liaisons auprès des opérateurs locaux. En France, on peut s'adresser à France Télécom et, bientôt, à d'autres, pour toutes nos LS. Il suffit d'indiquer à l'opérateur le débit souhaité, les adresses des sites à connecter et, si cela est proposé, le type d'interface désirée : V.35 ou X21/V11 la plupart du temps.

Figure 11-1.
Réseau intersite reposant sur des LS.



Pour la liaison internationale Strasbourg-Londres, le problème se complique puisque vous avez affaire à deux opérateurs, BT et France Télécom, par exemple. La liaison est administrativement découpée en **demi-circuits**, chacun étant géré par l'opérateur situé aux tenants et aboutissants de la LS.

Votre correspondant à Londres doit donc s'adresser à BT pour la fourniture du demi-circuit anglais, et vous devez faire de même auprès de France Télécom pour le demi-circuit français. Dans le cas où la LS traverse plusieurs autres pays, aucune autre démarche n'est à effectuer.

Les opérateurs ont passé des accords de coopération multilatéraux et ont mis en place des structures de coordination pour que les demi-circuits soient regroupés en une seule liaison internationale. Cette procédure est transparente pour vous. Néanmoins, le délai de mise en service est généralement de dix semaines, voire plus dans certains pays.

La mise en service de la LS se concrétise par l'installation d'un modem numérique CSU (*Channel Service Unit*) dans vos locaux et par un test BERT (*Bit ERror Tests*). Ce dernier consiste à envoyer un train continu de bits pendant une durée minimale de 24 heures, et à mesurer le taux d'erreur qui ne doit pas dépasser 10^{-6} à 10^{-10} (une erreur tous les 10 milliards de bits).

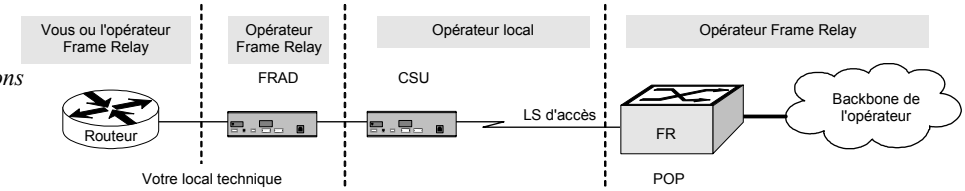
Généralement, l'opérateur vous accorde un délai de quelques jours avant de déclarer la liaison opérationnelle et de débuter la facturation. Cela vous permet de tester la liaison avec vos routeurs.

Mettre en place un réseau Frame Relay

Faire appel à un opérateur permet de ne s'adresser qu'à un seul interlocuteur, quels que soient les pays concernés. Même en retenant un opérateur étranger pour vos sites français, celui-ci se chargera de commander les liaisons d'accès (en fait, des LS aux technologies E1 ou xDSL) auprès de l'opérateur local. Il suffit de lui indiquer les adresses de vos sites pour qu'il commande les LS entre vos sites et ses POP les plus proches.

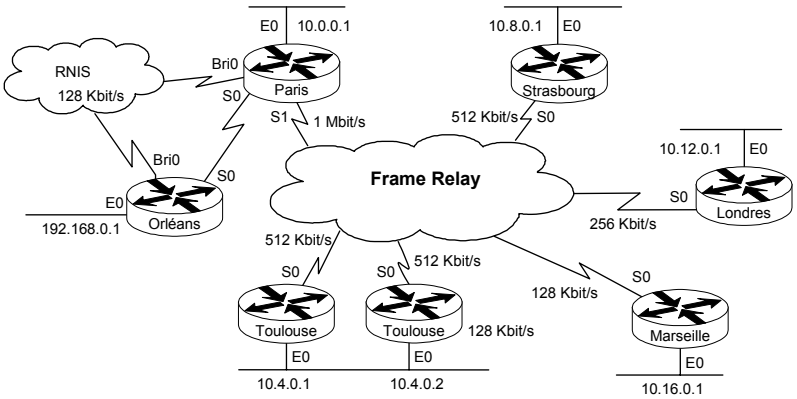
Comme pour la LS, l'opérateur local installe son modem numérique (le CSU) dans vos locaux. L'opérateur retenu pour le réseau Frame Relay installe ensuite un commutateur appelé **FRAD** (*Frame Relay Access Device*) qu'il va connecter au modem. L'interface série du routeur sera ensuite connectée à un des ports du FRAD.

Figure 11-2. *Responsabilités des configurations de l'accès Frame Relay.*



Le FRAD est un équipement de conversion entre des protocoles d'entrée (interfaces E1, X.21/V11, voix, données, etc.) et le protocole Frame Relay. Cet équipement prend les paquets IP issus d'un routeur, ou les canaux voix issus d'un PABX, et les encapsule dans les trames Frame Relay. Le FRAD est relié *via* la liaison d'accès au commutateur Frame Relay situé dans le POP de l'opérateur. Plusieurs flux sont ainsi multiplexés sur une même liaison.

Figure 11-3. *Réseau intersite reposant sur un réseau opérateur Frame Relay.*



Les avantages sur la première solution sont immédiatement perceptibles :

- Il n'y a plus qu'une seule LS par site. Nous avons besoin de moins d'interfaces série, ce qui diminue d'autant le coût de nos routeurs.
- Les LS sont locales entre nos sites et les POP de l'opérateur, ce qui diminue également leur coût.

Critère	Solution opérateur	Solution privée
Caractéristiques de l'offre	Un réseau fédérateur et des liaisons locales.	Des liaisons longue distance entre les sites.
Déménagement d'un site	Facile : une liaison locale à changer. Coûts réduits.	Difficile : plusieurs liaisons longue distance à changer. Coûts élevés dus aux changements et à la période de recouvrement.
Modification des débits	Très souple : débit des liaisons locales et/ou des CIR.	Assez difficile (surtout à l'international) : changement du débit des liaisons longue distance.
Redondance de site (second site en secours du premier)	Possible : une liaison locale sur le site principal et une sur le site de secours.	Impossible, sauf à doubler les liaisons longue distance (très cher).
Exploitation du réseau	Prise en charge par l'opérateur.	Prise en charge par le client *via* des contrats de maintenance.

Nous avons conservé notre LS entre Paris et Orléans, car le réseau opérateur était plus cher dans ce cas précis. Cette situation permet de montrer l'usage de trois réseaux d'opérateurs : liaison spécialisée, RNIS et Frame Relay. Par ailleurs, deux liaisons d'accès ont été conservées à Toulouse car le site est stratégique.

Qualité de service et facturation

Un opérateur s'engage toujours sur une qualité de service et base son offre commerciale sur les mécanismes offerts par les protocoles qu'il utilise.

Par exemple, Frame Relay permet de **gérer la congestion** du réseau, de diminuer le débit des LS et donc les coûts. Comment ? La réponse est intimement liée au fonctionnement du protocole Frame Relay.

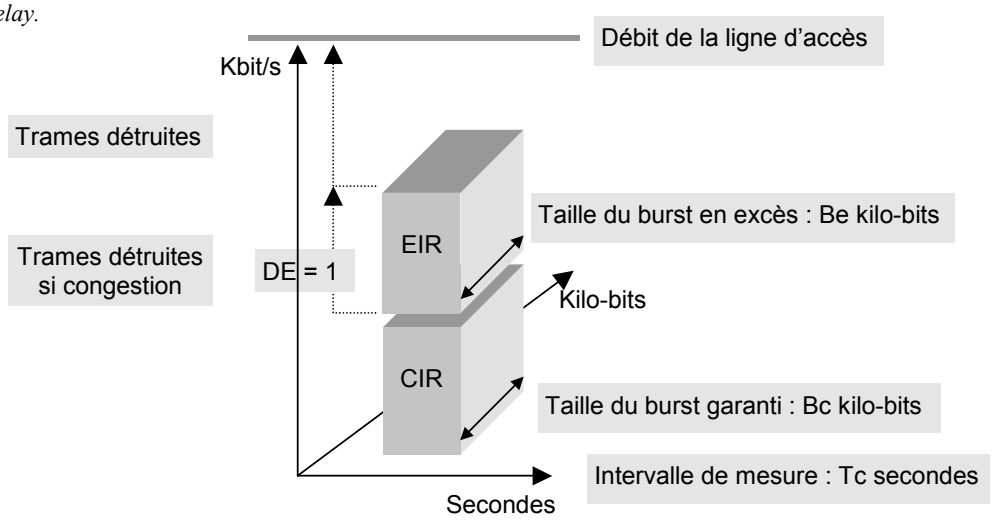
Débit garanti

Tout d'abord, le réseau garantit à l'utilisateur un volume de Bc (*committed burst size*) kilobits pendant une période de Tc (*committed rate measurement interval*) secondes, définissant ainsi un débit garanti **CIR** (*Committed Information Rate*) : CIR = Bc / Tc. Lorsque le commutateur voit passer plus de Bc kilobits pendant Tc secondes (c'est-à-dire lorsque le débit dépasse le CIR), le bit **DE** (*Discard Eligibility*) des trames en dépassement est positionné à « 1 ». Cela signifie que ces trames seront détruites en priorité en cas de congestion du réseau.

Ensuite, le réseau autorise à l'utilisateur un volume supplémentaire de Be (*excess burst size*) kilobits pendant une période de Tc secondes, définissant ainsi un débit en excédent **EIR** (*Excess Information Rate*) : EIR = Be / Tc. Lorsque le commutateur voit passer plus de Bc+Be kilobits pendant Tc secondes (quand le débit dépasse l'EIR), toutes les trames en excès sont détruites.

Généralement, la période de mesure Tc est fixée à une seconde, et l'**AIR** (*Allowed Information Rate* = CIR + EIR) ne dépasse pas le débit de la liaison d'accès.

Figure 11-4.
Qualité de service Frame Relay.



Le débit des liaisons internes au réseau peut être inférieur à la somme des AIR des clients. L'opérateur se fonde sur des calculs de probabilité qui montrent que tous les clients n'utiliseront pas leur AIR en même temps (surbooking), et compte sur les mécanismes de contrôle de congestion offerts par le protocole pour résoudre les problèmes. L'opérateur ne facturera ainsi que le CIR (le débit garanti), l'EIR étant gratuit ou presque.

Dès lors, tout le monde « joue » sur les CIR : le client pour payer le moins cher possible, et l'opérateur pour dépenser le moins possible.

Le débit de la liaison d'accès peut ainsi être de 512 Kbit/s, et celui du CIR de 64 Kbit/s ; ce qui permet d'obtenir un débit maximal de 512 Kbit/s tout en ne payant que pour 64 Kbit/s. Inversement, pour un site central, la somme des CIR peut être égale à 200 % du débit de la liaison d'accès (on espère alors que toutes les applications n'utiliseront pas le réseau en même temps).

Cependant, le débit offert au-delà du CIR n'étant pas garanti, les temps de réponse risquent d'être mauvais et erratiques. Il est donc conseillé de dimensionner suffisamment les CIR afin d'obtenir une bonne qualité de service, notamment pour les applications que vous considérez comme étant critiques et surtout pour les flux voix.

Connecter un routeur au réseau de transport

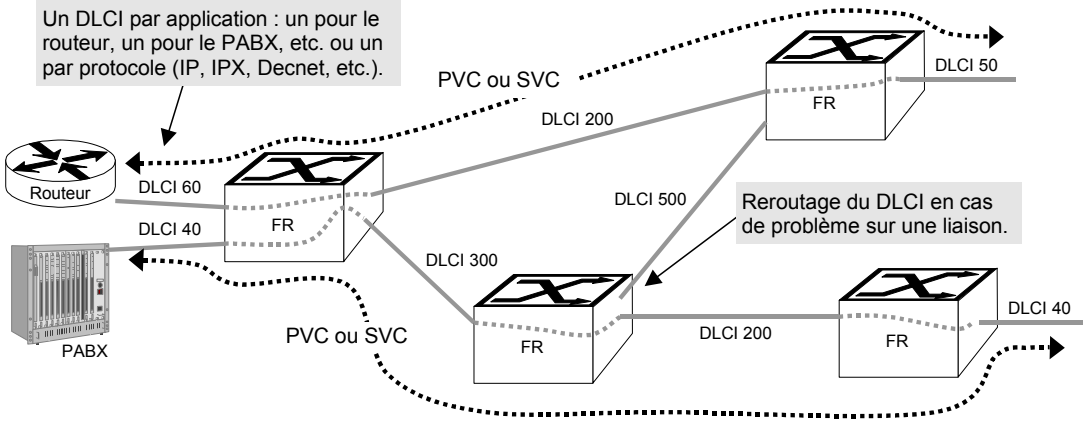
La norme Frame Relay définit uniquement une **interface d'accès** pour l'utilisateur. Bien qu'il soit possible de réaliser un réseau 100 % Frame Relay, le réseau interne de l'opérateur repose souvent sur un protocole propriétaire ou ATM. Pour le client, cela n'a pas d'importance : il faut simplement que l'interface soit de type Frame Relay.

Chaque trame est identifiée par un **DLCI** (*Data Link Connection Identifier*), une adresse locale partagée par deux commutateurs. Il n'y a pas d'adressage de bout en bout, mais uniquement un adressage point à point entre deux commutateurs. Un commutateur recevant une trame avec un DLCI donné la routera sur un autre port et l'enverra avec un autre DLCI, et ainsi de suite.

La connexion entre deux commutateurs s'effectue par l'ouverture de circuits virtuels permanents (**PVC**, *Permanent Virtual Circuit*) ou commutés (**SVC**, *Switched Virtual Circuit*), c'est-à-dire ouverts à la demande *via* un protocole de signalisation (qui utilise quelques Kbit/s dans le DLCI 0).

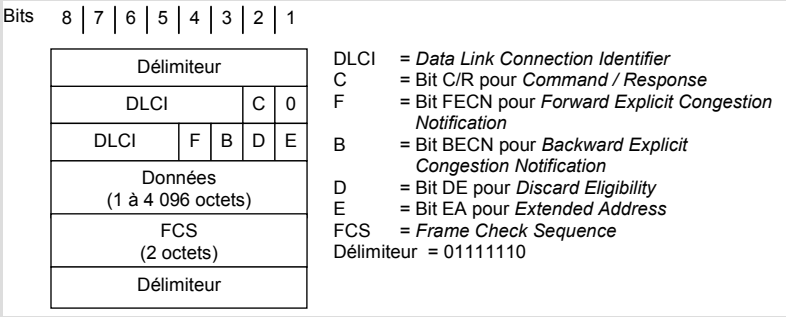
L'intérêt des SVC est que la qualité de service (CIR et EIR) peut être spécifiée à la demande, permettant ainsi de réduire (encore) les coûts. Sur un PVC, le CIR est fixé une fois pour toutes et engendre un coût fixe. De plus, la signalisation Q.933 utilisée pour établir les SVC permet de demander un délai de transit maximal, alors que cette fonction n'est pas prévue pour les PVC.

Figure 11-5.
Circuits virtuels et DLCI Frame Relay.



LE POINT SUR FRAME RELAY (ITU Q.922 ANNEXE A, ANSI T1.618, FRF 1.1)

Le Frame Relay (relais de trames) est un protocole de niveau 2 multipoint, qui définit uniquement l'interface d'accès au réseau (**UNI**, *User Network Interface*). Les nœuds intermédiaires relaient les trames sans réaliser le moindre contrôle de flux ni aucune reprise sur erreur : les trames peuvent ainsi être transportées par n'importe quel protocole. Seuls les nœuds d'extrémité sont tenus de respecter la norme Q.922, appelée « *Frame Relaying Bearer Service* ».



Le champ FCS est un code de contrôle d'erreur de type CRC (*Cyclic Redundancy Check*). Si une erreur est détectée, la trame est détruite. Les pertes de données et les reprises sur erreur sont laissées à l'initiative des couches supérieures (TCP dans les cas Internet/intranet). Le bit EA permet d'étendre les 10 bits du DLCI à 16 ou 23 bits.

Lorsque le débit des trames atteint le **CIR** (*Commited Information Rate*), le commutateur positionne le bit **DE** à " 1 ". Si, dans le réseau, une **congestion** est décelée, les trames marquées DE seront détruites en priorité. Autrement dit, le débit peut donc dépasser le CIR, mais sans garantie.

Les congestions sont détectées au niveau des files d'attente. Lorsqu'un seuil est dépassé, le commutateur avertit explicitement l'émetteur du flux (en positionnant à " 1 " le bit **BECN** des trames circulant dans l'autre sens) et le récepteur du flux (en positionnant le bit **FECN** à " 1 ").

La recommandation Q.922 suggère que le commutateur émetteur réduise son flux de 30 % si, pour plus de S % des trames qu'il reçoit, le bit BECN est positionné à " 1 ". La valeur S est calculée dynamiquement en fonction du débit, du délai de transit, des paramètres Bc et Be, etc. Si le bit BCEN est toujours positionné à " 1 " dans les trames qui continuent d'arriver, le flux est réduit de 50 %, puis de 75 % si le problème persiste.

La recommandation Q.922 suggère également que le commutateur récepteur réduise son flux de 25 % si, pour plus de la moitié des trames qu'il reçoit, le bit FECN est positionné à " 1 ". Si la proportion s'inverse, il peut augmenter son flux par paliers de 1/16.

Les équipements terminaux (les routeurs, par exemple) peuvent également interpréter le bit BECN, et réagir de la même façon que les commutateurs. De même, ils peuvent détecter implicitement des problèmes de congestion lorsque des trames sont perdues après qu'un certain nombre d'entre elles aient été reçues avec les bits DE à " 1 ".

Cependant, la norme Q.922 préconise que les commutateurs avertissent les équipements terminaux des congestions *via* le protocole **CLLM** (*Consolidated Link Layer Management*). Les messages CLLM (envoyés dans le DLCI 1023) contiennent la liste des DLCI pouvant causer des congestions à court, moyen et long termes. Cette signalisation de niveau 3 est plus fiable que l'interprétation des bits BECN et DE, car elle évite l'attente de trames (seuls vecteurs des bits BECN et DE) et évite aux couches supérieures d'avoir affaire à des informations internes au niveau 2.

Un PVC peut être géré comme un SVC au sein du réseau de l'opérateur. Cette facilité lui permet d'offrir le reroutage du PVC en cas de panne. Le client ne voit cependant qu'un PVC, appelé *soft* PVC, *switched* PVC ou encore *shadowed* PVC.

Si le routeur ne supporte pas Frame Relay

La manière la plus simple de raccorder nos routeurs est de les connecter à des FRAD (*Frame Relay Access Device*) de l'opérateur *via* leur interface série.

Une liaison Frame Relay peut être configurée en point à point ou en multipoint. Dans tous les cas, elle permet au routeur de joindre directement n'importe quel autre routeur. Elle doit donc être considérée comme étant un réseau IP, au même titre qu'un segment Ethernet, ce qui implique d'affecter une adresse IP à l'interface série du routeur.

```
interface serial 1
ip address 172.16.0.1 255.255.255.252
encapsulation hdlc
```

QU'EST-CE QUE LA SIGNALISATION ?

Dans le jargon de l'ITU, la signalisation est un protocole qui permet de gérer l'état des connexions : ouverture et fermeture, négociation des paramètres, gestion de la qualité de service, etc. Vis-à-vis de l'utilisateur, la signalisation est appelée **UNI** (*User Network Interface*). Entre les équipements réseau, la signalisation est désignée sous le nom générique de **NNI** (*Network to Network Interface)*.

Dans le monde TCP/IP, la séparation signalisation/données utilisateur n'est pas si nette : le trafic de service est souvent mêlé au trafic des données. Quelques protocoles utilisent cependant le principe de séparation, tel que FTP : le client et le serveur ouvrent deux ports TCP, le 21 pour envoyer les commandes, le 20 pour transférer les données.

Dans tous les cas, il s'agit d'une séparation logique, car la bande passante est partagée entre les données de service et les données utilisateur. Dans les réseaux des opérateurs, la signalisation peut cependant emprunter un chemin différent de celui des données.

En fait, cette séparation n'a réellement de sens que pour les protocoles de niveau 2 agissant en mode connecté, tels que Frame Relay et ATM. Une signalisation de niveau 3 permet ainsi de gérer les circuits virtuels de bout en bout à l'aide d'un adressage global de niveau 3.

Norme ITU	Description
Q.2931 et suivants	UNI ATM pour les réseaux publics = DSS2 (*Digital Subscriber Signalling System No.* 2)
Q.2100 et suivants	UNI ATM pour les réseaux privés = SAAL (*Signalling ATM Adaptation Layer*)
Q.931	UNI RNIS = DSS1 (*Digital Subscriber Signalling System No.* 1)
Q.933	UNI Frame Relay = DSS1 (*Digital Subscriber Signalling System No.* 1)

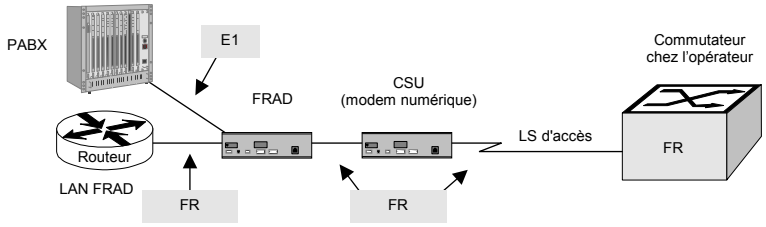
Toutes ces normes reposent sur le même protocole (format des messages, procédures, paramètres généraux). Seuls changent les paramètres propres à chaque réseau.

Si le routeur supporte Frame Relay

La solution la plus souple consiste cependant à configurer le routeur en FRAD : les trames LAN (Ethernet dans notre cas) sont converties en trames Frame Relay.

Dans ce cas, le FRAD de l'opérateur n'est utile que si d'autres équipements sont connectés, comme un PABX.

Figure 11-6. *Connexion d'un routeur à un FRAD.*



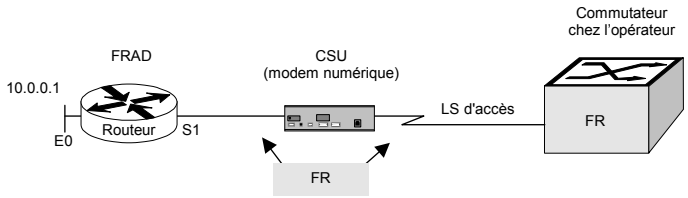
Le FRAD de l'opérateur se comporte comme un commutateur Frame Relay vis-à-vis du routeur, et comme un FRAD voix (voice-FRAD) vis-à-vis du PABX.

Si vous choisissez un service de niveau 3, l'opérateur prend en charge le routeur qui peut alors être intégré au FRAD (ou inversement, le routeur peut supporter des cartes voix, permettant de configurer en voice-FRAD). Les constructeurs télécoms proposent des FRAD intégrant cartes voix et cartes routeur, tandis que les constructeurs informatiques proposent des routeurs intégrant des cartes voix.

Si, comme dans notre cas, seuls des réseaux locaux doivent être connectés, le routeur peut directement être raccordé au commutateur situé chez l'opérateur (dans son POP) :

```
interface serial 1
ip address 172.16.0.1 255.255.255.252
encapsulation frame-relay ietf
```

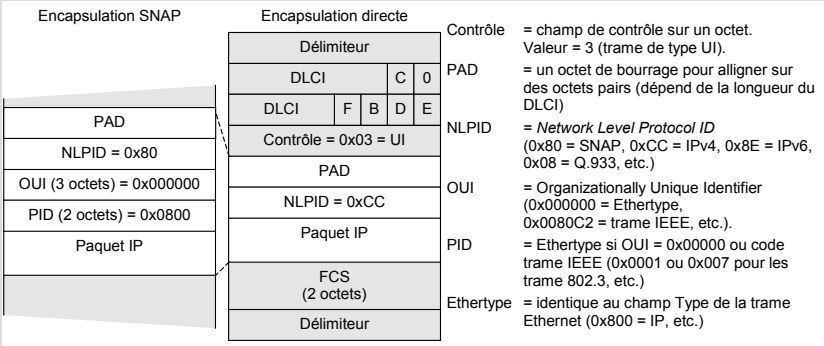
Figure 11-7. *Connexion d'un routeur à un commutateur Frame Relay.*



L'option « ietf » indique au routeur de respecter le RFC 2427 au lieu d'utiliser le format propre à Cisco.

ENCAPSULATION DANS FRAME RELAY (RFC 2427, ANSI T1.617A, FRF 3.1)

Deux méthodes sont définies pour encapsuler les protocoles dans une trame Frame Relay. La plupart le sont par la méthode **SNAP** (*Sub Network Access Protocol*) tandis que certains, dont IP, peuvent l'être soit *via* SNAP, soit directement dans la trame *via* l'identificateur **NLPID** (*Network Level Protocol ID*).



Si la fonction pont est activée sur le routeur, les trames Ethernet, Token-Ring, FDDI, etc., sont encapsulées dans une trame Frame Relay selon la méthode SNAP.

Gérer les circuits virtuels

La gestion des CV (circuits virtuels) implique l'activation de trois mécanismes :

1. un protocole de signalisation Q.933 qui permet d'ouvrir et de fermer les SVC et de gérer les CV (PVC et SVC) ;
2. l'utilisation de la procédure LAP-F (décrite dans la norme Q.922) qui permet de transporter les messages Q.933 de manière sûre (reprise sur erreur, contrôle de flux, etc.) ;
3. un adressage global de bout en bout E.164 ou X.121 pour les SVC.

Ainsi, les routeurs peuvent échanger des informations avec les commutateurs en utilisant la signalisation LMI (*Local Management Interface*). Notre routeur supporte trois formats de messages : celui de l'ANSI (T1.617), celui de l'ITU (Q.933) et un autre, propre à Cisco, qui utilise le DLCI 1023 (en principe réservé à CLLM). Nous préférons utiliser le mode de fonctionnement défini par l'ITU.

```
interface serial1
frame-relay lmi-type q933a
```

" a " comme Annexe A de la norme Q.933

Les paramètres par défaut peuvent être utilisés pour la signalisation Q.933 et la procédure LAP-F. La configuration en devient simplifiée. Si la liaison est de mauvaise qualité (souvent suite à un problème survenu sur la liaison d'accès), il peut être intéressant de réduire la taille initiale de la fenêtre à 8, voire à 1 :

```
interface serial1
frame-relay lapf k 8
```

LA SIGNALISATION FRAME RELAY (ITU Q.933, ANSI T1.617)

Les commutateurs Frame Relay utilisent un PVC dédié (DLCI 0) pour véhiculer les messages de signalisation Q.933. Ce protocole s'appuie sur celui utilisé par le RNIS (Q.931).

La première fonction offerte permet à l'équipement d'extrémité (un routeur, par exemple) de demander l'ouverture et la fermeture des **SVC** (*Switched Virtual Circuit*). Pour l'ouverture, les messages Setup et Connect contiennent le numéro d'appel (adresse aux formats **E.164** ou **X.121**), le DLCI affecté, etc., ainsi que le délai maximal de transit négocié. L'ouverture du SVC se traduit par l'affectation des DLCI entre les commutateurs et les équipements terminaux.

La deuxième fonction a trait à la procédure **LMI** (*Local Management Interface*) dont le rôle est de surveiller l'état des **PVC** (*Permanent Virtual Circuit*). Des messages *Status Enquiry* sont périodiquement envoyés par l'équipement terminal (un routeur, par exemple) pour connaître l'état du PVC. Le commutateur répond par un message *Status* (PVC actif, inactif, etc., ou nouveau PVC). En cas de défaillance, le routeur peut ainsi rerouter les flux. Le commutateur peut également interroger l'équipement terminal.

Les messages **ELMI** (*Enhanced LMI*) permettent au commutateur de communiquer à l'équipement terminal (un routeur, par exemple) la valeur des CIR, Be, Bc, etc. Le support de cette norme évite d'avoir à configurer les paramètres en double, à la fois sur le commutateur et sur le routeur.

Combien de circuits virtuels ?

La première question à se poser est de savoir combien de PVC ou de SVC utiliser : un circuit virtuel (CV) par protocole, ou un seul circuit virtuel pour tous les protocoles ?

- Un CV par protocole permet de bénéficier d'un débit garanti, ce qui assure que les transferts de fichiers ne perturberont pas les flux conversationnels. Mais cette solution coûte plus cher, car les opérateurs facturent chaque CV en fonction du CIR affecté.
- Un seul CV est plus économique, mais tous les flux de réseaux locaux sont mêlés. Le contrôle du flux doit donc être reporté au niveau du routeur à travers un système d'affectation de priorité par protocole ou de réservation de ressources (voir chapitre 14). Ces mécanismes fonctionnent généralement bien.

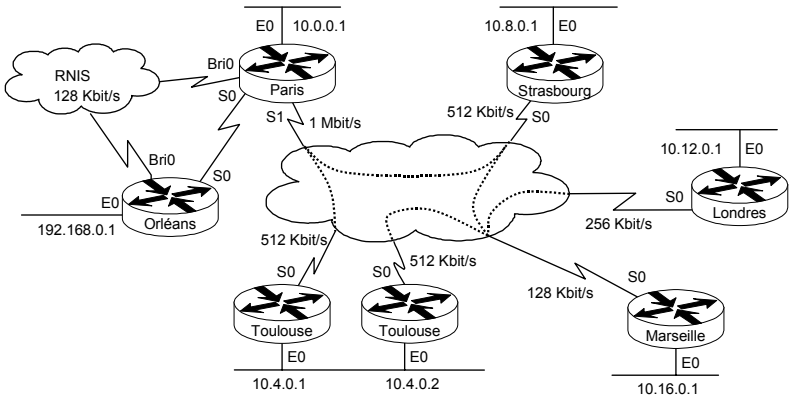
Nous choisirons donc la seconde solution.

Un site peut communiquer avec plusieurs autres sites, ce qui implique d'utiliser un CV par connexion, les CV Frame Relay étant de type point à point.

Si nous voulons construire un réseau parfaitement maillé, chaque routeur doit voir un CV par site distant. Cette configuration est envisageable, mais risque d'être onéreuse, car l'opérateur facture chaque CV.

Si vous voulez réduire les coûts, il vaut mieux définir des CV là où les flux sont les plus importants. Les autres communications transiteront éventuellement par plusieurs routeurs (par exemple, Marseille ↔ Paris dans notre cas).

Figure 11-8.
Exemples de circuits virtuels.



La question du choix entre PVC et SVC ayant été discutée précédemment, nous commencerons par configurer des PVC.

LA COMMUTATION FRAME RELAY (ITU Q.922)

La norme définit un second mode de fonctionnement, appelé " *Frame Switching Bearer Services* ", qui offre des mécanismes de reprise sur erreur et de contrôle de flux. Il repose sur l'utilisation de la procédure **LAP-F** (*Link Access Procedure-Frame*) inspirée du LAP-D du RNIS (Q.921), lui-même issu du LAP-B de HDLC. Le format des trames diffère légèrement, mais les principes restent les mêmes.

Ce *mode* de fonctionnement n'est actuellement utilisé que pour le PVC véhiculant la signalisation **Q.933** (dans le DLCI 0).

Octets	Bits 8	7	6	5	4	3	2	1
	Délimiteur							
1	DLCI						C	0
2	DLCI			F	B	D	E	
	Contrôle (1 ou 2 octets)							
	Données (1 à 2 048 octets)							
	FCS (2 octets)							
	Délimiteur							

DLCI = *Data Link Connection Identifier*
C = Bit C/R pour *Command / Response*
F = Bit FECN pour *Forward Explicit Congestion Notification*
B = Bit BECN pour *Backward Explicit Congestion Notification*
D = Bit DE pour *Discard Eligibility*
E = Bit EA pour *Extended Address*
FCS = *Frame Check Sequence*
Délimiteur = 01111110

•••

LA COMMUTATION FRAME RELAY (SUITE)

Le champ " Contrôle " a été ajouté à la trame *standard*. Il permet à la procédure LAPF de fournir un mécanisme de **reprise sur erreur** et de **contrôle de flux**. Il existe trois formats de trames (I, S, U) définissant autant de formats du champ de contrôle.

Bits	8	7	6	5	4	3	2	1
Trame I	N(S)							0
	N(R)							P/F
Trame S	0	0	0	0	Su	Su	0	1
	N(R)							P/F
Trame U	M	M	M	P/F	M	M	1	1

Trame I = Information
Trame S = Supervision
Trame U = Unnumbered
N(S) = Numéro séquence émission
N(R) = Numéro séquence réception
P/F = Poll/final (trame commande/réponse)
Su = Supervision
M = Modifier function
Les bits Su et M distinguent les 11 types de trames (SABME, UI, DISC, XID, RR, RNR, etc.)

Format de trame	Type	Description
I	--	Transport des données utilisateur
S	RR	*Receive Ready* : accusé de réception, et prêt à recevoir
	RNR	*Receive Not Ready* : accusé de réception, et non prêt à recevoir
	REJ	Reject : trame rejetée ; retransmettre à partir de N(R)
U	SABME	*Set Asynchronous Balanced Mode Extended* : initialisation de la procédure
	FRMR	*Frame Reject* : trame rejetée, erreur non récupérable
Etc.		

Afin de réduire le trafic de service, une **fenêtre d'émission** d'une taille *N* (entre 1 et 127) indique que le *commutateur* enverra *N* trames d'affilée et attendra un acquittement en retour. La taille de la fenêtre varie au cours du temps selon la qualité de la transmission (meilleure elle est, plus *N* est grand).

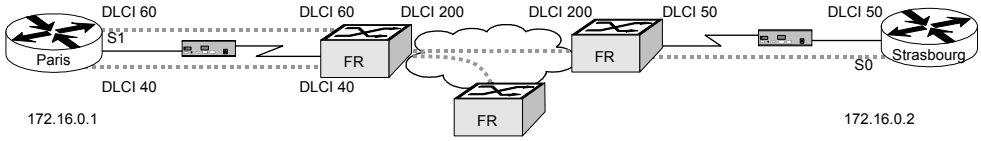
Le champ $N_A(S)$ est le **numéro de séquence** de la trame envoyée par le commutateur A ; $N_A(R)$ est celui de la trame dernièrement reçue du commutateur B, augmenté de 1. Si, après *N* trames envoyées, le commutateur A reçoit un $N_B(R)$ inférieur à son $N_A(S)$, cela signifie que des trames ont été perdues ou endommagées. Il retransmet alors toutes les trames à partir du $N_B(R)$ reçu, et $N_A(S)$ devient égal à $N_B(R)$.

Configurer les PVC

Dans notre réseau, un site peut communiquer avec plusieurs autres sites : les communications sont dites multipoint. L'utilisation des interfaces non numérotées (*unnumbered*) est toujours possible, mais il vaut mieux considérer le réseau Frame Relay comme étant un réseau IP à part entière : il sera plus évolutif et pourra être administré. Notre plan d'adressage (voir chapitre 5) prévoit d'affecter le réseau 172.16.0.0/16.

De plus, l'interface du routeur doit être configurée avec plusieurs DLCI (un PVC par site distant). À Paris, notre opérateur nous a affecté les DLCI 40 et 60, respectivement pour Toulouse et Strasbourg :

Figure 11-9. *Exemples de DLCI.*




```
interface serial1
ip address 172.16.0.1 255.255.255.248
encapsulation frame-relay ietf
frame-relay lmi-type q933a
frame-relay interface-dlci 40 broadcast
frame-relay interface-dlci 60 broadcast
```

Interface multipoint : jusqu'à 6 sites sur ce subnet

L'option broadcast indique que les broadcasts IP seront transmis sur la ligne série, permettant ainsi à des protocoles de routage comme OSPF de fonctionner (cf. chapitre 12).

Correspondance entre adresses IP et DLCI

Sur le réseau intersite, les routeurs utilisent le protocole **Inverse ARP** pour découvrir les adresses IP des routeurs distants et les associer aux DLCI.

LE POINT SUR INVERSE ARP (RFC 1293)

Le protocole **ARP** (*Address Resolution Protocol* — RFC 826) permet à une station IP de connaître l'adresse physique (MAC ou autre) d'une autre station en connaissant son adresse IP. Le protocole **RARP** (*Reverse ARP* — RFC 903) permet à une station d'obtenir, à partir de **son** adresse MAC et auprès d'un serveur d'adresses, l'adresse IP qui lui a été affectée.

InARP (*Inverse ARP*) est le mécanisme inverse d'ARP : ce protocole permet à une station (typiquement un routeur) de **connaître l'adresse IP du routeur se trouvant à l'autre bout d'un circuit virtuel** (Frame Relay ou ATM). Les paquets envoyés sont identiques à ceux d'ARP. Le routeur envoie une requête InARP, et attend une réponse contenant l'adresse IP du routeur distant. Le routeur associe alors l'adresse IP reçue au DLCI local du circuit virtuel. Plusieurs routeurs distants peuvent répondre si les circuits virtuels sont de type multipoint. InARP fonctionne sur le même principe pour d'autres protocoles, tels que Decnet, Apple Talk ou IPX.

Si vous rencontrez des problèmes d'incompatibilité avec ce protocole, ou si InARP n'est pas pris en charge par le routeur distant, vous devez associer manuellement l'adresse IP de destination avec le DLCI :

```
interface serial 1
ip address 172.16.0.1 255.255.255.248
encapsulation frame-relay ietf
frame-relay lmi-type q933a
frame-relay map ip 172.16.0.6 40 broadcast
frame-relay map ip 172.16.0.2 60 broadcast
```

L'opérateur configure les mêmes DLCI sur son équipement (FRAD ou commutateur) relié au routeur.

Configurer les SVC

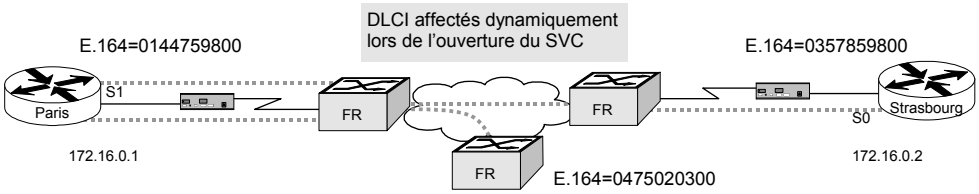
Le choix de circuits virtuels commutés peut être intéressant à double titre :

- Du point de vue du coût, tout d'abord. L'opérateur facture le SVC uniquement lorsqu'il est ouvert (facturation à la durée et/ou au volume).
- Du point de vue des performances ensuite. Si de nombreux CV doivent être utilisés (cas d'un réseau parfaitement maillé, par exemple), fermer ceux qui sont inutiles permet de libérer des ressources mémoire et CPU dans les routeurs et les commutateurs.

En plus de la signalisation Q.933 et de la procédure LAP-F, les SVC nécessitent l'emploi d'un troisième mécanisme, celui de **l'adressage** (voir la fin de ce chapitre pour plus de détails). Frame Relay utilise soit celui du RNIS (E.164) dans le cas des réseaux publics, soit X.121 dans le cas des réseaux privés et publics. Cet adressage de niveau 3 est utilisé par le protocole de signalisation Q.933 pour ouvrir les SVC, c'est-à-dire affecter dynamiquement les DLCI d'entrée et de sortie dans chaque commutateur traversé. Les trames sont ensuite acheminées en fonction des DLCI qui réalisent un adressage de niveau 2.

Cette fois, l'activation des SVC nécessite une configuration manuelle, non des DLCI, mais des adresses E.164 ou X.121, ainsi que l'association des SVC avec les adresses IP.

Figure 11-10. *Affectation dynamique des DLCI.*



```
interface ser 1
ip address 172.16.0.1 255.255.255.248
encapsulation frame-relay ietf
frame-relay svc
frame-relay lmi-type q933a
map group strasbourg
map group toulouse

map-list strasbourg source-addr E.164 0144759800 dest-addr E.164
0357859800
ip 172.16.0.2 class operateur broadcast ietf

map-list toulouse source-addr E.164 0144759800 dest-addr E.164
0475020300
ip 10.16.0.6 class operateur broadcast ietf

map-class frame-relay operateur
frame-relay traffic-rate 256000 384000
```

Encapsulation selon la RFC 2427

Paramètre optionnel

CIR = 256 Kbps

AIR = CIR + EIR

Les CIR et AIR sont utilisés lors de la négociation qui a lieu à l'ouverture du SVC. Ces paramètres doivent être identiques à ceux configurés dans le commutateur.

Gérer la qualité de service

Quel que soit le type de CV, le routeur peut moduler les flux de données selon l'état du réseau (en cas de congestion). Sur les routeurs Cisco, la qualité de service est gérée en activant la fonction *traffic shaping* (voir chapitre 14). À cette occasion, le routeur met en place une file d'attente par DLCI, et adapte le flux en fonction des informations envoyées par le commutateur :

```
interface serial 1
ip address 172.16.0.1 255.255.255.248
encapsulation frame-relay ietf
frame-relay lmi-type q933a
frame-relay traffic-shaping
```

De plus, l'activation de la fonction **ELMI** (*Enhanced Local Management Interface*) permet au routeur de connaître automatiquement les paramètres Frame Relay en recevant les messages LMI du commutateur (sur le DLCI 0). Le commutateur communique ainsi au routeur les valeurs des paramètres CIR, Be et Bc :

```
frame-relay qos-autosense
```

Enfin, les commutateurs peuvent informer les routeurs de l'état du réseau de deux manières : *via* les messages **CLLM** (*Consolidated Link Layer Management*) ou *via* le bit **BECN** (*Backward Explicit Congestion Notification*). Nous préférons le mode CLLM (appelé *foresight* chez Cisco) au mode BECN, moins fiable (voir encadré « Le point sur Frame Relay »).

```
interface s 1
ip address 172.16.0.1 255.255.255.248
encapsulation frame-relay ietf
frame-relay lmi-type q933a
frame-relay traffic-shaping
frame-relay qos-autosense
frame-relay class operateur
!
map-class frame-relay operateur
  frame-relay adaptive-shaping foresight
```

Pour les SVC, ajoutez :
frame-relay svc
map group ….
map-list …
ip … class operateur …

ELMI

…et supprimez cette commande.

CLLM

Il se peut que les fonctions CLLM et ELMI ne soient pas disponibles sur le FRAD de l'opérateur ou que l'on y rencontre des incompatibilités. La solution est donc de configurer manuellement tous les paramètres. L'équivalent de la configuration précédente est, de ce fait, plus complexe :

```
interface s 1
bandwidth 512
ip address 172.16.0.1 255.255.255.248
encapsulation frame-relay ietf
frame-relay lmi-type q933a
frame-relay traffic-shaping
frame-relay class operateur

map-class frame-relay operateur
  frame-relay traffic-rate 256000 384000
```

Remplace ELMI

La commande *traffic-rate* indique le CIR et l'AIR, que nous avons respectivement positionnés à 256 Kbit/s et 384 Kbit/s. Si la valeur de l'AIR est omise, la valeur par défaut est celle du débit de la ligne indiquée par la commande *bandwidth*.

Il est possible de définir plus finement les paramètres de qualité de service décrits dans la norme Q.922. Le routeur Cisco permet même de le faire dans les deux sens, bien que, généralement, les opérateurs proposent des valeurs identiques afin de simplifier la configuration des commutateurs et la grille tarifaire. Le profil « personnalisé » présenté ici se substitue alors au profil opérateur précédemment décrit :

```
map-class frame-relay personnalise
frame-relay cir in 1280000
frame-relay bc in 256000
frame-relay be in 256000
frame-relay cir out 2560000
frame-relay bc out 256000
frame-relay be out 128000
frame-relay idle-timer 30
```

CIR = Bc/Tc

EIR = Be/Tc

AIR = CIR + EIR = (Bc+Be) / Tc

À partir des 3 valeurs CIR, Bc et Be, on peut déduire celle de Tc (2 secondes en entrée et 1 seconde en sortie).

Toujours dans l'optique de gérer la qualité de service, vous pouvez affecter une priorité plus faible à certains protocoles en positionnant le bit DE dans les trames qui les véhiculent :

```
int s 1
frame-relay de-group 1 50
                                  Liste de DLCI (rien = tous les DLCI)
frame-relay de-list 1 protocol ip characteristic tcp 20
frame-relay de-list 1 interface e 0 characteristic list 100 access-list
101
...
```

Par exemple, des protocoles comme FTP pourront être marqués de cette manière afin de privilégier la voix sur IP en cas de congestion du réseau.

Les sous-interfaces

Dans certains cas, il peut être intéressant d'utiliser le principe des sous-interfaces proposé par Cisco, afin d'activer un secours RNIS individuellement, par PVC, et non pas sur la chute de l'interface série. Les sous-interfaces permettent également de configurer le traffic shaping, ainsi que d'autres paramètres Frame Relay PVC par PVC.

Sur notre routeur Cisco, l'interface physique est configurée en Frame Relay et est associée à des sous-interfaces logiques (une par DLCI). Nous avons choisi un DLCI par site distant, donc un mode point à point pour chaque sous-interface, ce qui permet de ne pas utiliser d'adresse IP sur la liaison WAN.

```
interface s 1
encapsulation frame-relay ietf
frame-relay lmi-type q933a
frame-relay traffic-shaping
frame-relay qos-autosense

interface serial1.1 point-to-point
ip address 172.16.0.1 255.255.255.252
frame-relay interface-dlci 40
  class operateur

interface serial1.2 point-to-point
ip address 172.16.0.5 255.255.255.252
frame-relay interface-dlci 60
  class personalise
```

Pour terminer, il est possible d'optimiser l'utilisation des liaisons série en compressant les données (norme FRF.9 du Frame Relay Forum).

```
frame-relay map payload-compress frf9 stac
```

Méthode Stacker

Cette commande peut s'appliquer à une interface physique ou à une sous-interface.

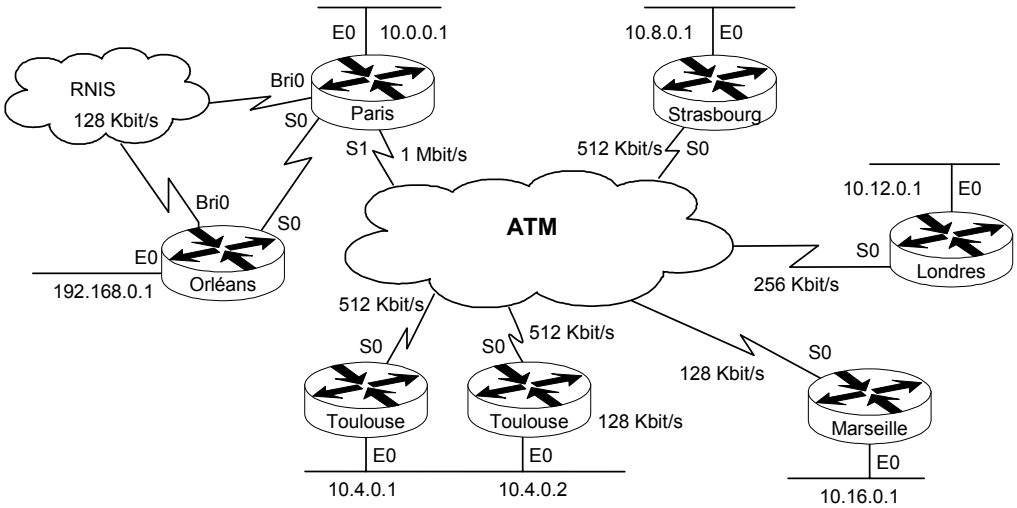
Mettre en place un réseau ATM

Les réseaux ATM sont réservés aux hauts débits : à partir de 34 Mbit/s en France, et de 45 Mbit/s aux États-Unis. La facturation associée est donc très élevée.

Nous nous plaçons ici dans cette situation.

Comme précédemment, l'opérateur local installe son modem numérique (le CSU) dans vos locaux. L'opérateur retenu pour le réseau ATM installe éventuellement un commutateur ATM qu'il connecte au modem. L'interface série du routeur sera ensuite connectée à l'un des ports du commutateur.

Figure 11-11.
Réseau intersite reposant sur un réseau opérateur ATM.



Qualité de service et facturation

Avec ATM, les opérateurs peuvent contrôler précisément la qualité de service. Ils élaborent leur offre commerciale à partir des mécanismes offerts par le protocole.

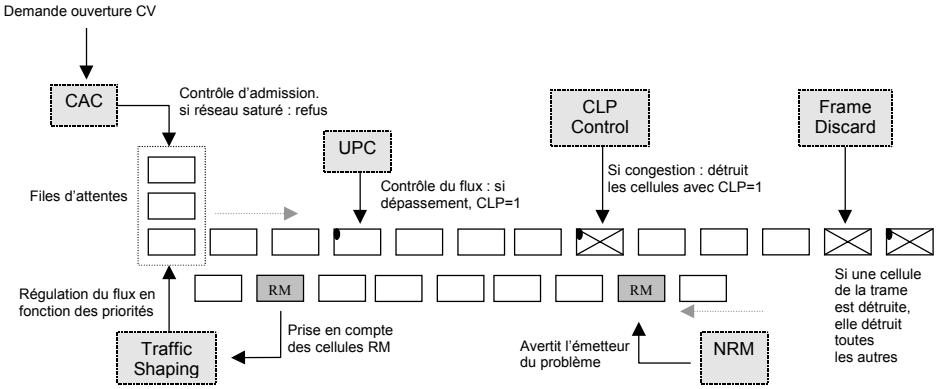
Par exemple, ATM permet de **gérer la congestion** du réseau et d'offrir différentes **classes de service**, ce qui permet de proposer une facturation adaptée à chaque type de situation (flux voix et/ou données et/ou vidéo). Comment ? La réponse est intimement liée au fonctionnement du protocole ATM.

La gestion du trafic : TMS 4.0 (ATM Forum af-tm-0056.000)

La spécification TMS (*Traffic Management Specification*) définit les paramètres et procédures relatifs à la gestion du trafic ainsi qu'à la qualité de service. Elle reprend les normes de l'ITU en précisant leur fonctionnement (paramètres et procédures pris en charge ou non) et, surtout, en les modifiant. Elle définit également les procédures et algorithmes permettant aux commutateurs ATM d'assurer la qualité de service demandée.

Procédure	Description
CAC (*Connection Admission Control*)	Avant d'autoriser l'ouverture d'un CV, chaque commutateur vérifie qu'il pourra bien assurer la qualité de service demandée sans remettre en cause celles déjà accordées.
UPC (*Usage Parameter Control*)	Chaque commutateur vérifie que le flux émis dans un CV respecte bien la QOS demandée. En cas de dépassement, les cellules sont marquées avec le bit **CLP** positionné à « 1 » et peuvent être détruites.
CLP Control (*Cell Loss Priority Control*)	L'équipement terminal (par exemple, le routeur) peut positionner le bit CLP à « 1 » dans les cellules qui ne sont pas prioritaires. En cas de congestion ou de dépassement de la QOS, les commutateurs détruiront ces cellules en priorité.
NRM (*Network Resource Management*)	Si des signes de congestion apparaissent (remplissage d'une file d'attente, par exemple), les commutateurs peuvent émettre des cellules **RM** à destination de l'émetteur d'un flux (retour sur la qualité de service - feedback) l'invitant à réguler son trafic *via* le mécanisme de traffic shaping.
Traffic Shaping	Les équipements terminaux et les commutateurs peuvent modifier les caractéristiques du flux émis (réduction du débit, régulation d'un trafic erratique, etc.).
Frame Discard	Si une cellule est détruite, le commutateur peut détruire toutes les autres cellules appartenant aux mêmes données (par exemple, toutes les cellules d'un paquet IP).

Figure 11-12. *Gestion de la qualité de service par ATM.*



Les classes de service ATM Transfer Capabilities (ITU I.371 et TMS)

Les procédures détaillées ci-dessus ont pour but d'assurer une certaine qualité de service aux applications. Celles-ci ont le choix entre **cinq classes de service** adaptées à divers types de flux. Par exemple, les cellules véhiculant un canal voix d'une conversation téléphonique doivent être transmises avec une grande régularité (les horloges de l'émetteur et du récepteur doivent être synchronisées). En revanche, un flux de type réseau local, qui est par nature erratique (trafic par rafales), ne nécessite pas la même qualité de service.

Ces classes de service sont accessibles *via* différentes interfaces d'accès à ATM, appelées **AAL** (*ATM Adaptation Layer*).

Classe de service	Caractéristiques du trafic	Applications
CBR (*Constant Bit Rate*)	Le débit est constant et garanti.	Voix non compressée en émulation de circuit (accès de préférence *via* AAL-1)
rt-VBR (*Real-Time Variable Bit Rate*)	Le débit est variable et est garanti. Le délai de transit est garanti et varie peu.	Données (accès de préférence *via* AAL-2)
nrt-VBR (*Non Real-Time Variable Bit Rate*)	Le débit est variable et est garanti.	Vidéo + voix compressées (accès de préférence *via* AAL-2 ou AAL-5)
ABR (*Avaible Bit Rate*)	Le débit est variable et peut être modifié si le réseau le demande (en cas de congestion, par exemple).	Données (accès de préférence *via* AAL-5)
UBR (*Unavaible Bit Rate*)	Acheminement sans garanties, au mieux des capacités du réseau (*best effort*) .	Données (accès de préférence *via* AAL-5)

La classe de service UBR n'offre aucune garantie de service, tandis que l'ABR est la plus utilisée, notamment par l'UNI 4.0.

LA SIGNALISATION UNI 4.0 (ATM FORUM AF-SIG-0061.000)

L'UNI (*User Network Interface*) reprend les normes de l'ITU en précisant leur fonctionnement (paramètres et procédures pris en charge ou non), et surtout en les modifiant :

- Appels point à point (**Q.2931**) sans les fonctions OAM (*Operations, Administration and Maintenance*).
- Adressage **NSAP** (ISO 8348, ITU X.213, RFC 1629).
- Couche **SAAL** (Q.2100, Q.2110, Q.2130) ; VPI/VCI = 0/5.
- Appels point à multipoint (**Q.2971**) avec extensions pour que les feuilles puissent accepter de nouveaux participants à la conférence en cours (dans Q.2971, une conférence à plusieurs est gérée sous forme d'arbre, dont seule la racine permet à d'autres utilisateurs de rejoindre une conférence).
- Direct Dialling In (Q.2951).

L'ATM Forum a ajouté les extensions suivantes à Q.2931 :

- Gestion des **adresses de groupe** ATM (adresses *anycasts*) identifiées par le préfixe C50079. Par exemple, le groupe C50079.00000000000000000000.00A03E000001.00 permet de joindre le LECS (*LAN Emulation Configuration Server*)
- Négociation des caractéristiques de la connexion (CBR, VBR, rt-VBR, ABR, UBR).
- Présentation individuelle des paramètres de qualité de service.
- Paramètres décrivant la classe de service **ABR**.
- **Signalisation proxy** : un nœud ATM peut gérer la signalisation à la place d'un autre nœud qui ne la supporte pas. Permet également à un serveur disposant de plusieurs interfaces ATM de ne posséder qu'une seule adresse NSAP.

En l'état actuel de la norme, les cellules RM (*Resource Management*) ne sont utilisées que pour réguler le trafic ABR.

Connecter le routeur au réseau de transport

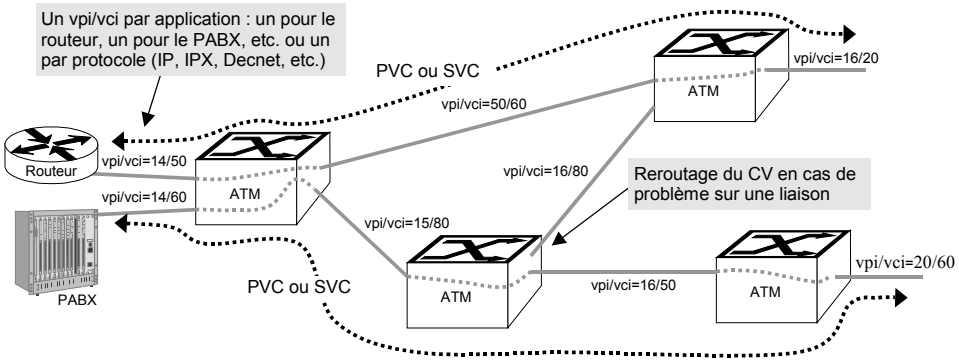
La norme ATM définit un ensemble de protocoles reposant sur la commutation de petites cellules de taille fixe (58 octets dont 5 d'en-tête).

Chaque cellule est identifiée par un couple d'adresses **VPI** (*Virtual Path Identifier*) et **VCI** (*Virtual Channel Identifier*), adresse locale partagée par deux commutateurs. Il n'y a pas d'adressage de bout en bout, mais uniquement un adressage point à un point entre deux commutateurs. Un commutateur recevant une trame avec un VPI/VCI donné la routera sur un autre port et l'enverra avec un autre VPI/VCI.

La connexion entre deux commutateurs ATM s'effectue par l'ouverture de circuits virtuels permanents (**PVC**, *Permanent Virtual Circuit*) ou commutés (**SVC**, *Switched Virtual Circuit*), c'est-à-dire ouverts à la demande *via* le protocole de signalisation Q.2931 (qui utilise quelques Kbit/s du VPI/VCI = 0/5).

L'intérêt des SVC est que la qualité de service peut être spécifiée à la demande permettant ainsi de réduire (encore) les coûts. Sur un PVC, le débit est fixé une fois pour toutes et engendre un coût fixe.

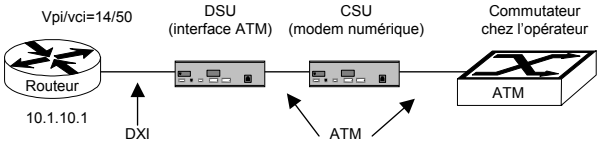
Figure 11-13.
Circuits virtuels et VPI/VCI ATM.



Si le routeur ne dispose pas d'interface ATM

La manière la plus simple de connecter nos réseaux locaux est de configurer l'interface série du routeur en mode **DXI** (*Data eXchange Interface*). Ce mode de fonctionnement implique l'utilisation d'un DSU (*Data Service Unit*) externe fourni par l'opérateur. Le DSU se charge d'adapter le flux série émis et reçu par le routeur en cellules ATM.

Figure 11-14.
Connexion d'un routeur à un commutateur ATM.



L'exemple suivant montre un PVC configuré sur notre routeur parisien à destination du réseau de Strasbourg :

```
interface serial 0
ip address 172.16.0.1 255.255.255.248
encapsulation atm-dxi
dxi pvc 14 50 mux
dxi map ip 172.16.0.2 14 50 broadcast
```

VPI = 14 / VCI = 50

Le paramètre « mux » indique que nous avons choisi l'encapsulation de type multiplexage par circuit virtuel : un seul protocole (IP dans notre cas) utilisera le PVC identifié par le VPI 14 et le VCI 50. L'encapsulation LLC/SNAP (paramètre « snap » à la place de « mux ») n'apporterait aucun avantage et ajouterait un overhead de 8 octets par paquet IP.

L'INTERFACE DXI (ATM FORUM AF-DXI-0014.000)

La norme **DXI** (*Data eXchange Interface*) définit le protocole de niveau 2 utilisé pour échanger les données entre un équipement non ATM (*via* une **interface série** V.35 ou HSSI) et un DSU ATM. Trois modes de fonctionnement sont possibles.

Dans le **mode 1a**, le routeur place ses données dans une trame DXI (SDU de 9 232 octets), puis le DSU réalise l'encapsulation AAL-5, la segmentation SAR AAL-5 et l'accès UNI. Ce mode supporte 1 023 CV simultanés.

Dans le **mode 1b**, le routeur réalise l'encapsulation AAL-3/4 puis place le résultat dans une trame DXI (SDU de 9 224 octets). Le DSU réalise la segmentation SAR AAL-3/4 et l'accès UNI. Le mode 1a est également supporté, mais le mode 1b doit être utilisé pour au moins 1 CV. Ce mode supporte 1 023 CV simultanés.

Dans le **mode 2**, le routeur réalise l'encapsulation AAL-3/4, puis place le résultat dans une trame DXI (SDU de 65 535 octets). En cas de besoin, le DSU peut convertir l'encapsulation AAL-3/4 en AAL-5 (changement d'encapsulation). Le DSU réalise ensuite l'assemblage SAR (AAL-3/4 ou AAL-5) et l'accès UNI. Ce mode supporte 16 millions de CV simultanés.

Si le routeur supporte ATM

La seconde solution consiste à insérer une carte ATM dans le routeur. Fonctionnellement, celle-ci se comporte comme un DSU et peut être directement connectée au CSU (c'est-à-dire le modem numérique) de l'opérateur ou à un commutateur.

Pour la même interconnexion Paris-Strasbourg et Paris-Toulouse, la configuration devient la suivante :

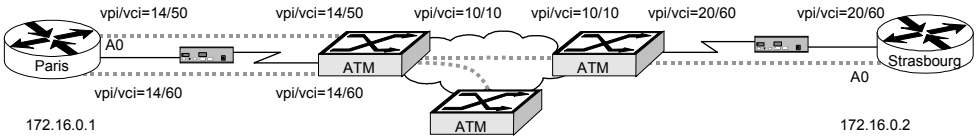


Figure 11-15. *Configuration des VPI/VCI ATM.*

```
interface atm 0
ip address 172.16.0.1 255.255.255.248
```

```
atm pvc 1 14 50 aal5mux ip
atm pvc 2 14 60 aal5mux ip
map-group operateur

map-list operateur
ip 172.16.0.2 atm-vc 1 broadcast
ip 172.16.0.6 atm-vc 2 broadcast
```

Identifiant du PVC propre à Cisco (interne au routeur)

L'encapsulation choisie, « aal5mux », est la même que précédemment : un CV transporte le protocole IP. Cependant, l'encapsulation LLC/SNAP, utilisée conjointement avec l'interface ATM, nous donne la possibilité supplémentaire de pouvoir exécuter **Inverse ARP**, et donc de simplifier la configuration :

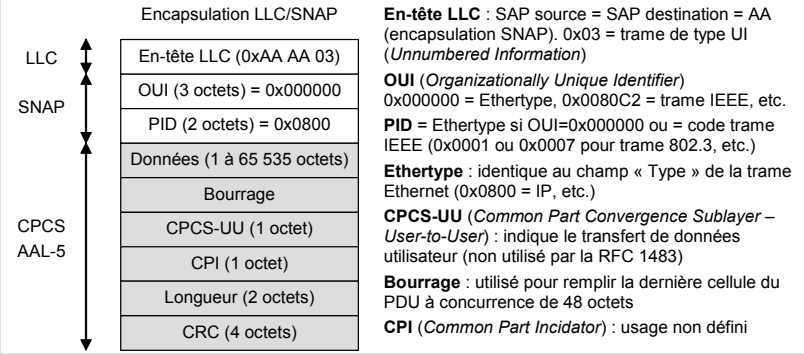
```
interface atm 0
ip address 172.16.0.1 255.255.255.248
atm pvc 1 14 50 aal5snap inarp 5
atm pvc 2 14 60 aal5snap inarp 5
```

La valeur 5 affectée à « inarp » indique que la procédure inverse ARP est lancée toutes les 5 minutes. La correspondance manuelle entre adresse IP et circuit virtuel n'est donc plus nécessaire.

L'ENCAPSULATION DES PROTOCOLES DANS LES CELLULES ATM (RFC 1483)

Le transport d'un protocole dans les cellules ATM requiert l'utilisation de l'interface **AAL-5** (*ATM Adaptation Layer 5*), accessible *via* la couche **CPCS** (*Common Part Convergence Sublayer*). L'encapsulation peut être réalisée de deux manières : soit en dehors d'ATM, soit au niveau d'ATM.

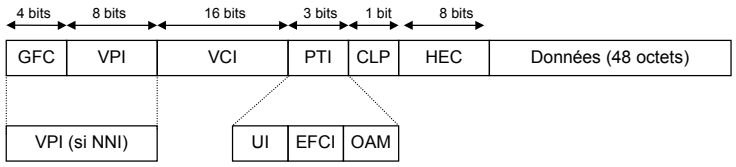
La première méthode permet de transporter plusieurs protocoles dans un seul circuit virtuel ATM ; elle utilise pour cela l'encapsulation **LLC/SNAP** (*Logical Link Control/Sub Network Access Protocol*).



La seconde méthode, appelée **multiplexage par circuit virtuel**, consiste à encapsuler le protocole directement dans le PDU AAL-5, ce qui implique l'utilisation d'un circuit virtuel ATM par protocole.

LE POINT SUR ATM (ITU I.361)

ATM (*Asynchronous Transfer Mode*) découpe la bande passante en tranches de temps fixe appelées **cellules**.



GFC (*Generic Flow Control*). Priorité de la cellule (0 = la plus basse).

VPI (*Virtual Path Identifier*). Identifie le chemin virtuel (255 possibilités).

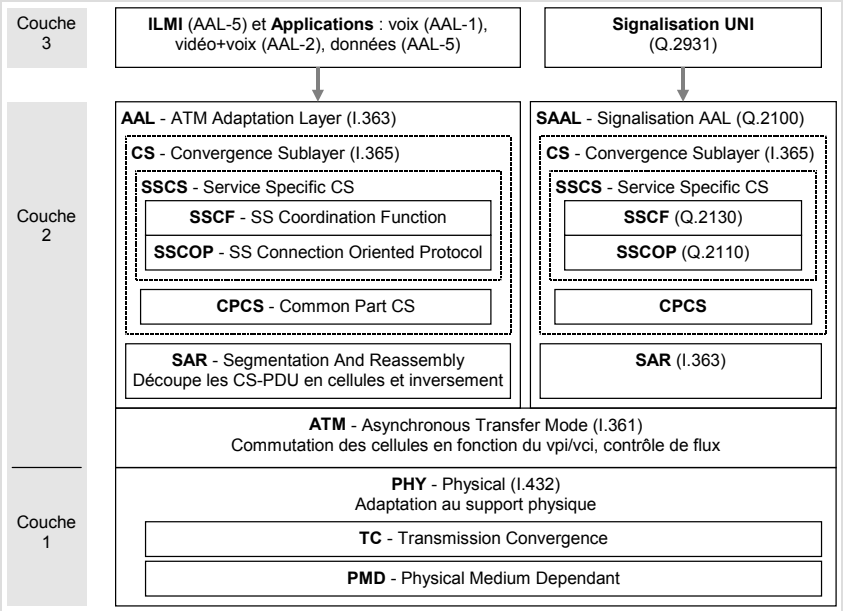
VCI (*Virtual Channel Identifier*). Identifie la voie virtuelle au sein du chemin virtuel (65 535 possibilités).

PTI (*Payload Type Indicator*). Le premier bit indique si la cellule transporte des données de contrôle ou des données utilisateur. Dans ce dernier cas, le deuxième bit, appelé **EFCI** (*Explicit Forward Congestion Indication*), indique à l'application qu'il faut prévoir des délais d'acheminement pour les cellules à venir (suite à une congestion par exemple). Positionné à "1", le troisième bit indique que le champ d'information contient des données utilisées par les applications d'administration **OAM** (*Operations, Administration and Maintenance*).

CLP (*Cell loss Priority*). Positionné à "1", ce bit indique que la cellule peut être détruite par le commutateur en cas de congestion.

HEC (*Header Error Control*). Cet octet permet à la couche TC (*Transmission Convergence*) d'opérer un contrôle d'erreur sur l'en-tête de la cellule.

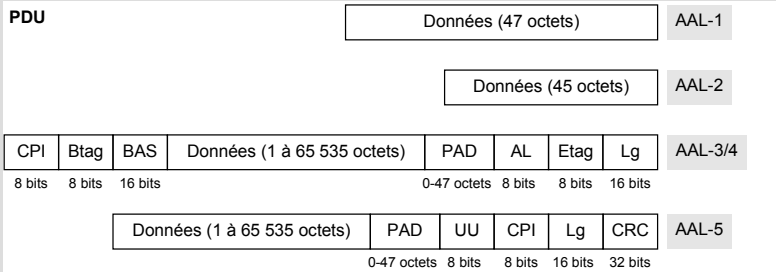
Les applications transmettent leurs données à la couche **AAL** (*ATM Adaptation Layer*) qui se charge de les convertir en cellules, puis de les envoyer en respectant le niveau de service demandé (AAL-1 à AAL-5 et SAAL).



• • •

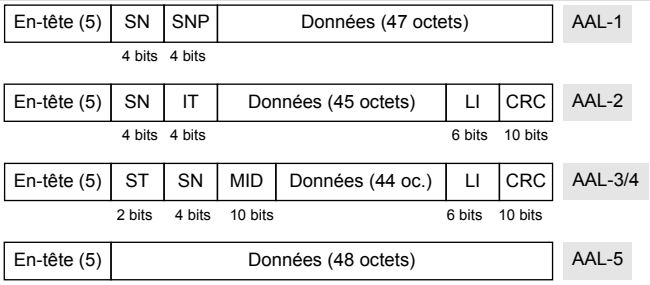
LE POINT SUR ATM (ITU I.361 – SUITE)

La couche AAL, et donc les sous-couches qui la composent, est adaptée à chaque type de service (AAL-1, AAL-2, AAL-3/4, AAL-5 et SAAL). Par exemple, la couche **CS** accepte des données au format **CS-PDU** (*Convergence Sublayer-Protocol Data Unit*).



CPI = *Common Part Indicator* : fonction non encore définie
Btag = Délimiteur de début
BAS = *Buffer Allocation Size*
PAD = Bourrage pour remplir la dernière cellule à 48 octets
AL = Alignement
Etag = Délimiteur de fin
Lg = Longueur du champ de données
UU = *User-to-User indicator*
CRC = *Cyclic Redundancy Check* : code de contrôle d'erreur portant sur le champ de données

De même, la couche **SAR** structure différemment les 48 octets du champ de données des cellules ATM (appelé SAR-PDU).



SN = *Sequence Number* : détecte les cellules manquantes ou erronées
SNP = *Sequence Number Protection* : code autocorrecteur portant sur le SN
IT = *Information Type* : début, continuation ou fin d'un CS-PDU
ST = *Segment Type* : début, fin, continuation ou segment simple
MID = *Multiplexing Identifier* : partage d'un circuit virtuel par plusieurs applications de la couche SAR
LI = *Length Indicator* : nombre d'octets significatifs dans le cas d'une cellule partiellement remplie
CRC = *Cyclic Redundancy Check* : code de contrôle d'erreur portant sur le champ de données

En comparant les deux schémas précédents, on constate que les PDU des AAL-1 et AAL-2 sont directement insérés dans une cellule ATM. La couche AAL n'utilise donc pas obligatoirement les mécanismes de segmentation et d'assemblage.

Il existe également une couche SSCS pour Frame Relay (I.365.1).

Configurer les SVC

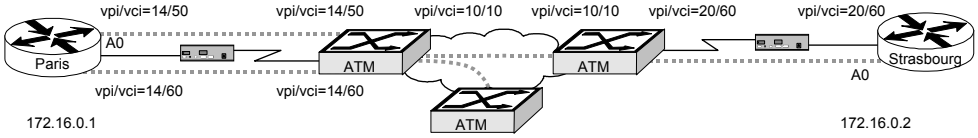
Si l'opérateur nous en donne la possibilité, la mise en place de SVC permet, comme pour Frame Relay, de réaliser des économies : l'opérateur nous facture en fonction de l'utilisation (à la durée et/ou au volume). De plus, si les CV sont nombreux, la fermeture des SVC, s'ils ne sont pas utilisés, permet de libérer des ressources CPU et mémoire dans les routeurs et les commutateurs.

L'utilisation des SVC implique l'activation de deux nouveaux mécanismes :

- Un PVC pour le protocole SAAL (*Signaling ATM Adaptation Layer*) qui gère les SVC (ouverture, fermeture, etc.). Le VPI/VCI utilisé est 0/5.
- L'utilisation d'un adressage global (de niveau 3) permettant d'identifier les nœuds du réseau. L'adressage utilisé par ATM est de type NSAP ; trois encapsulations d'adresses sont possibles : DCC, ICD et E.164 (reportez-vous à la fin de ce chapitre pour plus de détails).

La commande **map-list** permet de configurer manuellement la correspondance entre adresses NSAP et adresses IP :

Figure 11-16. *Configuration des VPI/VCI ATM.*

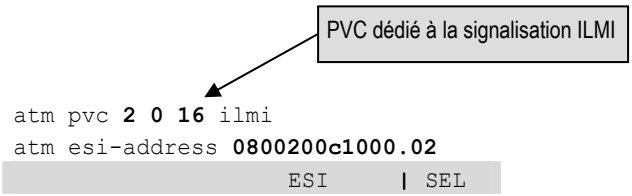



```
interface atm 0
ip address 172.16.0.2 255.255.255.248
atm nsap-address
 47.0091.81.000000.0061.705b.7701.0800.200c.1A2B.01
AFI|ICD |       Préfixe DSP      |     ESI      |SEL
atm pvc 1 0 5 qsaal
map-group operateur
!
map-list operateur
ip 172.16.0.2 atm-nsap
47.0091.81.000000.0061.705b.7701.0800.200c.1000.02 broadcast
ip 172.16.0.6 atm-nsap
47.0091.81.000000.0061.705b.88a7.0900.20ab.2000.01 broadcast
```

PVC dédié à la signalisation Q.2931

Adresses ATM

Une autre manière d'affecter l'adresse NSAP à l'interface est d'utiliser la signalisation **ILMI** (*Integrated Local Management Interface*). Elle permet au routeur d'obtenir le préfixe de l'adresse, le DSP (champs ESI et SEL) étant toujours affecté par l'équipement terminal (le routeur). Le PVC dédié à ILMI est : VPI/VCI = 0/16 :



La première partie du chiffre correspond au champ « ESI » (6 octets, l'équivalent d'une adresse MAC) ; la seconde au champ « Selector » (1 octet) de l'adresse NSAP. Le reste de l'adresse (13 octets) est affecté automatiquement par le commutateur ATM *via* ILMI.

L'activation de cette signalisation permet également au routeur et au commutateur de surveiller l'état des circuits virtuels.

LA SIGNALISATION ILMI 4.0 (ATM FORUM AF-ILMI-0065.00)

La signalisation ILMI (*Integrated Local Management Interface*)* est un protocole permettant aux commutateurs ATM d'échanger des informations avec les équipements terminaux (stations ATM, routeurs, etc.), telles que :

- la configuration et l'état des circuits virtuels ;
- les préfixes des adresses NSAP ;
- les services et options supportés par les équipements.

Le protocole utilisé par ILMI est **SNMP** (*Simple Network Management Protocol* — RFC 1157) qui utilise les services de **AAL-5** (VPI/VCI = 0/16). La base de données **MIB** (*Management Information Base*) contient des informations relatives aux couches physique et ATM, aux circuits virtuels (état des VPI/VCI), aux adresses, ainsi qu'aux services supportés et pouvant être négociés (version UNI, nombre maximal de VPI/VCI, qualité de service, etc.).

Par exemple, ILMI permet de configurer les LEC (*LAN Emulation Client*) et de trouver le LECS (*LAN Emulation Configuration Server*).

La bande passante utilisée par ILMI ne doit pas dépasser 1 % du débit de la ligne, et 5 % en pic.

*Note : le 1er « I » signifiait *Interim* car l'ATM Forum attendait la normalisation de l'ITU. Mais celle-ci s'étant fait attendre, l'ATM Forum a définitivement entériné sa proposition de norme.

Gérer la qualité de service

Les normes prévoient de nombreux et complexes mécanismes pour gérer la qualité de service sur ATM. Seuls certains d'entre eux sont disponibles sur nos routeurs.

Pour les PVC, seules les caractéristiques du flot de données peuvent être spécifiées :

```
atm pvc 1 14 50 aal5snap 384 256 inarp 5
```

Cette commande précise qu'un maximum de 384 Kbit/s sera alloué à notre PVC et que le débit moyen du trafic sera de 256 Kbit/s.

Les possibilités de paramétrage sont plus étendues en ce qui concerne les SVC puisque l'on peut demander l'activation de **classes de service**. Celles-ci sont décrites implicitement par des combinaisons de paramètres :

```
map-list operateur
ip 172.16.0.2 atm-nsap
47.0091.81.000000.0061.705b.7701.0800.200c.1000.02 broadcast class
traficUBR
ip 172.16.0.6 atm-nsap
47.0091.81.000000.0061.705b.88a7.0900.20ab.2000.01 broadcast class
traficNrtVBR

map-class traficUBR
atm forward-peak-cell-rate-clp1 384
atm backward-peak-cell-rate-clp1 256

map-class traficNrtVBR
atm forward-peak-cell-rate-clp1 384
atm forward-sustainable-cell-rate-clp1  256
atm forward-max-burst-size 128
atm backward-peak-cell-rate-clp1 384
atm backward-sustainable-cell-rate-clp1  256
atm backward-max-burst-size 128
```

Cette combinaison de paramètres active la classe de service nrt-VBR.

La qualité de service peut être différente dans chaque sens.

Le suffixe « clp1 » indique que le paramètre s'applique aux cellules dont le bit CLP est positionné à 1 ou 0. Le suffixe « clp0 » permet d'appliquer les mêmes paramètres aux cellules dont le bit CLP est à 0 (c'est-à-dire non marquées en suppression).

Dans le cas de la classe de service « traficUBR », aucune bande passante n'est réservée pour le SVC. C'est le mode de fonctionnement par défaut si aucun paramètre n'est spécifié.

Enfin, il est possible d'activer la procédure de contrôle CAC (*Connection Admission Control*) au niveau du routeur.

```
atm sig-traffic-shaping strict
```

Avec cette commande, l'ouverture d'un SVC ne sera possible que si les commutateurs ATM sont capables d'assurer la qualité de service demandée.

Les paramètres décrivant les classes de service (ITU I.356 et ATM Forum TMS 4.0)

Chaque classe de service (UBR, ABR, etc.) est définie par un ensemble de paramètres qui décrivent les caractéristiques du flux qui sera généré, ainsi que la qualité de service demandée. Ces paramètres peuvent, par exemple, être configurés sur nos routeurs.

Paramètre	Description du trafic
PCR (*Peak Cell Rate*)	Débit maximal autorisé en pointe (nombre maximal de cellules par secondes).
SCR (*Sustainable Cell Rate*)	Débit moyen autorisé (nombre moyen de cellules par secondes).
MBS (*Maximum Burst Size*)	Nombre de cellules autorisées pendant le débit en pointe. PCR/MBS = durée pendant laquelle le débit en pointe est autorisé.
MCR (*Minimum Cell Rate*)	Débit minimal demandé (nombre minimal de cellules par secondes).
Paramètre	**Description de la qualité de service**
CDV (*Cell Delay Variation*)	Demande d'une variation maximale du délai de transit (la gigue - *jitter*). Le flux doit être le plus constant possible (± CDV ms). La fonction UPC utilise pour cela l'algorithme GCRA (*Generic Cell Rate* Algorithm) de type *Leaky-bucket* (1) .
MCTD (*Maximum Cell Transfer Delay*)	Délai maximal de transit des cellules entre l'UNI de l'émetteur et l'UNI du récepteur.
CLR (*Cell Loss Ratio*)	Pourcentage acceptable de cellules pouvant être perdues (appliqué aux cellules ayant le bit CLP positionné à « 0 »).
Autre paramètre	**Description**
RM (*Resource Management*)	Traitement des cellules RM permettant d'adapter le trafic en fonction de l'état du réseau reporté par les commutateurs (*feedback*).

(1) *Leaky-bucket* signifie littéralement « seau (d'eau) percé ». Voir le chapitre 14 à ce sujet.

Le tableau suivant indique les combinaisons autorisées. Ainsi, le paramètre PCR seul indique implicitement la classe de service UBR. En revanche, la configuration des paramètres

SCR et MBS seuls ne correspond à aucune classe de service ; elle représente donc une combinaison invalide.

Classe de service	Description du trafic				Qualité de service			Autre paramètre
	PCR	SCR	MBS	MCR	CDV	MCTD	CLR	RM
CBR	X				X	X	X	
rt-VBR	X	X	X		X	X	X	
nrt-VBR	X	X	X		X		X	
ABR	X			X	X		X	X
UBR	X							

D'autres paramètres relatifs à la gestion de la qualité de service sont prévus par la norme :

- CER (*Cell Error Ratio*). Taux maximal de cellules pouvant être en erreur ;
- SECBR (*Severely Errored Cell Block Ratio*). Taux maximal de cellules consécutives (par blocs de *N* définis dans ITU I.610) pouvant être en erreur ;
- CMR (*Cell Misinsertion Rate*). Taux maximal de cellules pouvant être mal insérées (c'est-à-dire dont les erreurs portant sur l'en-tête n'ont pas été détectées).

La détection des erreurs et la surveillance des performances sont réalisées par la couche OAM (*Operations, Administration and Maintenance* – ITU I.610).

L'adressage

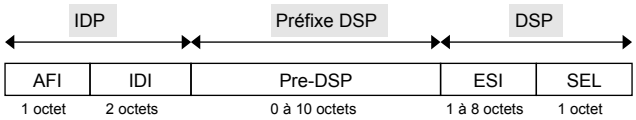
L'adressage NSAP (ISO 8348, ITU X.213, RFC 1629)

NSAP (*Network Service Access Point*) définit un **adressage global de niveau 3** permettant d'identifier les utilisateurs d'un réseau ATM, Frame Relay ou RNIS. Dans le cas du RNIS, il s'agit tout simplement du numéro de téléphone.

	Adresse locale	Adresse globale
Protocole	Niveau 2	Niveau 3
Frame Relay	DLCI	E.164 ou X.121
ATM	VPI/VCI	NSAP (encapsulation DCC, ICD ou E.164) ou E.164
RNIS	SAPI/TEI	E.164

L'apparent paradoxe d'un adressage de niveau 3 utilisé par des protocoles de niveau 2 s'explique par le fait que ces protocoles agissent en **mode connecté**. L'adressage global est utilisé par les **protocoles de signalisation** (situés au niveau 3) pour établir les communications de niveau 2 (les circuits commutés RNIS ou virtuels ATM) ; la commutation de circuits RNIS ou de cellules ATM n'utilise ensuite que des adresses locales de niveau 2.

Figure 11-17.
Format générique d'une adresse NSAP.



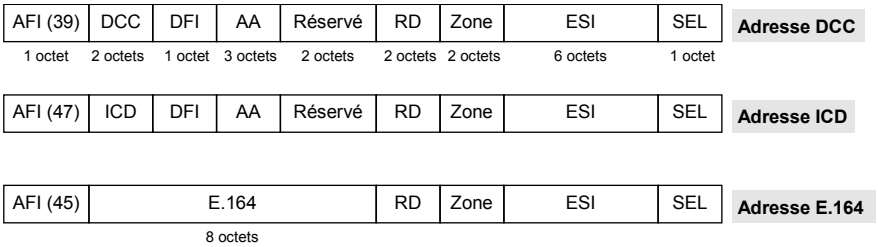
IDP = *Initial Domain Part* : format de l'adresse (37=X.121, 45=E.164, etc.)
AFI = *Authority and Format Identifier*
IDI = *Initial Domain Identifier*
Préfixe ou HO-DSP pour *High Order DSP* : partie de l'adresse structurée en fonction de l'IDI
DSP = *Domain Specific Part* : partie de l'adresse affectée localement
ESI = *End System Identifier* : identifiant du nœud du réseau
SEL = *Selector* : multiplexage si le nœud comporte plusieurs interfaces ou protocoles

L'adressage ATM

ATM peut utiliser trois types d'adresses encapsulées au format NSAP :

- DCC pour les réseaux publics et privés ;
- ICD pour les réseaux privés ;
- E.164 pour les réseaux publics, soit de manière native, soit encapsulée dans une adresse NSAP.

Figure 11-18.
Format des adresses ATM.



AFI = *Authority and Format Identifier* : type d'adresse (39 = DCC, 47 ICD, 45 E.164, etc.)
DCC = *Data Country Code* : indique le code associé au pays (ISO 3166)
DFI = *DSP - Domain Specific Part Format - Identifier* : format de la suite de l'adresse
AA = *Administration Authority* : autorité administrative en charge du domaine d'adressage
RD = *Routing Domain* : numéro de domaine de routage
Zone = identifiant de zone
ESI = *End System Identifier* : adresse MAC
SEL = *Selector* : généralement, le numéro de l'interface ATM
ICD = *International Code Designator* : code attribué par le *British Standards Institute*
E.164 = numéro de téléphone RNIS sur 8 octets (15 chiffres + ½ octet de bourrage)

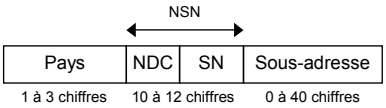
Dans les réseaux publics, ATM utilise une adresse E.164 native.

L'adressage Frame Relay

Les réseaux Frame Relay peuvent utiliser deux types d'adresses :

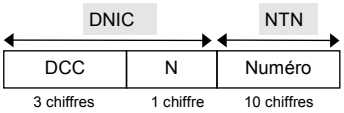
- **E.164**, identique aux numéros de téléphone RNIS ;
- **X.121**, identique aux réseaux X.25.

Figure 11-19. *Format d'une adresse E.164.*



Pays 33 = France, 44 = UK, 1 = USA, etc.
NSN = *National Significant Number*
NDC = N*ational Destination Code* : numéro du réseau ou numéro de région.
SN = *Subscriber Number* : numéro de téléphone de l'abonné
Longueur maximale du numéro E.164 = 15 chiffres

Figure 11-20. *Format d'une adresse X.121.*



DNIC = *Data Network Identification Code*
NTN = *Network Terminal Number*
DCC D = code d'échappement, CC = code pays
N = numéro de réseau dans le pays

Interopérabilité entre Frame Relay et ATM

Les réseaux Frame Relay et ATM peuvent être utilisés conjointement de deux manières : soit par encapsulation de Frame Relay dans les cellules ATM, soit par conversion de protocole. L'opérateur peut de cette manière offrir un réseau avec une interface Frame Relay tout en utilisant ATM au sein de son réseau fédérateur.

La seconde méthode consiste à convertir une trame Frame Relay en un PDU AAL-5.

Frame Relay	ATM
Encapsulation NLPID ou SNAP (RFC 2427)	Encapsulation SNAP/LLC sur AAL-5 (RFC 1483)
Bit DE	Bit CLP
Bit FECN	Bit EFCI
LMI	ILMI
Inverse ARP (PVC)	Classical IP : Inverse ARP (PVC) et serveur ATM (SVC)

12

Commutation et routage LAN / MAN / 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 mettre en place un réseau fédérateur puis un réseau de campus ;
- à configurer des VLAN ;
- à établir le lien entre commutateurs LAN et routeurs WAN ;
- à configurer les protocoles de routage OSPF et BGP ;
- le fonctionnement de MPLS.

Mettre en place un réseau fédérateur

Les données du problème

Le réseau de 800 postes dont nous avons décrit l'installation au chapitre 3 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 2.

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 interfacer les réseaux Ethernet avec le 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 Mbit/s 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.

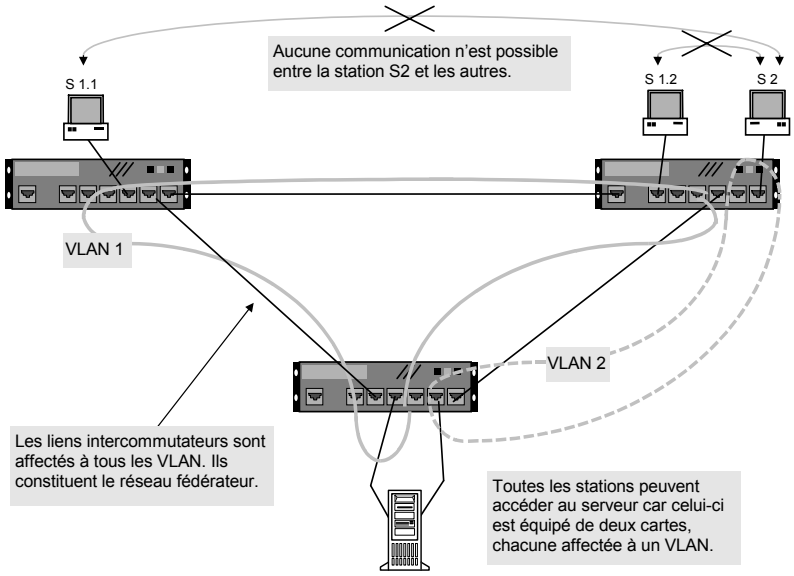
Mais, de nos jours, la traditionnelle répartition 80/20 (80 % du trafic local sur le réseau et 20 % vers d'autres réseaux) n'est plus valable. Par exemple, la constitution de groupes de travail pluridisciplinaires amène des personnes dispersées au sein de l'immeuble à établir des liens de communication 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 12-1.
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 3 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 broadcasts (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 3. 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.

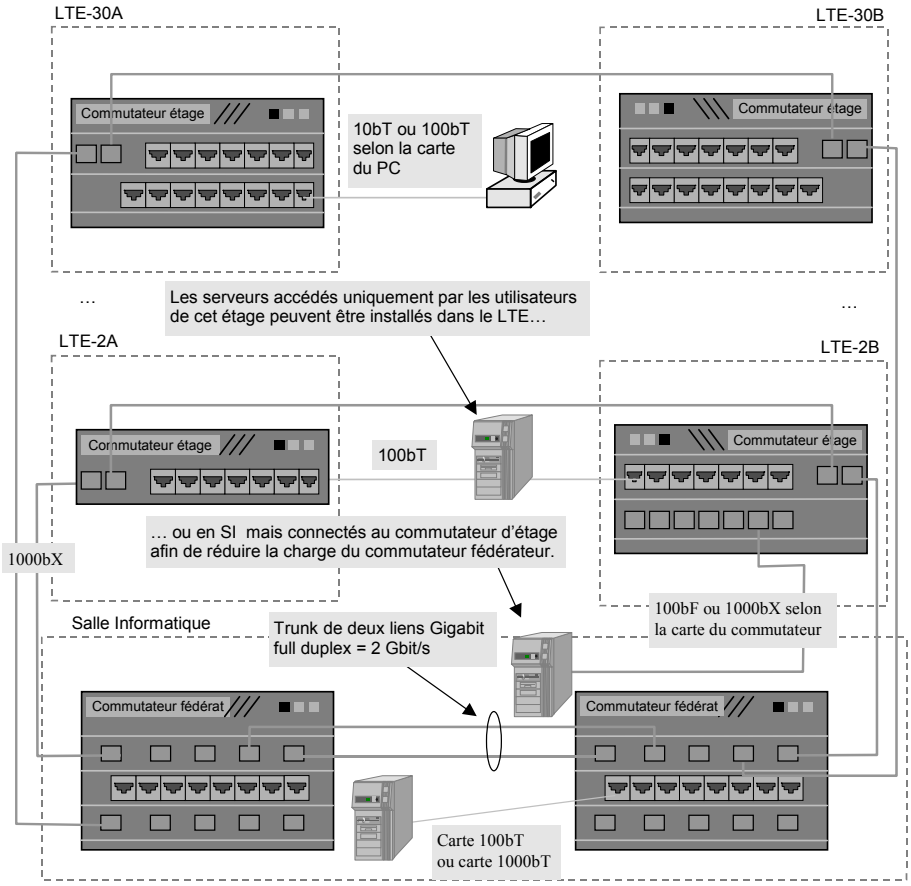
Pour le reste, nous appliquerons les recettes indiquées au chapitre 3.

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 informatiques. 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 12-2.
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 3.

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 ac-

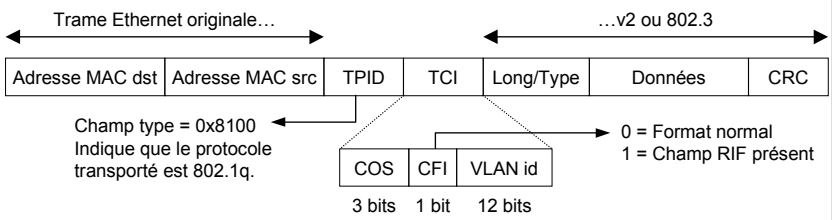
tivé que sur les ports qui raccordent des commutateurs entre eux, qu'on appellera des ports **trunk** (ports de liaison) :

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

Active le protocole 802.1q sur ces ports

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, généralement véhiculée qu'entre les commutateurs, permet d'étendre un VLAN 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 à ce que ces derniers 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**.

La norme spécifie également l'utilisation du protocole **GVRP** (*GARP VLAN registration Protocol*), qui permet de configurer les VLAN sur un seul commutateur, puis de les diffuser aux autres commutateurs participant au spanning tree. L'attribut échangé par les GID via le protocole GIP (cf. chapitre 3) est l'identifiant de VLAN sur 12 bits.

Le protocole **MSTP** (*Multiple Spanning Tree Protocol*), issu du groupe de travail IEEE **802.1s**, permet de créer plusieurs spanning tree (cf. chapitre 3). La répartition des VLAN sur plusieurs spanning tree présente l'avantage de pouvoir activer des liens qui sont désactivés par d'autres spanning tree. Il est donc possible de configurer ce protocole indépendamment pour chaque VLAN :

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

Le paramètre 100 correspond au numéro de VLAN

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.

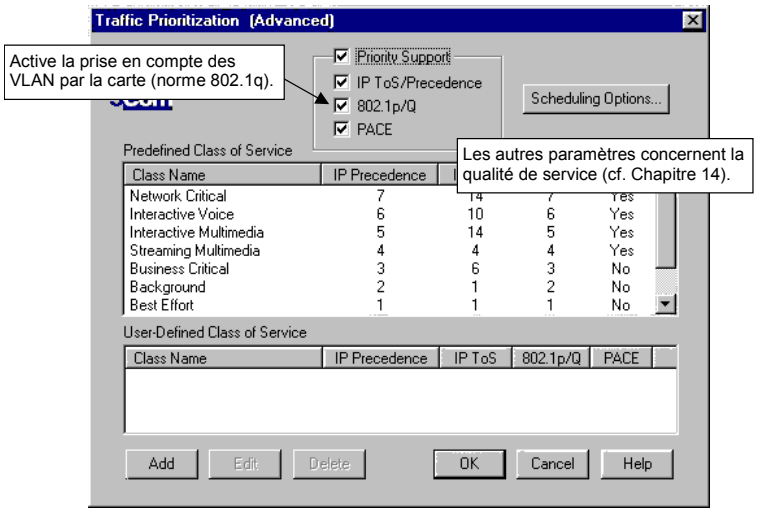
Dans notre architecture, il est également prévu d'agréger deux liens gigabit entre les deux commutateurs fédérateurs :

```
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.

Mettre en place un réseau de campus

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. L'enjeu 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 ne soit associée à ce type de réseau. Il s'agit simplement d'une dénomination conceptuelle désignant les réseaux métropolitains à haut débit.

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 autorisation pour en poser entre nos bâtiments.

Pour ceux qui en doutaient encore, le Gigabit Ethernet convient parfaitement à ce type de besoins (voir les différents réseaux Ethernet au chapitre 3). Certains opérateurs proposent même ce service sur quelques centaines de kilomètres. Tout dépend de la fibre optique utilisée.

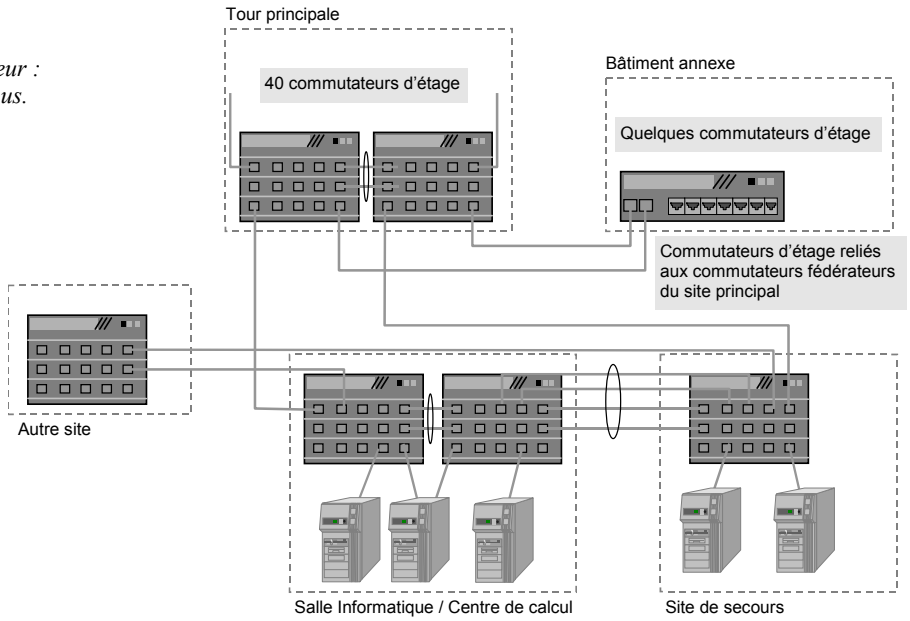
	Support de transmission	Distance maximale	Longueur d'onde
1000bSX	Fibre multimode 62,5µ (qualité standard)	220 m	850 nm
	Fibre multimode 62,5µ 160/500 MHz-km	275 m	850 nm
	Fibre multimode 62,5µ 200/500 MHz-km	500 m	850 nm
	Fibre multimode 50µ 400/400 MH-km	550 m	850 et 1300 nm
1000bLX	Fibre multimode 62,5µ (∀ la bande passante)	550 m	1 300 nm
	Fibre multimode 50µ	550 m	1 300 nm
	Fibre monomode 9µ	5 km	1 550 nm
	Fibre monomode 9µ à faible dispersion	70 à 100 km	1 550 nm

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 12-3.
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.

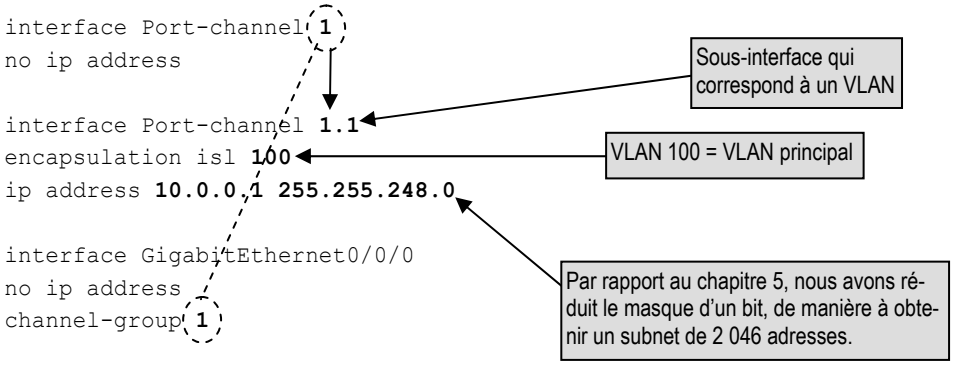
L'extension du réseau fédérateur se fait donc de manière très simple, sans remettre en cause les choix technologiques et l'architecture.

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 5 :



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 9. 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.

La 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 5). 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) :

```
#Commutateur 1
interface Port-channel 1.1
encapsulation isl 100
ip address 10.0.0.2 255.255.248.0
standby 1 priority 110
standby 1 preempt
standby 1 ip 10.0.0.1

#Commutateur 2
interface Port-channel 1.1
encapsulation isl 100
ip address 10.0.0.3 255.255.248.0
standby 1 priority 100
standby 1 preempt
standby 1 ip 10.0.0.1
```

Adresse de la carte de routage

Adresse HSRP, virtuelle, seule connue des PC et des serveurs

Groupe HSRP

La priorité la plus basse indique que la carte est en attente d'une éventuelle défaillance de l'autre.

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 la 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)

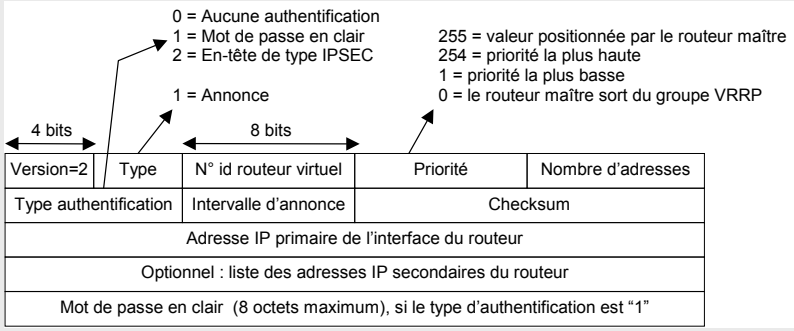
Le protocole **VRRP** (*Virtual Router Redundancy Protocol*) reprend les principes du protocole **HSRP** (*Hot Standby Router Protocol* – RFC 2281) spécifié par Cisco. Les formats des paquets sont en revanche différents, ce qui rend ces deux protocoles incompatibles.

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.

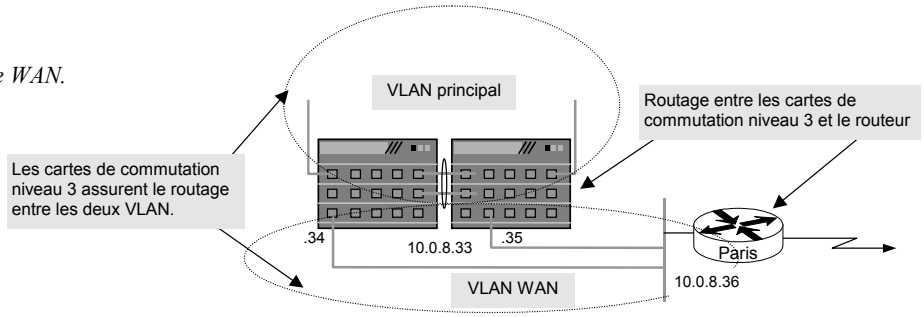
Les mécanismes classiques d'icmp-redirect et de proxy ARP sont toujours opérants.

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 12-4.
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
```

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

```
# 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
```

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 - OSPF

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.

Configurer le routage OSPF

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 de routage 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*), **OSPF** (*Open Shortest Path First*) et **EIGRP** (*Enhanced Interior Gateway Routing Protocol*). Les protocoles spécialisés dans le routage inter AS sont de type **EGP** (*Exterior Gateway Protocol*) dont 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.

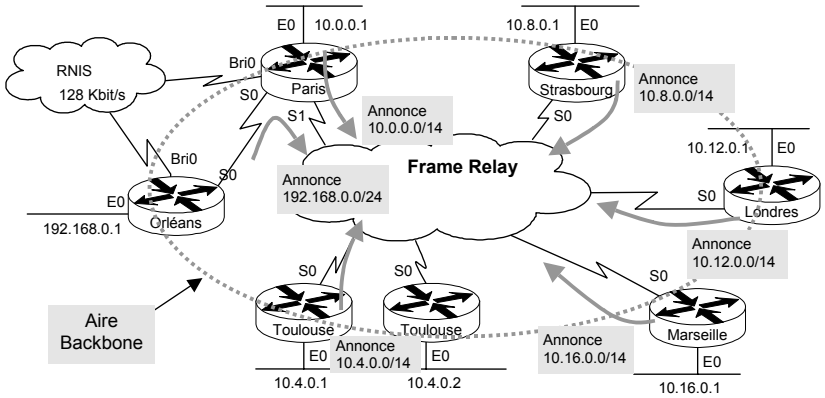
Étant donné que notre plan d'adressage défini au chapitre 5 prévoit l'affectation d'un subnet complet à l'ensemble des liaisons WAN, une seule commande sur chaque routeur est nécessaire pour affecter l'aire 0 :

```
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 12-5.
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
```

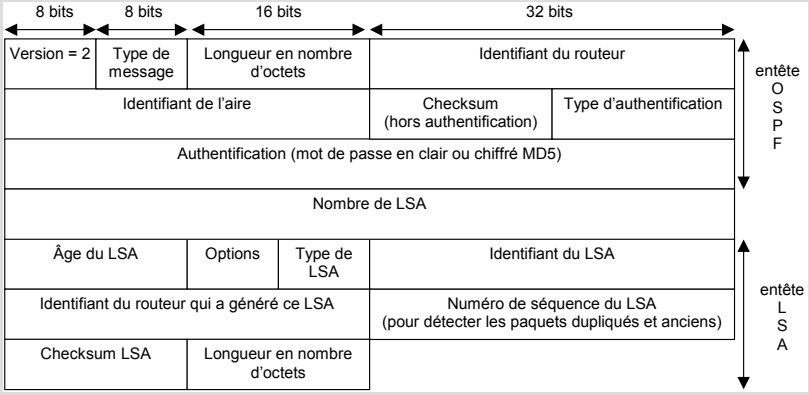
Redondance en cas de panne

En reprenant le réseau Frame-Relay que nous avons construit au chapitre 11, 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 supportant le 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 multicasts 224.0.0.5.

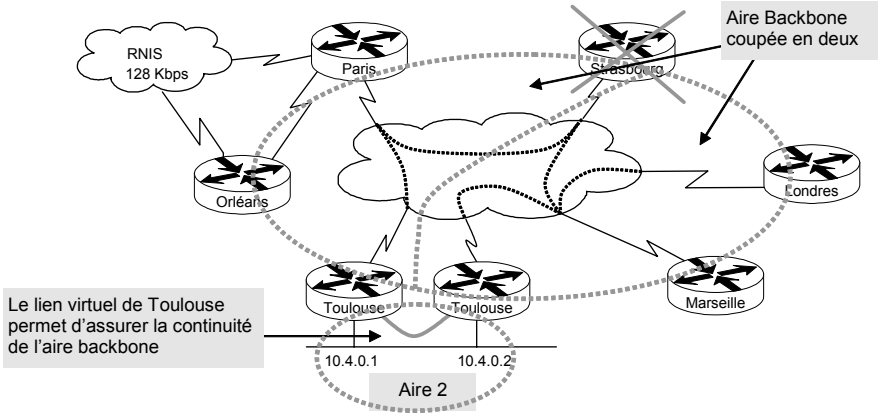
Dans la *backbone area*, les routeurs de bordure s'échangent les bases de données des aires auxquelles ils sont rattachés. Ils calculent les meilleures routes qui sont ensuite diffusées aux routeurs intra-aire. Les routeurs intra-aire calculent la meilleure route pour sortir de l'aire (*via* un routeur 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 12-6.
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 à ce 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 sites, 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  0.0.0.255
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  0.0.0.255
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
```

Envoie 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

La 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	

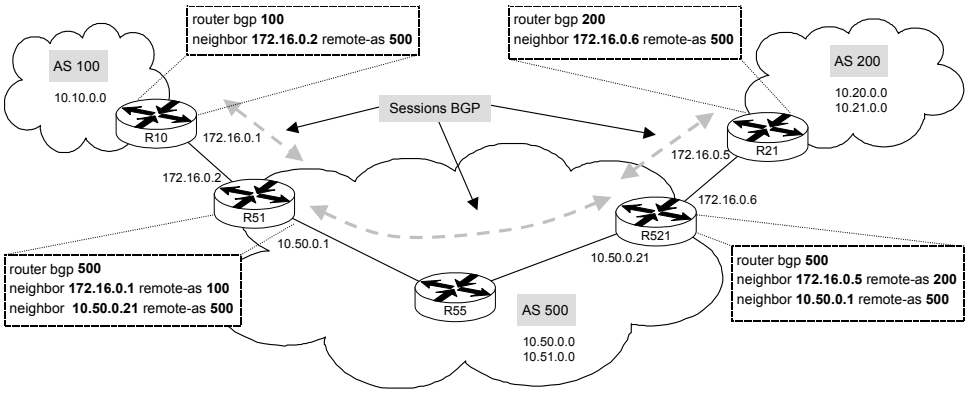
Le temps CPU pour calculer l'arbre du plus court chemin (algorithme de Dijsktra) est de l'ordre de n*log(n) pour *N* routes et 200 routeurs, soit environ 15 millisecondes pour un processeur de 10 Mips. En découpant un système autonome en aires, la charge CPU est réduite, car il y a moins de routeurs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

Le routage entre systèmes autonomes - BGP

Avec OSPF, votre réseau fonctionne de manière indépendante avec son système autonome OSPF, tout comme le fait une autre entreprise ou un ISP avec le leur.

Dans certaines entreprises, on peut envisager de découper son réseau en domaines administratifs distincts, chacun d'eux étant sous la responsabilité d'une équipe d'exploitation indépendante. Ce type de besoin s'exprime également avec l'Internet où chaque ISP ne gère que sa portion de réseau. L'autre intérêt de ce découpage est qu'il limite la taille des tables de routage OSPF.

Arrive un moment où il faut interconnecter ces systèmes autonomes, que ce soit en entreprise ou sur Internet. L'échange des informations de routage entre AS est alors réalisé à l'aide du protocole BGP (*Border Gateway Protocol*).



Un routeur BGP peut injecter automatiquement toutes les routes apprises de BGP dans OSPF, afin qu'elles soient diffusées dans le reste de l'AS. En retour, il ne faut pas réinjecter les routes OSPF dans BGP, mais seulement y injecter les routes de l'AS, ce qui implique une configuration manuelle. Le routeur R51 doit donc être configuré comme suit :

```
router bgp 500
neighbor 172.16.0.1 remote-as 100
neighbor 10.50.0.21 remote-as 500
network 10.50.0.0 mask 255.255.0.0
network 10.51.0.0 mask 255.255.0.0

router ospf 1
network 10.50.0.0 0.0.255.255 area 0
network 10.51.0.0 0.0.255.255 area 0
redisbribute bgp 500
```

Numéro de l'AS

R51 apprend les routes annoncées pas ses voisins…

…et leur annonce les routes de son AS.

Toutes les routes apprises de BGP (donc les routes des autres AS) sont injectées dans OSPF afin qu'elles soient diffusées dans l'AS 500 .

Il est noter que pour BGP la commande « *network* » indique la liste des réseaux à annoncer, tandis que pour OSPF, elle spécifie les interfaces prises en compte par le processus.

Les configurations des autres routeurs sont similaires, par exemple, celle de R21 :

```
router bgp 200
neighbor 172.16.0.6 remote-as 500
network 10.20.0.0 mask 255.255.0.0

router eigrp 65020
network 10.20.0.0
network 10.21.0.0
redisbribute bgp 500
```

Par décision de l'administrateur, seul le réseau 10.20 doit être vu, et donc accessible des autres AS.

Ce système autonome fonctionne avec EIGRP (*Enhanced Interior Gateway Routing Protocol*), protocole propriétaire Cisco.

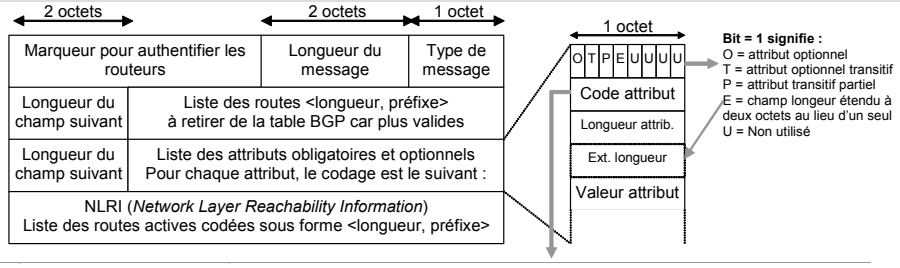
LE POINT SUR BGP (RFC 1771 À 1774)

L'objectif de BGP (*Border Gateway Protocol*) est de diffuser les routes entre des systèmes autonomes tels que ceux d'OSPF. Le principe consiste à injecter manuellement les routes d'un AS dans BGP ; ce dernier annonce ses routes aux routeurs BGP des autres AS qui les réinjectent ensuite dans le routage interne à leur AS respectif.

À leur initialisation, les routeurs BGP ouvrent des sessions TCP (port 179) avec leurs partenaires (*peers*) et échangent l'intégralité de leur table de routage, puis uniquement les mises à jour par la suite.

Quatre types de messages sont ainsi échangés :

- **Open** contient des informations relatives au routeur BGP qui envoie ce message : l'AS dans lequel il est, le délai par défaut du keepalive, son identifiant (généralement son adresse IP), ainsi que des paramètres optionnels relatifs à l'authentification.
- **Keepalive** est envoyé régulièrement aux partenaires pour indiquer que le routeur est toujours actif.
- **Notification** est envoyé lorsqu'une erreur a été détectée (valeur non conforme dans le message, réponse non reçue dans les temps, erreur de format du message, erreur de procédure, etc.).
- **Update** contient d'éventuelles routes précédemment annoncées et qui ne sont plus valides, ainsi que les routes actives associées à l'adresse du routeur de bordure et la liste des AS à traverser pour y aller.



Code	Nom	Description
1	ORIGIN	Origine de l'information (IGP, EGP ou incomplète)
2	AS_PATH	Liste des AS traversés par ce message. Le numéro d'AS y est ajouté en sortie.
3	NEXT_HOP	Adresse IP du saut suivant pour aller vers les réseaux listés dans le champ NLRI
4	MULTI_EXIT_DISC (optionnel)	Métrique permettant de choisir entre plusieurs points de sortie vers un autre AS
5	LOCAL_PREF	Degré de préférence pour une route de sortie vers un autre AS
6	ATOMIC_AGGREGATE	Indique qu'une route moins spécifique a été sélectionnée parmi celles indiquées dans la liste
7	AGGREGATOR (optionnel)	Dernier n° d'AS et adresse IP du dernier routeur BGP qui a formé la route agrégée.
8	COMMUNITY (optionnel)	Indique comment une route a été apprise et permet de décider à qui la diffuser (cf. RFC 1997)
…	Actuellement, 18 attributs	cf. www.iana.org/assignments/bgp-parameters

Les messages " *Update* " sont propagés de routeur en routeur : à chaque sortie d'AS, le routeur BGP ajoute son numéro d'AS dans la liste AS_PATH, permettant ainsi de reconstituer le chemin d'AS. Chaque routeur BGP communique ainsi à son partenaire la liste des réseaux atteignables (NLRI) via tel routeur indiqué dans NEXT_HOP.

S'il existe plusieurs routes entre deux AS, les routeurs BGP de chaque AS calculent la meilleure route vers l'autre AS à l'aide des attributs MULTI_EXIT_DISC et LOCAL_PREF.

Les informations de routage (NLRI) sont échangées sous le format **CIDR** (*Classless Inter-Domain Routing*) qui consiste à ne coder que le préfixe des adresses (la partie réseau des adresses IP) et à agréger les routes contiguës.

Les routeurs BGP partenaires (*internal peers* si dans le même AS ou *external peers* si dans deux AS différents) doivent avoir au moins une de leur interface dans le même subnet IP.

Le protocole BGP requiert un maillage complet de toutes les paires de routeurs au sein d'un AS (entre *Internal peers*). Afin de simplifier cette configuration pour les grands réseaux, la solution consiste à regrouper les routeurs dans des clusters BGP et à désigner un leader au sein de chaque cluster, appelé *Route Reflector*. Les échanges avec les routeurs d'autres AS (entre *External peers*) sont alors réalisés via les *Routes Reflectors* qui redistribuent les routes apprises aux routeurs BGP de leur AS respectif. Utilisé au sein d'un AS, le même protocole est appelé I-BGP (*Internal BGP*) tandis qu'utilisé entre AS, il est appelé E-BGP (*External BGP*). La différence réside dans les règles de diffusion des routes apprises : un préfixe d'un voisin I-BGP ne peut pas être annoncé à un autre voisin I-BGP, ce qui est en revanche possible avec des voisins E-BGP.

La commutation sur le WAN - MPLS

Une alternative au routage classique est la **commutation MPLS** (*Multiprotocol Label Switching*). Dans le premier cas, chaque routeur analyse l'en-tête des paquets IP et décide du routage en fonction d'une table représentant la topologie du réseau. Dans le second cas, l'analyse de l'entête est réalisée par le premier routeur qui ajoute un label, ce qui permet aux autres routeurs de commuter le paquet d'après ce label sans analyser le paquet. L'avantage immédiat est de réduire la taille des tables de routage et d'améliorer les performances.

L'idée originale était effectivement de réduire le temps de traitement des paquets en combinant le routage OSPF à la périphérie du réseau et la commutation ATM au cœur du réseau. Plusieurs constructeurs ont ainsi proposé des solutions : Cisco avec *Tag Switching*, IBM avec ARIS (*Aggregate Route-Based IP Switching*), Ipsilon et Cascade. Ces approches propriétaires ont conduit l'IETF à se pencher sur le problème en 1997 puis à proposer le standard MPLS en janvier 2001, sorte de synthèse des précédentes solutions. Entre temps, les soucis de performances ont perdu de leur actualité grâce à la montée en puissance des routeurs, mais la principale technique mise en œuvre dans MPLS, c'est-à-dire la définition de chemins IP commutés, s'est avérée pouvoir remplir d'autres fonctions.

En effet, le concept de **circuit virtuel au niveau IP**, appelé **LSP** (*Label Switched Path*), permet de créer des chemins en fonction de différents critères et de leur appliquer des traitements différenciés. Le critère, qui permet d'affecter un paquet IP à un LSP, peut ainsi correspondre :

- à une adresse IP et/ou à un port TCP ou UDP ;
- à une classe de service ;
- à des choix d'ingénierie (dénommés communément *traffic engineering*), tels la répartition de trafic entre deux liaisons parallèles (*load balancing*), une route de secours en cas de panne du lien principal, etc.

Ces possibilités ouvrent la voie à trois applications :

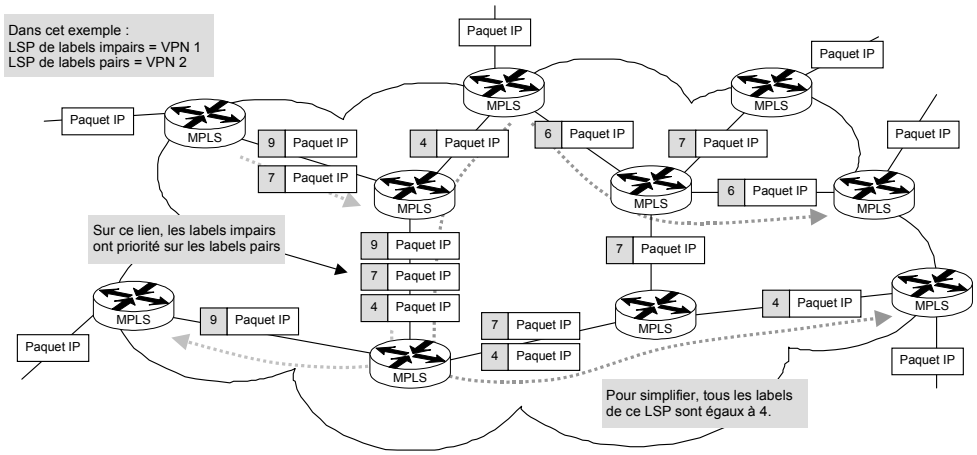
- la création de réseaux logiques séparés, appelés VPN-MPLS ;
- le traitement différencié des flux de paquets en fonction de la classe de service qui leur est affectée ;

- l'orientation du trafic en fonction de l'état du réseau.

L'intérêt de reporter un maximum de fonctions au niveau 3 est qu'IP est présent partout, quel que soit le protocole de transport. On peut ainsi gérer de bout en bout et uniformément les VPN, la qualité de service et le trafic, que ce soit dans les réseaux privés en entreprise ou sur le réseau public Internet. En effet, MPLS est multiprotocole, car il repose sur n'importe quel protocole de niveau 2 (ATM, Frame Relay, PPP et Ethernet à ce jour) et peut être utilisé par n'importe quel protocole de niveau 3 (IPv4 et IPv6 à ce jour).

Fonctionnement de MPLS

Un LSP (*Label Switched Path*) est matérialisé par une succession de routeurs qui assignent les mêmes labels aux paquets IP ayant les mêmes caractéristiques (même adresse destination, même classe de service diffserv, etc.). Comme dans un réseau IP, les paquets sont mélangés, mais dans un domaine MPLS, les paquets d'un même LSP ne peuvent pas « voir » les autres paquets IP et sont traités différemment des autres LSP.

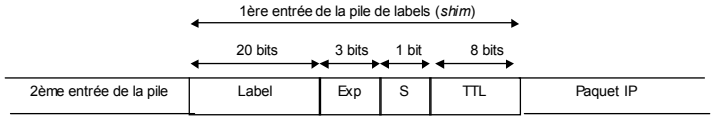


Le label peut correspondre au VPI/VCI d'un CV ATM (RFC 3035) ou au DLCI d'un CV Frame Relay (RFC 3034). Dans ce cas de figure, les commutateurs ATM et Frame Relay se comportent comme des routeurs MPLS.

Comparaison	ATM et de Frame Relay	MPLS
Circuit virtuel	CV	LSP
Identification du CV	VPI/VCI et DLCI	Label
Création des CV	Statique	Dynamique via LDP
Critères de création d'un circuit virtuel	Aucun	Adresses, ports tcp/udp, QoS…
Hiérarchie de CV	Non	Oui
VPN	Niveau 2	Niveau 3

LE POINT SUR MPLS (RFC 3031 ET 3032)

Le protocole MPLS (*Multiprotocol Label Switching*) créé des circuits virtuels au niveau IP, appelés **LSP** (*Label Switched Path*), sur la base de critères, appelés **FEC** (*Forwarding Equivalent Class*), tels que les adresses destinations, les classes de services, le type d'application, etc. Tous les paquets correspondant à un même FEC emprunteront ainsi le même LSP et seront identifiés par les mêmes **labels**.



Plusieurs labels peuvent être empilés en tête du paquet IP.

Label Numéro de label.
0 = *Ipv4 explicit null label* (valable pour le dernier label de la pile). Indique que le label doit être retiré de la pile et que le routage doit être réalisé sur la base de l'entête IP.
1= *Router Alert Label* (non valable pour le dernier label de la pile).
2 = *Ipv6 explicit null label*. Même signification que la valeur 0, mais appliquée à Ipv6.
3 = *Implicit null label*. Utilisé par le protocole de distribution des labels.
4-15 : valeurs réservées à un usage futur. Le reste des valeurs est libre d'utilisation.

Exp (*Experimental*).

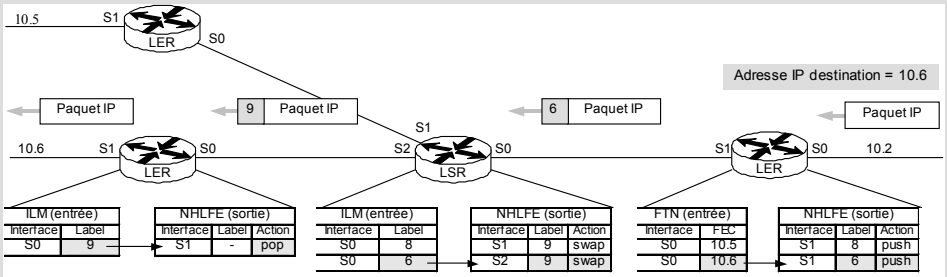
S (*Stack*). 1 = dernier label de la pile.

TTL (*Time To Live*). Est décrémenté de 1 par les LSR afin d'éviter les boucles infinies. En entrée d'un LSP, le LER copie le TTL du paquet IP dans ce champ et inversement en sortie.

À la frontière d'un **domaine MPLS**, le paquet IP est pris en charge par un routeur de bordure (**LER** – *Label Edge Router*) à qui il ajoute (en entrée) ou retire (en sortie) un label. Les paquets sont ensuite commutés par des routeurs MPLS (**LSR** – *Label Switch Router*) qui ont la possibilité de modifier les labels.

Alors qu'un routeur classique traitera chaque paquet individuellement sur la base de sa table de routage et de l'entête IP, un LSR se contentera de commuter les paquets uniquement sur la base des labels, sans s'occuper de quoi que ce soit d'autre.

Les labels sont affectés indépendamment par chaque routeur MPLS. Ils n'ont donc qu'une portée locale, entre deux LSR adjacents.



NHLFE (*Next Hop Label Forwarding Entry*) contient les informations permettant de router un paquet MPLS tels que le saut suivant, l'opération à réaliser sur le label (*swap*, *pop*, *push*, *aggregate*, *untag*), la manière de coder le label, etc. Plusieurs NHLFE peuvent être associés à un label entrant, comme dans le cas du multicast, par exemple.

ILM (*Incoming Label Map*) est une table de correspondance entre un label et un ou plusieurs NHLFE

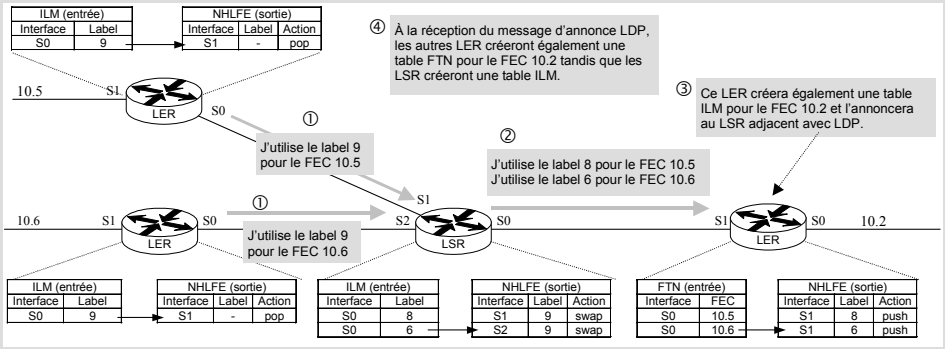
FTN (*FEC-to-NHLFE*) est une table de correspondance entre un FEC et un ou plusieurs NHLFE

• • •

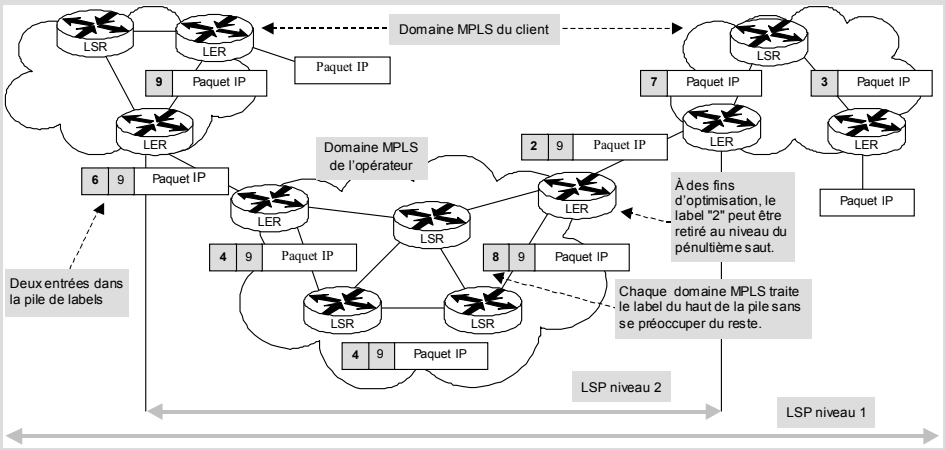
LE POINT SUR MPLS (SUITE)

Chaque routeur MPLS indique au routeur situé en amont le label associé à tel FEC grâce au protocole de distribution des labels (**LDP** – *Label Distribution Protocol* – RFC 3036 à 3038 et 3215). À partir des tables de routage et des informations échangées entre LSR, LDP constitue une table de correspondance appelée **LIB** (*Label Information Base*) qui associe un FEC (dans le cas de LDP, un numéro de réseau ou une adresse de machine) à un ou plusieurs couples " Identifiant d'espace de label / Label ".

MPLS peut cependant utiliser d'autres protocoles pour distribuer les labels. Ainsi, des extensions de BGP (RFC 3107) et de RSVP (RFC 3209) ont été spécifiées à cet effet.



Il est possible de créer des **hiérarchies de LSP**, par exemple pour agréger des LSP dans un seul LSP sur une partie du chemin (création d'un tunnel) à l'instar d'ATM qui agrège plusieurs VC dans un VP (cf. chapitre 11), mais sur un seul niveau alors que MPLS n'est pas limité. Dans l'entête MPLS, les labels sont ainsi empilés par les LER de chaque domaine sur le mode LIFO (*Last In, First Out*). Il est alors possible d'imbriquer plusieurs domaines MPLS, chaque domaine ne traitant que le label situé au dessus de la pile.



À la différence d'un domaine MPLS, un tunnel LSP ne concerne que deux routeurs qu'ils soient d'ailleurs adjacents (tunnel routé saut par saut) ou non (tunnel routé explicitement).

• • •

LE POINT SUR MPLS (FIN)

Le routage MPLS peut être traité de deux manières :

- **saut par saut**, chaque LSR choisissant indépendamment le prochain saut pour chaque FEC sur la base des informations fournies par un protocole de routage tel qu'OSPF.
- **explicitement**, le LER d'entrée ou de sortie indiquant tout (routage explicite strict) ou partie (routage explicite lâche) des LSR à emprunter pour tel FEC.

Les protocoles de distribution de label permettant de créer des LSP saut par saut sont LDP et une extension de BGP (RFC 3107). Les protocoles de distribution de label permettant de créer des LSP explicites sont actuellement **RSVP-TE** (*RSVP Traffic Engineering*, RFC 3209) ou **CR-LDP** (*Constraint-Based LDP*, RFC 3212 à 3214).

Les LSP peuvent être créés avec LDP selon le **mode de contrôle indépendant** ou **ordonné**. Dans le premier cas, la diffusion des labels est réalisée par chaque LSR dès que possible, on pourrait dire de manière désordonnée, tandis que dans le second cas, les LSR ne diffusent les labels d'un FEC que s'il a reçu le label correspondant à ce FEC en provenance d'un routeur amont.

Plusieurs FEC peuvent être **agrégés** en un seul FEC, seul ce dernier étant alors lié à un label. Cela peut, par exemple, être le cas de FEC partageant les mêmes préfixes d'adresse. De même, plusieurs labels entrant peuvent être **fusionnés** en un seul label, ce qui revient à fusionner plusieurs flux auparavant distincts. Les flux fusionnés ne peuvent alors plus être séparés.

Les VPN-MPLS sont fonctionnellement équivalents aux VLAN Ethernet décrits au début de ce chapitre. Les premiers agissent au niveau 2, tandis que les seconds agissent au niveau 3.

L'exemple suivant est proposé à titre indicatif, car les implémentations actuelles sont encore propriétaires, notamment au niveau du protocole de diffusion de labels et de l'utilisation du champ « Exp ». Le lecteur pourra se tenir informé des évolutions de MPLS sur le site www.ietf.org/html.charters/mpls-charter.html.

```
int e0/1
tag-switching advertise-tags
```

Active MPLS

```
ip vrf  VPN_client
rd 100:1
route-target both 100:1
```

Crée un VPN

```
router bgp 500
neighbor 10.5.0.1 remote-as 400
neighbor 10.5.0.1 activate
```

Crée un domaine BGP pour ce VPN

QUATRIÈME PARTIE

La gestion avancée des flux IP

13

Les flux multimédias

L'objectif de ce chapitre est de présenter les caractéristiques des flux multimédias (essentiellement la voix et la vidéo), afin de montrer comment leurs particularités influent sur un réseau de paquets tel que TCP/IP.

Il n'a pas pour objet de faire de vous des experts en audiovisuel, mais plutôt de vous donner les informations essentielles vous permettant de savoir de quoi on parle, ce que l'on manipule et de comprendre l'impact de ces flux audio et vidéo sur votre réseau.

Vous découvrirez ainsi :

- comment sont transportés le son et l'image sous forme numérique ;
- ce qu'est un codec audio et vidéo ;
- les problèmes posés par les délais de transit et la gigue ;
- quels débits ces types de flux engendrent sur votre réseau.

Les caractéristiques des flux multimédias

Initialement, un réseau IP, tel que l'Internet, était conçu pour véhiculer des données entre deux machines : transfert de fichiers, connexion Web, messagerie, etc. Depuis, la voix et l'image ont fait leur apparition et ont étendu le champ d'utilisations du réseau : téléphonie, diffusion de films à la demande et conférences à plusieurs. Dans ce dernier domaine, on distingue l'audioconférence (voix uniquement), la visioconférence (voix + vidéo) et, d'une manière générale, la téléconférence (voix + vidéo + données).

Ces flux multimédias induisent un certain nombre de contraintes nouvelles :

- Les signaux audio et vidéo doivent être numérisés, ce qui veut dire convertir les signaux analogiques en bits numériques.
- La voix nécessite une bonne synchronisation entre l'émetteur et le récepteur.
- La vidéo engendre une augmentation du volume des données transférées.
- La téléconférence nécessite de diffuser un flux entre un émetteur et plusieurs récepteurs qui peuvent devenir, à leur tour, émetteurs.

La numérisation des signaux nécessite d'échantillonner la voix, de la quantifier, de la coder et, de plus en plus souvent, de la compresser :

- L'**échantillonnage** consiste à prélever des échantillons du signal à intervalles réguliers, à l'instar du cinéma qui utilise 24 images par seconde pour traduire le mouvement. Ainsi, plus la fréquence d'échantillonnage sera élevée, plus le nombre de bits utilisés pour représenter un échantillon sera grand et plus la forme numérique sera fidèle au signal original. En contrepartie, le débit réseau généré sera plus important. La plupart du temps, l'échantillon porte sur 8 bits, alors que, pour la haute fidélité (requise, par exemple, pour un Compact Disc), l'échantillon porte sur 16 bits.
- La **quantification** fait correspondre une valeur à l'amplitude d'un échantillon par rapport à des valeurs-étalons appelées niveaux de quantification. L'amplitude des échantillons prélevés peut varier de façon illimitée, mais doit pouvoir être représentée par un nombre fini de valeurs binaires. La valeur résultant, généralement codée sur un octet, sera égale au niveau de quantification le plus proche de celui mesuré.
- Le **codage** consiste à transmettre un flux d'informations binaires correspondant à l'échantillon représenté par un octet.
- De plus en plus souvent, le codage est associé à un algorithme de **compression**.

Par exemple, la voix génère des signaux à une fréquence oscillant entre 300 Hz et 3 300 Hz, valeur arrondie à 4 000 Hz par les équipements numériques. Un chercheur, appelé Shannon, a montré que la fréquence d'échantillonnage devait être égale au double de la fréquence du signal à numériser. Un signal analogique de 4 000 Hz nécessite donc 8 000 échantillons par seconde, qui sont représentés sur 8 bits, soit un débit de 64 Kbit/s pour coder la voix. Cette unité, appelée **DS0** (*Digital Signaling 0*), a longtemps été la référence et continue encore de l'être pour les liaisons d'accès E1/T1 proposées par les opérateurs, aussi bien pour la voix que pour les données.

LES CODEC AUDIO

Il existe une multitude de codec, que l'on peut ranger en trois catégories : temporel, source et hybride. Cette différenciation repose sur la manière dont est réalisée la quantification.

La quantification différentielle consiste à mesurer la différence entre deux échantillons. L'échelle utilisée peut être fixe, on parle alors de **quantification différentielle scalaire** (de type **linéaire** ou **logarithmique**), ou varier en fonction du signal précédent, qui est appelée **quantification différentielle adaptative**.

Au lieu de travailler échantillon par échantillon, la **quantification vectorielle** travaille sur un groupe d'éléments qui sont codés à l'aide d'un dictionnaire (*codebook*). Ce type de quantification peut être appliqué à des échantillons de voix ou à des paramètres quelconques.

Les codeurs temporels (w*aveform*)

Les codeurs de ce type codent directement le signal d'entrée, échantillon par échantillon. Cela est le cas des codec de type **PCM** (*Pulse Code Modulation*) qui utilisent une quantification différentielle logarithmique (loi A ou loi μ) et des codec de type **ADPCM** (*Adaptative Differential PCM*) qui reposent sur une quantification différentielle adaptative.

Les codeurs sources

Cette catégorie est constituée par des algorithmes qui transmettent les paramètres de la voix (intensité et fréquence principalement) au lieu d'une représentation du signal lui même. **La voix est modélisée par une fonction d'excitation** choisie parmi plusieurs. À la réception, la voix est reconstituée en utilisant cette fonction et les paramètres qui ont été transmis.

Les plus répandus de cette famille sont les **algorithmes à prédiction linéaire** (***LPC*** – *Linear Prediction Coding*). Ils utilisent une combinaison des échantillons précédents pour construire une valeur prévisible approchant le plus possible l'échantillon à coder. Ils sont spécifique au codage de la voix (tout autre son serait restitué avec une très mauvaise qualité). C'est pour cela que les codec utilisant ce type d'algorithme sont appelés **vocoder** (*voice coder*). Ils produisent une voix d'aspect synthétique, mais le débit réseau généré est très bas. Par exemple, LPC10 (10 coefficients) code des trames de 20 ms sur 54 bits, ce qui donne un débit de 2,4 Kbit/s.

Un autre type d'algorithme fréquemment utilisé est le **LTP** (*Long Term Predictor*), qui s'attache à construire une valeur prévisible de la fréquence du son.

Les codeurs hybrides

Ces codeurs combinent l'utilisation des technologies temporelles et source, en plus de l'utilisations de méthodes de quantification vectorielles.

La méthode la plus utilisée est l'**analyse par synthèse** (*ABS – Analysis By Synthesis*), technique sur laquelle se basent les codec de type **MPE** (*MultiPulse Excited*), **MP-MLQ** (*MultiPulse Maximum Likelihood Quantization*), **RPE** (*Regular-Pulse Excited*), et surtout **CELP** (*Code-Excited Linear Predictive*). Elle consiste à analyser le signal d'entrée en synthétisant différentes approximations déduites des signaux précédents. Citons également les variantes **LD-CELP** (*Low-Delay CELP*) et **CS-ACELP** (*Conjugate-Structure Algebraic CELP*).

Un codec ABS travaille sur des **groupes d'échantillons appelés trames**, typiquement longues de 10, 20 ou 30 ms. Pour chaque trame, le codec détermine les coefficients d'un codage source (appelé ici filtre), puis choisit une fonction d'excitation qui sera appliquée à ce filtre. Les codes transmis correspondent aux paramètres tels que ceux nécessaires à LPC, RPE et LTP.

• • •

LES CODEC AUDIO (SUITE)

Quelques techniques complémentaires peuvent être utilisées par les codec.

Le codage en sous bande

Nombre de codeurs récents utilisent, en plus, une technique de codage en sous bande (*SBC - Sub-Band Coding*) également appelée codage psychoacoustique. Cette technique consiste à découper le signal d'entrée en sous-fréquences et à utiliser plus de bits pour coder les fréquences les plus audibles. En d'autres termes, l'algorithme concentre les efforts de codage sur les fréquences auxquelles notre oreille est la plus sensible.

Le premier algorithme à employer cette technique a été G.222, un codec de type temporel. Il utilise deux sous-bandes sur lesquelles sont appliquées un codage ADPCM, tandis des codec hybrides plus récents tels que MPEG utilisent jusqu'à 32 sous-bandes.

Le codage des silences

Par ailleurs, les codec intègrent des détecteurs de silence (*VAD – Voice Activity Detection*). Lors d'une conversation, les silences représentent, en effet, 35 à 40% du temps global.

Les quatre opérations constituant la numérisation d'un signal analogique (audio ou vidéo) sont réalisées par des processeurs spécialisés appelés **DSP** (*Digital Signaling Processing*). Les algorithmes qui définissent la manière de réaliser ces opérations sont appelés **Codec** (codeur/décodeur).

Les codec étant normalisés ITU-T, série G, leurs noms commencent par cette lettre. Le premier d'entre eux et le plus utilisé sur les liaisons E1/T1 ainsi que sur le RTC, répond à la norme **G.711**. Le codage est de type **PCM** (*Pulse Code Modulation*) ; il n'utilise aucun algorithme de compression. En France, la norme est appelée **MIC** (modulation par impulsions codées), d'où le nom des liaisons à 2 Mbit/s proposées par France Télécom (32 canaux DS0, dont 2 pour la signalisation).

Le codec G.711 existe en deux variantes de codage : A-law (Europe) et μ-law (Amérique du Nord et Japon).

Application	Échantillonnage	Bande passante transmise	Convertisseurs Analogique / Numérique	Débit brut généré
Téléphonie	8 KHz	300 Hz à 3,4 KHz	8 à 13 bits	64 à 104 Kbit/s
Radio FM	16 KHz	50 Hz à 7 KHz	14 ou 15 bits	224 ou 240 Kbit/s
NICAM (son stéréo pour la TV)	32 KHz	30 Hz à 15 KHz	16 bits	512 Kbit/s
Compaq Disc	44,1 KHz	20 Hz à 20 KHz	16 bits	706 Kbit/s
DAT	48 KHz	10 à 22 KHz	24 bits	1 152 Kbit/s

Choisir un codec audio

Pour notre oreille, la qualité d'un codec se résume à sa qualité acoustique. Celle-ci dépend de la fréquence d'échantillonnage, de la bande passante transmise (faculté à reproduire les aigus et les graves) et de la manière dont est codé le signal. L'ensemble de ses paramètres techniques se concrétise par l'appréciation subjective de l'oreille. Celle ci est évaluée par sondage sur une population d'utilisateurs et s'exprime par un indice de satisfaction moyen, le MOS (*Mean Opinion Score*).

Pour la voix sur IP, la qualité d'un codec dépend d'autres facteurs :

- du délai des opérations de codage et de décodage ;
- du débit réseau généré ;
- de la sensibilité aux erreurs. Par exemple, le codec G.726 est moins sensible que G.711, tandis que G.723 tolère un taux d'erreur de près de 3%. Généralement, la perte d'un paquet IP a un impact très important sur les codec PCM et ADPCM (la resynchronisation de l'émetteur sur les échantillons est difficile) et moins sur les codec hybrides ;
- de sa complexité, donc de la puissance processeur requise pour le codage et le décodage. Pour certains codec, l'utilisation de DSP s'avère souvent préférable.

Caractéristiques

D'autres paramètres, très techniques, caractérisent cependant les codec.

Codec (standard ITU)	Algorithme de codage	Durée de l'échantillon	Nombre d'échantillons par trame	Longueur de la trame	Lookahead	Délai algorithmique
		x	=	+	=	
G.711	PCM	0,125 µs	8	1 ms	0 ms	1 ms
G.722	ADPCM	0,0625 µs	16	1 ms	1,5 ms	2,5 ms
G.723.1	ACELP MP-MLQ	0,125 µs	240	30 ms	7,5 ms	37,5 ms
G.726	ADPCM	0,125 µs	80	10 ms	0 ms	10 ms
G.728	LD-CELP	0,125 µs	5 x 4	2,5 ms	2,5 ms	2,5 ms
G.729	CS-ACELP	0,125 µs	80	10 ms	5 ms	15 ms
GSM 6.1	RPE-LTP	0,125 µs	160	20 ms	0 ms	20 ms
GSM 6.2	VSELP	0,125 µs	160	20 ms	4,4 ms	24,4 ms
GSM 6.6	CD-ACELP	0,125 µs	160	20 ms	0 ms	20 ms

Le *lookahead* (analyse en avant) représente le temps moyen d'analyse d'une partie de l'échantillon N+1 permettant de prédire l'évolution du signal et de coder l'échantillon N.

Le délai algorithmique, ou délai de traitement, d'un codec correspond donc à la taille de la trame augmentée du *lookahead*.

Cependant, les codec temporels ne travaillent pas sur des trames mais uniquement échantillon par échantillon. Transportés dans des paquets IP, ces échantillons de 125 microsecondes sont regroupés pour former l'équivalent d'une trame.

Au délai de traitement du codec, il faut donc ajouter le temps mis pour remplir une trame. Par exemple, le codec G.723 génère une trame contenant 30 ms de voix. Elle comporte 240 échantillons compressés à 189 bits (24 octets) ou à 158 bits (20 octets), les deux premiers bits indiquant respectivement le type de codec et la taille de la trame.

Performances

Quelle que soit sa complexité, un codec doit coder et décoder les échantillons selon un temps fixe déterminé par la fréquence d'échantillonnage. Selon la complexité de l'algorithme, un codec requerra donc plus ou moins de puissance, évaluée en Mips (millions d'instructions par secondes). En outre, la puissance requise par un codec est plus importante si le calcul est effectué en virgule flottante que s'il est effectué en virgule fixe.

Les processeurs spécialisés que sont les DSP exécutent en une seule instruction et en un seul cycle horloge des calculs répétitifs comme, par exemple, « (a*b)+résultat précédent », calculs que des processeurs généralistes comme le Pentium exécutent en plusieurs instructions et plusieurs dizaines de cycles. Les DSP intègrent également les dictionnaires (*codebook*) utilisés par les codec Hybrides.

Les valeurs présentées ci-après sont les moyennes indiquées par un constructeur pour ses DSP en virgule fixe. Elles ont été arrondies aux valeurs les plus proches. Pour les codec temporels, la puissance de calcul est quasi identique en codage et en décodage, alors que pour les codec hybrides, une puissance plus importante est nécessaire au codage.

Codec	Débit réseau généré	Longueur de la trame	Nombre d'octets par trame	Nombre de trames par paquet RTP	Complexité codage / décodage
		x	÷ 8 =		
G.711	64 Kbit/s	1 ms	8	20	0,1 / 0,1 Mips
G.722 (1)	48, 56 et 64 Kbit/s	1 ms	6 / 7 / 8	20	13 / 13 Mips
G.723.1 (1)	5,3 et 6,4 Kbit/s	30 ms	20 24	1	15,5 / 2 Mips
G.726 **G.727** (2)	16, 24, 32 ou 40 Kbit/s	10 ms	20 / 30 / 40 / 50	2	4 / 4,5 Mips

Codec	Débit réseau généré	Longueur de la trame	Nombre d'octets par trame	Nombre de trames par paquet RTP	Complexité codage / décodage
G.728	16 Kbit/s	2,5 ms	5	8	17,5 / 12 Mips
G.729	8 Kbit/s	10 ms	10	1	16 / 3,5 Mips 8,5 / 2 Mips
GSM 6.1	13 Kbit/s	20 ms	32,5	1	2,5 Mips
GSM 6.2	5,6 Kbit/s	20 ms	14	1	17,5 Mips
GSM 6.6	12,2Kbit/s	20 ms	30,5	1	15,4 Mips

(1) Les codec G.722 et G.723 sont des codec combinés (*embedded*) : ils changent automatiquement de débit.
(2) G.727 est la version combinée de G.726.

À titre de comparaison, un processeur Pentium cadencé à 100 MHz fournit une puissance d'environ 166 Mips (source Intel).

Qualité

En dehors des performances intrinsèques d'un codec résultant des paramètres qui viennent d'être décrits, sa qualité dépend essentiellement de l'échantillonnage et du taux de compression : plus ce dernier est élevé, plus la voix apparaît métallique à celui qui l'écoute.

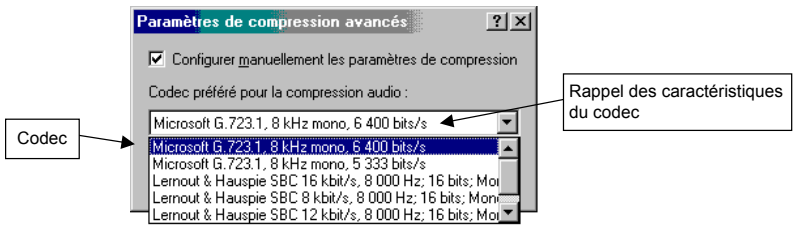
Sans compression, le codec G.711 sert de référence pour la qualité de la voix sur le téléphone, référence qui est appelée *toll quality*. La perception de la qualité de la voix est subjective. C'est pour cela qu'elle est mesurée à travers un indice moyen de satisfaction appelé **MOS** (*Mean Opinion Score*), sur une échelle allant de 0 (très mauvais) à 5 (très bon).

Codec	Algorithme de codage	Échantillonnage	Bande passante transmise	MOS	Applications
G.711	PCM	8 KHz	3,4 KHz	4,2	RTC, RNIS
G.722	ADPCM	16 KHz	7 KHz	4 / 4 / 4,3	RNIS
G.723.1	ACELP MP-MLQ	8 KHz	3,4 KHz	3,7 / 3,9	VoIP
G.726	ADPCM	8 KHz	3,4 KHz	2 / 3,2 / 4 / 4,2	DECT
G.728	LD-CELP	8 KHz	3,4 KHz	4	RNIS
G.729	CS-ACELP	8 KHz	3,4 KHz	4	Frame-Relay
GSM 6.1	RPE-LTP	8 KHz	3,4 KHz	3,6 – 3,8	Téléphone cellulaire

Codec	Algorithme de codage	Échantillonnage	Bande passante transmise	MOS	Applications
GSM 6.2	VSELP	8 KHz	3,4 KHz	3,5 – 3,7	Téléphone cellulaire
GSM 6.6	CD-ACELP	16 KHz		4,1	Téléphone cellulaire

Les voix ayant une sonorité métallique obtiennent un MOS inférieur à 3 ; celles acceptables pour une communication téléphonique, un MOS compris entre 3 et 4. Au-delà de 4, la qualité est excellente.

Le choix d'un codec résulte donc d'un compromis qualité/débit sur le réseau utilisé. Ainsi, le programme Netmeeting permet-il de sélectionner le plus approprié à votre contexte (menu « Outil→Options→Audio→Avancé »).



Les codec vidéo

Une image vidéo affichée sur un écran de PC représente 640 × 480 points en 16,7 millions de couleurs (24 bits) × 25 images par seconde, ce qui correspond à un débit de 23 Mbit/s (640 × 480 × 3 × 25).

Afin de diminuer le débit généré sur le réseau, une première solution consiste à diminuer le nombre d'images par seconde. La meilleure solution consiste cependant à compresser les images, à l'instar de la voix.

Pour les données, aucun bit ne doit être altéré dans les phases de compression et de décompression : la décompression d'un fichier doit reproduire exactement le fichier original. Les taux de compression sont, en moyenne, de 1 pour 2, voire plus pour les images bitmap.

Mais, pour la vidéo, l'image décompressée peut être différente de la même image avant compression. Ainsi, une légère variation de couleur ou la suppression d'un pixel n'est pas perceptible lorsqu'ils affectent 1 image sur 25 par seconde. Ce principe permet d'atteindre des ratios de compression de 1 pour 100, voire 1 pour 300. Une perte de qualité est acceptable à partir du moment où elle est à peine perceptible.

Normes	Bande passante requise	Remarque
MPEG-1	De 1,8 à 80 Mbit/s	Compression préalable, puis décompression en temps réel
MPEG-2	De 1,8 à 80 Mbit/s	Vidéo à la demande par des câblo-opérateurs, HDTV. Codec en temps réel. Interfacé à AAL5.
MPEG-4	De 64 Kbit/s à 1,5 Mbit/s	Stockage des données multimédias et vidéo haute qualité et bas débit (DivX)
M-JPEG	De 8 à 10 Mbit/s	Agrégation d'images JPEG.
H.261	p x 64 Kbit/s (1 < p < 30) QCIF	Utilisé par les normes H.32x.Visioconférence.
H.263	p x 64 Kbit/s (1 < p < 30) Sub-QCIF et QCIF	Utilisé par les normes H.32x. Amélioration de H.261.

Principes de compression des images

La compression peut être **spatiale** (c'est-à-dire porter sur une image — mode intraframe) ou **temporelle** (c'est-à-dire porter sur plusieurs images dans le temps — mode interframe).

Dans le mode **intraframe**, l'image est divisée en blocs de 8×8 pixels : les pixels et leurs couleurs associées sont convertis en fréquence de changement de couleurs et amplitude des variations de couleurs. La moyenne est calculée sur le résultat, de manière à ce que la perte d'informations qui en résulte ne soit pas perceptible par l'œil. Les données sont ensuite compressées par différents algorithmes.

Dans le mode **interframe**, seule la différence entre deux images est transportée à partir d'une image de référence (*key frame*) ou à partir de l'image précédente. Une technique de compensation de mouvement est, de plus, utilisée : le delta est calculé non pas, pixel par pixel, mais par bloc de pixels en mouvement.

Parce qu'image après image, le mode interframe dégrade l'image de manière trop importante, il est nécessaire de transmette régulièrement l'image complète en mode intraframe.

Par exemple, M-JPEG (*Motion JPEG*) utilise un algorithme spatial (série d'images JPEG), alors que les normes MPEG et H.26x utilisent une combinaison des deux.

LE POINT SUR MPEG (ISO CD 11172-2)

La norme MPEG (*Moving Pictures Expert Group*) est structurée en **trois couches** : la première, dite **système**, a trait à la synchronisation et au multiplexage des flux, la seconde à la compression de la **vidéo** et la troisième à la compression de l'**audio**. La norme utilise également le terme « couche » pour désigner les unités de travail sur lesquelles s'appliquent les fonctions de compression d'image et également pour désigner les trois niveaux de qualité audio (I, II et III). MP3 tire son nom de la qualité de niveau III.

Couche vidéo	Description
Séquence	Train de bits produits par la compression MPEG
Groupe d'images	Trames I (*Intra frame*), P (*Predictive frame*) et B (*Bidirectional frame*)
Image	Unité primaire sur laquelle s'applique le codage
Tranche (slice)	Groupe de macroblocs sur laquelle s'applique la synchronisation
Macrobloc	Groupe de 6 blocs sur laquelle s'applique la compensation de mouvement
Bloc	Bloc de 8x8 pixels sur lequel est appliqué l'algorithme DCT (*Discrete Cosine Transform*)

Il existe 3 versions de MPEG : MPEG-1, MPEG-2 et MPEG-4. La version 3 n'a jamais vu le jour, car une simple évolution de MPEG-2 s'est avérée suffisante pour atteindre les objectifs assignés à MPEG-3.

La norme **MPEG-1** produit des images de toutes tailles et à tous les débits, jusqu'à 80 Mbit/s, définies par des paramètres appelés CPB (*Constrained Parameters Bitstream*). Les formats les plus couramment utilisés sont le **SIF** (*Standard Interchange format*) et le **CCIR-601** (Comité Consultatif International Radio). Les débits générés dépendent de la résolution et du nombre d'images par seconde.

Pixels	Images / sec.	Format	Débit maximal	Application
352 x 240	30	SIF	4 Mbit/sec	VHS, NTCS
352 x 288	25	SIF	4 Mbit/sec	VHS, PAL/SECAM
720 x 480	30	CCIR	15 Mbit/sec	HDTV, DVD
1440 x 1152	25		60 Mbit/sec	HDTV, DVD

En plus des trames **I** (*Intra frame*) et **P** (*Predicted frame*), l'algorithme MPEG produit des trames **B** (*Bidirectional*) résultant du codage opéré sur l'interpolation des trames précédentes et futures. La séquence de trames produites ressemble ainsi à la suivante : IBBPBBPBBPBBIBBPBBPBBPBBI... (une trame I toutes les douze trames).

La compression résulte d'une série d'opérations appliquée à l'image complète (trames I) ou à des portions d'image (trames P), celles qui ont changé par rapport à l'image précédente. Dans les deux cas, l'image ou la portion d'image est codée avec l'algorithme DCT, quantifiée, puis successivement codée par RLE et par entropie (selon la fréquence des valeurs).

Le codage **DCT** (*Discrete Cosine Transform*) permet de représenter l'image par des fréquences, seuls les coefficients d'une fonction sinusoïdale étant générés. La **quantification** a pour objectif de réduire le nombre de coefficients DCT à coder (comme pour la voix, certaines informations sont alors perdues). Le codage **RLE** (*Run Length Encoding*) remplace les valeurs successives identiques par un code de quelques octets. Le **codage par entropie** consiste à trier les valeurs par fréquence d'apparition puis à les remplacer par des codes plus petits.

• • •

LE POINT SUR MPEG (SUITE)

Concernant l'audio, les débits vont de 32 Kbit/s à 448 Kbit/s selon la qualité du son, s'il est stéréophonique, etc.

Couche audio	Débit	Trame	Application
I	384 Kbit/s	384 échantillons PCM 48 KHz, 8 ms	Satellite
II	64 à 224 Kbit/s	1 152 échantillons PCM 48 KHz, 24 ms	MiniDisc de Sony
III	64 à 128 Kbit/s	1 152 échantillons	MP3

Par exemple, pour une qualité CD (*Compaq Disc*) avec un échantillonnage à 44,1 KHz, la couche III génère un débit de 128 Kbit/s (trames de 417 octets).

La norme **MPEG-2** utilise, quant à elle, des algorithmes améliorés offrant des performances de 10 à 20 % supérieures à celles MPEG-1 tout en étant compatible avec ce dernier. Les paramètres CBP sont remplacés par une combinaison de quatre **profils** (algorithmes utilisés et fonctions activées) et de quatre **niveaux** définissant les paramètres (résolutions minimale et maximale, échantillonnage, couches audio et vidéo supportées, débit maximal, etc.).

La principale caractéristique de **MPEG-4** (ISO 14496) est son orientation objet. Appliqué aux images de synthèse, l'algorithme de codage décrit les objets, leur forme, leur texture, leur dimension (2D ou 3D) en s'appuyant sur les spécifications VRML (*Virtual Reality Modeling Language*). Son objectif est également de produire des vidéos à bas débit, de 64 Kbit/s à 1,5 Mbit/s. Les applications visées par cette norme sont, par exemple, **DivX** (*Digital Video Express*), les bases de données multimédias et interactives, le vidéophone, l'audiovisuel mobile, etc.

La norme **H.261** promue par l'ITU-T repose sur les mêmes principes que MPEG, mais consomme moins de CPU pour la compression. De plus, l'algorithme permet d'augmenter la compression, au détriment cependant de la qualité des images qui bougent vite (s'il existe plus de 15 pixels de différence entre deux images, l'image devient floue).

La norme **H.263** est une amélioration de H.261 : elle permet de consommer 50% de bande passante en moins et de récupérer plus facilement les erreurs de transmission, et ce, pour une qualité équivalente.

Le CIF (*Common Intermediate Format*) est le format de base d'une image respectant les normes H.26x.

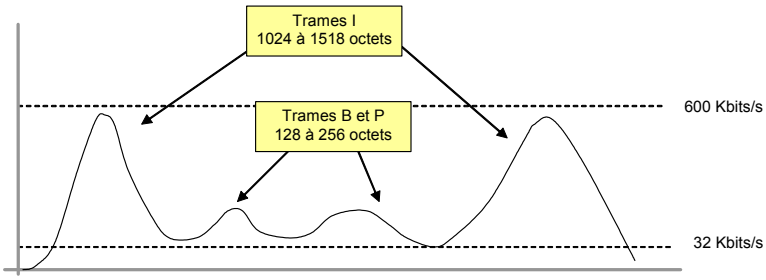
Format de l'image	Nombre de pixels pour la luminance	Nombre de lignes pour la luminance	Nombre de pixels pour la chrominance	Nombre de lignes pour la chrominance
sub-QCIF	128	96	64	48
QCIF	176	144	88	72
CIF	352	288	176	144
4CIF	704	576	352	288
16CIF	1 408	1 152	704	576

Généralement, le format QCIF est utilisé lorsque la bande passante est inférieure à 192 Kbit/s (p <= 3). H.261 et H.263 travaillent sur des GOB (*Group Of Block*) de 33 macroblocs (soit 198 blocs). Le format QCIF correspond ainsi à 1 GOB.

Les problèmes posés par les transmissions audio et vidéo

Il ressort de cette présentation que le trafic réseau généré par les codec vidéo est par nature erratique, les pointes de débit correspondant aux transmissions de trames I (50% du volume du trafic), et les débits les plus faibles correspondant aux trames P et B.

Figure 13-1.
Profil de trafic généré par les flux vidéo



En revanche, le trafic audio est plutôt continu, et surtout beaucoup plus faible que les flux vidéo.

Les flux audio et, dans une moindre mesure, les flux vidéo sont sensibles à plusieurs phénomènes :

- l'écho ;
- le délai de transit (encore appelé latence), c'est-à-dire le temps qui s'écoule entre la prononciation d'un mot et sa restitution côté récepteur ;
- la variation du délai de transit, appelée gigue (*jitter*, en anglais).

Le phénomène d'**écho** provient de la réflexion du signal sur le câble, surtout au niveau des convertisseurs 2 fils/4 fils présents en nombre sur le RTC. Il est surtout perceptible sur de longues distances et est amplifié si les délais de transit sont importants. Si ces derniers sont assez bas (moins de 50 ms), l'écho n'est pas perceptible et est masqué par la conversation. Étant donné que les délais de transit sont, la plupart du temps, supérieurs à 50 ms, des appareils spécifiques, appelés annulateurs d'écho (ITU G.165), sont systématiquement utilisés sur le RTC.

Le **délai de transit** affecte une conversation téléphonique : si le temps qui s'écoule entre la fin d'une phrase et sa réception complète par le récepteur est trop long, les personnes commencent à parler en même temps puis s'arrêtent en même temps, se coupent la parole, etc. La norme **G.114** préconise un délai de transmission maximal de 400 ms pour le RTC. Dans les faits, la dégradation de la qualité de la voix est nettement perceptible lorsque les délais dépassent 300 ms. De plus, l'écho devient perceptible, ce qui ajoute à la mauvaise qualité.

Les **variations des délais de transit** (gigues) sont essentiellement dues à la charge du réseau qui doit traiter différents types de flux. Le RTC fonctionne en mode commutation de circuit : un canal de 64 Kbit/s est réservé pendant toute la durée de la conversation. Il y a donc très peu de gigue, voire pas du tout. L'inconvénient est que la bande passante du réseau n'est pas optimisée : les 64 Kbit/s restent accaparés même lorsqu'ils ne sont pas utilisés, pendant les silences ou en l'absence de transmission.

En revanche, un réseau de paquets, de trames ou de cellules véhicule simultanément une multitude de flux et exploite au mieux ses ressources en les partageant entre plusieurs utilisateurs. Le délai de traitement varie donc en fonction de sa charge.

La variation du délai crée ainsi des interruptions inattendues au milieu d'une phrase ou d'un mot (un paquet voix arrive plus tard que les autres), qui peuvent rendre une conversation inintelligible. Pour compenser la gigue, on utilise des tampons mémoire. Les trames arrivant en retard par rapport à la moyenne sont restituées immédiatement, mais celles arrivant en avance par rapport à cette même moyenne restent plus longtemps dans la mémoire tampon.

L'inconvénient est donc que le délai de transit est augmenté proportionnellement à la taille du tampon. Généralement, cette taille correspond à un délai égal à environ deux fois celui du traitement du codec. Dans la pratique, les équipements la font varier dynamiquement.

Concernant le téléphone, notons également une spécificité liée aux fonctionnalités offertes par le **DTMF** (*Dual-Tone MultiFrequency*). Ce signal permet à un utilisateur de dialoguer avec un serveur audiotel, par exemple pour accéder au rappel sur occupation (touche « 5 » avec France Télécom) ou pour consulter son compte bancaire. La compression de la voix empêche la transmission de ce type de signal. Celui-ci est donc codé, puis régénéré à son arrivée (G.729).

Estimation du temps de transit

Le délai de transit sur le réseau est la somme des délais induits par tous les équipements traversés : câbles, routeurs, commutateurs, etc.

Il faut ainsi additionner :

- le délai de sérialisation déterminé par la vitesse de la ligne et la taille du paquet ;
- les délais de traitement propres au codec ;
- le délai de transit dans un nœud (routeur ou passerelle), déterminé par l'empaquetage et le dépaquetage des données, augmentés des temps de traitement dus aux protocoles réseau (interprétation des entêtes, routage, etc.).

La **sérialisation** est l'action d'envoyer les bits sur le câble. Plus le débit de la liaison est élevé, plus le temps de sérialisation est court. Par exemple, il faut 125 microsecondes pour envoyer un octet sur une LS à 64 Kbit/s contre 0,05 microsecondes sur une liaison à 155 Mbit/s. De même, plus les sites sont rapprochés, plus le délai de propagation est court. Par exemple, il faut compter 32 ms de délai de propagation sur une LS 64 Kbit/s entre l'Europe et les États-Unis.

À titre d'exemple, le tableau suivant dresse la liste des délais nécessaires pour envoyer une trame de 30 ms générée par le codec G.723.

Opération	Délai
Traitement codec émission	7,5 ms (G.723)
Remplissage de la trame	30 ms de voix (20 ou 24 octets)
Encapsulation de la trame dans RTP/UDP/IP	5 ms (dépend de l'OS et de la pile IP)
Encapsulation IP/Frame Relay	2 ms (traitement du routeur)
Sérialisation de 31 octets (20 + 5 + 6)* sur une ligne à 128 Kbit/s	2 ms (4 ms sans compression)
Sous-total 1	46,5 ms
Décapsulation IP/Frame Relay	2 ms (traitement du routeur)
Décapsulation de la trame dans RTP/UDP/IP	5 ms (dépend de l'OS et de la pile IP)
Tampon de gigue en réception	60 ms (deux fois le délai du codec)
Traitement codec réception	7,5 ms (G.723)
Sous-total 2	74,5 ms
Temps de transit réseau **	De 30 à 150 ms
Total	De 151 à 271 ms

(*) L'encapsulation dans un paquet IP ajoute 2 à 5 octets (avec compression des en-têtes RTP/UDP/IP, voir chapitre 16). Il faut ajouter à cela les entêtes Frame Relay (6 octets) ou ATM (5 octets) ou encore Ethernet, etc.

(**) Ce temps est égal à celui induit par les routeurs et/ou commutateurs et/ou passerelles, éventuellement augmenté de celui induit par la propagation des signaux sur les liaisons internes au réseau de l'opérateur.

Le transport des données multimédias

À la différence des données applicatives, la voix et la vidéo — ainsi que, dans une moindre mesure, la télécopie — acceptent l'altération des données. Il en résulte une dégradation de la qualité du son, de l'image ou de la page, qui peut toutefois être tolérée par les personnes qui les reçoivent. En plus du délai de transit et de sa variation, des paquets peuvent donc être perdus, mais bien sûr dans une certaine limite.

Qualité	Voix	Fax	Vidéo
Bonne (dégradations non perceptibles)	Délai < 200 ms Gigue < 15 ms Perte < 5 %	< 200 ms Gigue < 100 ms Perte < 2 %	< 200 ms Gigue < 15 ms Perte < 5 %
Limite de l'acceptable	Délai = 400 ms Gigue = 30 ms Perte = 10 %	Délai = 300 ms Gigue < 1 000 ms Perte = 4 %	Délai = 400 ms Gigue = 20 ms Perte = 10 %

En outre, les liens réseau doivent être correctement dimensionnés, afin de supporter les débits générés par les flux multimédias. Ces derniers sont variables, car ils dépendent de la qualité voulue pour le son et l'image.

Application	Bande passante requise
Audio (voix qualité téléphone)	64 Kbit/s sans compression (*toll quality*) De 11 à 30 Kbit/s selon la compression
Vidéoconférence	384 à 768 Kbit/s (qualité visio professionnelle) 10 Mbit/s (qualité DVD)
Tableau partagé	De 10 à 64 Kbit/s

Sur le réseau téléphonique à commutation de circuit (RTC, RNIS), la voix occupe un débit fixe réservé entre l'appelant et l'appelé pendant toute la durée de la communication. Sur un réseau à commutation de paquets, tel qu'IP, aucun débit n'est réservé, et celui-ci peut varier dans le temps (par exemple, les silences ne sont pas transmis, etc.).

La voix peut ainsi être transmise sur IP — on parle alors des produits VoIP (*Voice Over Internet Protocol*) —, directement sur Frame Relay (VoFR) ou sur ATM (VoATM).

Dans tous les cas, le transport de la voix et de la vidéo requiert des mécanismes spécifiques, afin d'assurer la **qualité de service** requise (débit, délai de transit, variation du délai de transit, perte de paquets tolérée).

Un autre problème, qui cette fois n'est pas nouveau, est de dimensionner les liens réseau en fonction du nombre d'utilisateurs. On parlera ici de canaux, un canal correspondant à une session voix ou vidéo.

Une règle simple, mais peu économique, est de prévoir autant de canaux que d'utilisateurs : le débit total est alors égal au débit d'un canal VoIP multiplié par le nombre d'utilisateurs. Une règle plus complexe, mais plus rationnelle, consiste à s'appuyer sur le nombre de communications simultanées et à utiliser des modèles statistiques permettant de prévoir le nombre de canaux nécessaires en fonction de différents paramètres, tels que le nombre d'utilisateurs, la durée moyenne des communications, le taux de débordement acceptable, etc.

Les modèles standards s'appuient sur des distributions de Poisson qui permettent de convertir la nature aléatoire des appels en probabilité, sur une unité de mesure appelée Erlang ou CCS (*Centum Call Seconds*).

En définitive, un travail d'ingénierie important est nécessaire pour préparer son réseau IP au multimédia.

14

La qualité de service sur IP

Par défaut, un réseau IP se contente d'acheminer les paquets au mieux de ses possibilités, et sans distinction. Tant que la bande passante (c'est-à-dire le débit) est suffisante, il n'y a pas de problème. Mais, en cas de saturation, les routeurs sont obligés de rejeter des paquets, invitant tous les émetteurs à réduire leur flux. En conséquence, l'utilisateur constate une dégradation des performances du réseau.

La notion de qualité de service (QoS, *Quality of Service*) introduit la possibilité de partager le plus équitablement possible une ressource devenant de plus en plus rare, car partagée par un grand nombre de flux applicatifs qui peuvent interférer les uns avec les autres. Elle introduit également la possibilité de déterminer différents niveaux de service en fonction de la nature de ce flux (une visioconférence, un transfert de fichier, etc.).

Au chapitre 11, vous avez sans doute remarqué que la gestion de la qualité de service est déjà prise en compte par des protocoles de niveau 2, tels qu'ATM et Frame Relay. Alors pourquoi gérer la QoS sur IP ? Parce que, lorsqu'une application génère des flux sur un réseau Ethernet, qui traversent ensuite un réseau ATM ou Frame Relay pour arriver sur un autre réseau local, le seul dénominateur commun est IP.

Dans ce chapitre, vous apprendrez ainsi :

- comment améliorer les performances de votre réseau ;
- les différents moyens permettant de gérer la qualité de service ;
- quelle politique de qualité de service choisir ;
- comment configurer votre réseau pour gérer cette qualité de service.

Améliorer les performances du réseau

Notre réseau est de plus en plus sollicité par de nouvelles applications aux besoins très divers : connexions Telnet pour se connecter à une machine Unix, transferts de fichiers, bases de données en mode client-serveur et, maintenant, flux audio et vidéo.

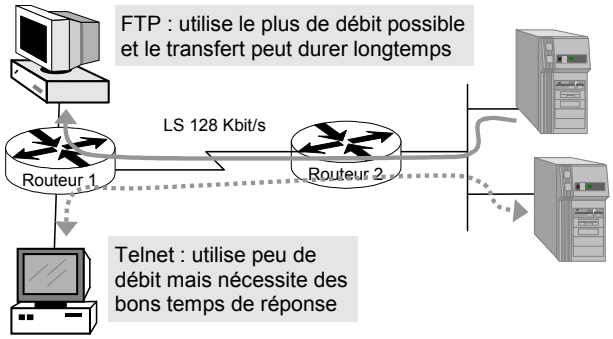
La charge du réseau augmente, et les flux générés se perturbent mutuellement : quel utilisateur n'a pas expérimenté des temps de réponse très longs alors que son voisin a lancé le transfert d'un gros fichier ?

Affecter des priorités aux files d'attente

Le moyen le plus rudimentaire d'améliorer la qualité du service rendu par votre réseau est de définir manuellement, et sur tous vos routeurs, des priorités. Cette approche consiste surtout à améliorer les performances du réseau en favorisant des applications au détriment d'autres. La qualité est améliorée, mais non garantie.

Le cas le plus couramment rencontré est celui d'une liaison WAN qui doit véhiculer des connexions Telnet (de type conversationnel) et des transferts de fichiers FTP. Les temps de réponse Telnet se dégradent dès qu'un transfert FTP est lancé. Aucune qualité de service n'étant gérée, la bande passante du réseau est, en effet, accaparée par celui qui en consomme le plus, cela bien sûr au détriment des autres.

Afin d'éviter que les applications Telnet ne soient gênées par un transfert de fichiers, le moyen le plus simple est de définir des priorités sur la base des ports TCP qui identifient les applications (23 pour Telnet et 20 pour le canal de données FTP).



En classant les paquets Telnet en priorité haute, les paquets IP comportant le port TCP 23 seront placés dans la file d'attente à priorité haute, tandis que les paquets FTP seront stockés dans la file d'attente à priorité basse. Nos routeurs, qui sont de marque Cisco, gèrent ainsi quatre files d'attente correspondant à quatre priorités : haute, moyenne, normale et basse (*high*, *medium*, *normal*, *low*).

LE RÔLE DES FILES D'ATTENTE

Lorsqu'il arrive plus de paquets qu'il n'en sort du routeur, celui-ci les garde en mémoire en attendant que les plus anciens soient envoyés. Si d'autres paquets continuent d'arriver, la file d'attente se sature. Le débordement de la file d'attente se traduit par le rejet des paquets continuant d'arriver (le routeur les ignore).

Un algorithme de traitement des files a donc deux rôles essentiels :

- **traiter** en priorité tel ou tel paquet en cas de **congestion** ;
- **rejeter** en priorité tel ou tel paquet en cas de **saturation** du lien réseau.

La priorité de traitement d'un paquet dépend des paramètres de qualité de service qui peuvent être prédéfinis dans l'algorithme, soit définis statiquement dans le routeur, soit définis dynamiquement par l'application.

Les paquets rejetés signifient pour TCP qu'il doit réduire sa fenêtre d'émission et donc son flux. Si le même paquet est rejeté plusieurs fois, l'utilisateur attend, et si l'attente se prolonge, la session risque d'être interrompue à l'expiration d'un timeout. Le choix des paquets à rejeter dépend de l'**algorithme** choisi pour traiter la file d'attente.

Il existe un jeu de files d'attente par interface réseau. La priorité doit donc être activée sur les interfaces série qui émettent les flux FTP perturbateurs. Étant donné que le transfert de fichiers peut avoir lieu dans les deux sens, la même configuration doit être appliquée aux deux routeurs :

```
priority-list 1 protocol ip high tcp 23
priority-list 1 protocol ip low tcp 20
priority-list 1 protocol ip low tcp 21
int s 0
priority-group 1
```

Telnet

FTP

En positionnant une priorité basse pour les flux FTP, les transferts de fichiers prendront plus de temps mais ne perturberont pas les sessions Telnet. Lorsque plusieurs paquets sont en file d'attente, le routeur enverra de préférence davantage de paquets Telnet que de paquets FTP.

Afin d'offrir le même comportement sur l'ensemble du réseau, il faut configurer de manière identique tous les routeurs. De même, les interfaces Ethernet pourraient être traitées de manière identique si les réseaux locaux étaient chargés. En fonctionnement normal, le débit de 10 Mbit/s suffit, en effet, à absorber les flux FTP et Telnet. Ce n'est qu'en cas de charge que la perturbation se manifeste et que l'activation des priorités permet de conserver le même niveau de service que celui obtenu en fonctionnement non chargé.

Il est intéressant de spécifier plus en détail un flux en utilisant une access-list :

```
priority-list 1 protocol ip low list 10
access-list 10 permit 10.0.0.1 0.255.255.255
```

Par défaut, tous les autres paquets seront traités de la manière suivante :

```
priority-list 1 default medium
```

Il est également possible de contrôler indirectement le débit en définissant le nombre maximal de paquets pouvant être en attente :

```
priority-list 1 queue-limit 20 40 60 80
```

Les quatre chiffres indiquent le nombre maximal de paquets pouvant être stockés dans les files d'attente haute, moyenne, normale et basse (les valeurs indiquées sont celles par défaut). Si la file d'attente est pleine, les paquets en excès sont rejetés, et un message ICMP *source quench* est envoyé à l'émetteur pour lui indiquer de ralentir le flux.

Cette stratégie a des limites, car elle est à double tranchant lorsque le flux de paquets est important : si la file d'attente est de trop grande dimension, les paquets s'accumulent, ce qui a pour conséquence d'augmenter le temps de réponse global. Si, en revanche, elle est de trop petite taille, des paquets peuvent être perdus, notamment entre deux réseaux de débits différents (depuis un réseau Ethernet à 10 Mbit/s vers un lien série, par exemple) : le routeur rejette tout paquet entrant tant que ses files d'attente sont saturées.

Dans notre exemple, en abaissant ainsi à 60 le nombre de paquets dans la file d'attente basse, le débit maximal du flux FTP sera encore abaissé.

Agir sur les files d'attente

L'étape suivante vers une gestion de la qualité de service consiste à agir sur le comportement des files d'attente des routeurs.

L'algorithme FIFO – Un fonctionnement simple

Le moyen le plus simple est de gérer les files d'attente sur le mode **FIFO** (*First In, First Out*) : le routeur traite les paquets au fil de l'eau dans leur ordre d'arrivée, au mieux de ses capacités. C'est le principe du *best effort*, comportement par défaut des routeurs. Aucun paramétrage n'est possible, si ce n'est de définir des priorités ce qui a pour effet de créer une file d'attente par niveau de priorité.

Gérer les congestions

Une méthode plus efficace est de traiter les files d'attente à l'aide de l'algorithme **WFQ** (*Weighted Fair Queueing*). Celui-ci identifie dynamiquement les flux et veille à ce que les applications générant peu de trafic ne soient pas perturbées par celles générant beaucoup de données.

Le principe de l'algorithme étant figé, la possibilité de paramétrage est donc limitée au seuil de rejet et au nombre de files d'attente :

```
int s 0
bandwidth 512
fair-queue 64 256 0  (valeurs par défaut)
```

LA GESTION DES FILES D'ATTENTE

Comme dans la vie de tous les jours, une file d'attente se forme lorsque le flux entrant est plus important que le rythme de sortie des paquets, par exemple depuis une interface Ethernet à 10 Mbit/s vers une liaison série à 512 Kbit/s. Le routeur peut alors adopter différentes stratégies pour traiter les paquets en attente. Même si tout se passe en une fraction de seconde, le choix de l'une ou de l'autre d'entre elles peut avoir une grande influence sur le comportement général du réseau.

Le mode le plus simple est de type **FIFO** (*First In First Out*) qui consiste à traiter les paquets dans leur ordre d'arrivée. C'est celui qui consomme le moins de CPU et qui engendre le moins de latence pour les paquets. En cas de congestion, il se révèle, en revanche, moins performant car il n'y a aucune régulation du trafic.

Le principe du **WFQ** (*Weighted Fair Queueing*) repose quant à lui sur un contrôle de flux dynamique en fonction de discriminants propres à chaque protocole (le DLCI Frame Relay, les adresses IP et le port TCP, une priorité, etc.), le principe étant de privilégier le trafic à faible volume sur celui à fort volume. Pour cela, chaque flux est identifié (à partir des adresses IP source et destination et des port TCP/UDP source et destination), et son débit est mesuré.

Le principe de l'algorithme **RED** (*Random Early Detection*) consiste à prévenir les congestions. Lorsque la file d'attente commence à être saturée, des paquets correspondant à des flux sélectionnés aléatoirement sont rejetés. Les variantes **WRED** (*Weighted RED*) et **ERED** (*Enhanced RED*) permettent de sélectionner les flux, en fonction de priorités qui déterminent le rejet des paquets.

Le **traffic shaping** repose sur l'algorithme du **leaky-bucket** (littéralement le seau percé), dont le principe est de lisser un flux erratique afin d'obtenir un débit stable en sortie. Pour ce faire, les pics de trafic sont temporisés dans la file d'attente avant d'être retransmis. Le **policing** repose, quant à lui, sur l'algorithme du **token-bucket**, qui consiste à écrêter le trafic dépassant un certain seuil : le trafic en débordement est ignoré. Selon les implémentations, le *traffic shaping* et le *policing* peuvent opérer seuls ou conjointement avec une file FIFO ou WFQ.

Dans l'ordre, les paramètres indiquent :

- le seuil au-delà duquel les paquets vont commencer à être rejetés (c'est-à-dire la taille de la file d'attente moins une petite marge), ici 64 paquets par file ;
- le nombre de files d'attente pouvant être créées dynamiquement pour les flux sans qualité de service (de type *best effort*), ici 256 files pour traiter simultanément 256 flux ;
- le nombre de files d'attente pouvant être réservées par RSVP (voir plus loin), dans notre cas aucune.

Les paquets qui disposent des mêmes adresses IP source et destination, des mêmes ports source et destination et du même champ TOS (voir plus loin) correspondent à un même flux. Une fois identifiés, tous les paquets du même flux sont placés dans la même file d'attente.

La commande « bandwidth », qui indique le débit du lien réseau en Kbit/s, permet à l'algorithme de définir le nombre nécessaire de tampons d'émission associés à l'interface (généralement quelques-uns).

Le mot « *weighted* » dans WFQ indique que l'algorithme prend en considération la priorité indiquée dans le champ « IP precedence » du paquet (voir le paragraphe suivant à ce sujet).

Prévenir les congestions

Alors que l'algorithme WFQ permet de gérer les situations de congestion, l'algorithme **WRED** (*Weighted Random Early Detection*) permet de les prévenir. Dès qu'une congestion est détectée, l'algorithme rejette des paquets, ce qui contraint l'émetteur, au niveau de la couche TCP, à ralentir son flux :

```
int s 0
bandwidth 512
random-detect
```

Liaison à 512 Kbit/s

Une approche récente, appelée **ECN** (*Explicit Congestion Notification*), consiste à indiquer explicitement à TCP qu'il y a une congestion, plutôt que de laisser cette couche le deviner en constatant que des paquets sont rejetés dans le réseau. Les congestions peuvent ainsi être évitées plus rapidement.

```
int s0
service-policy input flux

policy-map flux
class generique
bandwidth percentage 35
random-detect ecn
```

35 % du trafic est alloué à cette politique de gestion de la QoS.

L'introduction d'une composante aléatoire par WRED permet de remédier au problème de la **synchronisation globale**. En cas de congestion, toutes les sessions TCP réagissent en même temps et de la même manière en réduisant leur flux, puis constatant que le réseau est désengorgé, augmentent à nouveau leur flux. Le réseau oscille ainsi entre des périodes de congestion et de non congestion. En rejetant aléatoirement des paquets, WRED oblige les sessions les moins prioritaires à réduire leur flux avant que les autres sessions n'expérimentent des congestions.

Réguler le trafic

Une autre alternative est de réguler le flux selon le principe du *traffic shaping* : le but de cet algorithme est de donner un caractère prévisible aux flux en convertissant un flux erratique en un flux à peu près constant.

Sur nos routeurs, l'algorithme peut être appliqué à l'ensemble du trafic, ou à une portion de celui-ci définie par une *access-list* :

```
access-list 101 permit ftp any any
int s 0
traffic-shape group 101 128000 64000 32000

int e1
traffic-shape rate 5000000 10000000 2000000
```

Taille du *burst* en bits

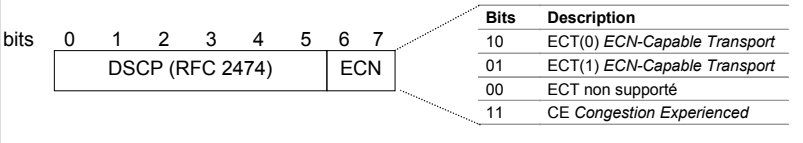
Définitions identiques à celles du Frame Relay décrit au chapitre 11.

Burst en excès = nombre de bits pouvant dépasser le débit moyen.

LA NOTIFICATION DE CONGESTION (RFC 3168)

Dans un réseau TCP/IP, les congestions sont détectées par l'émetteur lorsque des paquets sont perdus (acquittements dupliqués ou timeout). La couche TCP répond à cette situation en réduisant la fenêtre d'émission (voir fin du chapitre 5) ou, plus rarement, en réagissant au messages ICMP " *source quench* " (voir début du chapitre 6).

Associé à une file d'attente **RED** (voir encadrés précédents), le mécanisme **ECN** (*Explicit Congestion Notification*) permet de détecter la congestion avant qu'elle ne survienne et de l'indiquer à l'émetteur grâce à deux bits situés dans l'ancien champ TOS des paquets IP.



Bits	Description
10	ECT(0) *ECN-Capable Transport*
01	ECT(1) *ECN-Capable Transport*
00	ECT non supporté
11	CE *Congestion Experienced*

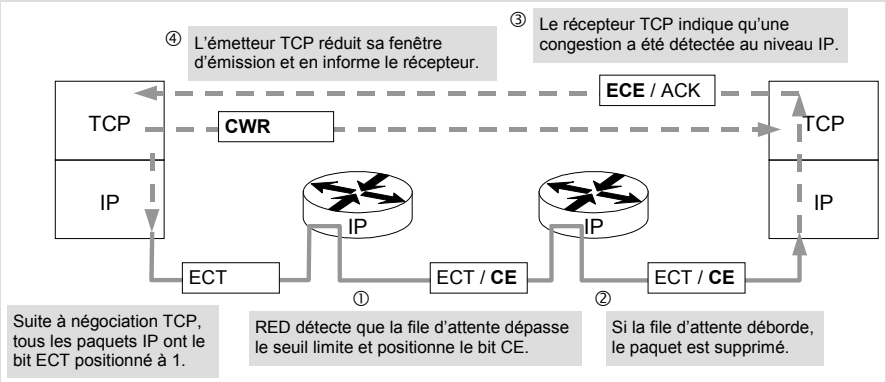
Par ailleurs, la RFC ajoute deux bits de contrôle dans l'**en-tête TCP**.

Bits de contrôle TCP

Réservé (4 bits)	ECE	CWR	URG	ACK	PSH	RST	SYN	FIN

Bits	Description
ECE	ECN-Echo : le récepteur indique à l'émetteur qu'une congestion a été détectée au niveau IP
CWR	*Congestion Window Reduced* : l'émetteur indique au récepteur qu'il a pris en compte la congestion en réduisant sa fenêtre
…	Autres bits : voir fin du chapitre 5

Lors de l'ouverture d'une session, les deux couches TCP impliquées doivent négocier le support ou non d'ECN, puis réagir lorsque le bit CE est positionné par la couche IP.



Lorsque la file d'attente RED dépasse un seuil prédéfini (seuil d'alerte), le champ ECN est positionné à ' 11 '. Si un autre seuil est dépassé (seuil de saturation), les paquets peuvent être détruits. Grâce à ECN, la couche TCP peut donc réagir dès le seuil d'alerte, c'est-à-dire avant la congestion, alors qu'auparavant, elle réagissait après la congestion, lorsque des paquets étaient perdus suite à une saturation. La RFC 2884 discute des avantages à utiliser la notification de congestion au niveau IP.

La première commande indique que le débit moyen est limité à 128 Kbit/s, mais qu'un dépassement de 32 kilo-bits est autorisé pendant une demi-seconde (64/128). Cela signifie que le débit peut atteindre 192 Kbit/s (128 + 32/0,5) pendant une demi-seconde. Cette régulation s'applique aux flux FTP comme cela est indiqué dans l'access-list 101.

La deuxième commande est appliquée à une interface Ethernet afin de limiter le débit vers ce réseau à 5 Mbit/s en autorisant une pointe à 6 Mbit/s (5 + 2/(10/5)) pendant 2 secondes (10/5).

Quelle file d'attente choisir pour son réseau ?

Chacune des fonctions qui viennent d'être décrites présente des avantages et des inconvénients. Le problème est alors de savoir dans quelle situation et à quel endroit du réseau utiliser l'une ou l'autre de ces fonctionnalités.

Nous l'avons vu précédemment, une **file d'attente FIFO** ne permet pas de gérer des situations de congestion. En conséquence, une file d'attente de ce type ne peut être utilisée qu'avec des réseaux non saturés et offrant suffisamment de débit par rapport au trafic.

L'affectation de **priorités sur les files d'attente**, qui a été l'objet de notre première démarche, consiste en une programmation arbitraire, statique, et qui ne distingue qu'individuellement les types de paquets. Lorsqu'une congestion survient, les paquets prioritaires sont traités tant qu'il y en a dans la file d'attente correspondante. L'effet de bord qui en résulte est que les autres paquets restent bloqués dans la file d'attente, ce qui peut entraîner un ralentissement conséquent, voire la coupure des sessions correspondantes à ces paquets non prioritaires. C'est néanmoins la méthode d'affectation de priorité qui utilise le moins de CPU, car l'algorithme est simple. L'utilisation de cette fonction sera donc limitée à des liens à bas débit sur lesquels le trafic est bien identifié.

La **file d'attente WFQ** identifie chaque flux, et traite tous les paquets d'un même flux de la même manière. En cas de congestion, l'algorithme traite équitablement tous les flux en privilégiant ceux à faible volume, mais pas au détriment de ceux à fort volume. Bien qu'efficient, cet algorithme complexe consomme de la ressource CPU, ce qui le destine plutôt à des liens à bas et moyens débits. Du fait qu'il prend en compte les priorités indiquées dans les paquets, on activera cet algorithme au sein du réseau et non dans sa périphérie.

La **file d'attente WRED** permet de prévenir les congestions en rejetant aléatoirement des paquets. À partir d'un certain seuil, le taux de rejet de paquets augmente à mesure que la file d'attente se remplit. Cet algorithme est de ce fait particulièrement bien adapté au protocole TCP qui est prévu pour réduire sa fenêtre d'émission en cas de perte de paquet (ce qui n'est pas le cas des protocoles IPX et AppleTalk).

Enfin, le ***traffic shaping*** et le ***policing*** sont des mécanismes qui permettent de réguler le trafic et de fixer un débit à chaque type de flux. Ils sont donc bien adaptés aux routeurs situés en bordure du réseau, c'est-à-dire sur les réseaux locaux, là où utilisateurs et serveurs émettent leurs données. Le trafic erratique peut ainsi être lissé ou écrêté lorsqu'il entre dans le réseau longue distance, ce qui permet ensuite aux routeurs situés au sein de ce réseau d'utiliser les files d'attente WFQ et WRED.

La file d'attente de type…	… doit être utilisée
FIFO	Lorsque le réseau n'est pas saturé
Priorité sur file d'attente	Sur des liens à bas débit
WFQ	Au sein du réseau, sur des liens à bas et moyen débits (< 2 Mbit/s), pour gérer les congestions
WRED	Au sein du réseau, sur des liens à haut débit, pour prévenir les congestions
Traffic shaping	En entrée ou au sein du réseau, pour lisser un flux erratique afin de limiter les risques de congestion ou pour adapter le flux entre des réseaux à débits différents
Policing	En entrée du réseau, pour adapter le flux entre des réseaux à débits différents

Gérer la qualité de service

À la base, un réseau IP comme l'Internet ne garantit pas que tous les paquets émis seront délivrés au destinataire. Il assure simplement que les paquets effectivement remis le seront sans erreur ni duplication en routant les paquets au mieux de ses possibilités (principe du *best effort*).

C'est également le principe de la poste qui s'efforce d'acheminer le courrier dans les meilleurs délais. Mais, pour s'assurer qu'un colis arrivera à destination en temps et en heure, il faut payer un service supplémentaire. De même, pour les réseaux, il faut mettre en œuvre des moyens supplémentaires afin de garantir ce niveau de service.

Alors que les flux de données classiques (Telnet, FTP, etc.) se contentent du service de base (celui du meilleur effort), le transport des flux audio et vidéo nécessite plus que cela. Il faut en effet garantir :

- qu'une application disposera du minimum de débit réseau nécessaire à son bon fonctionnement ;
- un temps de réponse ;
- que le temps de réponse variera peu dans le temps.

Afin de gérer la qualité de service nécessaire au traitement des flux multimédias, trois modèles sont actuellement proposés :

- La **préséance** (***precedence*** en anglais), historiquement la première et en voie de disparition, qui consiste à traiter le paquet IP en fonction de la priorité indiquée dans son en-tête.
- La **différenciation de service** (modèle appelé **DiffServ**) repose sur l'affectation de priorités et de classes de service dont les valeurs sont transportées dans les paquets IP. Le flux est formaté (classé) à l'entrée du réseau, puis la qualité de service est appliquée de la même manière dans tous les routeurs en fonction de la valeur indiquée dans les paquets.
- L'**intégration de service** (modèle appelé **IntServ**) consiste à réserver les ressources tout le long du chemin qu'emprunteront les paquets, puis à appliquer à tout le flot de paquets qui suivent la qualité de service demandée lors de la réservation.

La qualité de service selon la préséance

Dès l'origine, les concepteurs des protocoles TCP/IP ont pensé à intégrer la notion de qualité de service dans un champ du paquet IP, appelé TOS (*Type of Service*) . Ce champ était destiné à transporter des informations relatives à la priorité (appelée *precedence*) et à la classe de service, mais il n'a jamais été réellement utilisé, jusqu'à ce que la qualité de service soit d'actualité. La préséance IP est cependant toujours supportée par les routeurs, et les nouveaux modèles de QoS sont compatibles avec le champ « IP Precedence ».

Les algorithmes WFQ et WRED prennent en compte le champ « Precedence » (d'où le « W » pour *Weighted*). L'activation de ces files d'attente est donc particulièrement judicieuse au sein du réseau, car elles permettent de réguler le flux en fonction des priorités fixées à l'entrée de celui-ci.

L'exemple suivant, qui montre quelques valeurs par défaut, permet d'expliquer comment WRED interagit avec les priorités :

```
random-detect precedence 0 109 218 10
random-detect precedence 1 122 218 10
...
random-detect precedence 7 194 218 10
```

La première valeur (de 0 à 7) correspond à la priorité indiquée dans le champ « Precedence ». Les trois paramètres suivants indiquent quant à eux :

- le seuil minimal, en nombre de paquets dans la file d'attente, à partir duquel les paquets commenceront à être rejetés ;
- le seuil maximal, en nombre de paquets dans la file d'attente, à partir duquel les paquets seront rejetés au rythme indiqué dans le paramètre suivant ;
- la fraction de paquets rejetés lorsque le seuil maximal est atteint, par défaut un sur dix.

Afin de favoriser les paquets prioritaires, le seuil minimal est d'autant plus élevé que la priorité est grande. L'algorithme WRED calcule la taille moyenne de la file d'attente et la compare aux seuils. En cas de congestion, les paquets des files d'attentes les moins prioritaires seront rejetés en premier.

Le seuil maximal est déterminé automatiquement en fonction du débit de l'interface et de la mémoire disponible, tandis que le seuil minimal correspond à une fraction du seuil maximal dépendant de la priorité (par exemple, 1/2 pour 0, 10/18 pour 1, 11/18 pour 2, etc.). Les commandes « *sh queueing* » et « *sh int random-detect* » permettent de visualiser les paramètres actifs.

LE CHAMP TOS (RFC 791 ET 1349)

Le champ **TOS** (*Type of Service*) consiste en un unique octet :

IP precedence	TOS	0
3 bits	4 bits	

Le champ **IP Precedence** indique au routeur la manière de traiter le paquet sur la base d'une priorité allant de 0 à 7 (RFC 1812), les deux dernières étant réservées au trafic de service (routage, etc.) :

0 = Routine	2 = Immediate	4 = Flash override	6 = Internetwork control
1 = Priority	3 = Flash	5 = Critical	7 = Network control

Les bits **TOS** (la RFC utilise malencontreusement le même terme TOS pour l'octet entier ainsi que pour ces 4 bits) indiquent la classe de service souhaitée :

Bit	Description	Signification des valeurs
3	minimize Delay	(0 = Normal delay ; 1 = Low delay)
4	maximize Throughput	(0 = Normal throughput ; 1 = High throughput)
5	maximize Reliability	(0 = Normal reliability ; 1 = Reliability)
6	minimize monetary cost	(ajouté par la RFC 1349)

Ces 4 bits positionnés à 0 indiquent tout simplement le service standard (le *best effort*).

Maintenir la qualité de service

À la différence du *traffic shaping* qui lisse un flux erratique, le ***policing*** veille au respect du flux conformément à un profil prédéfini. Il peut alors **rejeter** les paquets non conformes ou les **marquer** pour une régulation ultérieure sur d'autres équipements au sein du réseau :

```
int e0
rate-limit output 5000000 32000 45000 conform-action transmit exceed-
action drop
```

Le mot clé « *output* » indique que la classification du flux est appliquée en sortie de l'interface :

- Le premier paramètre indique le débit moyen autorisé, ici 5 Mbit/s.
- Le deuxième précise la taille du burst, ici 32 Ko.
- La troisième valeur est le burst en excès, ici 45 Ko.

Si le flux est conforme à ce profil (*conform-action*), le paquet est placé dans la file d'attente tel quel. Dans le cas contraire (*exceed-action*), il est rejeté.

La fonction de *policing* peut également marquer le paquet, c'est-à-dire lui affecter une priorité en fonction du respect ou non du profil de flux. On peut augmenter la granulométrie de cette dernière en précisant le protocole (et éventuellement les réseaux, voire les adresses source et destination) sur lequel va s'appliquer le marquage :

```
rate-limit output access-group 101 5000000 24000 32000 conform-action
set-prec-transmit 1 exceed-action set-prec-transmit 0
access-list 101 permit tcp any any eq www
```

Nous avons décidé ici de limiter le flux des navigateurs Web (protocole HTTP), à 5 Mbit/s. La priorité est fixée à 1 (prioritaire) s'il se conforme à la QoS, et à 0 (*best effort*) dans le cas contraire.

Sur notre backbone WAN à 2 Mbit/s, connecté en 100bT au LAN de notre site central, nous décidons de limiter le débit à 1 Mbit/s pour les paquets de priorité 1. Pour ceux de priorité 0, il est limité à 500 Kbit/s. Le burst autorisé est de 500 Ko, et de 500 Ko en excès.

```
int e 1/0
rate-limit input access-group rate-limit 100 1000000 500000 500000 con-
form-action transmit exceed-action drop

rate-limit input access-group rate-limit 101 500000 250000 250000 con-
form-action transmit exceed-action drop

access-list rate-limit 100 1
access-list rate-limit 101 0
```

Priorité (0 à 7) écrite dans le champ *IP Precedence*.

La commande « *access-list rate-limit* » est ici utilisée pour affecter une priorité basse si le volume du flux est important, et haute dans le cas contraire. Les différents équipements au sein du réseau se chargeront d'appliquer la QoS, par exemple, au niveau de leurs files d'attente WRED ou WFQ.

À la place de l'action « drop », il est possible de marquer le paquet, par exemple, en lui baissant sa priorité :

```
rate-limit input access-group rate-limit 101 500000 250000 250000 con-
form-action transmit exceed-action set-prec-transmit 1
```

Une file d'attente WRED, appliquée cette fois à l'ensemble des flux, rejettera en priorité les paquets de plus faible priorité. Ce double mécanisme cascadé, permet au flux de dépasser le débit autorisé, s'il y a de la bande passante disponible, et dans le cas contraire, de rester dans les limites autorisées.

La qualité de service selon DiffServ

Depuis la notion de préséance définie dans la RFC 791 en 1981, les besoins ont évolué, ce qui a conduit différents groupes de travail au sein de l'IETF à modifier la signification du champ TOS. Avec les bits ECN, DiffServ achève donc de remanier ce champ qui devrait donc rapidement devenir obsolète.

Configuration des routeurs

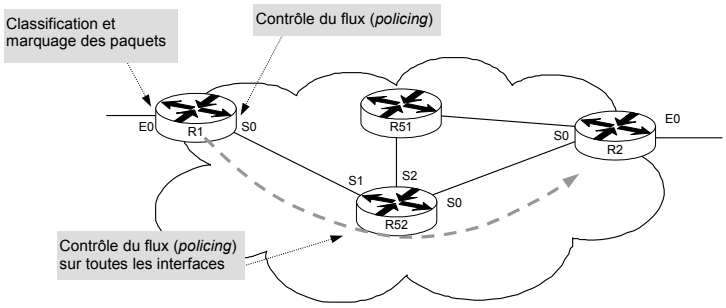
Prenons notre réseau sur lequel nous devons véhiculer des applications aux exigences très diverses.

Applications	Classe de service	Limitation	PHB / DSCP
VoIP	Premium	512 Kbit/s	*Expedited forwarding* / 46
HTTP	Gold	25 % de la bande passante	*Assured forwarding* / 18
Autres applications	Best effort	256 Kbit/s	Best effort / 0

La classe de service *Expedited forwarding*, telle que définie par DiffServ, nous semble tout indiquée pour traiter la voix sur IP, tandis que les applications devant faire l'objet d'un traitement particulier, tel que l'intranet, tombent dans la classe *Assured forwarding*. Les autres flux, a priori non identifiés, seront traités sur le mode du meilleur effort.

En respectant en cela les principes de DiffServ, les flux seront **classés** et **marqués** à l'entrée du réseau, tandis qu'au sein du réseau sera maintenue la gestion de la QoS (***policing***).

Figure 14-1.
Classificatoin, marquage et policing au sein d'un réseau



Sur les interfaces d'entrée des routeurs R1 et R2, les paquets sont donc classés et marqués…

```
int e0
service-policy input set_dscp

policy-map set_dscp
   class EF
    set ip dscp 46
   class AF21
    set ip dscp 18
   class CS0
    set ip dscp 0

class-map EF
   match ip rtp 16384 32768
class-map AF21
   match access-group 102
class-map CS0
   match access-group 103

access-list 102 permit tcp any any eq www
access-list 103 permit ip any any
```

Définit les règles de marquage. Les codes DSCP sont affectés selon les recommandations des RFC et du IANA.

Le DSCP 46 correspond au PHB *Expedited forwarding.*

Définit les règles de classification selon nos critères.

Voir chapitre 16 pour le transport de la voix sur IP.

Les flux HTTP sont arbitrairement affectés au PHB AF21.

…tandis que sur les interfaces de sortie de ces mêmes routeurs, nous devons nous assurer que les flux sont conformes à notre politique de qualité de service (*policing*).

```
int s0
service-policy output maintien_qos

policy-map maintien_qos
   class premium
    priority 512
   class gold
    bandwidth percent 25
   class best_effort
    police 256000 0 0 conform-action transmit exceed-action drop

class-map premium
   match ip dscp 46
class-map gold
   match ip dscp 18
class-map best-effort
   match ip dscp 0
```

L'algorithme WFQ traite en priorité la VoIP et lui garantit au moins 512 Kbit/s de bande passante.

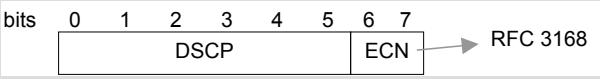
25% de la bande passante est allouée au PHB AF21.

Nous retrouvons ici nos flux déjà marqués, simplement identifiés par leur DSCP.

LE POINT SUR DIFFSERV (RFC 2474 ET 2475)

Le modèle de différentiation de service repose sur le transport, dans chaque paquet IP, de la **classe de service**. Il se propose en plus de définir la manière d'implémenter cette qualité de service.

Le champ TOS est renommé **DSCP** (*Differentiated Service Code Point*) et est structuré différemment, ce qui ne va pas sans poser des problèmes de compatibilité pour les routeurs traitant ce champ sur l'ancien mode (discussion objet de la RFC 2873).



Le champ DSCP permet de définir 64 **codepoints**, réparties en trois pools :

xx0 32 codepoints pour les actions standard ;
x11 16 codepoints pour un usage expérimental et une utilisation locale ;
x01 16 codepoints pour un usage expérimental, une utilisation locale ou une extension du premier pool.

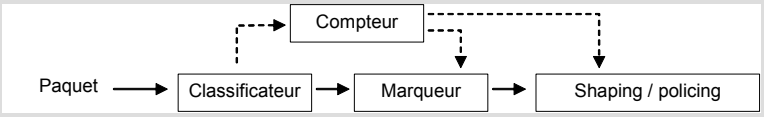
Un codepoint permet de sélectionner une classe de service de type **PHB** (*Per-Hop Behavior*) à appliquer au paquet, c'est-à-dire la manière de traiter le paquet au sein de chaque routeur. Le codepoint xxx000 (issu du pool 1) permet d'assurer la compatibilité ascendante avec le champ *IP Precedence*. En particulier, le DSCP 000000 définit le PHB par défaut, le *best effort*, tel que défini dans la RFC 1812.

Les 3 premiers bits du champ DSCP définissent ainsi la classe de service (**sélecteur de classe**), sur le modèle du champ *IP precedence*, tandis que les 3 derniers bits définissent la ***Drop Precedence***, c'est-à-dire la priorité avec laquelle les paquets seront rejetés en cas de congestion.

Cette gestion de la QoS implique la mise en œuvre des fonctions décrites ci-après.

Le **classificateur** affecte une priorité et une classe de service au paquet en fonction de règles paramétrables (adresses IP source et destination, port TCP/UDP, etc.).

Le **conditionneur** applique la qualité de service en fonction du champ DSCP. Il analyse le flux de données (fonction de comptage — ***metering***) et le compare à la QoS demandée. Il peut alors redéfinir la classe de service (fonction de marquage — ***marking***) ou réguler le trafic, soit par lissage des pointes de trafic (***traffic shaping***), soit par écrêtage des pics de trafic (***policing***).



À l'heure actuelle, trois types de PHB sont définis :

- **AF** - *Assured Forwarding* (RFC 2597), pour les flux nécessitant une bande passante limitée, le trafic en excès pouvant être rejeté progressivement selon un mécanisme de priorité à 12 niveaux (4 classes × 3 priorités de rejet).
- **EF** - *Expedited Forwarding* (RFC 3246), également appelé " Premium service " (DCSP = 101110), pour les flux requérant une bande passante garantie avec des faibles taux de perte, de gigue et de latence.
- **CS** - *Class Selector* (RFC 2474), qui assure la compatibilité avec le champ IP Precedence : CS0=routine, CS1=priority, CS2=immediate, CS3=flash, CS4=Flash override, CS5=Critical, CS6=Internetwork control, CS7=Network control.

Sur les routeurs du cœur de réseau, comme R52, seules les files d'attente WRED sont appliquées.

```
int s0 ... s1 ... S2
service-class ctrl_qos
...
policy-map ctrl_qos
   class gold
    bandwidth percent 25
    random-detect ecn
    random-detect dscp-based
    random-detect dscp 18 20 40 10
...
class-map gold
   match ip dscp 18
```

L'algorithme WRED prend en compte les priorités définies au format DiffServ et applique la notification de congestion.

Est appliqué au DSCP 18 :
- un seuil d'alerte de 20 paquets
- un seuil de rejet de 40 paquets
- le rejet d'un paquet sur 10 si congestion

Bien qu'il n'y ait pas de standardisation en la matière, la classification des flux généralement admise est la suivante :

Type de flux	Classe de service Diffserv	Code DSCP	Nom commercial
Trafic réseau (routage…)	CS7, CS6	56, 48	Critical / network
Voix	EF	46	Premium
Visio	AF41, AF42, AF43	34, 36, 38	Platinium
Signalisation voix (optionnel)	AF31, AF32, AF33	26, 28, 30	Gold
Critique	AF21, AF22, AF23	18, 20, 22	Silver
Prioritaire	AF11, AF12, AF13	10, 12, 14	Bronze
Standard (*best effort*)	CS0	0	Best effort

Les 8 classes de service sont définies par les 3 premiers bits du champ DSCP (**sélecteur de classe**) qui reprend le principe du champ « IP Precedence ». La plus prioritaire est CS7, et la moins prioritaire CS0.

Dans chaque classe de type *Assured Forwarding*, le PHB AFx1 est davantage prioritaire que AFx2 qui lui-même, l'est plus que AFx3. Cette priorité est caractérisée par la préséance de rejet (***Drop Precedence***), définies par les 3 derniers bits du champ DSCP.

Les PHB de type CS correspondent à une « Drop Precedence » égale à 0.

Le tableau, ci-après, permet de mieux visualiser la signification des bits, des pools et des codes DSCP par rapport aux PHB ainsi que les correspondances avec le champ « IP Precedence » (CS0=IPprec0=routine, CS1=IPprec1=priority, etc.).

Class selector

Drop precedence

Pool	Bits du champ DSCP						Valeur déci-male	PHB	Bits du champ DSCP						Valeur déci-male	PHB
	0	1	2	3	4	5			0	1	2	3	4	5		
1	0	0	0	0	0	0	0	CS0	1	0	0	0	0	0	32	CS4
3	0	0	0	0	0	1	1		1	0	0	0	0	1	33	
1	0	0	0	0	1	0	2		1	0	0	0	1	0	34	AF41
2	0	0	0	0	1	1	3		1	0	0	0	1	1	35	
1	0	0	0	1	0	0	4		1	0	0	1	0	0	36	AF42
3	0	0	0	1	0	1	5		1	0	0	1	0	1	37	
1	0	0	0	1	1	0	6		1	0	0	1	1	0	38	AF43
2	0	0	0	1	1	1	7		1	0	0	1	1	1	39	
1	0	0	1	0	0	0	8	CS1	1	0	1	0	0	0	40	CS5
3	0	0	1	0	0	1	9		1	0	1	0	0	1	41	
1	0	0	1	0	1	0	10	AF11	1	0	1	0	1	0	42	
2	0	0	1	0	1	1	11		1	0	1	0	1	1	43	
1	0	0	1	1	0	0	12	AF12	1	0	1	1	0	0	44	
3	0	0	1	1	0	1	13		1	0	1	1	0	1	45	
1	0	0	1	1	1	0	14	AF13	1	0	1	1	1	0	46	EF
2	0	0	1	1	1	1	15		1	0	1	1	1	1	47	
1	0	1	0	0	0	0	16	CS2	1	1	0	0	0	0	48	CS6
3	0	1	0	0	0	1	17		1	1	0	0	0	1	49	
1	0	1	0	0	1	0	18	AF21	1	1	0	0	1	0	50	
2	0	1	0	0	1	1	19		1	1	0	0	1	1	51	
1	0	1	0	1	0	0	20	AF22	1	1	0	1	0	0	52	
3	0	1	0	1	0	1	21		1	1	0	1	0	1	53	
1	0	1	0	1	1	0	22	AF23	1	1	0	1	1	0	54	
2	0	1	0	1	1	1	23		1	1	0	1	1	1	55	
1	0	1	1	0	0	0	24	CS3	1	1	1	0	0	0	56	CS7
3	0	1	1	0	0	1	25		1	1	1	0	0	1	57	
1	0	1	1	0	1	0	26	AF31	1	1	1	0	1	0	58	
2	0	1	1	0	1	1	27		1	1	1	0	1	1	59	
1	0	1	1	1	0	0	28	AF32	1	1	1	1	0	0	60	
3	0	1	1	1	0	1	29		1	1	1	1	0	1	61	
1	0	1	1	1	1	0	30	AF33	1	1	1	1	1	0	62	
2	0	1	1	1	1	1	31		1	1	1	1	1	1	63	

Les 3 bits du champ IP Precedence

Configuration des commutateurs de niveau 2

En fait, il est bien plus judicieux de marquer les paquets à la source, c'est-à-dire sur les équipements directement en contact avec les PC. Il s'agit bien sûr des commutateurs Ethernet 10/100bT.

Les postes de travail sont, en effet, généralement connectés à des commutateurs de niveau 2, alors que l'on réserve la commutation de niveau 3 pour le réseau fédérateur en raison de son coût. Cependant, ils gèrent rarement les priorités. De ce fait, les commutateurs transforment les trames Ethernet en trames 802.1q pour gérer la qualité de service et les VLAN. Lorsqu'elles sont envoyées vers un port de sortie, ces trames sont reformatées en trames Ethernet 802.3.

Étant donné que toutes les trames ne sont pas obligatoirement de type 802.1q, par défaut le commutateur ne prend donc pas en compte le champ CoS (*Class of Service*), et la priorité doit donc être fixée au niveau du port.

Par défaut, le port est en mode *untrusted* : la CoS est fixée à la valeur par défaut du port, c'est-à-dire à 0. Il est possible de changer la valeur par défaut associée à chaque port :

```
set qos enable
set port qos 1/4 trust untrusted
set port qos 1/4 cos 6
```

Valeur par défaut du champ COS = 0

Fixe la valeur par défaut du champ COS

Toutes les trames arrivant par le port 1/4 auront un champ COS fixé d'office à 0, à moins qu'une autre valeur par défaut ne soit affectée au port.

Il n'est ainsi pas nécessaire de gérer la qualité de service au niveau du poste de travail et des serveurs. Chacun d'eux étant raccordé à un port du commutateur, celui-ci se charge de marquer les paquets. L'inconvénient est que ce marquage ne distingue pas le type de flux.

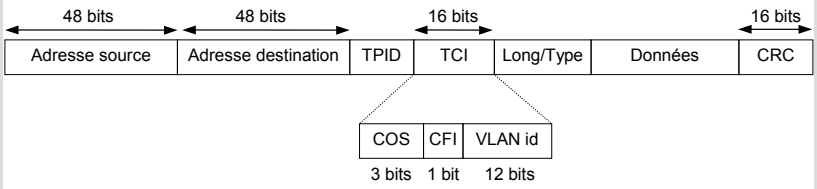
Si la priorité est déjà fixée par un équipement terminal ou un autre commutateur (donc générant une trame 802.1q), il est possible de prendre en compte le champ COS :

```
set qos enable
set port qos 1/3 trust trust-cos
```

Le champ COS des trames 802.1q arrivant par le port 1/3 sera ainsi pris en compte.

LA QUALITÉ DE SERVICE SUR ETHERNET (IEEE 802.1P)

La norme initiale 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 des **classes de service** (802.1p) .



Cette trame n'est généralement véhiculée qu'entre les commutateurs, bien que les PC soient de plus en plus souvent capables de la gérer. Si cela n'est pas le cas, les commutateurs enlèvent le champ 802.1q lorsqu'ils transmettent la trame à un équipement terminal (PC, serveur, etc.).

Sept classes de service (**CoS**) sont ainsi définies :

0 = Best effort	4 = Applications à contrôle de charge (streaming audio/vidéo)
1 = Tâche de fond (batch)	5 = Vidéo
2 = Réservé	6 = Voix
3 = Excellent effort	7 = Network control

Certains pilotes de carte réseau permettent de gérer la priorité directement au niveau du PC.

Configuration des commutateurs de niveau 3

Équipé d'une carte de commutation de niveau 3, le commutateur peut assurer toutes les fonctions DiffServ : classification, marquage et *policing*.

Par défaut, une politique de qualité de service est associée à un port, mais il est possible de l'associer à un VLAN :

```
set qos enable
set port qos 1/1 port-based
set port qos 1/2 vlan-based
```

La configuration de nos commutateurs consiste à définir une **politique de qualité de service** contenant une règle de **marquage**, une règle de ***policing*** et une règle de **classification** (l'ordre de définition ne respecte pas celui du séquencement des opérations).

Définir une règle de marquage

Le commutateur associe aux ports une **politique par défaut** dont il est possible de modifier la règle de marquage :

```
set qos acl default-action ip dscp 0
```

Marquage

Cette politique ne contient qu'une règle de marquage, qui consiste à marquer le champ DSCP (*Differentiated Service Code Point*) de tous les paquets avec la valeur indiquée (0 par défaut).

D'autres marquages sont possibles en remplaçant le mot clé *dscp* et sa valeur :

- *trust-dscp* : le champ DSCP n'est pas modifié et est accepté tel quel.
- *trust-ipprec* : le champ DSCP est modifié en prenant la valeur du champ « IP precedence ».
- *trust-cos* : le champ DSCP est modifié en fonction de la valeur contenue dans le champ COS de la trame Ethernet 802.1q.

Définir une règle de policing

Il est également possible d'associer plusieurs règles de *policing* à la politique par défaut :

```
set qos acl default-action ip dscp 0
microflow micro_de_base aggregate aggr_de_base
```

Marquage

Policing

La création d'une règle de *policing* s'effectue comme suit :

```
set qos policer microflow micro_de_base rate 1000 burst 1000 drop

set qos policer aggregate aggr_de_base rate 1000 burst 1000 policed-dscp
```

Ces deux règles limitent le débit moyen à 1 Mbit/s et le dépassement (*burst*) à 1 Mo.

Le mot-clé *microflow* précise que la règle s'applique à chaque flux considéré individuellement, tandis que le mot-clé *aggregate* précise qu'elle s'applique à l'ensemble des flux.

Vient ensuite le type de traitement à appliquer au paquet lorsque le flux ne correspond pas au profil : le rejeter (*drop*) ou le marquer à nouveau (*policed-dscp*).

Dans ce dernier cas, la commande suivante permet de définir les correspondances, soit par plage de valeurs, soit individuellement :

```
set qos policed-dscp-map 63-62:62 61-60:60 ... 1-0:0
set qos policed-dscp-map 41:40
```

Nouvelle valeur

Anciennes valeurs

Les valeurs montrées ici sont celles par défaut qui sont conservées si elles ne sont pas explicitement redéfinies.

Définir une règle de classification

La politique par défaut qui vient d'être décrite ne contient pas de règle de classification puisqu'elle s'applique à tous les paquets IP. En revanche, nous pouvons créer une règle s'appliquant à un certain type de trafic :

```
set qos acl ip politique_ip_de_base trust-cos
microflow micro_de_base aggregate aggr_de_base
10.0.0.0 255.0.0.0 precedence routine
```

Marquage
Policing
Classification

La politique ainsi définie s'applique aux paquets IP provenant du réseau 10.0.0.0 et ayant leur champ « Precedence » à 0 (mot-clé *routine*). Les paquets se conformant aux profils « micro_de_base » et « aggr_de_base » seront affectés de la priorité dérivée du champ COS de la trame 802.1q (mot clé *trust-cos*).

Les valeurs qui suivent le mot-clé *precedence* sont celles décrites dans le champ TOS : *routine*, *priority*, *immediate*, *flash*, *flash-override*, *critical*, *internet* et *network*.

À la place de l'adresse IP et de son masque, on peut mettre le mot-clé *any* ou le mot clé *host* suivi d'une adresse IP.

Nos commutateurs nous offrent la possibilité de décrire d'autres types de politiques s'appliquant aux trafics TCP, UDP, ICMP, IGMP, etc. Par exemple, il est ainsi possible d'affecter une règle à chaque application (qui se distingue par son port TCP ou UDP) :

```
set qos acl ip politique_web trust-ipprec
microflow micro_de_base
any tcp www
```

Marquage
Policing
Classification

Associer une politique à un port

La dernière étape consiste à associer la politique que nous venons de créer à un port du commutateur ou à un VLAN :

```
set qos acl map politique_ip_de_base 1/1
set qos acl map politique_web VLAN 1
```

Les ports ou le trafic associés au VLAN 1 se verront appliquer la politique Web que nous venons de définir.

Affecter des valeurs au champ DSCP

Par défaut, tous les ports fonctionnent sur le mode *untrusted*, ce qui signifie que les priorités (COS et DSCP) des paquets y entrant sont ignorées.

La carte de commutation de niveau 3 peut néanmoins prendre en compte les priorités des trames et paquets pour un port auquel est connecté une machine capable de générer des trames 802.1q et/ou des paquets IP DiffServ.

La manière de prendre en compte ces valeurs peut être définie soit au niveau de la règle de marquage de la politique qui est ensuite affectée au port comme nous l'avons déjà fait, soit être définie directement par port :

```
# Soit définie au niveau de la politique
set qos acl ip politique_ip_de_base trust-cos ...
set qos acl map politique_ip_de_base 1/1

# Soit définie directement au niveau du port
set qos acl ip politique_par_port dscp 0 ...
set port qos 1/4 trust untrusted
set port qos 1/3 trust trust-cos
set port qos 1/2 trust trust-ipprec
set port qos 1/1 trust trust-dscp
```

Valeur par défaut si le port est marqué *untrusted*

La signification des mots-clés est la suivante :

- *untrusted* : le champ DSCP est fixé à 0 ou par la valeur indiquée dans la politique de qualité de service.
- *trust-cos* : le champ DSCP est fixé à la valeur du champ COS de la trame Ethernet.
- *trust-ipprec* : le champ DSCP est fixé à la valeur du champ « Precedence » du paquet IP.
- *trust-dscp* : le champ DSCP du paquet est conservé tel quel.

Les correspondances COS (0 à 7) → valeur DSCP (0 à 63) et IP Precedence (0 à 7) → DSCP par défaut conviennent, mais il est possible de les modifier comme suit :

```
set qos cos-dscp-map 0 8 16 24 32 40 48 56
set qos ipprec-dscp-map 0 8 16 24 32 40 48 56
```

Chaque valeur correspond à un DSCP compris entre 0 et 63. Les valeurs indiquées dans cet exemple sont celles par défaut.

Lorsque le paquet est finalement envoyé sur le port de sortie, le champ COS de la trame est dérivé de la valeur DSCP (conservée ou marquée à nouveau) :

```
set qos dscp-cos-map 0-7:0 8-15:1 ... 48-55:6 56-63:7
```

Là encore, les valeurs indiquées sont celles par défaut.

Configuration des postes de travail

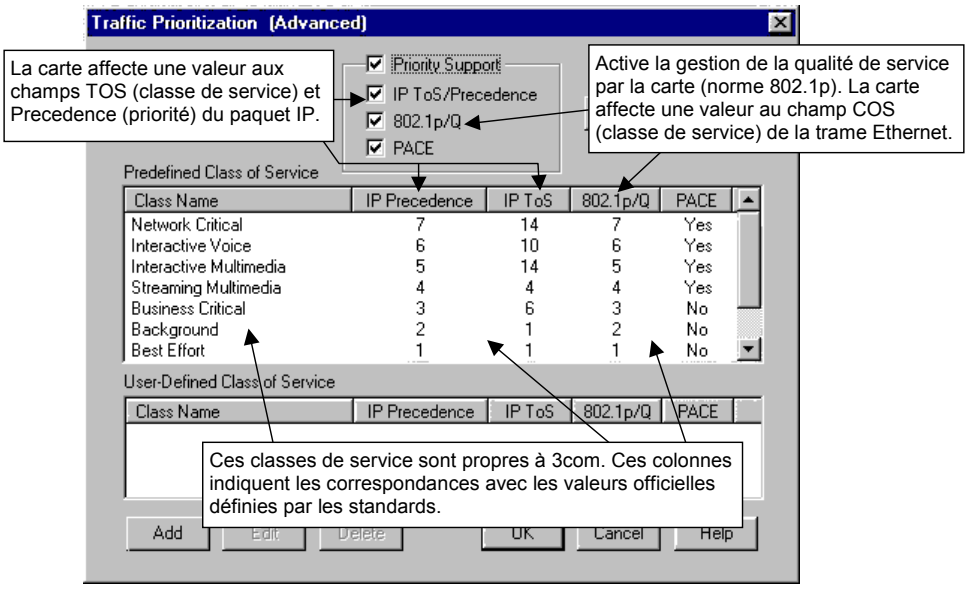
Nous venons de le voir, il est possible de gérer la qualité de service au niveau du WAN et du LAN. Cette gestion s'appuie sur des valeurs positionnées par les routeurs et les commutateurs de niveau 3 dans les paquets IP (champs TOS et Precedence), ainsi que par les commutateurs de niveau 2 dans les trames Ethernet (champ COS).

La question se pose alors de savoir où positionner ces paramètres. On peut le faire :

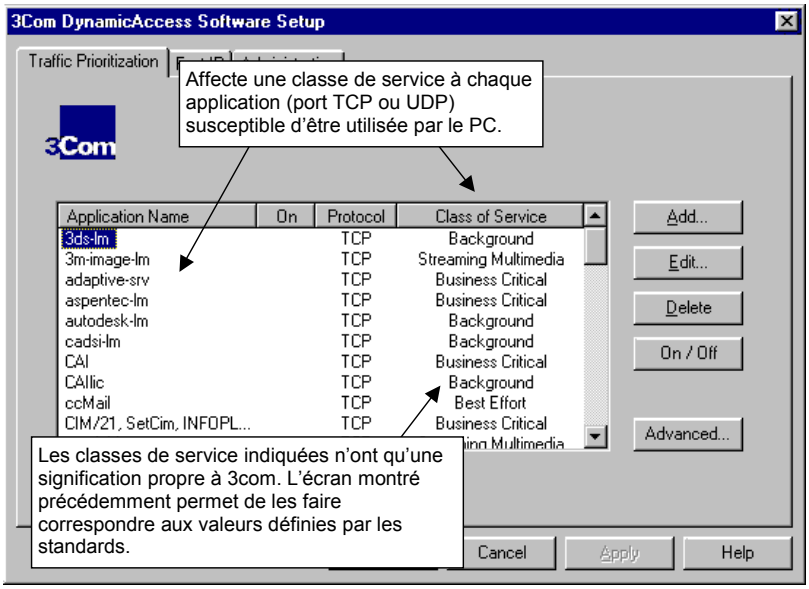
- au niveau des commutateurs de niveau 2 afin de gérer la qualité de service sur le LAN uniquement ;
- au niveau des commutateurs de niveaux 2 et 3 afin de gérer la qualité de service sur le LAN et le WAN ;
- au niveau des routeurs pour gérer la qualité de service sur le WAN.

Cette dernière solution est préférable si les sites ne sont pas tous équipés de cartes de commutation de niveau 3. Par ailleurs, la bande passante est généralement suffisante sur le LAN pour permettre de se passer d'une gestion élaborée de la qualité de service.

Une dernière solution est proposée par les constructeurs : l'affectation des priorités au niveau des cartes Ethernet des postes de travail et des serveurs, c'est-à-dire à la source d'émission des trames et paquets.



Il faut donc affecter une classe de service à chaque application que sont susceptibles d'utiliser les postes de travail et les serveurs. Et il faut, bien évidemment, que toutes les cartes soient configurées de manière identique.



Cette solution est cependant difficile à mettre en œuvre pour plusieurs raisons :

- Tous les PC et serveurs doivent être équipés de cartes supportant cette fonctionnalité.
- Toutes les cartes doivent, de préférence, provenir du même constructeur, afin de simplifier l'exploitation et garantir l'homogénéité des paramètres.
- Toutes les configurations doivent être identiques.
- Le paramétrage du driver doit pouvoir être réalisé à distance.
- Le nombre de nœuds à configurer est considérablement plus important que le nombre de commutateurs sur lesquels ce même paramétrage peut être défini.

En outre, si les utilisateurs disposent des droits administrateurs de leur poste de travail, ils peuvent changer les paramètres de leur carte réseau, ce qui fausse la politique de qualité de service que vous vouliez mettre en place. Toutes ces contraintes rendent donc difficile la gestion de la qualité de service au niveau des postes de travail.

La pratique la plus courante consiste en fait à classer et à marquer les paquets au niveau 3, par les commutateurs fédérateurs, voire les commutateurs d'étage s'ils sont de niveau 3. Si le marquage est réalisé au niveau 2, les commutateurs de niveau 3 ou routeurs doivent marquer les paquets IP en faisant correspondre le champ COS avec les 3 premiers bits du champ DSCP et en ajoutant une précédence de rejet selon une classification plus précise.

La qualité de service selon IntServ

À la différence de DiffServ qui repose sur le transport de la qualité de service dans les paquets IP, le modèle IntServ distribue ces informations (débit, temps de réponse, etc.) à tous les routeurs concernés par le flux, afin d'assurer la même qualité de service de bout en bout, c'est-à-dire entre les applications émettrices et destinataires. Cette cohérence nécessite deux fonctions distinctes :

- la définition des spécificateurs (paramètres) qui permettent de caractériser la qualité de service ;
- le protocole de signalisation permettant aux routeurs d'échanger les spécificateurs ainsi que des informations relatives à l'état du réseau.

La première fonction fait l'objet des RFC 2211 à 2216 qui définissent les éléments d'un réseau à intégration de service. La seconde est couverte par les RFC 2205 à 2209 qui définissent le protocole RSVP (*Resource reSerVation Protocol*).

La RFC 2210 établit la liaison entre les deux fonctions : elle traite du fonctionnement de RSVP au sein d'un réseau à intégration de service.

RFC	Sujet traité
2211	Classe de service par contrôle de charge
2212	Classe de service par garantie de service
2213	MIB de l'intégration de service
2214	MIB de l'intégration de service — Extension à la garantie de service
2215	Définition des paramètres permettant de calculer une QoS
2216	Définitions de la QoS pour les éléments du réseau (définit un cadre général sur lequel s'appuient les RFC 2211 et 2212)
2210	Utilisation de RSVP dans un réseau à intégration de service
2205	Spécifications fonctionnelles de RSVP
2206	MIB de RSVP
2207	Extensions pour les flux IPSEC
2208	Préconisations pour la mise en œuvre de RSVP
2009	Règles de traitement des messages

La réservation des ressources

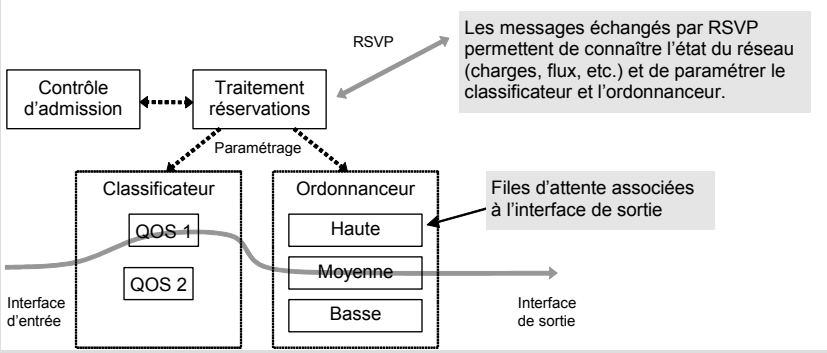
La signalisation RSVP est le protocole qui permet de négocier et de mettre en place une qualité de service dans des routeurs. Dans la pile TCP/IP, il est situé au même niveau que ICMP et IGMP (paquet IP de type 46, adresse unicast ou multicast identique à celle du flux). Il est cependant possible de l'encapsuler dans un paquet UDP (ports 1698 et 1699, adresse 224.0.0.14).

LE POINT SUR INTSERV (RFC 1633)

Il existe différentes façons d'assurer une qualité de service : affectation des priorités, choix de l'algorithme pour gérer les files d'attente, etc. Mais, si chaque routeur gère un niveau de priorité qui lui est propre ou gère des files d'attente selon des paramètres différents, la qualité de service demandée par une application sera interprétée différemment par les routeurs et ne pourra donc pas être assurée de bout en bout.

Pour que tout le monde parle le même langage, la RFC 1633 définit un cadre fonctionnel, appelé **intégration de service**. Quatre nouveaux composants doivent être intégrés aux routeurs :

- Le **contrôle d'admission** dont le rôle est de vérifier que la QoS demandée pour un nouveau flux ne remettra pas en cause les QoS accordées aux autres flux existants.
- Le **classificateur** de paquets dont le rôle est d'affecter une classe de service aux paquets entrants.
- L'**ordonnanceur** de paquets, qui gère les files d'attente pour les paquets sortants.
- Un mécanisme de **réservation** des ressources permettant de paramétrer le classificateur et l'ordonnanceur.



Une demande de réservation est traitée par le module de contrôle d'admission qui vérifie que l'application a le droit d'effectuer une telle requête et que le réseau peut offrir la QoS demandée sans remettre en cause celles déjà accordées.

L'implémentation logicielle de l'ordonnanceur dépend du routeur. La plupart d'entre eux utilisent néanmoins un algorithme standard de type WFQ (*Weighted Fair Queue*) ou WRED (*Random Early Detection*) pour la gestion des files d'attente.

De même, le classificateur peut utiliser différents éléments pour classer un paquet dans tel ou tel niveau de service : l'adresse IP et le port TCP/UDP sont des informations de base qui peuvent être utilisées.

Enfin, le mécanisme de réservation repose sur le protocole de signalisation RSVP (*Resource reSerVation Protocol*) qui permet de demander à tous les routeurs concernés de réserver des ressources.

Les paquets RSVP sont routés saut par saut, c'est-à-dire de routeur à routeur : à chaque routeur traversé, les paquets remontent donc à la couche RSVP, puis sont réémis. Le routage des paquets est assuré par un protocole de routage quelconque, unicast ou multicast selon le type d'adresse.

Le protocole doit être activé sur chaque interface :

```
int s 0
ip rsvp bandwidth
```

Par défaut, 75% de la bande passante est gérée par RSVP

Un émetteur de flux multimédia envoie périodiquement une annonce de chemin à un destinataire unicast ou à un groupe multicast (message *Path*). Chaque routeur, situé entre l'émetteur et le (ou les) destinataire(s), garde une trace de cette annonce avec les interfaces d'entrée et de sortie du paquet.

Chaque destinataire peut demander à réserver des ressources correspondant à une qualité de service telle que définie dans la RFC 2210 (message *Resv*).

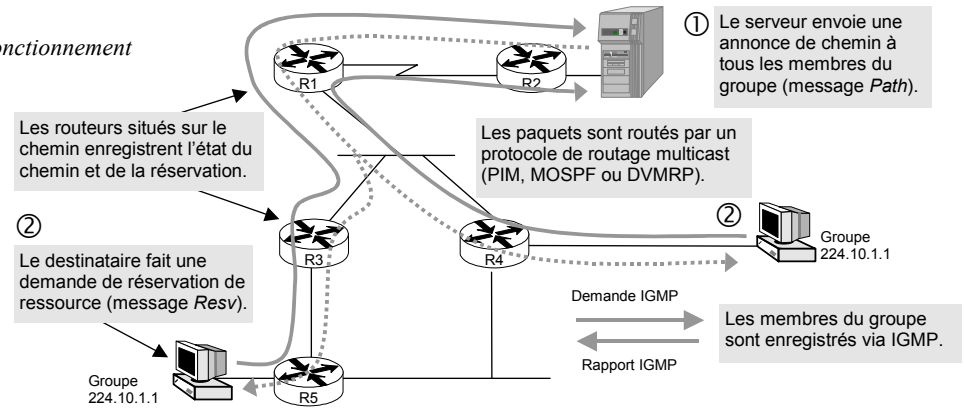
Le logiciel RSVP, qui reçoit une demande de réservation, procède aux traitements tels que spécifiés par la RFC 1633 :

- contrôle d'admission afin de garantir que la demande ne remettra pas en cause la QoS d'autres flux (l'application a-t-elle le droit de faire une telle demande ? Est-elle raisonnable ? etc.) ;
- affectation du flux à une classe de service permettant d'obtenir la QoS désirée.

Si ces contrôles sont positifs, le message *Resv* est envoyé à l'émetteur en suivant la même route que le message *Path* émis. Chaque routeur situé entre le destinataire et l'émetteur du flux garde une trace de cette demande, et procède au même contrôle que décrit précédemment. Chaque routeur peut donc accepter ou refuser la demande (voir figure ci-après).

RSVP ne définit que le moyen d'établir une qualité de service entre un émetteur et un destinataire en garantissant que les routeurs réserveront suffisamment de ressources (mémoire, CPU, files d'attente, bande passante sur les liaisons, etc.) pour assurer cette qualité de service.

Figure 14-2.
Principe de fonctionnement de RSVP.



Sur nos routeurs, il est possible de définir quelle bande passante va pouvoir être gérée par RSVP :

```
int s 0
ip rsvp bandwidth 200 20
fair-queue 64 256 10
```

200 Kbit/s gérée par RSVP

20 Kbit/s par flux

10 files d'attente pour RSVP

L'algorithme WFQ a également été activé pour gérer la qualité de service demandée. Dix files d'attente sont ainsi dédiés à RSVP : c'est, en effet, le nombre de flux auquel nous nous attendons (10 × 20 Kbit/s = 200 Kbit/s). Les valeurs par défaut, 64 et 256, ont été conservées pour les autres files d'attente.

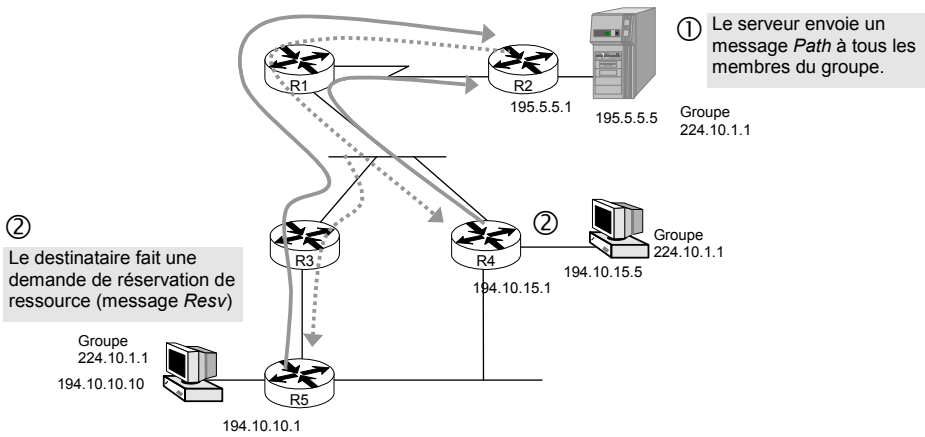
Ce qui vient d'être expliqué implique deux choses :

- que les piles TCP/IP des PC et des serveurs supportent le protocole RSVP ;
- et que les applications sachent décrire et demander une QoS.

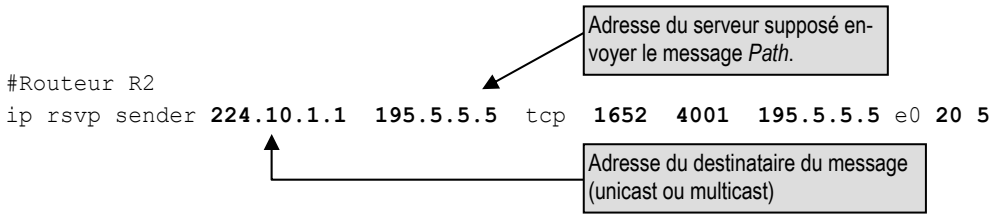
Le premier point est résolu avec Windows 98, NT, 2000 et XP. Le second est, quant à lui, résolu par les API RSVP disponibles dans Winsock 2.0. et sous la plupart des versions d'Unix. Ainsi, certaines applications, telles qu'Internet Explorer ou Netmeeting, prennent directement en charge ce protocole pour les flux audio et vidéo.

Mais peu d'applications prennent actuellement en charge RSVP et les spécificateurs de QoS. Heureusement, nos routeurs offrent une fonctionnalité intéressante qui consiste à simuler l'émission des messages *Path* et *Resv* par un serveur ou un PC.

Dans notre cas, qui sert d'exemple, un serveur doit diffuser une vidéoconférence à différents utilisateurs *via* le protocole Proshare d'Intel. L'architecture est la suivante :



La commande suivante permet de simuler la réception de messages *Path par notre serveur vidéo* :



Le routeur agit comme si une application sur le serveur d'adresse IP 195.5.5.5 lui envoyait un message *Path* pour l'adresse de groupe multicast 224.10.1.1.

Les paramètres qui suivent sont :

- les ports destination (tcp 1652 = vidéoconférence Proshare) et source (4001) – les mots-clés « udp » ou un port IP quelconque sont également acceptés ;
- l'adresse du saut précédent, donc celle du serveur ou du routeur le plus proche du serveur ;
- l'interface physique par laquelle est censé transiter le message émis par le serveur ;
- la bande passante, en Kbit/s, à réserver (ici 20 Kbit/s) ;
- le dépassement autorisé (burst), en Ko (ici 5 Ko).

En retour, il est nécessaire de simuler l'envoi d'un message *Resv* par le destinataire du flux :

```
#Routeur R5
ip rsvp reservation 224.10.1.1 195.5.5.5 tcp 1652 4001 194.10.10.10 e0
SE rate 20 5
```

Le routeur agit comme si une application sur le PC envoyait un message *Resv* pour l'adresse de groupe multicast 224.10.1.1 et vers le serveur 195.5.5.5.

Les paramètres qui suivent sont :

- les ports destination (1652 = vidéoconférence Proshare) et source – le mot clé « udp » ou un port IP quelconque sont également acceptés ;
- l'adresse du saut précédent, donc celle du PC ou du routeur le plus proche du PC ;
- l'interface physique par laquelle est censé transiter le message émis par le serveur ;
- le style de réservation : FF (Fixed-Filter), WF (Wildcard-Filter) ou SE (Shared-Explicit) que nous avons retenu pour notre flux vidéo ;
- la classe de service demandée (voir paragraphe suivant) : « rate » (QoS garantie) ou « load » (contrôle de charge) ;
- la bande passante, en Kbit/s, à réserver (ici 20 Kbit/s) ;
- le dépassement autorisé (burst), en Ko (ici 5 Ko).

Nous devons également nous assurer que les flux Telnet disposeront toujours de la bande passante suffisante et des temps de réponse corrects :

```
#Routeur R2
Message Path     Émetteur  |Récepteur     |Flux Telnet |Émetteur
ip rsvp sender 195.5.5.5 194.10.10.10 tcp 23 1025 195.5.5.1 e0

#Routeur R5
Message Resv          Émetteur  |Récepteur     |Flux Telnet|Récepteur
ip rsvp reservation 195.5.5.5 194.10.10.10 tcp 23 1025 194.10.10.1 e0 FF
load 32 1
```

Toutes les interfaces de sortie (par rapport au flux du serveur vers le PC) de tous les routeurs situés sur le chemin emprunté par le flux vidéo doivent pouvoir gérer la qualité de service demandée, ce qui est réalisé *à travers* la file d'attente WFQ. Si l'on ne considère que notre flux vidéo, cela donne :

```
# Routeur
int e1        (int e0 pour R2, in s0 pour R1, int e2 pour R4)
ip rsvp bandwidth 200 20
fair-queue 64 256 10
```

200 Kbits/s gérés par RSVP
et 20 Kbits/s affectés par flux

Pour les routeurs situés au sein de notre réseau qui utiliseraient l'algorithme WRED, il est possible de modifier les seuils de rejet appliqués par défaut à RSVP :

```
random-detect precedence rsvp 205 218 10
```

Les paramètres montrés sont ceux par défaut ; leur signification est identique à celle donnée en page 352.

LE POINT SUR RSVP (RFC 2205 À 2210)

L'émetteur d'un flux (vidéo, par exemple) envoie régulièrement aux destinataires un message **Path** leur permettant, ainsi qu'aux routeurs situés sur le chemin emprunté par le paquet RSVP, de déterminer la route empruntée par les paquets IP du flux. Les destinataires renvoient régulièrement à l'émetteur un message **Resv**, qui est pris en compte par tous les routeurs situés sur le chemin qui vient d'être déterminé. Cette méthode permet de s'assurer que les ressources seront réservées sur le chemin emprunté par les paquets du flux multimédia. Au sein d'un réseau IP, le chemin retour peut, en effet, être différent du chemin aller.

Une demande de réservation consiste en un descripteur de flux composé d'un spécificateur de flux, **flowspec** (c'est-à-dire la QoS désirée qui permet de paramétrer l'ordonnanceur de paquets) et d'un filtre de flux, **filterspec** (c'est-à-dire les caractéristiques du trafic qui permettent de paramétrer le classificateur de paquet).

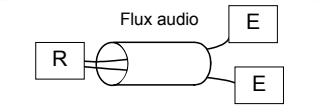
Un filtre consiste, au minimum, en une adresse IP et en un numéro de port TCP/UDP. Une forme plus évoluée peut inclure des données issues des couches applicatives ou le champ TOS du paquet IP.

• • •

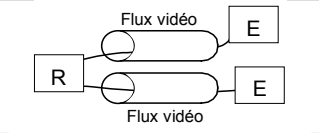
LE POINT SUR RSVP (SUITE)

Il existe plusieurs **styles de réservation** selon que la demande est partagée ou non par plusieurs flux et qu'elle concerne explicitement ou non les émetteurs :

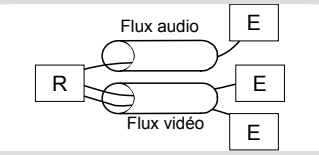
WF (*Wildcard-Filter*) : une seule réservation est faite pour les flux de tous les émetteurs (par exemple, pour une audioconférence où le nombre de personnes parlant en même temps est limité).



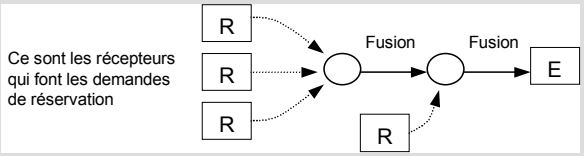
FF (*Fixed-Filter*) : une réservation est faite par flux (par exemple, une réservation par flux vidéo émis par chaque participant d'une visioconférence).



SE (*Shared Explicit*) : une seule réservation est faite pour les flux de quelques émetteurs (elle est adaptée, par exemple, à une visioconférence associée à une conférence audio).



Un routeur peut recevoir des demandes de réservation en provenance de différents destinataires qui concernent un même flux émis par un même émetteur. Les QoS demandées seront alors fusionnées en une seule qui sera remontée, et ainsi de suite jusqu'à l'émetteur du flux. La réservation qui est remontée jusqu'à l'émetteur du flux dépend alors du style des réservations émises par les destinataires.



Un message RSVP comprend plusieurs objets, eux-mêmes structurés en plusieurs champs. Par exemple, un message *Resv* comprend les adresses IP de l'émetteur et du destinataire du message RSVP, la périodicité d'envoi du message et le style de réservation, suivi des descripteurs de flux.

Afin de tenir compte des changements de topologie (apparition et disparition des routeurs et des membres de groupe), les messages *Path* et *Resv* sont envoyés périodiquement.

4 bits	4 bits	8 bits	16 bits	
Version	Options	Type de message	Checksum	
TTL du paquet IP		Réservé	Longueur du message	
Longueur de l'objet 1			Classe de l'objet	Type de l'objet
Contenu de l'objet (valeurs)				
Longueur de l'objet 2			Classe de l'objet	Type de l'objet
….				

La description de la qualité de service

RSVP est un protocole qui permet aux applications et aux routeurs de mettre en place une qualité de service en échangeant des informations. Mais, comment une application peut-elle décrire la qualité de service dont elle a besoin ? Quels paramètres doit-elle fournir aux routeurs ? Comment les routeurs configurent-ils leur classificateur et leur ordonnanceur ?

La réponse à ces questions passe par le respect d'un langage commun, qui est décrit dans les RFC 2210 à 2216. Ce langage commun repose sur trois types de paramètres qui permettent de caractériser la qualité de service :

- la classe de service demandée ;
- les paramètres décrivant la classe de service ;
- les paramètres décrivant les caractéristiques du flux de données pour lequel cette demande de qualité de service est faite.

Ces paramètres sont regroupés dans des objets transportés dans les messages *Resv*.

Les classes de service

Le support d'une classe de service implique que les routeurs paramétreront leur classificateur et leur ordonnanceur de manière identique et que, en définitive, ils se comporteront de la même manière, afin d'assurer une qualité de service uniforme, celle qu'attend l'application.

Deux classes de service sont actuellement définies :

- Un service de **contrôle de charge** (RFC 2211). En prenant en charge ce type de service, le routeur garantit que la plupart des paquets seront routés sans erreur et que le délai de transit n'excédera pas un certain seuil pour la plupart des paquets routés.
- Un **service garanti** (RFC 2212). En prenant en charge ce type de service, le routeur garantit que le délai de transit des paquets ne dépassera pas un seuil fixé et que l'application disposera de la bande passante demandée.

Description des classes de service

La RFC 2215 définit, quant à elle, les paramètres utilisés pour calculer les classes de service :

- *non-is_hop* : information décrivant qu'un routeur ne gère aucune QoS (routeurs sans intégration de service) ;
- *number_of_is_hops* : nombre de sauts à intégration de service (c'est-à-dire de routeurs gérant une QoS) ;
- *available_path_bandwidth* : bande passante disponible le long d'une route (qui traverse plusieurs routeurs et plusieurs réseaux) ;
- *minimum_path_latency* : délai minimal de transit d'un paquet le long d'une route (dépend notamment de la bande passante et du délai de traitement des routeurs). ;
- *path_mtu* : taille maximale des paquets le long d'une route : c'est la plus grande MTU (*Maximum Transmission Unit – 1 500 octets pour Ethernet*) autorisée par les différents types de réseaux empruntés par un flot de données.

Caractéristiques des flux

La même RFC 2215 définit également les paramètres que l'application communique aux routeurs *via* RSVP (regroupés dans une structure appelée *token_bucket_tspec*). Elle décrit les caractéristiques du flux pour lequel l'application fait une demande de classe de service. Il s'agit :

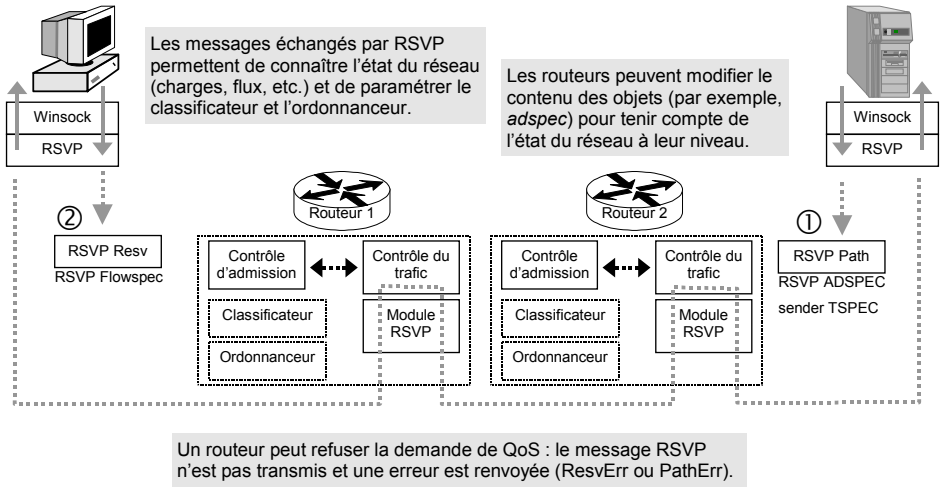
- du débit moyen du flux (*token rate*), en octets par seconde ;
- du dépassement de débit (*bucket size*) autorisé pendant une courte période, en d'autres termes le *burst*, exprimé en nombre d'octets ;
- du débit en pointe (*peak rate*), en octets par seconde ;
- de l'unité de contrôle minimale, en octets, qui correspond à la taille minimale des paquets. Le routeur contrôle en permanence le débit, et l'ajuste par rapport au *token-bucket* (*rate* + *size*) et à la taille des paquets. Les paquets dont la taille est inférieure à cette valeur seront traités, dans le calcul de la QoS, comme ayant cette valeur ;
- de la taille maximale des paquets en nombre d'octets.

Objets RSVP

Ces paramètres sont regroupés dans des objets transportés dans les messages *Resv* de RSVP. Ces objets sont décrits dans la RFC 2210 :

- *flowspec* : est généré par le récepteur vers le ou les émetteurs d'un flux afin d'indiquer la QoS désirée. Les routeurs peuvent modifier les informations *flowspec* en cours de chemin, notamment lors de la fusion RSVP.
- *sender_tspec* : indique les caractéristiques des flux que le ou les émetteurs vont générer. Les routeurs intermédiaires prennent en compte cette information sans la modifier.
- *rsvp_adspec* : est généré par le ou les émetteurs d'un flux, mais modifié par les routeurs intermédiaires. Cet objet contient des informations collectées dans le réseau, relatives aux délais de transit dans chaque nœud, à la bande passante mesurée sur les liaisons, à la classe de service supportée, etc. Chaque routeur modifie ou ajoute ses propres informations dans ce paquet.

Figure 14-3.
RSVP avec l'intégration de service (RFC 2210).



L'objet *flowspec* contient les paramètres suivants :

- la classe de service demandée : garantie ou contrôle de charge ;
- tspec (traffic specification) : correspond aux paramètres token_bucket_tspec ;
- rspec (request specification) : si la classe de service demandée est de type garantie. Cet objet décrit les caractéristiques de la classe de service, c'est-à-dire le débit (exprimé en paquets par seconde) et la gigue maximale autorisée (exprimée en microsecondes).

L'objet *sender_tspec* contient uniquement les paramètres *token_bucket_tspec*.

L'objet *rsvp_adspec* contient tout ou partie des paramètres de description de classe de service. Ils sont modifiés ou ajoutés par les routeurs pour décrire des caractéristiques locales et, en définitive, décrire les caractéristiques globales pour la route.

Déployer une politique de qualité de service

L'application manuelle d'une politique de qualité de service sur l'ensemble des routeurs et des commutateurs peut se révéler fastidieuse et être source d'erreur.

Il est possible d'automatiser ce processus grâce au protocole **COPS** (*Common Open Policy Service*). La politique de qualité de service est alors définie à l'aide d'une interface graphique au niveau d'un serveur COPS, puis diffusée à partir de celui-ci aux routeurs du réseau. La création des règles et leur diffusion sont ainsi simplifiées.

Il faut avant tout indiquer aux routeurs de télécharger leur politique en utilisant le protocole COPS à partir d'un serveur :

```
# Sur un routeur :
ip rsvp policy cops servers 10.0.20.2

# Sur un commutateur :
set qos policy-source cops
set port qos 1/1 roles regulation
set cops server 10.0.20.2 diff-serv

# Sur un autre commutateur :
set qos policy-source cops
set cops server 10.0.20.2 rsvp
set cops domain-name domaine_multimedia
```

Indique l'adresse IP du serveur COPS

Profil à télécharger pour ce port

QoS DiffServ ou IntServ

Nom du domaine défini au niveau du serveur COPS

Concernant le serveur lui-même et son interface graphique, il existe différents produits, tels que *QoS Policy Manager* (Cisco), *Orchestream* (Network Computing), *PolicyXpert* (HP) et *RealNet Rules* (Lucent).

La qualité de service sur MPLS

Il a été indiqué, au chapitre 12, que MPLS pouvait gérer la qualité de service en intégrant les modèles DiffServ et IntServ.

Le RFC 3270 spécifie deux manières d'intégrer DiffServ avec MPLS :

- Dans le mode **E-LSP** (*EXP-Inferred-PSC LSP*), le champ « Exp » du label MPLS est utilisé pour véhiculer la CoS et le « Drop Precedence », les trois premiers bits du champ DSCP étant copiés dans ce champ. Un LSP (*Label Switched Path*) peut donc véhiculer des paquets ayant différentes classes de service.
- Dans le mode **L-LSP** (*Label-Only-Inferred-PSC LSP*), la CoS est codée dans le label tandis que le champ « Exp » contient le « Drop Precedence ». Un LSP ne véhicule alors que des paquets ayant la même CoS. C'est la seule option possible pour les LSR ATM.

Le terme PSC signifie *PHB Scheduling Class*. Le « Drop Precedence » indique que faire du paquet en cas de congestion, la manière la plus simple consistant à détruire le paquet à l'instar du bit CLP d'ATM.

Le mode E-LSP est actuellement le plus utilisé, car les opérateurs ne proposent généralement pas plus de quatre CoS.

Concernant IntServ, une extension de RSVP, appelée **RSVP-TE** (*RSVP Traffic Engineering*) et spécifiée par la RFC 3209, permet de distribuer les labels en même temps qu'un chemin est réservé pour une CoS donnée.

Les travaux en cours sur le sujet peuvent être trouvés sur le site www.ietf.org/html.charters/mpls-charter.html.

15

Le routage des flux multimédias

Qui dit multimédia dit également diffusion d'un flux audio ou vidéo à plusieurs destinataires, dans le cadre d'une conférence, par exemple. Un même film vidéo peut ainsi être dupliqué en autant de flux qu'il y a de destinataires.

Afin de limiter la charge induite sur le réseau, il est bien plus judicieux de dupliquer ce flux au plus près des destinataires. Pour ce faire, trois mécanismes doivent être mis en place sur notre réseau IP :

- un adressage permettant de désigner des groupes de machines plutôt qu'une seule ;
- un mécanisme permettant d'identifier les groupes actifs au sein du réseau ;
- un mécanisme permettant de router les paquets en les dupliquant le moins possible.

Sans cela, notre réseau IP serait vite engorgé, surtout sur les liaisons WAN pour lesquelles le débit est compté.

Dans ce chapitre, vous apprendrez ainsi :

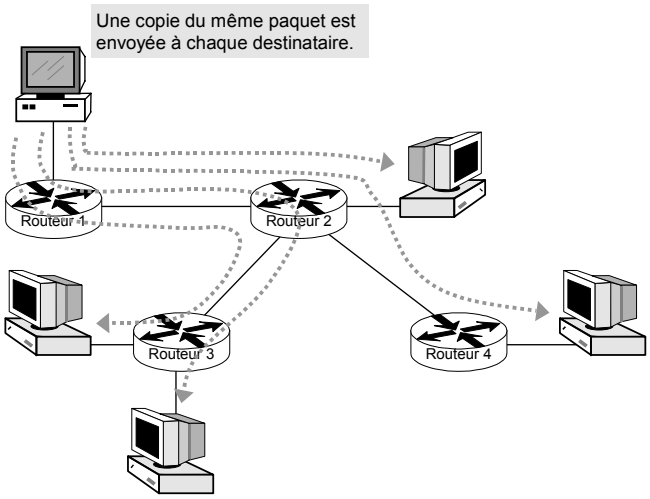
- à utiliser l'adressage multicast ;
- à gérer les groupes de diffusion ;
- à configurer les algorithmes de routage multicast DVMRP, MOSFP et PIM ;
- à choisir l'un de ces algorithmes en fonction de vos besoins.

La diffusion sur un réseau IP

Un participant à une audioconférence parle à tous les autres participants : sa voix est numérisée, puis découpée en paquets IP qui sont ensuite envoyés sur le réseau.

Au premier abord, il est possible de transmettre ces paquets sur un réseau IP classique : une copie de ce paquet est alors transmise à chaque destinataire. Chacun d'eux est, en effet, identifié par une adresse IP unique qui est insérée dans le champ destination du paquet. Les routeurs se servent de cette adresse pour acheminer le paquet jusqu'au destinataire.

Figure 15-1.
Diffusion des paquets unicasts.



Cet exemple montre que ce type de fonctionnement n'est pas adapté à une diffusion : le réseau est inondé par des paquets dupliqués dès la source d'émission. Les adresses utilisées (classes A, B ou C) sont, en effet, de type **unicast**, car à une adresse est associée une machine cible. Par extension, les paquets sont dits unicasts.

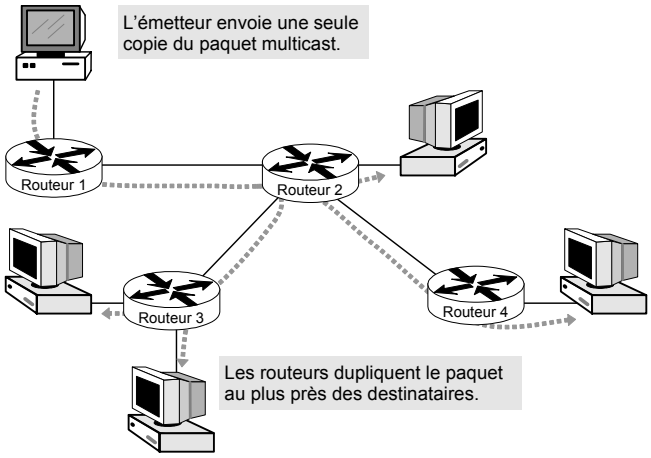
L'adressage IP propose une autre classe d'adresses, la classe D. Ces adresses sont dites **multicasts**, car elles désignent un groupe de machines. Il ne s'agit pas d'une adresse de broadcast, car seules les machines qui sont configurées pour accepter une adresse multicast prendront en compte le paquet.

La plage réservée pour la classe D s'étend de 224.0.0.0 à 239.255.255.255. L'adresse 224.0.0.0 n'est attribuée à aucun groupe ; l'adresse 224.0.0.1 permet d'adresser toutes les machines sur un réseau (le réseau local sur lequel est émis le paquet). L'adresse 224.0.0.2 permet d'adresser, plus spécifiquement, tous les routeurs sur un réseau.

Des adresses de groupes permanents sont attribuées officiellement par l'IANA (*Internet Assigned Number Authority* — www.iana.org).

Groupe concerné	Exemple d'adresse attribuée
Toutes les machines sur le réseau local	224.0.0.1
Tous les routeurs sur le réseau local	224.0.0.2
Tous les routeurs DVMRP	224.0.0.4
Tous les routeurs OSPF	224.0.0.5
Tous les routeurs OSPF désignés	224.0.0.6
Tous les routeurs RIP	224.0.0.9
Tous les routeurs PIM	224.0.0.13
Messages RSVP encapsulés dans UDP	224.0.0.14
NTP (*Network Time Protocol*)	224.0.1.1
Artificial Horizons — Aviator	224.0.1.5
Music-Service	224.0.1.16
IETF-2-Video	224.0.1.15
Microsoft-ds	224.0.1.24

Figure 15-2.
Diffusion d'un paquet multicast.

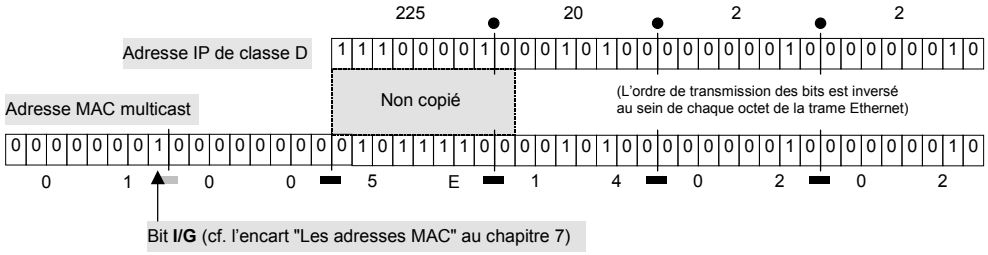


Sans les mécanismes du multicast, le serveur devrait, tout d'abord, déterminer les destinataires (leur adresse IP ou leur nom), puis envoyer autant d'exemplaires du paquet qu'il y a de destinataires.

L'adresse IP source est toujours celle de la station qui envoie le paquet (une adresse unicast). Aucun paquet n'est émis avec une adresse source multicast. Une adresse de classe D identifie toujours un groupe de destinataires.

Sur les réseaux qui prennent en charge ce type d'adressage, le multicast IP utilise les fonctions de multicast du niveau 2, c'est-à-dire des trames multicasts. Sur un réseau Ethernet, les paquets multicasts IP sont envoyés dans des trames multicasts dont l'adresse MAC de destination commence par « 01-00-5E ». Les 23 derniers bits de cette adresse correspondent aux 23 derniers bits de l'adresse IP.

Figure 15-3.
Correspondance entre les adresses multicasts IP et Ethernet.



Un PC envoie donc tous ses paquets multicasts IP dans une trame multicast Ethernet. Le mécanisme de routage par défaut (voir chapitre 9) ne s'applique, en effet, qu'aux paquets unicasts.

Sur les réseaux ne disposant pas de la fonction multicast de niveau 2 (X.25, Frame Relay, ATM, par exemple), les paquets multicasts sont transportés dans les trames unicasts de niveau 2.

La gestion des groupes de diffusion

Un groupe de diffusion (groupe multicast) est dynamique : ses membres peuvent adhérer au groupe ou le quitter à tout moment, être dispersés à travers le monde, adhérer à plusieurs groupes en même temps. Aucune restriction quant au nombre de participants n'est également appliquée. Le groupe peut être permanent ou non. Un participant peut être actif ou non (la machine est éteinte).

La première tâche pour un participant (une machine connectée sur le réseau) est donc de se faire connaître. Pour cela, il dispose du protocole **IGMP** (*Internet Group Membership Protocol*) défini par la RFC 1112, datant de 1989, et mis à jour en 1997 par la RFC 2236 (IGMP v2).

En principe, toutes les piles TCP/IP, et en particulier celle de Windows, supportent IGMP. Aucune configuration spécifique n'est nécessaire, car la gestion des groupes n'est pas réalisée manuellement mais directement par les applications *via* des API (*Application Programming Interface*).

Par exemple, l'interface Winsock offre les primitives permettant d'entrer et de sortir d'un groupe :

- *JoinHostGroup* (adresse IP du groupe, numéro de l'interface réseau) ;
- *LeaveHostGroup* (adresse IP du groupe, numéro de l'interface réseau).

Ces fonctions permettent à la pile TCP/IP d'accepter d'une part les paquets dont l'adresse IP de destination lui correspond et, d'autre part, les paquets dont l'adresse IP de destination est celle du ou des groupes multicasts auxquels l'interface réseau est non jointe . La pile TCP/IP maintient ainsi une liste d'appartenance pour chaque interface réseau de la machine.

Ainsi, plusieurs logiciels utilisent déjà le multicast à votre insu :

- les serveurs WINS de Windows NT, qui se placent d'office dans le groupe 224.0.1.24 qui est utilisé pour dialoguer avec des partenaires de réplication ;
- le logiciel de visioconférence Netmeeting ;
- le logiciel de diffusion audio et vidéo RealG2 Player.

La plupart des routeurs prennent également en charge ce protocole. La configuration d'un routeur Cisco consiste simplement à activer la fonction multicast d'IP :

```
int e 0
ip multicast-routing
```

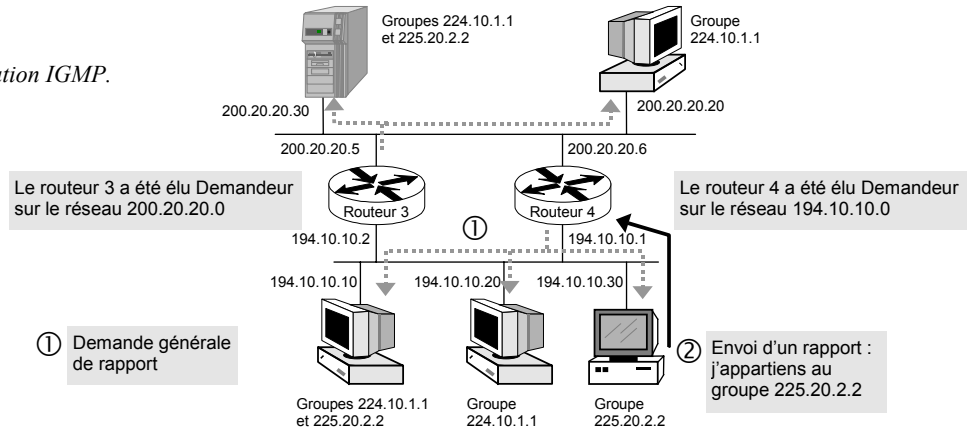
Si le routeur dispose de plusieurs interfaces sur le même réseau local, une seule doit être configurée avec IGMP.

On peut indiquer au routeur de devenir membre d'un groupe. Cela permet aux exploitants de savoir si un groupe est joignable en utilisant la commande ping. Si personne n'est actif sur ce groupe, le routeur pourra au moins répondre au ping :

```
ip igmp join-group 224.10.1.1
```

Considérons l'exemple suivant : chacun des deux routeurs est élu Demandeur pour un réseau en fonction de son adresse IP.

Figure 15-4.
Exemple de configuration IGMP.

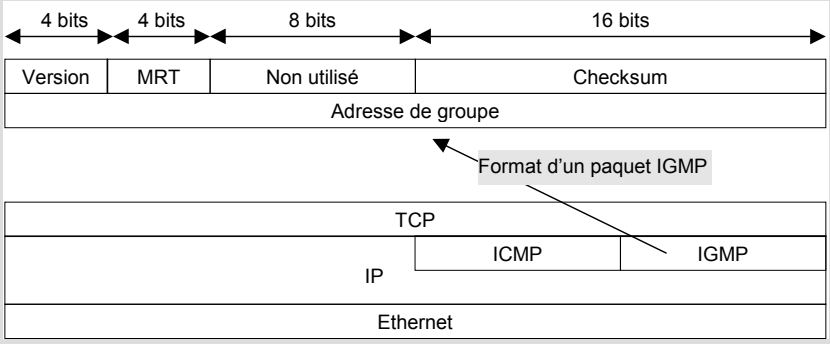


LE POINT SUR IGMP (RFC 1112 ET 2236)

IGMP (*Internet Group Membership Protocol*) est un protocole qui fonctionne conjointement avec IP, au même titre qu'ICMP. Il permet aux membres d'un groupe de signaler leur présence au routeur le plus proche, celui connecté au réseau local. Si plusieurs routeurs sont connectés à un même réseau local, celui qui a la plus petite adresse IP est élu Demandeur.

Le routeur Demandeur émet périodiquement une **demande générale de rapport** (adresse de destination 224.0.0.1). Les membres de tous les groupes actifs sur ce réseau local lui répondent en envoyant un r**apport** pour chaque groupe auquel ils appartiennent.

Le routeur maintient à jour une liste des groupes dont au moins un membre est actif. Les routeurs non demandeurs n'émettent pas de demande de rapport, mais lisent les rapports et mettent à jour leur table.
Il existe trois types de paquets IGMP : demande de rapport, sortie d'un groupe (demandée par un membre) et rapport d'activité. Un quatrième type, le rapport IGMP version 1, est supporté à des fins de compatibilité.



Le champ MRT (*Max Response Time*) est utilisé dans les paquets de demande de rapport pour indiquer, en dixièmes de secondes, le délai maximal autorisé pour envoyer un rapport d'activité. Au-delà de ce délai, le destinataire est considéré comme n'étant pas actif dans le groupe.

Un membre peut également demander à sortir du groupe (message à destination de 224.0.0.2). Le routeur Demandeur émet alors une **demande spécifique de rapport** (adresse de destination identique à celle du groupe) pour s'assurer qu'un membre au moins est encore actif.

Les routeurs sont à l'écoute de tout message multicast, et prennent donc en compte tous les rapports. La commande suivante montre le résultat obtenu sur un routeur Cisco :

```
Routeur3# show ip igmp interface
Ethernet0 is up, line protocol is up
```

```
  Internet address is 194.10.10.2, subnet mask is 255.255.255.0
  IGMP is enabled on interface
  IGMP query interval is 120 seconds
  Inbound IGMP access group is not set
  Multicast routing is enabled on interface
  Multicast TTL threshold is 0
  Multicast designated router (DR) is 194.10.10.1
  Multicast groups joined: 224.10.1.1 225.20.2.2
Ethernet1 is up, line protocol is up
  Internet address is 200.20.20.5, subnet mask is 255.255.255.0
  IGMP is enabled on interface
  IGMP query interval is 120 seconds
  Inbound IGMP access group is not set
  Multicast routing is enabled on interface
  Multicast TTL threshold is 0
  Multicast designated router (DR) is 200.20.20.5
  Multicast groups joined: 224.10.1.1 225.20.2.2
```

Le routeur 3 n'est pas Demandeur pour le réseau 194.10.10.0 ; il n'envoie donc pas de requête. En revanche, il prend en compte tous les rapports qu'il voit passer :

```
Routeur4# show ip igmp groups
IGMP Connected Group Membership
Group Address    Interface       Uptime    Expires   Last Reporter
224.10.1.1       Ethernet0       17:15:15  0:02:05   194.10.10.20
224.10.1.2       Ethernet1        1:01:01  0:01:05   220.20.20.30
224.10.1.1       Ethernet1       17:15:40  0:01:40   220.20.10.30
225.20.2.2       Ethernet0        1:00:45  0:01:50   194.10.10.10
```

De même, le routeur 4 n'est pas Demandeur pour le réseau 220.20.20.0 ; il n'envoie donc pas de requête. En revanche, il prend également en compte tous les rapports qu'il voit passer.

La périodicité d'envoi des demandes générales de rapports peut être paramétrée comme suit :

```
ip igmp query-interval 120
```

Toutes les 120 secondes

La version 3 d'IGMP, en cours d'étude, permettra aux machines d'indiquer au routeur l'adresse IP source pour laquelle elle accepte de recevoir des paquets multicasts ; ainsi, la diffusion du paquet prendra en compte l'adresse de l'émetteur, en plus du groupe destination.

Le routage des flux multicasts

Le protocole IGMP permet aux routeurs de détecter les groupes situés sur leurs réseaux locaux (les réseaux auxquels une interface est connectée).

Mais les routeurs ne savent pas où sont situés les autres membres du groupe, puisque IGMP n'a qu'une portée locale. Pour cela, il faut utiliser des protocoles de routage spécifiques au multicast. Trois standards sont disponibles :

- DVMRP (*Distance Vector Multicast Routing Protocol*), analogue au protocole RIP adapté au multicast ;
- MOSPF (*Multicast Open Shortest Path First*), une extension d'OSPF ;
- PIM (*Protocol Independent Multicast*), spécialement dédié au multicast.

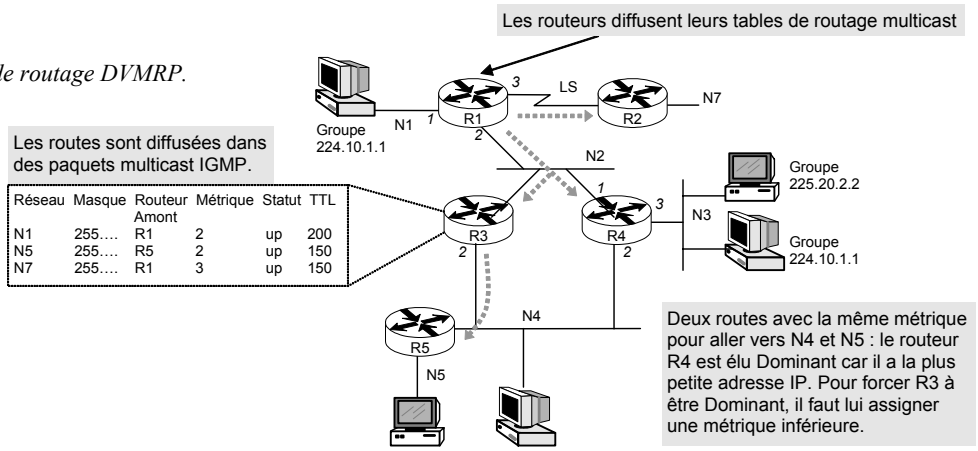
Le choix de l'un de ces protocoles dépend de nombreux paramètres. Les paragraphes suivants décrivent donc leurs principes de fonctionnement, afin de mieux cerner leurs conséquences sur l'architecture de notre réseau.

Le routage à l'aide de DVMRP

Le premier protocole de routage multicast a été DVMRP (*Distance Vector Multicast Routing Protocol*). Comme son nom l'indique, il repose sur un algorithme de calcul du plus court chemin, basé sur le plus petit nombre de routeurs à traverser pour atteindre une destination (nombre de sauts, appelé métrique).

Le principe est en cela identique à RIP (*Routing Information Protocol*) : les routeurs s'échangent l'intégralité de leurs tables de routage. DVMRP doit donc être utilisé en supplément d'un protocole de routage unicast (RIP, OSPF, etc.).

Figure 15-5.
Diffusion des tables de routage DVMRP.



La commande suivante permet d'activer DVMRP sur notre routeur R1, qui est un Netbuilder de marque 3com. Elle doit être utilisée pour chaque interface gérant le trafic multicast. La commande MIP permet d'activer le protocole IGMP :

```
setdefault -MIP control = enable
setdefault !1 -DVMRP control = enable
setdefault !2 -DVMRP control = enable
setdefault !3 -DVMRP control = enable
```

Lorsque plusieurs routeurs coexistent sur le même réseau local, seul l'un d'eux a la charge de diffuser les paquets multicasts, afin d'éviter leur duplication. Comme pour RIP, la route choisie repose sur le plus petit nombre de sauts (la métrique). En cas d'égalité de métrique, le routeur dominant élu est celui qui possède la plus petite adresse IP.

Pour forcer R3 à être le routeur dominant sur le réseau N4, il faut lui attribuer une métrique inférieure à celle de R4 :

```
#Routeur R3
setdefault !2 -DVMRP metric = 5
#Routeur R4
setdefault !2 -DVMRP metric = 10
```

La commande suivante permet de visualiser la table de routage du routeur R4 :

```
show -dvmrp routetable long
```

```
SourceSubnet SubnetMask    FromGateway Metric Status TTL InPort OutPorts
200.10.10.0  255.255.255.0 200.20.20.1 2      Up     200 1      2 3*
200.30.30.0  255.255.255.0             1      Up     150 3      1 2
...
```

La colonne « FromGateway » indique le routeur le plus proche qui mène à la source (un champ vide indique que le réseau est directement connecté au routeur R4).

La colonne « InPort » (*Incoming Port*) indique l'interface par laquelle arrivent les paquets multicasts émis par la source précisée dans la colonne « SourceSubnet ».

La colonne « OutPorts » (*Outgoing Ports*) donne la liste des ports vers lesquels seront diffusés par défaut les paquets multicasts issus de la source indiquée dans la colonne « SourceSubnet ». Un astérisque indique que le port conduit à une feuille de l'arbre, c'est-à-dire qu'aucun routeur ne se trouve en dessous.

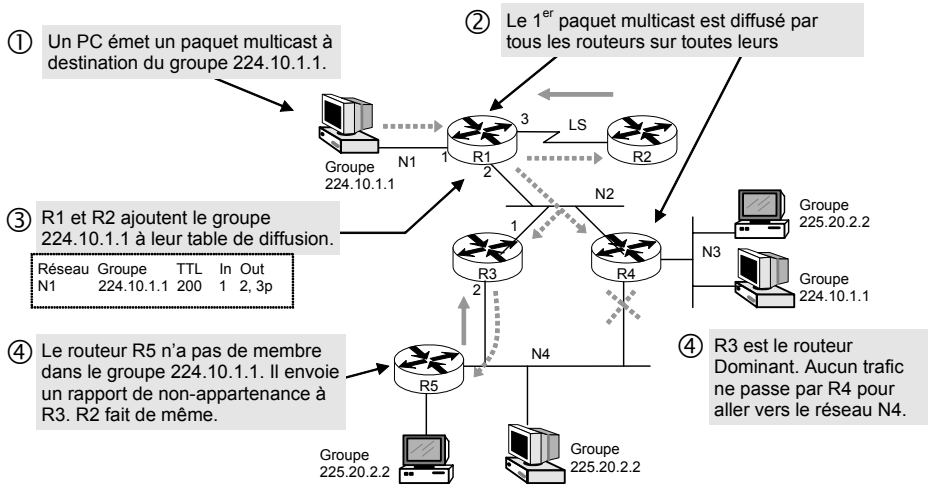
En plus de la table de routage, le routeur gère une table de diffusion construite lorsque les premiers paquets multicasts transitent par le routeur. Elle permet d'enregistrer les groupes identifiés pour chaque réseau source.

La commande suivante affiche la table de diffusion pour le routeur R4 :

```
show -dvmrp forwardtable

SourceSubnet MulticastGroup TTL InPort OutPorts
200.10.10.0  224.10.1.1     200 1      1 2p
200.30.30.0  224.10.1.1     150 3      2p 3
             225.20.2.2     200 3      2p 3
...
```

La colonne « MulticastGroup » indique la liste des groupes de diffusion des paquets en provenance de la source indiquée dans la colonne « SourceSubnet ». Les autres colonnes ont la même signification que celles de la table de routage. L'indicateur « p » (*prune*) indique qu'un message de non-appartenance a été reçu ; par conséquent, aucun paquet multicast ne sera envoyé vers cette interface. Dans notre cas, cela indique que R3 est dominant.

Figure 15-6. *Routage d'un paquet multicast.*



Dans notre réseau, le routeur R1 a diffusé le paquet multicast vers les interfaces 2 et 3. Il met alors à jour sa table de diffusion. L'interface 3 a par la suite été marquée « p » (*pruned*) car le routeur R2 lui a renvoyé un rapport de non-appartenance. En effet, ce dernier n'a détecté (*via* IGMP) aucun membre actif pour le groupe 224.10.1.1.

De même, le routeur R5 renvoie un message de non-appartenance au routeur R3, qui ne lui transmettra alors plus aucun paquet pour le groupe 224.10.1.1. Le routeur R3 fait ensuite de même vis-à-vis de R1 et R4.

LE POINT SUR DVMRP (RFC 1075)

DVMRP (*Distance Vector Multicast Routing Protocol*) utilise son propre protocole de routage de paquets multicasts, qui est analogue à celui de RIP : les routeurs s'échangent l'intégralité de leurs tables de routage et calculent les routes sur la base d'une **métrique** (nombre de sauts en termes de routeurs).

Il existe plusieurs implémentations de DVMRP qui diffèrent par l'algorithme utilisé pour construire la table de diffusion : **TRPB** (*Truncated Reverse Path Broadcasting)* ou **RPM** (*Reverse Path Multicasting*), ce dernier étant le plus répandu car plus performant. Les routeurs 3com et le démon Unix *mrouted* (à partir de la version 3.8) utilisent RPM.

Le principe du TRPB est le suivant : un routeur recevant un paquet multicast par une interface, dont la route de retour vers l'émetteur est la plus courte, le transmet sur toutes les autres interfaces. Les routeurs situés en aval reçoivent donc le paquet. S'ils ne disposent d'aucun membre déclaré ni d'aucun autre routeur en aval, ils renvoient un rapport de non-appartenance au routeur situé en amont. Ce dernier ne leur transmettra alors plus de paquets multicasts. Afin de tenir compte des changements de topologie, l'arbre RPM est périodiquement reconstruit (toutes les deux heures sur MBONE).

Si un nouveau membre s'enregistre sur un des routeurs situés en aval, celui-ci enverra au routeur situé en amont un message d'annulation, pour recevoir à nouveau les paquets multicasts destinés au groupe en question.

Le principe du RPM reprend celui du TRPB, mais pousse plus loin la remontée d'information : un routeur qui ne dispose pas de membre ni de routeur en aval ayant de membre envoie un rapport de non-appartenance aux routeurs situés en amont, de sorte que lui-même ne reçoive pas de paquets multicasts. Le rapport peut ainsi remonter jusqu'à la source si nécessaire.

Un paquet DVMRP est composé d'un en-tête IGMP auquel sont jointes des données de longueur variable (512 octets au maximum) formatées sur le mode : " Commande, Valeur " ou " Commande, nombre de valeurs, valeur 1, valeur 2, etc. ". L'exemple suivant montre un paquet d'annonce de la route 194.10.10.0.

4 bits	4 bits	8 bits	16 bits	
Version	Type	Sous-type	Checksum	
Cmde = AFI		Valeur = IP	Cmde = Métrique	Valeur = 2
Cmde = Intinity		Valeur = 6	Cmde = Subnet	Compte = 1
Valeur = 255.255.255.0				
Cmde = DA		Compte = 1	Valeur = 194.10	
.10.0				

Le champ " Cmde " contient le code d'une commande qui détermine la taille et la signification du champ Valeur. Plusieurs commandes se suivent dans un paquet. Le champ Sous-type détermine le type de message : requête, réponse, rapport de non-appartenance ou annulation d'un rapport de non-appartenance.

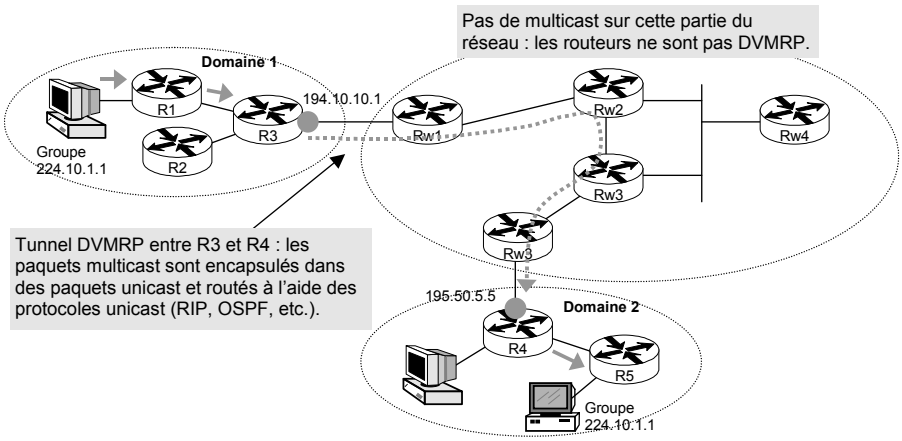
La commande AFI est toujours présente : elle indique simplement que la famille d'adresses est IP (seul supporté). La commande DA indique une adresse destination.

L'inconvénient de DVMRP est que les paquets multicasts doivent périodiquement être envoyés aux routeurs situés en aval pour tenir compte d'éventuels changements de topologie ou d'appartenance à un groupe (pour un groupe donné, les messages d'annulation ne remontent que si un multicast a été précédemment reçu et un message de non-appartenance, émis). Il en résulte une nouvelle cascade de rapports de non-appartenance ou d'annulation remontant vers les routeurs situés en amont.

Sur un réseau, tous les routeurs ne sont pas DVMRP. La mise en place du routage multicast est, en effet, progressive. En outre, le fait de ne pas installer de routeurs DVMRP partout permet de définir des domaines dans le but de circonscrire la diffusion des paquets multicasts.

Pour assurer néanmoins la diffusion des paquets multicasts, il est possible de créer un tunnel entre deux routeurs DVMRP séparés par des routeurs qui ne savent pas router les paquets multicasts.

Figure 15-7. *Principe du tunneling DVMRP.*



Sur nos routeurs 3com, la création du tunnel IP est réalisée simplement en indiquant les adresses des deux routeurs :

```
#Sur le routeur R3
setdefault !1 -DVMRP mon_tunnel = 194.10.10.1 195.50.5.5

#Sur le routeur R4
setdefault !2 -DVMRP mon_tunnel = 195.50.5.5 194.10.10.1

#Activation de DVMRP sur le tunnel
setdefault mon_tunnel -DVMRP control = enable
```

Afin d'atténuer les effets d'un trafic multicast important sur un réseau non dimensionné pour cela, il est possible d'en limiter le débit à 64 Kbit/s, par exemple :

```
# Sur les routeurs R3 et R4
setdefault mon_tunnel -dvmrp ratelimit = 64
```

De même, il est possible d'augmenter la périodicité d'échange des tables de routage entre les deux routeurs (60 secondes par défaut) :

```
setdefault -dvmrp updatetime = 120
```

Enfin, la durée de validité des entrées dans la table de diffusion peut également être allongée (dans notre cas, 3 600 secondes) :

```
setdefault -dvmrp cachetime = 3600
```

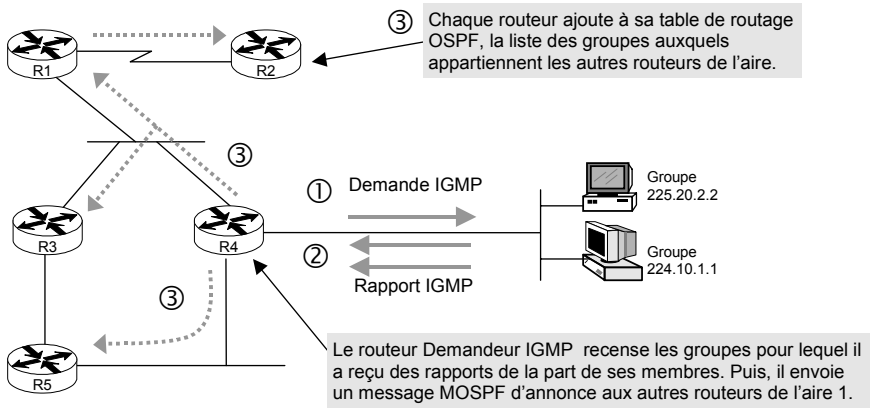
Le routage à l'aide de MOSPF

Le protocole MOSPF (*Multicast Open Shortest Path First*) est une extension d'OSPF permettant de router les paquets IP multicasts. La détection des membres actifs d'un groupe est, comme toujours, assurée par IGMP, mais MOSPF permet aux routeurs de diffuser les groupes auxquels ils appartiennent (c'est-à-dire pour lesquels au moins un membre est actif).

Reprenons l'exemple de notre réseau : chaque routeur est configuré avec IGMP et MOSPF. Localement, chaque routeur établit une liste de tous les groupes pour lesquels il existe un membre actif sur son (ou ses) interface(s) LAN.

En même temps, les routeurs s'échangent des messages d'annonce d'état des liens, suivant en cela le fonctionnement classique d'OSPF (voir chapitre 12). Cela leur permet de connaître la topologie du réseau.

Figure 15-8. *Diffusion d'un paquet multicast par MOSPF.*



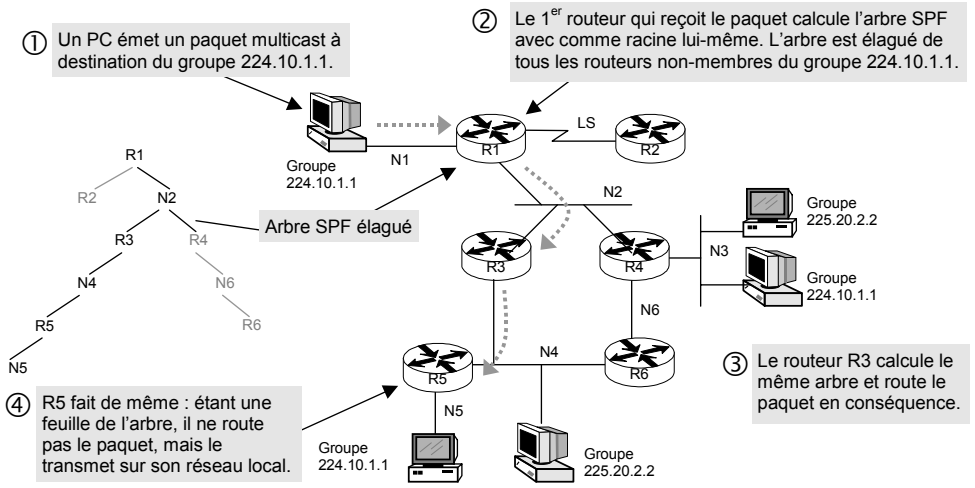
Dès le premier enregistrement d'un membre de groupe *via* IGMP, les routeurs sont considérés par OSPF comme étant membres du groupe. Ces routeurs s'échangent alors des messages d'annonce d'appartenance à un groupe. Tous les routeurs savent désormais quel routeur appartient à quel groupe.

Sur nos routeurs 3com, l'activation de MOSPF est effectuée grâce à la commande suivante :

```
#Activation d'igmp
setdefault -mip control = enable

#Activation de mospf sur chaque interface
setdefault !1 -mospf control = enable
setdefault !2 -mospf control = enable
```

Le routage des paquets multicasts au sein d'une aire est réalisé en fonction de la source (adresse IP unicast de la machine), de la destination (adresse IP multicast du groupe) et du TOS (*Type of Service*). Le routeur construit pour cela un arbre SPF (*Shortest Path First*) dont tous les routeurs non-membres du groupe indiqué dans le paquet ont été supprimés. Cet arbre est calculé à la demande lorsqu'un paquet multicast arrive sur le routeur.

Figure 15-9.
Calcul de l'arbre SPF élagué.



Dans cet exemple, il est à noter que le routeur Demandeur IGMP doit être un routeur MOSPF, seul capable d'envoyer des messages d'annonce d'appartenance à un groupe. Il faut donc affecter une priorité supérieure aux routeurs MOSPF pour qu'ils soient élus routeurs désignés (au sens OSPF).

LE POINT SUR MOSPF (RFC 1584)

Au sein d'une aire (voir encadré " Le point sur OSPF "), les routeurs s'échangent des **messages d'annonce d'appartenance à un groupe**, permettant à chacun d'eux de disposer d'une visibilité complète de la topologie du réseau.

Chaque routeur calcule un **arbre SPF** pour chaque triplet adresse source unicast/adresse destination multicast/TOS (*Type of Service*). Ce calcul est effectué à la demande, c'est-à-dire lorsque le premier paquet multicast arrive. L'arbre ainsi calculé est conservé dans une mémoire cache, puis détruit au bout d'un certain temps si aucun autre paquet de ce type n'a été reçu.

L'arbre est **élagué** : tous les routeurs qui ne sont pas membres du groupe indiqué dans le paquet sont supprimés. Les réseaux terminaux (*stub*) n'ont pas besoin d'être pris en compte : la diffusion locale au routeur est, en effet, assurée par IGMP.

Un chemin multicast est donc calculé en construisant un arbre élagué du plus court chemin dont la racine est l'émetteur du paquet (la RFC emploie l'expression : " *pruned shortest-path tree rooted at the packet's IP source* ").

Pour la diffusion des appartenances à un groupe, un nouveau paquet d'annonce a été ajouté : *group-membership-LSA*.

4 bits	4 bits	8 bits	16 bits
Link State Age		Options	Type = 6
Link State ID = adresse multicast du groupe			
Advertising router			
LS sequence number			
Checksum		Longueur	
Vertex ID			
Longueur du Vertex			

Par ailleurs, les modifications suivantes ont été apportées à OSPF :

- Le champ " Option " des paquets Hello, Description de la base et Annonce d'état de lien contient un nouvel indicateur, le bit MC, qui indique si le routeur prend en charge l'extension multicast d'OSPF (MOSPF).
- Le champ " Rtype " des paquets d'annonce d'état de liens contient un nouvel indicateur, le bit W, qui indique si le routeur accepte ou non les multicasts provenant de n'importe quelle source.

Un routeur calcule autant d'arbres qu'il y a de participants à une conférence, et calcule autant de variantes de cet arbre qu'il y a de groupes destinataires. La prise en compte du TOS augmente encore le nombre de combinaisons.

Le nombre d'arbres peut donc être très important. C'est pour cette raison qu'ils sont calculés à la demande et que le résultat est conservé en mémoire cache. Cela implique également que, plus le nombre de machines et de groupes est élevé, plus la CPU des routeurs est sollicitée pour calculer les arbres.

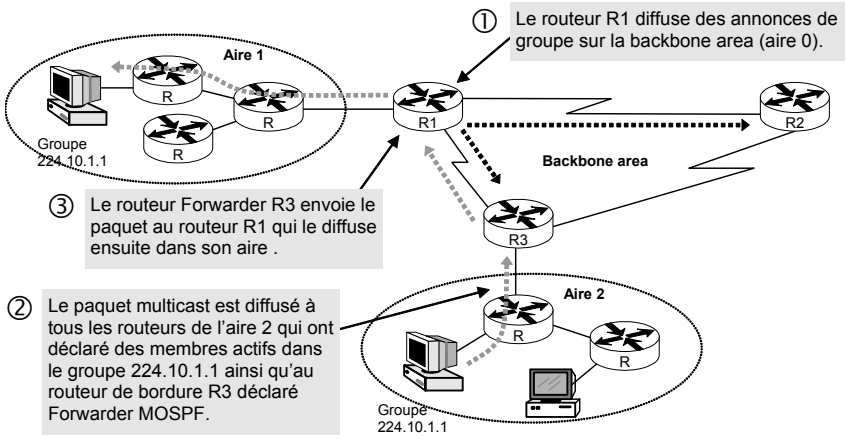
Le résultat du calcul d'un arbre est conservé aussi longtemps qu'un trafic entre le couple d'adresses IP unicast source/multicast destination existe, et jusqu'à ce qu'un changement de topologie intervienne. La taille mémoire requise est donc plus importante par rapport à un routeur OSPF classique (14 à 20 octets par triplet adresses source/groupe/TOS).

Le dimensionnement mémoire/CPU dépend du constructeur du routeur (architecture et puissance), de l'importance du réseau (nombre de machines multicasts et nombre de groupes) ainsi que des performances visées (utilisation intensive ou non des téléconférences, utilisation du réseau à d'autres fins que le multimédia, etc.).

Il est également recommandé de limiter le nombre de TOS, voire de ne pas activer cette fonctionnalité, afin de limiter le nombre de combinaisons d'arbres SPF.

La diffusion des paquets multicasts entre aires et systèmes autonomes utilise le principe appelé Multicast Forwader : les routeurs de bordure et les routeurs inter-AS déclarés comme tels ne sont jamais supprimés des arbres SPF. En conséquence, tous les paquets multicasts leur seront envoyés. De leur côté, ces routeurs calculent un arbre SPF pour chaque aire (ou AS) à laquelle ils sont rattachés.

Figure 15-10.
Routage des paquets multicast entre aires.



Les annonces d'appartenance à un groupe sont diffusées au sein de l'aire 0 (*backbone area*) par tous les routeurs de bordure. Tous les routeurs participant à la backbone area connaissent donc tous les groupes présents au sein du système autonome. Les informations provenant de la backbone area ne sont cependant pas transmises aux autres aires. Seuls les Forwarders conservent ces informations.

L'architecture et le dimensionnement des réseaux constituant la backbone area sont donc très importants. Il faut tenir compte du nombre de postes de travail et de leur répartition entre les aires. L'aire backbone concentre, en effet, la somme des flux circulant dans chacune des aires où les groupes sont actifs. Dans la mesure du possible, il est donc conseillé de limiter le nombre d'aires ou bien de répartir des groupes entre plusieurs aires.

La configuration pour déclarer Forwarder un routeur de bordure 3com est la suivante :

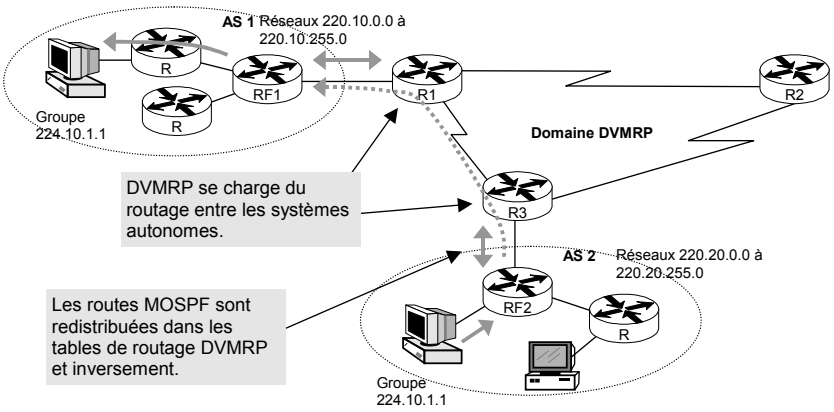
```
SETDefault -MOSPF MABR = Enable
```

Le mot clé MABR signifie *Multicast Area Border Router*.

Le principe est identique pour le routage entre systèmes autonomes : un routeur frontière est désigné Forwarder de système autonome, et reçoit tous les paquets multicasts circulant au sein du système autonome (et non plus seulement au sein d'une aire). La différence est importante, car la concentration de trafic est encore plus prononcée et les routeurs de ce type sont encore plus sollicités que les autres. Il est donc conseillé de dédier un routeur frontière de système autonome à cette seule tâche et de le relier à un réseau à haut débit.

Comme il a été mentionné au chapitre 12, OSPF ne permet pas de router les paquets entre AS. Il faut lui adjoindre un protocole de routage extérieur, tel que BGP4. Dans le cadre d'un réseau multicast, c'est DVMRP qui se charge de cette tâche au niveau des routeurs frontière.

Figure 15-11.
Routage DVMRP entre systèmes autonomes.



La configuration suivante doit donc être effectuée sur nos routeurs frontière RF1 et RF3 (ici, l'exemple concerne RF1) :

```
#Active la diffusion des routes MOSPF dans DVMRP
add -dvmrp mospf 220.10.0.0/16 aggregate 1
#Active le routage d'ospf vers dvmrp
setdefault -dvmrp policycontrol = mospf

#Active la diffusion des routes DVMRP dans ospf
add -mospf dvmrp 220.20.0.0/16 aggregate 5
#Active le routage de dvmrp vers ospf
setdefault -mospf policycontrol = dvmrp
```

L'option « Aggregate » permet de diffuser une seule route pour les routes comprises entre 220.10.0.0 et 220.10.255.0 au lieu de diffuser 256 routes.

Le système permet de mêler des routeurs OSPF et MOSPF. Cependant, aucune machine située derrière un routeur OSPF ne recevra les messages multicasts. Il doit donc y avoir continuité des routeurs MOSPF tout au long des routes menant vers les machines multicasts. À l'inverse de DVMRP, le tunneling n'est, en effet, pas supporté.

De plus, il faut toujours que le routeur désigné sur un réseau local soit un routeur MOSPF, sinon les requêtes IGMP n'alimenteront pas le protocole de routage multicast. Pour cela, il faut assigner une priorité supérieure aux routeurs OSPF également configurés en MOSPF.

Le routage à l'aide de PIM

Comme indiqué aux paragraphes précédents, le protocole MOSPF — et encore moins DVMRP — n'est pas adapté pour de grands réseaux, et notamment Internet (plusieurs dizaines de milliers d'aires et de systèmes autonomes).

Pour résoudre ce problème, l'IETF a défini le protocole PIM (*Protocol Independent Multicast*) qui, comme son nom le laisse penser, est indépendant du protocole de routage unicast utilisé. PIM fonctionne, en effet, avec RIP, OSPF, BGP, et même DVMRP.

PIM utilise deux modes de fonctionnement :

- *dense* (densité élevée), qui est adapté à des réseaux à haut débit et à des situations où les membres de groupes sont géographiquement proches ;
- *sparse* (clairsemé), qui correspond aux cas de figures où les réseaux ont de plus faibles débits et où les membres de groupes sont très dispersés.

Un routeur PIM bascule d'un mode à l'autre indépendamment de chaque groupe.

Le mode *dense* (PIM-DM) fonctionne de manière identique à DVMRP (algorithme *reverse path flooding*) sans toutefois utiliser un protocole de routage multicast dédié ; il utilise les protocoles de routage unicast en place (RIP, OSPF, etc.), d'ou le nom de « *Protocol Independant* ». PIM-DM ne fait pas encore l'objet d'une RFC.

Le mode *sparse* (PIM-SM) nécessite de convenir d'un routeur qui tienne lieu de point de rendez-vous pour chaque groupe. Un routeur peut être le point de rendez-vous pour plusieurs groupes. Il peut exister plusieurs points de rendez-vous pour un groupe, mais un seul doit être actif à un moment donné.

Principe de PIM-SM

Pour un couple réseau source/groupe multicast, un routeur PIM-SM bascule entre deux algorithmes de routage : un dont le chemin passe par le point de rendez-vous, et un second qui bénéficie d'un chemin direct entre la source et la destination. Pour le premier, le calcul de la route s'établit à partir d'un arbre partagé dont la racine est le point de rendez-vous (*Rendez-vous Point shared tree*). Pour le second, ce calcul se fonde sur un arbre du plus court chemin dont la racine est la source d'émission du paquet multicast (*source shortest path tree*).

Si plusieurs routeurs sont connectés au même réseau local, celui dont l'adresse IP est la plus élevée sera élu routeur désigné et aura à charge d'envoyer et de recevoir les messages Join/Prune. Pour découvrir ses voisins, un routeur émet périodiquement un message Hello dont la périodicité par défaut est de 30 secondes :

```
ip pim query-interval 30
```

Tous les routeurs désignés dont les membres actifs appartiennent à un même groupe s'enregistrent auprès du même routeur, appelé point de rendez-vous (RP). Les routeurs intermédiaires enregistrent également cette information et font de même auprès du RP. Ce dernier peut donc calculer un arbre de routage pour un couple adresse réseau source/adresse multicast de groupe. En définitive, le RP connaît tous les réseaux sources susceptibles d'émettre et de recevoir des paquets multicasts au sein d'un groupe donné.

Pour terminer cette introduction, plusieurs routeurs peuvent être candidats à l'élection du point de rendez-vous. L'un d'eux est élu routeur bootstrap ; il est chargé de désigner le RP actif pour chaque groupe. Il en résulte de nouveaux messages d'annonce entre les RP.

Principe du calcul des routes

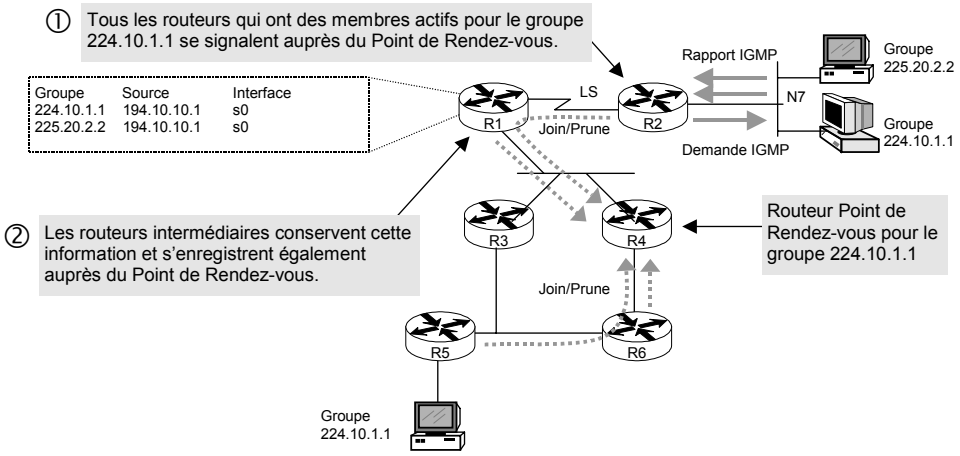
Un routeur PIM-SM ayant enregistré (*via* IGMP) des membres d'un ou plusieurs groupes envoie périodiquement un message Join/Prune (joindre/élaguer) au point de rendez-vous RP. Chaque routeur chargé de transporter ce message se trouvant sur le chemin enregistre les mêmes informations : le groupe, la source (l'adresse du réseau sur lequel ont été identifiés les membres du groupe) et l'interface d'entrée du message Join/Prune.

Un message Join/Prune contient, pour chaque adresse de groupe (multicast), la liste des adresses sources incluses dans l'arbre de routage (*joined*, qui ont rejoint l'arbre) et de celles qui ne le sont pas (*pruned*, élaguées de l'arbre).

Un routeur émet un message Join/Prune vers un routeur situé en amont (en direction de la racine de l'arbre). Il indique par là qu'il fait partie de l'arbre de routage pour les paquets multicasts du groupe G émis par la source S.

En retour, un routeur diffuse les paquets multicasts uniquement vers les interfaces à partir desquelles des messages Join/Prune ont été reçus (et qui correspondent à un couple adresse source unicast/groupe multicast).

Figure 15-12.
Calcul des routes par le routeur point de rendez-vous.



Les commandes suivantes permettent d'activer PIM sur nos routeurs Cisco :

```
ip multicast-routing
interface ethernet 0
ip pim sparse-dense-mode
```

Aucune autre configuration n'est requise pour les routeurs intermédiaires et le routeur point de rendez-vous.

En revanche, tous les routeurs susceptibles d'enregistrer des membres de groupe doivent pointer vers le routeur point de rendez-vous :

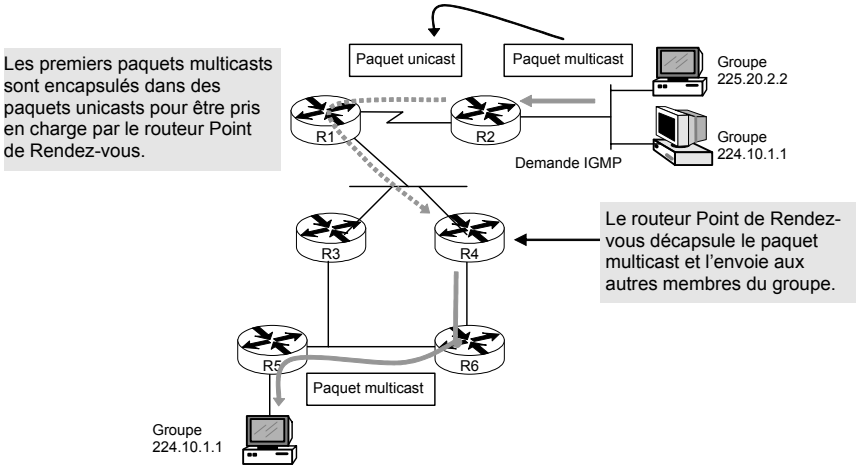
```
access-list 1 permit 224.10.1.1
access-list 1 permit 222.20.2.2
ip pim rp-address 220.20.20.6 1
```

L'access-list 1 précise la liste des groupes pour lesquels le routeur 220.20.20.6 sera élu point de rendez-vous.

Principe du routage

Le premier paquet multicast émis par un PC est envoyé dans un paquet multicast IP qui est pris en charge par le routeur PIM-SM désigné. Celui-ci encapsule ce paquet dans un paquet unicast à destination du RP. Le RP décapsule ce paquet et l'envoie tel quel (donc sous forme multicast) aux membres du groupe (les routeurs qui se sont déclarés).

Figure 15-13. *Principe du routage PIM-SM.*



Tous les messages passent donc par le RP, ce qui n'est pas forcément une route optimale.

Si le trafic multicast persiste, un routeur (la source, la destination ou le RP) peut alors décider que le flux emprunte un chemin direct entre la source et la destination. Les routeurs concernés basculent alors d'un arbre de routage dont la racine est le RP vers un arbre du plus court chemin dont la racine est la source.

Dans notre réseau, le routeur R1 crée la table de routage suivante :

```
Router4# show ip pim interface
Address        Interface    Mode    Neighbor  Query   DR
                                    Count     Interval
220.20.20.6    Ethernet0    sparse   2         30     220.20.20.6
197.10.10.0    serial0      sparse   1         10     0.0.0.0
```

La colonne *Address* indique l'adresse du routeur vers lequel envoyer les paquets à router *via* l'interface spécifiée dans la colonne *Interface*.

La colonne *Neighbor count* indique le nombre de voisins découverts sur cette interface.

La colonne *DR* indique l'adresse du routeur désigné.

Routage sur les liaisons WAN

Sur Ethernet, le paquet IP multicast est envoyé dans une trame Ethernet multicast. Mais, sur les réseaux qui n'offrent par cette fonction (X.25, Frame Relay, ATM, etc.), seules des trames unicasts peuvent être envoyées.

LE POINT SUR PIM-SM (RFC 2362)

Un routeur source qui reçoit des rapports IGMP indiquant que des membres du groupe G sont actifs ajoute une route (*, G) vers le routeur **point de rendez-vous** pour le groupe ; il lui envoie alors périodiquement un message **Join/Prune** (joindre/élaguer). Tous les routeurs situés sur le chemin qui relie ces deux routeurs ajoutent la même route (*,G) en notant l'adresse du réseau source et l'interface d'entrée, puis envoient un message Join/Prune au routeur point de rendez-vous. Ce dernier connaît donc le chemin le reliant au routeur source qui déclare que des membres du groupe G sont actifs. Les routeurs ont donc rejoint un **arbre partagé** qui a comme racine le routeur point de rendez-vous (*RP shared tree*).

Un routeur source ayant pris en charge un paquet multicast l'encapsule dans un paquet unicast, puis l'envoie au routeur point de rendez-vous, qui le décapsule et le renvoie vers tous les routeurs sources déclarés (qui sont ici des routeurs destinations) et ayant rejoint l'arbre partagé.

Au bout d'un moment, un routeur (source, destination ou point de rendez-vous) peut décider de quitter l'arbre partagé afin que les paquets multicasts soient échangés directement sans passer par le routeur point de rendez-vous. Il envoie alors un message Join/Prune au point de rendez-vous pour quitter l'arbre partagé, et un autre à la source pour rejoindre l'**arbre du plus court chemin** (SPT, S*hortest Path Tree* — dont la racine est la source) pour le couple routeur source S/groupe G. Tous les routeurs situés sur le chemin routeur source-routeur destination ajoutent la route (S,G) dans leur table de routage.

Le basculement vers le SPT s'opère lorsque le flux multicast est suffisamment important et de longue durée (en fait, au bout de quelques secondes).

Les messages Join/Prune contiennent, pour chaque groupe, la liste des sources (routeurs ou réseau) appartenant à l'arbre et celles qui en sont élaguées.

4 bits	4 bits	8 bits	16 bits
Version	Type	Longueur adresse	Checksum
Adresse du routeur upstream			
Réservé		Nombre de groupes	Durée de vie de l'état Join/Prune
1ère adresse multicast de groupe			
Nombre de sources dans l'arbre			Nombre de sources élaguées
Adresse de la source 1 ayant rejoint l'arbre (Join)			
…			
Adresse de la source 1 élaguée (Prune)			
…			
2ème adresse multicast de groupe			
…			

Le champ " Nombre de groupes " indique le nombre de groupes décrits dans le message (notés 1, 2, etc.).

Le champ " Adresse de la source " indique que le routeur transmettra (ou non, s'il est dans l'état Prune) un paquet multicast issu de cette source s'il provient de l'interface par laquelle il envoie ce message Join/Prune. Une adresse source peut être celle d'un routeur, d'un réseau ou d'un agrégat de réseaux.

Les routeurs diffusent les paquets multicasts uniquement vers les interfaces par lesquelles sont entrés des messages Join/Prune. Par ce biais, chaque routeur connaît la topologie du réseau et calcule un arbre de routage (soit partagé, soit du plus court chemin) pour chaque couple Source/Groupe. Les routeurs s'échangent, par ailleurs, des messages Hello pour découvrir leurs voisins.

Les messages PIM sont envoyés dans des paquets unicasts (*Register* et Register-*stop*) et multicasts 224.0.0.13 (Hello, etc.) en utilisant le protocole n° 103 au-dessus d'IP.

L'activation du mode NBMA (*NonBroadcast MultiAccess*) permet au routeur de garder une trace des adresses IP qui émettent les messages PIM Join arrivant par l'intermédiaire de l'interface WAN. Les paquets multicasts devant ainsi être diffusés vers cette interface seront dupliqués en autant de paquets encapsulés dans des trames unicasts.

C'est notamment le cas sur les interfaces série de nos routeurs R1 et R2 :

```
int s 0
ip pim nbma-mode
```

Il est à noter que le mode NBMA consomme de la mémoire puisqu'il garde trace de toutes les adresses IP sources arrivant par l'intermédiaire de l'interface WAN.

De même, le mode *source shortest path tree* requiert davantage de mémoire que le mode *RP shared tree*. En revanche, il réduit la charge réseau et optimise le chemin (il n'est plus besoin de passer par le RP).

Les routeurs Cisco basculent du mode *shared* vers le mode *shortest path* dès le premier paquet émis. Cependant, il est possible de n'activer le basculement qu'à partir d'un certain débit constaté, en d'autres termes, quand cela en vaut la peine.

```
ip pim spt-threshold 64 group-list 1
```

La commande précédente indique que le basculement s'opérera lorsque le flux multicast aura atteint 64 Kbit/s pour les groupes figurant dans l'access-list 1.

Quel protocole choisir ?

Pour l'enregistrement des membres d'un groupe auprès d'un routeur, IGMP est incontournable. Les informations collectées sont ensuite utilisées par les protocoles de routage multicast.

Trois protocoles de routage multicast sont proposés : trois solutions, trois approches différentes.

Critère	DVMRP	MOSPF	PIM
Mise à jour des tables de routage	Diffusion périodique de l'intégralité des tables	Diffusion des modifications par messages d'annonce	Centralisée au niveau du point de rendez-vous (PR)
Diffusion du premier paquet	Inondation du premier paquet	Routé directement vers les membres du groupe	Encapsulé dans un paquet unicast vers le PR, puis redifusé par le PR
Calcul du meilleur chemin (arbre SPT) pour le couple Source/Groupe	Par élagage des routeurs ne désirant pas recevoir le paquet	À la demande, par tous les routeurs recevant le premier paquet	Par remontée de messages Join/Prune de proche en proche jusqu'à la source
Protocole de routage	DVMRP	MOSPF s 'appuie sur OSPF	S'appuie sur les protocoles existants (RIP, OSPF, etc.)
Tunneling entre routeurs non multicasts	Oui	Non : les routeurs MOSPF doivent être contigus	Oui

DVMRP est le protocole le plus ancien et le moins performant (il présente les mêmes défauts que RIP). L'inondation régulière des paquets multicasts ainsi que la diffusion périodique des tables de routage ne font pas de DVMRP un protocole adapté aux grands réseaux.

MOSPF est le protocole le plus performant au sein d'une aire ; il est relativement performant entre aires. Entre systèmes autonomes, l'usage de DVMRP est requis pour le multicast (en plus de BGP pour l'unicast), ce qui peut présenter quelques inconvénients.

PIM est adapté aux grands réseaux et notamment Internet, là où les membres d'un même groupe sont très dispersés. Dans un petit réseau, cet avantage se transforme en inconvénient, car le routeur de point de rendez-vous n'est pas obligatoirement situé sur le meilleur chemin (à moins que le réseau soit très centralisé).

D'autres considérations, non techniques, entrent en jeu :

- Les RFC de DVMRP (datée de 1988) et de PIM (datée de 1998) n'en sont qu'au stade expérimental, alors que celle de MOSPF (datée de 1994) en est à l'état de proposition de standard.
- Cependant, Cisco, le principal fournisseur de routeurs sur Internet, ne prend en charge que PIM, et pour cause : les ingénieurs de la société ont participé à son élaboration mais pas à celle de MOSPF.
- En outre, DVMRP est le plus répandu au sein de MBONE (*Multicast Backbone*, une portion d'Internet expérimentant le multicast). Son utilisation est donc requise, au moins sur le routeur de frontière utilisé en interconnexion avec ce réseau.

En conclusion :

- Si vos routeurs sont de marque Cisco, PIM est le seul choix possible.
- Si votre réseau fonctionne avec OSPF et ne comprend pas de routeurs Cisco, le choix de MOSPF va de soi.
- Utilisez DVMRP entre les systèmes autonomes et pour vous raccorder à MBONE.

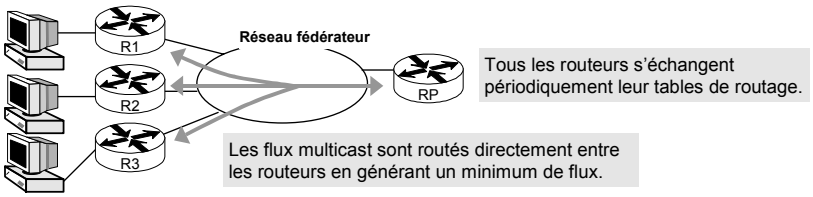
L'étape suivante consiste à définir l'architecture la mieux adaptée au protocole choisi, afin de limiter le trafic généré par les routeurs.

Si plusieurs routeurs sont présents sur le même réseau local, il faut déterminer quel est le meilleur candidat à l'élection du Demandeur. Le processus d'élection dépend du protocole de routage activé. Pour les routeurs DVMRP, le routeur Demandeur IGMP est celui qui a la plus petite adresse IP ou celui dont la métrique est la plus faible. Pour les routeurs MOSPF, il s'agit du routeur désigné OSPF (celui dont la priorité est la plus basse), tandis que pour PIM, il s'agit de celui dont l'adresse IP est la plus haute.

Architecture adaptée au protocole DVMRP

Il existe peu de solutions pour optimiser les flux DVMRP, si ce n'est l'emploi d'un réseau centralisé autour d'un réseau fédérateur.

Figure 15-14.
Optimisation des flux DVMRP.

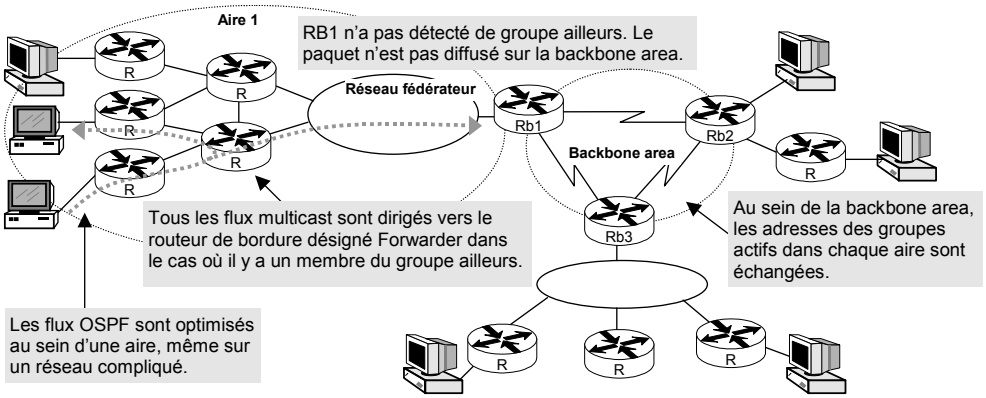


On le voit, DVMRP est adapté à un réseau localisé sur un site ou un campus avec un réseau fédérateur à haut débit.

Architecture adaptée au protocole MOSPF

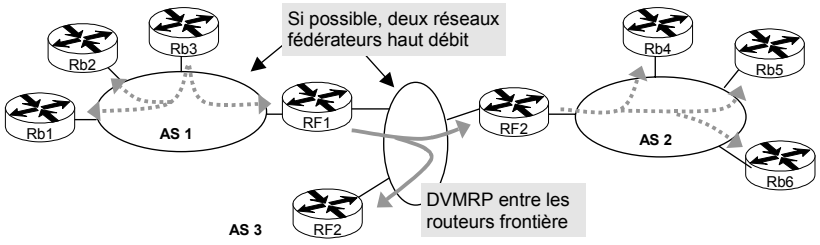
Le routage au sein des aires est bien contrôlé par MOSPF. Entre les aires, le routeur de bordure désigné Forwarder reçoit tous les flux multicasts ; il doit donc être positionné sur un réseau fédérateur à haut débit.

Figure 15-15.
Optimisation des flux MOSPF.



MOSPF est adapté aux grands réseaux mais limité à un seul système autonome. En effet, la concentration des flux entre AS est encore augmentée, à l'image de la concentration sur la backbone area.

Figure 15-16.
Optimisation des flux entre systèmes autonomes.



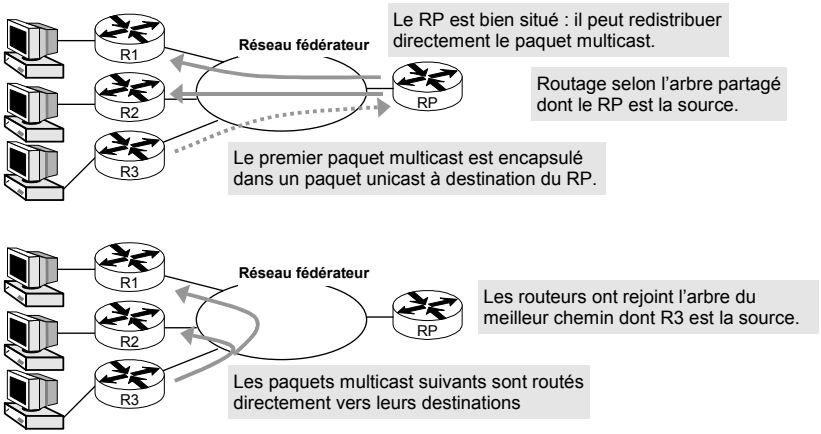
Entre systèmes autonomes, DVMRP doit être configuré en point à point, ou en multipoint sur un réseau fédérateur à haut débit, et ce afin d'optimiser le flux multicast.

Afin de limiter la taille des tables de routage, les routes redistribuées entre DVMRP et MOSPF doivent être agrégées. Les adresses réseau doivent donc être contiguës au sein d'un système autonome (voir chapitre 12).

Architecture adaptée au protocole PIM

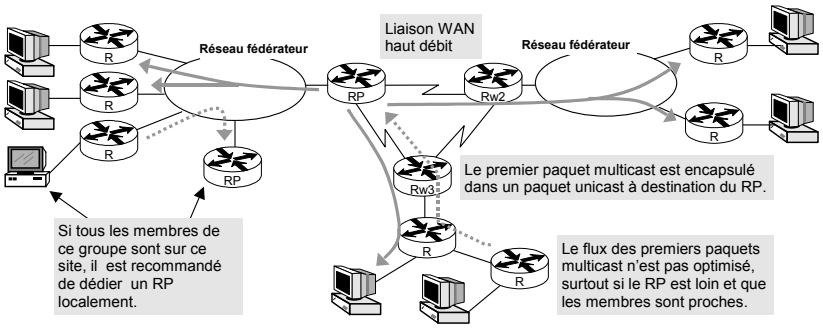
Le routeur point de rendez-vous de PIM doit être situé au centre du réseau, afin d'optimiser la route des premiers paquets multicasts.

Figure 15-17.
Optimisation des flux PIM.



Si les membres d'un groupe sont répartis entre plusieurs sites, la meilleure position du RP se situe sur un routeur WAN.

Figure 15-18.
Flux PIM sur un réseau intersite.



Il est conseillé de paramétrer les routeurs PIM pour basculer rapidement vers l'arbre du meilleur chemin, afin de ne plus passer inutilement par le RP. Les routeurs Cisco basculent dès le premier paquet : la commande *ip pim spt-threshold* ne doit donc pas être utilisée. Le revers de la médaille est un basculement inutile si le flux multicast est sporadique, ce qui entraîne une plus grande consommation de bande passante réseau (pour les messages Join/Prune) et de ressources CPU sur tous les routeurs qui calculent le nouvel arbre du meilleur chemin. En outre, ce dernier utilise beaucoup de ressources mémoire, d'autant plus qu'un arbre par couple adresse source/adresse de groupe est nécessaire.

Par ailleurs, des messages Bootstrap circulent entre les routeurs point de rendez-vous, ajoutant encore des flux de gestion propres à PIM.

Tout est donc affaire de dosage en fonction du nombre d'utilisateurs et du type de trafic. Du fait de la complexité du protocole PIM, la conception de l'architecture réseau n'en est que plus difficile.

Contrôler la diffusion sur son réseau

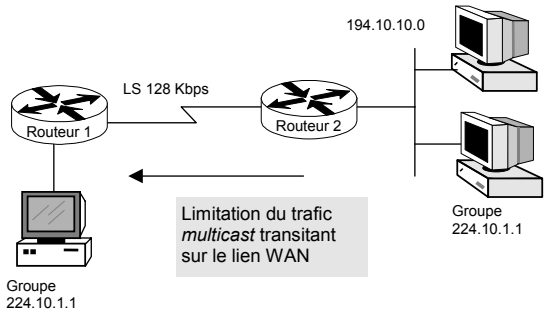
Les flux multicasts, tels que la vidéo, peuvent générer un volume de données important. Afin de contrôler ces flux, il est possible de filtrer les groupes multicasts sur chaque interface. Cela évite également de voir fleurir des groupes un peu partout sans que l'administrateur n'en soit tenu informé.

```
access-list 90 225.20.2.2 0.0.0.0
access-list 90 224.10.1.1 0.0.0.0
interface ethernet 0
ip igmp access-group 90
```

L'access-list 90 contient la liste des adresses de groupe qui seront autorisées à être diffusées sur l'interface Ethernet 0.

Il est également intéressant de contrôler le débit généré par les membres d'un groupe, pour ne pas pénaliser les autres applications et, également, pour ne pas inonder votre réseau.

Figure 15-19.
Contrôle du trafic multicast.



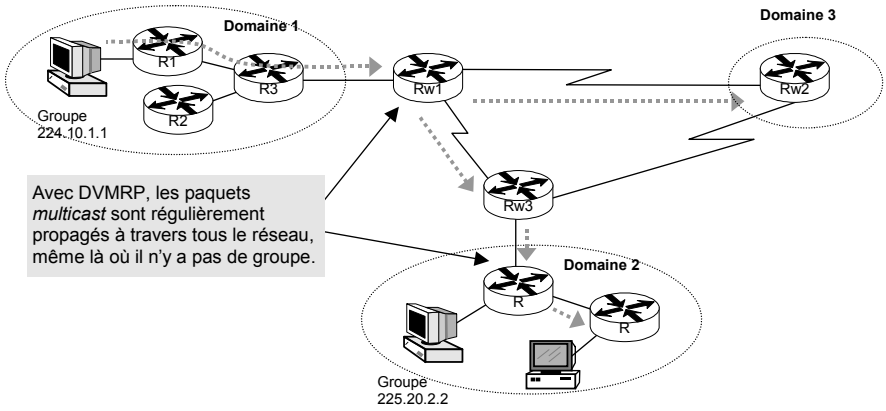
Dans cet exemple, les membres indiqués dans l'access-list 90 et qui émettent au sein des groupes précisés dans l'access-list 80 sont limités à un débit de 64 Kbit/s sur la liaison WAN. Cette limite peut être positionnée en entrée (*in*) ou en sortie (*out*) de l'interface.

```
interface serial 1
ip multicast rate-limit out group-list 80 source-list 90 64
access-list 80 permit 0.0.0.0 255.255.255.255
access-list 90 permit 194.10.10.0 0.255.255.255
ip multicast rate-limit in group-list 80 source-list 90 64
```

Dans l'exemple précédent, tous les multicasts issus du réseau 194.10.10.0 seront limités à 64 Kbit/s, laissant une bande passante de 64 Kbit/s disponible pour les autres applications.

Reprenons l'exemple de notre réseau intersite.

Figure 15-20.
Contrôler la diffusion.

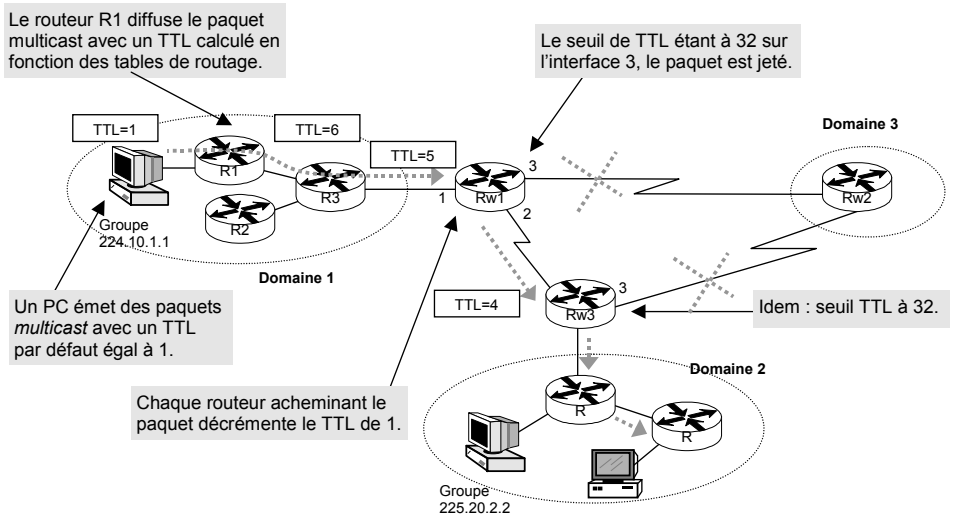


Si DVMRP est utilisé, tous les premiers paquets multicasts émis par une source pour un groupe donné sont diffusés à travers l'ensemble du réseau (tant qu'il y a des routeurs DVMRP). Comme on l'a vu, DVMRP est obligé de répéter cette procédure régulièrement afin de tenir compte des changements de topologie et d'appartenance à un groupe.

C'est également le cas de MOSPF, dans la backbone area, et de PIM, au niveau des réseaux auxquels est connecté le routeur point de rendez-vous.

Les routeurs permettent de limiter la propagation de ces paquets en contrôlant leur TTL (*Time To Live*). Ce mécanisme impose qu'un paquet IP ne soit pas routé si son TTL est égal à 1. Le mécanisme proposé permet de définir un seuil supérieur en dessous duquel le paquet ne sera pas routé.

Figure 15-21.
Limitation de la diffusion des paquets.



Seuls les paquets dont le TTL est supérieur au seuil de 32 seront routés par les routeurs Rw1 et Rw3 sur leur interface série 3. La commande Cisco est la suivante :

```
int s3
ip multicast ttl-threshold 32
```

La valeur par défaut est 0

En positionnant différents seuils à certains points stratégiques du réseau, il est possible de définir des domaines de diffusion. On peut ainsi limiter la portée du point de rendez-vous PIM ou la propagation des paquets *via* des routeurs DVMRP.

Par exemple, MBONE utilise la convention suivante :

TTL	Périmètre
0	Restreint à la machine (paquet boucle en local)
1	Restreint au même réseau local
32	Site
64	Région
128	Continent
255	Pas de restriction de diffusion

Des mécanismes propres à IGMP sont également prévus pour limiter le trafic local :

- Les rapports sont envoyés avec un TTL = 1, ce qui permet de supprimer le paquet rapidement : le rapport est censé être adressé uniquement au routeur du réseau local.
- Les rapports sont envoyés un à un pour chaque groupe d'appartenance, avec un intervalle de temps aléatoire.
- Au niveau du routeur, les temporisateurs *(timers)* sont armés indépendamment pour chaque groupe identifié.

Les groupes de paquets compris entre 224.0.0.0. et 224.0.0.255 ne sont jamais routés : de tels paquets transitent d'une machine vers un routeur, d'un routeur vers un autre routeur, et ce quelle que soit la valeur du TTL (en principe égale à 1).

Limiter les flux multicasts sur le réseau local

Une fois arrivés sur le réseau local, les flux multicasts sont diffusés par tous les commutateurs sur l'ensemble de leurs ports. Afin de limiter ces flux aux seuls ports connectant des PC abonnés, la norme 802.d spécifie le protocole **GMRP** (*GARP Multicast Registration Protocol*).

Les deux attributs échangés par les GID *via* le protocole GIP (cf. chapitre 3) sont l'adresse MAC multicast sur 48 bits ainsi qu'un type de service permettant au routeur d'indiquer vouloir recevoir tous les flux multicasts enregistrés ou non.

L'activation du protocole se réalise simplement comme suit :

```
set gmrp enable
set gmrp fwdall enable 1/1
```

La deuxième commande demande au commutateur de réaliser une copie de tout le trafic multicast sur le port 1/1, sur lequel est connecté un routeur ne supportant pas GMRP.

Certains constructeurs implémentent un protocole propriétaire, fonctionnellement équivalent à GMRP, sous le libellé **IGMP snooping**. Avec ce protocole, le commutateur est à l'écoute des requêtes IGMP (il opère donc une analyse au niveau 3), ce qui lui permet d'activer ou de bloquer les adresses MAC multicasts correspondantes. Dans la mesure du possible, le protocole GMRP doit lui être préféré.

16

La téléphonie et la vidéo sur IP

On entend souvent la conversation suivante entre un sceptique et un convaincu à propos de la téléphonie sur IP :

— Est-ce que ça marche ?

— Oui.

— Mais, est-ce que ça marche *bien* ? Est-ce que la qualité de service est bonne ?

La réponse est : oui, ça marche, et même très bien !

Maintenant que vous avez préparé votre réseau au multimédia, vous pouvez, en effet, mettre en place des solutions de téléphonie et de visioconférence.

Dans ce chapitre, vous apprendrez ainsi :

- à interconnecter des PABX *via* IP ;
- à raccorder un VoIP au réseau téléphonique classique ;
- le fonctionnement des protocoles multimédias.

Les protocoles H.323 de l'ITU-T sont ici expliqués dans le détail, car tous les produits du marché reposent sur cette norme. Mais, à l'avenir, cette dernière pourrait être remplacée par son concurrent SIP (*Session Initiation Protocol* — RFC 2543) promu par l'IETF. Les deux normes pourraient éventuellement cohabiter, SIP se posant en interface applicative, donc de plus haut niveau que H.323.

Présentation des protocoles multimédias

Historiquement le premier, le RTC (réseau téléphonique commuté) est le réseau qu'utilise toujours notre bon vieux téléphone. La version numérique de celui-ci, le RNIS (réseau numérique à intégration de service), est venue s'ajouter par la suite.

L'évolution des technologies a ensuite permis d'envisager d'autres supports pour transporter la voix : ATM, Frame Relay et, aujourd'hui, les réseaux IP (votre intranet et Internet). Toujours grâce à l'évolution des technologies, il est ensuite devenu possible d'y ajouter l'image, la vidéo et le partage de données.

Accompagnant le mouvement, les normes H.32x de l'ITU-T (*International Telecommunication Union — Telecommunication Standardization Sector*), sur lesquelles reposent le RTC et le RNIS, ont été adaptées à ces nouveaux supports de transmission et aux nouveaux besoins multimédias. Cette continuité dans la standardisation des protocoles téléphoniques permet ainsi d'assurer une cohabitation et une transition en douceur.

Réseau	Caractéristiques du réseau	Norme ITU-T
RNIS (réseau numérique à intégration de service)	Commutation de circuit Voix numérique	H.320
ATM (*Asynchronous Transfer Mode*)	Commutation de cellules Voix numérique	H.321
Réseaux WAN avec qualité de service	Frame Relay Voix numérique	H.322
Réseau VoIP (*Voice over Internet Protocols*)	Commutation de paquets IP Voix numérique	H.323
RTC (réseau téléphonique commuté)	Commutation de circuit Voix analogique	H.324

Ces normes décrivent un cadre général de fonctionnement et préconisent l'utilisation d'autres normes de l'ITU-T et de l'IETF (*Internet Engineering Task Force*) en fonction du support de transmission (IP, RTC, RNIS, etc.). Elles décrivent les codages audio et vidéo, l'adaptation au support de transmission, la signalisation, etc.

	H.320	H.321	H.322	H.323	H.324
Réseau	RNIS	ATM	WAN QoS	VoIP	RTC
Codec vidéo	H.261 H.263	H.261 H.263	H.261 H.263	H.261 H.263	H.261 H.263
Codec audio	G.711 G.722 G.728	G.711 G.722 G.728	G.711 G.722 G.728	G.711 G.722 G.728 G.723 G.729	G.723
Adaptation	H.221	H.222	H.221	H.225	H.223
Signalisation	H.242 H.230	H.242	H.242 H.230	H.245 H.225	H.245
Conférences multipoint	H.231 H.243	H.231 H.243	H.231 H.243	H.323	
Données	T.120	T.120	T.120	T.120	
Couche de transport	I.400	I.363 AAL	I.400	TCP/IP	Fils

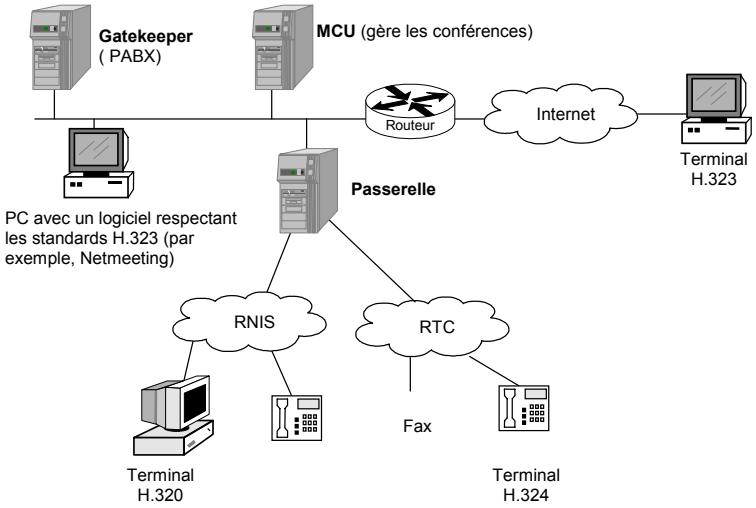
Ainsi, la norme **H.323** décrit un cadre général de fonctionnement pour un terminal multimédia fonctionnant sur IP. Elle implique l'utilisation d'autres normes décrites, cette fois-ci, par des RFC (*Request For Comments*) de l'IETF. Il s'agit notamment de **RTP** (*Real-time Transport Protocol*) qui décrit le format d'un paquet transportant des données (un échantillon sonore, une portion d'un écran vidéo, etc.).

Les composants d'un système H.323

La norme H.323 décrit le fonctionnement et l'interaction de quatre entités :

- un **terminal** qui supporte la voix et, optionnellement la vidéo et les données ;
- une passerelle qui permet l'interconnexion avec les autres réseaux H.32x tels que le RTC ;
- un serveur de conférence, appelé MCU (*Multipoint Control Unit*) ;
- un PABX IP, appelé **gatekeeper**, qui offre le routage, la conversion d'adresses ainsi que la coordination de l'activité de toutes les entités H.323.

Figure 16-1.
Entités décrites par la norme H.323.



La **passerelle** permet d'interconnecter le réseau téléphonique IP à d'autres réseaux tels que le RTC ou le RNIS. Elle assure la conversion des codec audio et vidéo, de la signalisation et du support de transmission.

Le **MCU** contrôle l'entrée et la sortie des participants à la conférence, rediffuse le flux entre émetteurs et récepteurs en minimisant le trafic réseau, et assure l'éventuelle conversion de codec (par exemple, si un participant est équipé d'un écran de moins bonne qualité que les autres). Même si son utilisation n'est pas nécessaire lorsque les terminaux disposent des fonctions multicasts et de négociation de paramètres, un MCU permet cependant de décharger le terminal de certaines fonctions.

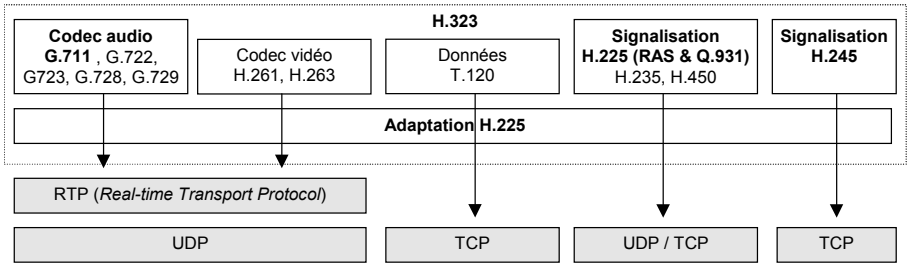
Le **gatekeeper** gère des services pour les entités de sa **zone** de couverture, tels que :

- la gestion des tables de correspondance entre les noms des terminaux (le nom de la personne), un numéro de téléphone E.164, une adresse e-mail et les adresses IP ;
- le contrôle d'admission : autorisation ou non de l'ouverture d'une communication ;
- la gestion de la bande passante sur le réseau, le nombre maximal de conférences, d'utilisateurs, etc. ;
- la localisation des passerelles et des MCU ;
- le routage des appels au sein de la zone et entre les zones.

Le **terminal**, enfin, se présente sous des formes diverses. Il peut prendre l'aspect d'un téléphone classique équipé d'une interface Ethernet dont la prise RJ45 est connectée à un commutateur de notre réseau local. Il peut également être un PC doté d'un logiciel de communication tel que Netmeeting de Microsoft. Le terminal intègre alors l'image, en plus de la voix, ce qui permet aux utilisateurs d'établir des visioconférences directement entre eux. Le PC devient ainsi un terminal multimédia traitant la voix, les données et l'image.

Le schéma suivant précise ce que la norme H.323 couvre (carrés blancs) et les éléments obligatoires (en gras). Les éléments grisés correspondent aux protocoles spécifiés par les RFC sur lesquelles s'appuie la norme H.323 pour transporter les informations sur un réseau IP.

Figure 16-2.
Protocoles utilisés par la norme H.323.



La norme H.235 recouvre les fonctions de sécurité (authentification, intégrité des données, chiffrement, etc.). La norme H.450 décrit, quant à elle, les services complémentaires (identification de l'appelant, transfert d'appel, rappel sur occupation, etc.).

Selon ce schéma, un terminal H.323 doit donc comprendre :

- un codec audio (au moins G.711) ;
- optionnellement, un codec vidéo (au moins H.261) ;
- optionnellement, l'échange des données (à la norme T.120) ;
- les protocoles de signalisation RAS, Q.931 et H.245 ;
- une couche H.225 qui assure l'adaptation des protocoles de l'ITU à ceux de l'IETF ;
- les protocoles TCP, IP, UDP et RTP.

Par ailleurs, une entité H.323 comprend deux composants fonctionnels :

- un **MC** (*Multipoint Controler*) pour négocier, *via* H.245, les niveaux de service entre les terminaux et les ressources utilisées au sein d'une conférence (canaux audio et vidéo, adresses multicasts, etc.) ;
- optionnellement, un **MP** (*Multipoint Processor*) pour traiter les flux multimédias (mixage, synchronisation de la parole et des images, diffusion, chiffrement, etc.).

En pratique, toutes les entités H.323 intègrent un MC pour participer à des conférences, tandis qu'un MCU intègre les deux composants pour gérer lesdites conférences.

L'établissement d'une communication

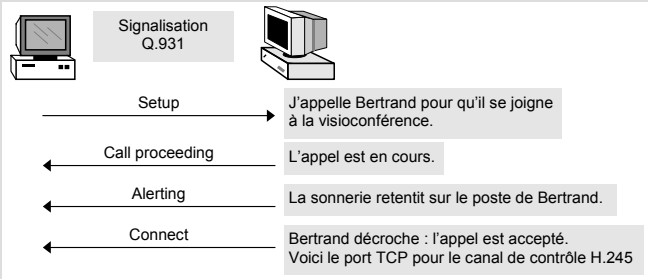
Pour communiquer entre elles, les entités H.323 ouvrent des canaux logiques (sessions TCP, UDP et RTP) soit dédiés à la signalisation, soit dédiés au transport des flux multimédias :

- un canal de signalisation RAS (*Registration, Admission and Status*) qui permet à un terminal de s'enregistrer auprès du gatekeeper de sa zone ;
- un canal de signalisation d'appel Q.931 (numérotation, sonnerie, etc.) qui permet à un terminal d'en appeler un autre ;
- un canal de contrôle H.245 qui permet d'échanger les fonctionnalités supportées par les entités (codec audio et vidéo, T.120), d'ouvrir et de fermer les canaux audio, vidéo et données ;
- les canaux audio, vidéo et données.

L'utilisation des services d'un gatekeeper, et donc l'ouverture d'un canal RAS, est optionnelle. Les entités H.323 peuvent, en effet, communiquer directement entre elles.

LA SIGNALISATION D'APPEL Q.931

Une entité désirant en appeler une autre doit ouvrir un **canal de signalisation d'appel Q.931** (dont l'utilisation au sein de H.323 est décrite par la norme H.225). Il s'agit d'une connexion sur le port TCP 1720 qui véhicule des messages Q.931.



Un message Q.931 comporte un en-tête, suivi d'un certain nombre d'éléments d'information obligatoires ou non. Par exemple, le message " *Connect* " contient, en plus de l'en-tête, les éléments d'information suivants : " *Bearer capability* ", " *Connect-UUIE* ", et optionnellement " *Progresss indicator* ", " *Notification Indicator* ", " *Facility* ", etc.

8 bits	16 bits	8 bits
Discriminator (=08)	Référence d'appel (identifiant local)	Type de message
Élément d'information	Champs de longueurs variables	
Élément d'information	Champs de longueurs variables	
etc.	etc.	

• • •

LA SIGNALISATION D'APPEL Q.931 (SUITE)

La signification des champs de l'en-tête Q.931 est la suivante :

- " Type de message " : 1 = Alerting, 2 = Call proceeding, 5 = Setup, 7 = Connect, etc.
- " Éléments d'information " : 4 = Bearer capability, 204 = numéro de téléphone de l'appelant, 224 = numéro de téléphone à appeler, 130 = temps de transit, etc.

Par exemple, le format de l'élément d'information " **Bearer Capability** " est le suivant :

8 bits	16 bits		8 bits
Bearer capability (=04)	Longueur (= 5 ou 6)	Transfer capbility	Transfer rate
Multiplier rate	Layer 1 protocol		

La signification des champs est la suivante :

- " Transfer capability " indique les fonctionnalités du terminal (audio, vidéo).
- " Transfer rate " indique le débit (mode paquet ou circuit, de 64 Kbit/s à 1 920 Kbit/s).
- " Multiplier rate " est uniquement présent si le champ précédent indique un débit " multirate ".
- " Layer 1 protocol " : G711 (A-law ou μ-law) ou H.225/H.245 (vidéophone H.323).

Voici un autre exemple montrant le format de l'élément d'information " **Called Number** " :

8 bits	16 bits	
Called number (=204)	Longueur	Type / Plan de num
Numéro de téléphone (2 à 131 octets)		

La signification des champs est la suivante :

- Type de numéro (3 bits) : 1 = international, 2 = national.
- Plan de numérotation (5 bits) : 1 = E.164, 8 = national, 9 = privé.
- Numéro de téléphone : par exemple 1#331836 ou 440553.

L'exemple suivant montre l'élément d'information utilisateur **Connect-UUIE** (*User-to-User Information Element*) pour le message " Connect " décrit selon la syntaxe ASN.1 (qui est utilisée dans toutes les normes de l'ITU-T) :

h323-message-body connect :{

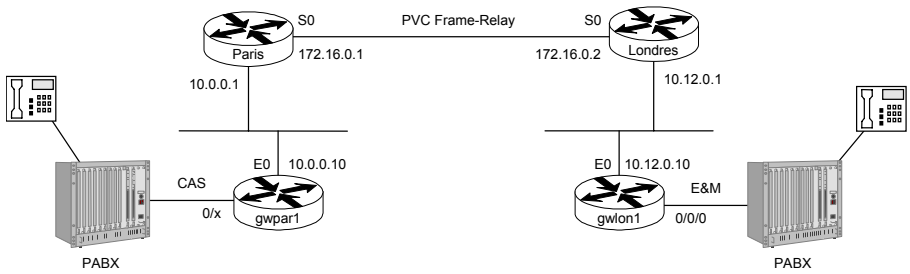
protocolIdentifier	{ 0 0 8 2250 0 3 },
h245Address	ipAddress : { ip '**0A0C0006**'h port 011026 } ,
destinationInfo	{ terminal { }, mc FALSE, undefinedNode FALSE },
conferenceID	5A 1E 55 55 8A B0 CF 72 00 00 00 00 17 00 86 B4,
callIdentifier	CallIdentifier,
h245SecurityMode	H245Security (champ optionnel),
tokens	Sequence of ClearToken (champ optionnel),
cryptoTokens	Sequence of CryptoH323Token (champ optionnel),
fastStart	Sequence of octet string (champ optionnel) }

Interconnecter les PABX *via* IP

Lors de l'introduction d'une nouvelle technologie, la première contrainte est bien souvent de prendre en compte l'existant, en l'occurrence les PABX de l'entreprise.

Dans notre exemple, nous souhaitons profiter de notre réseau WAN pour acheminer les appels internes entre deux sites sur notre réseau privé (on Net), mais également pour passer des appels vers l'extérieur (off Net), de manière à bénéficier des meilleurs coûts (par exemple, Paris-Angleterre en sortant à Londres pour le coût d'un appel local ou national).

Figure 16-3.
Utilisation des passerelles pour interconnecter les PABX via IP.



Nous avons choisi des routeurs Cisco comme passerelles, mais cela aurait pu être des PC équipés de cartes de même nature que celles des routeurs.

La première tâche est de configurer les liaisons entre les passerelles et les PABX, ce qui est réalisé de la manière suivante :

Configuration d'une liaison numérique CAS entre le PABX parisien et la passerelle gwpar1

```
#gwpar1
controler E1 0
 framing crc4
 linecode ami
 mode cas
 voice-group 1 timeslot 1-6 e&m-immediate
voice port 0/1
voice port 0/2
voice port 0/3
...
```

6 canaux sont utilisés

```
#gwlon1
voice-port 0/0/0
signal immediate
 operation 4-wire
 type 2
voice-port 0/0/1
 signal immediate
 operation 4-wire
 type 2
```

Configuration de deux liaisons analogiques E&M entre le PABX londonien et la passerelle gwlon1

Nous n'entrerons pas dans les détails, car ce domaine dépasse le cadre de cet ouvrage. Il faudrait, en effet, expliquer le fonctionnement des signalisations utilisées par les PABX classiques.

L'étape suivante consiste à configurer les liaisons entre les deux passerelles. Il s'agit cette fois de transporter la voix sur IP à l'aide des protocoles H.323.

Les utilisateurs des postes téléphoniques connectés aux PABX manipulent des numéros de téléphone au format E.164 (voir chapitre 11). Il faut donc associer des préfixes (c'est-à-dire la partie située le plus à gauche du numéro de téléphone) à des interfaces PABX d'un côté et à des interfaces IP de l'autre.

Passerelle	Préfixe	Correspondant
gwpar1 port 0/1-3	33	Numéro sur quatre chiffres
gwlon1 port 0/0/0	44	Numéro sur quatre chiffres

Chez Cisco, la connexion PABX est référencée par l'acronyme POTS (*Plain Old Telephone Service*), et la connexion IP est appelée VoIP (*Voice over IP*). Une connexion au sens large est appelée « dial-peer ».

```
# gwpar1
dial-peer voice 33 pots
port 0/1
destination-pattern 33....

dial-peer voice 44 voip
destination pattern 44....
session-target ipv4:10.12.0.10
codec g729r8
```

Tous les numéros d'appels commençant par 33 sont envoyés sur cette interface.

Quatre points indiquent que quatre numéros doivent suivre le préfixe 33.

Tous les numéros d'appel commençant par 44 sont envoyés vers la passerelle de Londres en H.323.

Ce codec utilise 8 Kbit/s de bande passante (cf. chapitre 13).

Sur la passerelle de Londres, nous devons disposer d'une configuration symétrique :

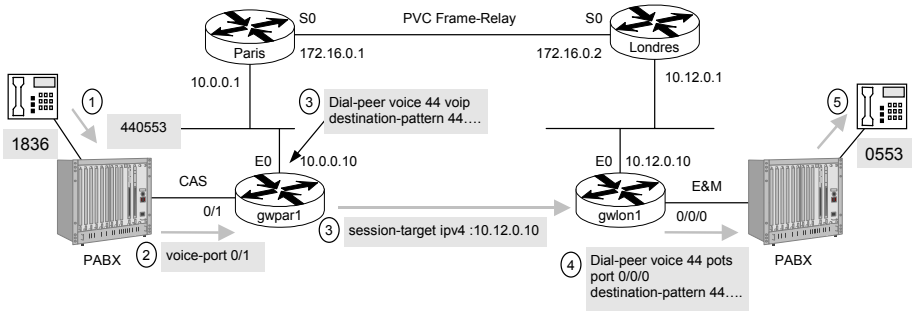
```
# gwlon1
dial-peer voice 44 pots
port 0/0/0
destination-pattern 44....

dial-peer voice 33 voip
destination pattern 33....
session-target ipv4:10.0.0.10
codec g729r8
```

Numéro arbitraire ayant uniquement une signification pour la passerelle locale. Nous avons choisi une notation mnémotechnique.

Active les protocoles H.323

Adresse IP de la passerelle parisienne.

Figure 16-4.
Établissement d'une communication VoIP via des passerelles.



Le correspondant de Paris compose le « 440553 ». Le PABX route ce numéro vers la liaison CAS établie avec la passerelle parisienne. Celle-ci compare le numéro à ses « dial-peer » et constate que le préfixe « 44 » correspond à une connexion H.323. Elle envoie alors la demande de connexion sur IP (signalisation Q.931) à la passerelle londonienne qui compare à son tour le numéro présenté à ses « dial-peer ». Celle-ci constate alors que le préfixe « 44 » correspond à un port physiquement connecté au PABX et y envoie le numéro. Le PABX déclenche la sonnerie du téléphone du correspondant recherché.

Nous souhaitons également acheminer des appels en off Net sans que les utilisateurs français changent leurs habitudes de numérotation : « 00 » pour l'international, « 44 » pour l'Angleterre, suivis de dix chiffres. De même, les utilisateurs anglais numérotent comme suit : « 00 » pour l'international, « 33 » pour la France, suivis de neuf chiffres. Il suffit de configurer autant de « dial-peer » qu'il y a de préfixes :

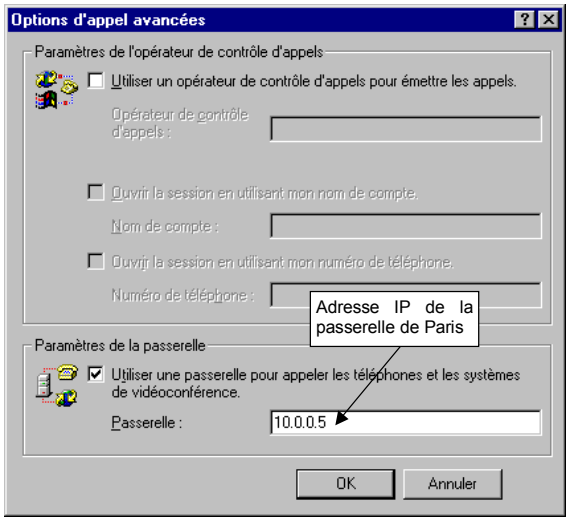
```
# gwpar1
dial-peer voice 0033 pots
port 0/4-6
destination-pattern 0033.........
dial-peer voice 0044 voip
destination pattern 0044..........
session-target ipv4:10.12.0.10
codec g729r8
```

Les neuf points correspondent aux neuf chiffres attendus.

Dix points = dix chiffres.

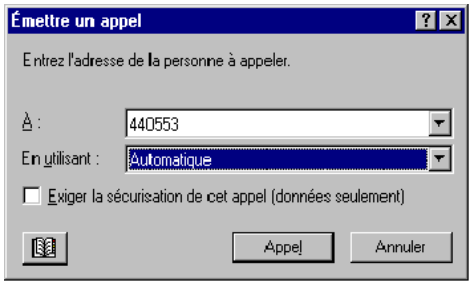
Un PC équipé de Netmeeting peut également profiter des services de nos passerelles pour appeler un correspondant situé sur le RTC classique. Il suffit d'indiquer son adresse IP dans le menu :

« Outil→Option→Appel Avancé ».



Il suffit ensuite de composer le numéro de téléphone de notre correspondant à Londres.

Nous utilisons ici la numérotation abrégée configurée dans nos passerelles.

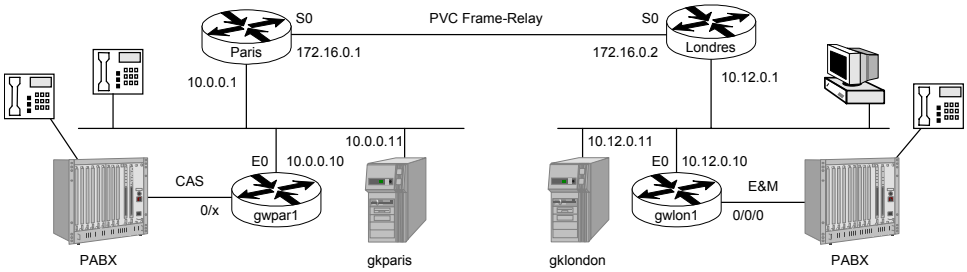


Mettre en place un gatekeeper

La complexité de notre installation augmente avec le nombre de sites à interconnecter. Nous devons de plus supporter des téléphones IP et des PC équipés de logiciels téléphoniques, tels que Netmeeting. Le réseau VolP prenant de l'importance, nous devons assurer des services plus évolués, comme le routage automatique, les statistiques ou encore la facturation. Nous devons donc assurer non seulement la reprise de l'existant mais également le développement de la voix tout IP.

La solution est de mettre en place un PABX IP, appelé **gatekeeper** dans la terminologie H.323. Nous choisissons donc d'en installer un par site, de marque Cisco (encore !), et de reprendre les deux sites précédents comme exemple.

Figure 16-5.
Utilisation des gatekeepers.



Afin de simplifier les configurations et la vie des utilisateurs, nous choisissons également d'utiliser les services d'un DNS (*Domain Name System*) privé, dont la mise en œuvre est expliquée au chapitre 7. Le DNS permet de manipuler des noms à la place des adresses IP en assurant la résolution de l'un vers l'autre.

Dès lors, la configuration des gatekeepers est la suivante :

```
#gkparis
gatekeeper
zone local gkparis fr.intranet
zone prefix gklondon 44....
zone prefix gklondon 0044..........
gw-type-prefix 1# default-tech
```

Active H.323, la fonction de gatekeeper et la signalisation RAS.

Le gatekeeper gère les terminaux IP situés dans le domaine DNS indiqué.

Le gatekeeper route les appels préfixés 44 et 0044 vers Londres.

Le préfixe technique, indiqué par la commande « gw-type-prefix » — à ne pas confondre avec le préfixe de numérotation — permet de désigner explicitement les services d'une passerelle, par exemple une passerelle offrant le routage onNet, une autre le routage offNet, le fax, etc. Par convention, le préfixe technique se termine par un dièse. Dans notre exemple, si aucun préfixe technique n'est indiqué dans le numéro appelé, l'appel sera routé vers la passerelle supportant le préfixe technique 1#.

Le gatekeeper de Londres est configuré de manière symétrique :

```
#gklondon
gatekeeper
zone local gklondon uk.intranet
zone prefix gkparis 33....
zone prefix gkparis 0033...........
gw-type-prefix 1# default-tech
```

Notre DNS privé est configuré avec un sous-domaine par pays, et nous installons un gatekeeper par zone.

Le préfixe technique par défaut est 1#.

Le routage des appels est désormais réalisé entre gatekeepers. Dans chacune des zones, les passerelles doivent par conséquent s'enregistrer auprès de leur gatekeeper de rattachement :

```
#gwpar1
gateway

int e0
h323-gateway voip interface

h323-gateway voip h323-id gwpar1@fr.intranet
h323-gateway voip id gkparis multicast
h323-gateway voip tech-prefix 1#
```

Active H.323 et la signalisation RAS.

Désigne explicitement l'interface Ethernet pour les fonctions H.323.

Nom de la passerelle dans la zone fr.intranet contrôlée par un gatekeeper

Recherche le gatekeeper gkparis à l'aide de la signalisation RAS dans sa version multicast.

À l'aide de ces commandes, la passerelle parisienne numéro 1 va rechercher le gatekeeper appelé « gkparis » et s'y enregistrer avec le préfixe technique 1# sous le nom « gwpar1 » dans la zone « fr.intranet ».

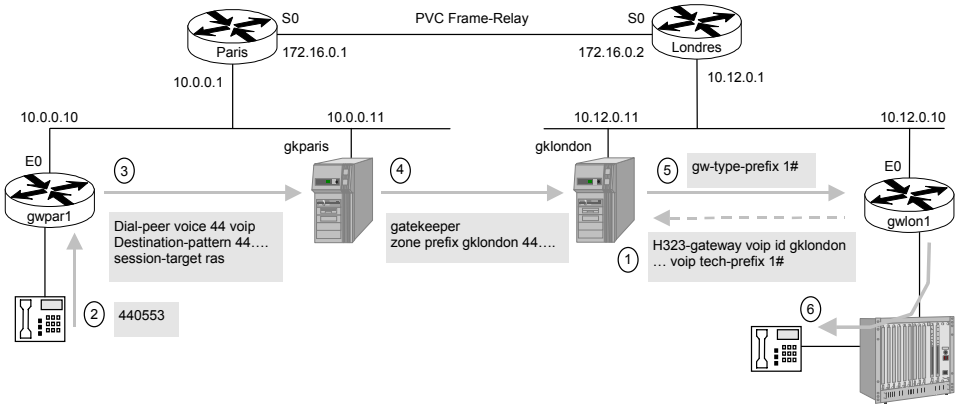
Toujours sur nos passerelles, la configuration des « dial-peer » change légèrement par rapport à une architecture sans gatekeeper :

```
dial-peer voice 44 voip
destination-pattern 44....
tech-prefix 1#
session-target ras
```

Optionnel : ajoute le préfixe 1# au numéro appelé, afin de désigner explicitement un groupe de passerelles.

Utilise les services du gatekeeper qui lui donnera l'adresse IP de la passerelle de Londres.

Lorsque ce « dial-peer » sera invoqué, il enverra l'appel au gatekeeper, qui le routera conformément à la configuration précédente.

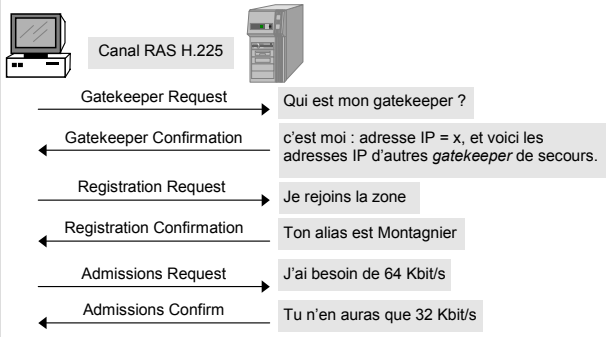


LE CANAL DE SIGNALISATION RAS (H.225)

La signalisation RAS (*Registration, Admission and Status*) permet à toutes les entités H.323 (un terminal, un MCU ou une passerelle) de communiquer avec un **gatekeeper** afin de bénéficier de ses services.

Les messages de **découverte** sont envoyés dans des paquets UDP, port 1718, soit à destination du gatekeeper s'il est connu (son adresse ayant été configurée manuellement), soit à destination du groupe multicast 224.0.1.41. Si plusieurs gatekeepers répondent, l'entité en choisit un.

L'étape suivante consiste à rejoindre la **zone** contrôlée par un gatekeeper en **s'enregistrant** auprès de ce dernier. Le canal RAS utilise maintenant des paquets UDP unicasts port 1719.



Désormais, l'entité H.323 s'adresse au gatekeeper pour appeler tous ses correspondants en les désignant par leur nom ou leur numéro de téléphone. Le gatekeeper se charge de convertir ce nom en une adresse IP, et de router l'appel :

- soit directement vers le correspondant si celui-ci se trouve dans la même zone ;
- soit vers un autre gatekeeper si le correspondant se trouve dans une autre zone ;
- soit vers une passerelle si le correspondant n'utilise pas un terminal IP (mais un téléphone, par exemple).

• • •

LE CANAL DE SIGNALISATION RAS (H.225 – SUITE)

Plusieurs types de messages peuvent être échangés : Découverte, Enregistrement/Annulation, Admission (demande d'autorisation d'utiliser le réseau), Changement de bande passante (demandé par le terminal), Localisation (conversion d'adresses), Statut et Information sur les ressources disponibles (pour les passerelles). Ces types regroupent 18 messages différents.

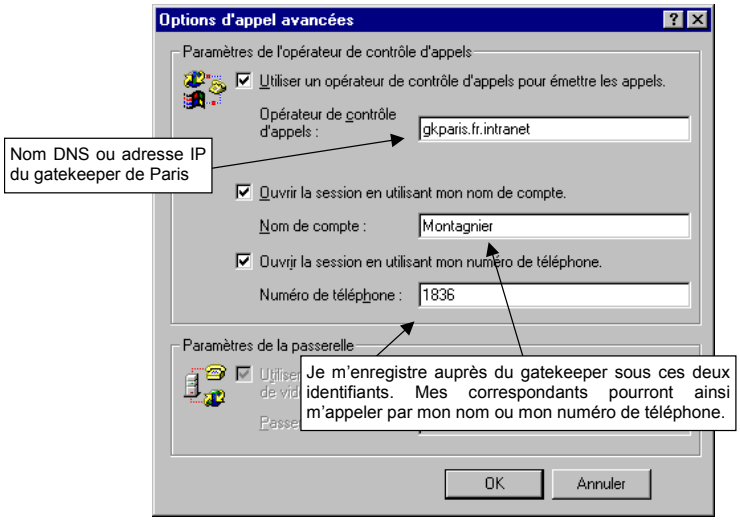
Voici, à titre d'exemple, le format, en syntaxe ASN.1, d'une demande d'enregistrement d'une passerelle auprès d'un gatekeeper :

RasMessage ::= registrationRequest : {
requestSeqNum 037001,
protocolIdentifier { 0 0 8 2250 0 3 },
discoveryComplete TRUE,
callSignalAddress { ipAddress : { ip '**A00C00A**'h, port 01720 } },
rasAddress { ipAddress : { ip A0 0C 00 0A, port 04520 } },
terminalType { gateway { protocol { voice : { supportedPrefixes { { prefixe 164 :"**1#**" } } } } } }, mc FALSE, undefinedNode FALSE },
terminalAlias { h323-ID :"gwlon1.fr.intranet" },
gatekeeperIdentifier { "gklondon.fr.intranet" },
endpointVendor { vendor { t35CountryCode 0181, t35Extension 00, manufacturerCode 018 } } }

ITU-T, recommandation série H, 225.0, (0), version 3

Q.931

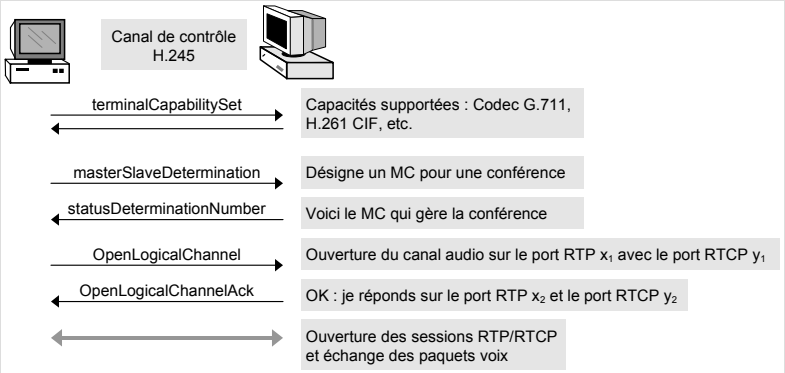
Un PC équipé de Netmeeting peut également utiliser les services d'un de nos gatekeepers. Il suffit d'indiquer son adresse IP, ou mieux, son nom DNS, dans le menu « Outil→Option→Appel Avancé ».



LE CANAL DE CONTRÔLE H.245

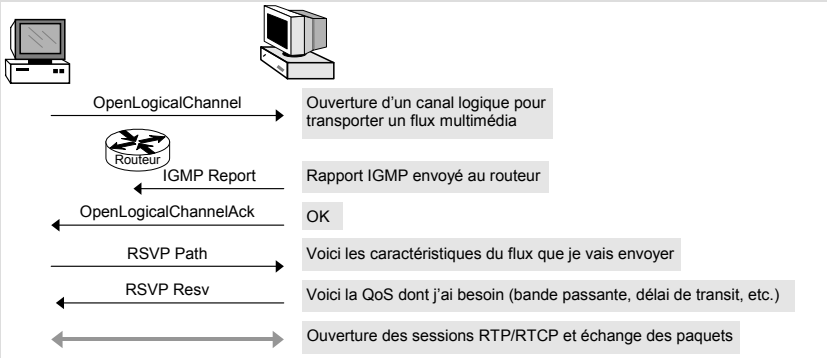
La dernière étape dans l'établissement d'une communication H.323 est l'ouverture du canal de contrôle H.245. Les messages échangés permettent de négocier les fonctions prises en charge par les terminaux (*Terminal Capability*) : choix des codec audio et vidéo, de la résolution d'image, etc., puis d'affecter dynamiquement les ports UDP supportant les **canaux audio et vidéo**.

Si une conférence à trois ou plus est demandée, la procédure désigne le MC (celui d'un terminal, ou le MCU s'il existe) qui sera responsable de la gérer.

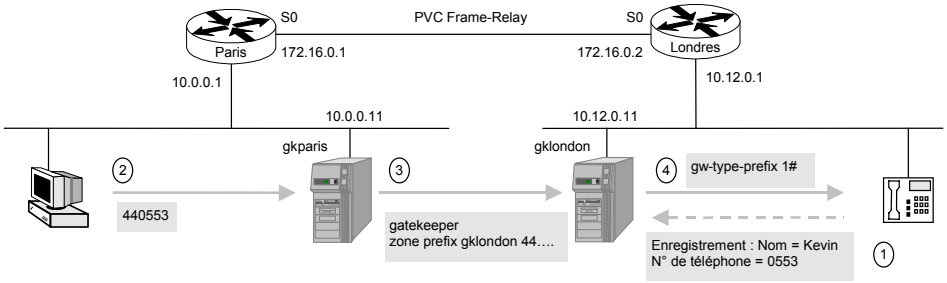


Le canal de contrôle H.245 véhicule ensuite les commandes permettant :

- d'ouvrir et de fermer les canaux audio et vidéo, c'est-à-dire les sessions **RTP** (*Real-time Transport Protocol*) ;
- de gérer les entrées et les sorties dans les conférences ;
- de gérer la **qualité de service**, grâce aux informations données par **RTCP** (*RTP Control Protocol*) et en faisant appel aux services de **RSVP** (*Resource Reservation Protocol*) pour réserver la bande passante nécessaire sur le réseau IP.

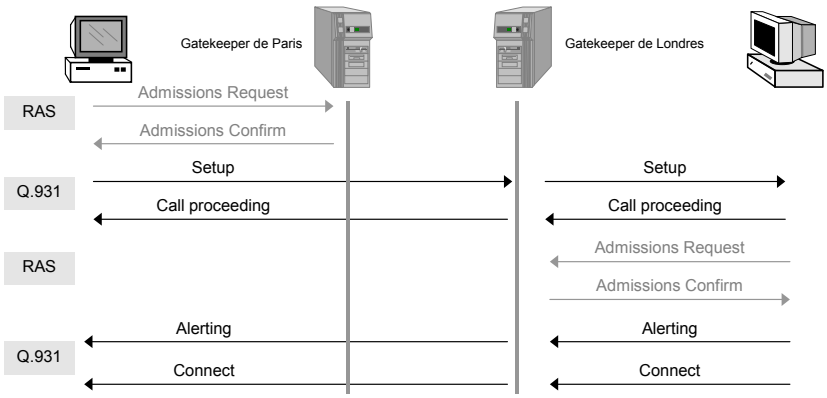


Si nécessaire, un terminal peut demander au gatekeeper l'autorisation de changer de bande passante en lui envoyant un message " Bandwidth Change Request ". S'il obtient l'accord, le canal de contrôle H.245 est fermé entre les deux terminaux, puis un nouveau est ouvert accompagné d'une nouvelle demande RSVP.



Sur chaque site, les terminaux s'enregistrent auprès de leur gatekeeper local (Paris et Londres dans notre cas). Lors d'un appel, le gatekeeper de Paris autorise l'ouverture directe du canal de signalisation d'appel entre le terminal et son correspondant, tandis que le gatekeeper de Londres impose que ce canal passe par lui.

Figure 16-6.
Gestion des appels via des gatekeepers.



Le gatekeeper de Paris peut trouver les informations du correspondant concernant un terminal auprès d'un autre gatekeeper en utilisant le message « *Resource Locator* ». Le gatekeeper de Londres lui renvoie alors son adresse ainsi que le port TCP sur lequel il répond. S'il avait décidé que la communication pouvait être directe, il aurait envoyé l'adresse IP du terminal.

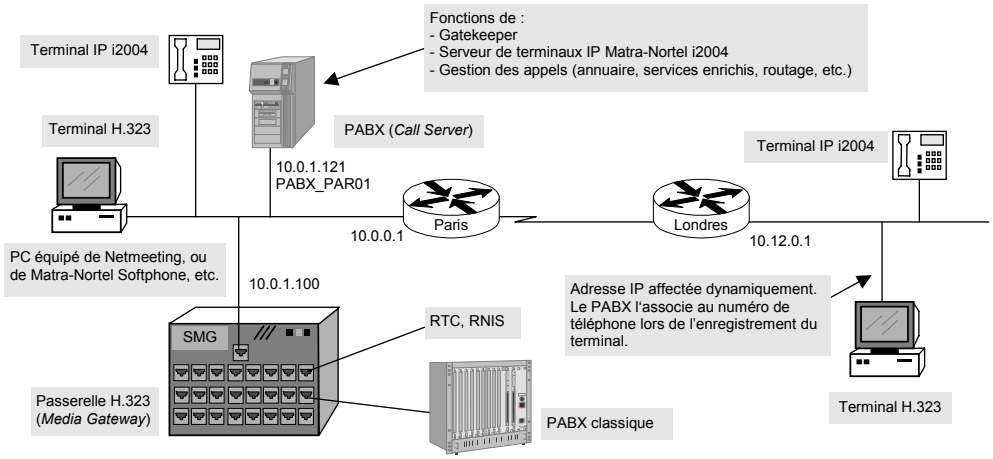
L'appel accepté, les deux entités peuvent ouvrir le canal de contrôle H.245 dans le but de négocier les paramètres du flux multimédia.

La voie vers le tout IP

Nous venons de mettre en place un système qui permet à nos terminaux IP (PC, téléphones IP) de communiquer entre eux (voix et image). Nous utilisons des gatekeepers pour assurer la conversion d'adresses et le routage des appels. Des passerelles sont également mises en place afin d'interconnecter notre réseau au monde téléphonique classique (PABX, RTC, RNIS, DECT, etc.).

Mais, par rapport à la téléphonie classique, il nous manque toute la gamme de services enrichis qu'un PABX peut offrir : plan de numérotation, messagerie vocale, transfert, filtrage, restriction et interception d'appels, etc.

L'aboutissement d'un réseau VoIP est donc la mise en place d'un PABX IP. Deux types d'offres sont proposées sur le marché : celles des constructeurs informatiques (Cisco, Lucent, etc.) et celles des constructeurs de PABX (Alcatel, Matra, Nortel, etc.). Les premiers ont l'avantage de bien connaître le monde IP ; les derniers offrent celui d'une plate-forme existante qu'il suffit de porter sur IP. C'est le cas de Matra qui a porté sous Windows sa gamme 6500 fonctionnant à l'origine sous un Unix propriétaire. Le résultat est le produit Succession qui offre le même niveau de fonctionnalité qu'un PABX traditionnel.



Le PABX peut être situé n'importe où sur le réseau. Les sites importants en sont équipés, tandis que les petits sites n'en ont pas besoin, les terminaux dialoguant alors avec le PABX du site principal.

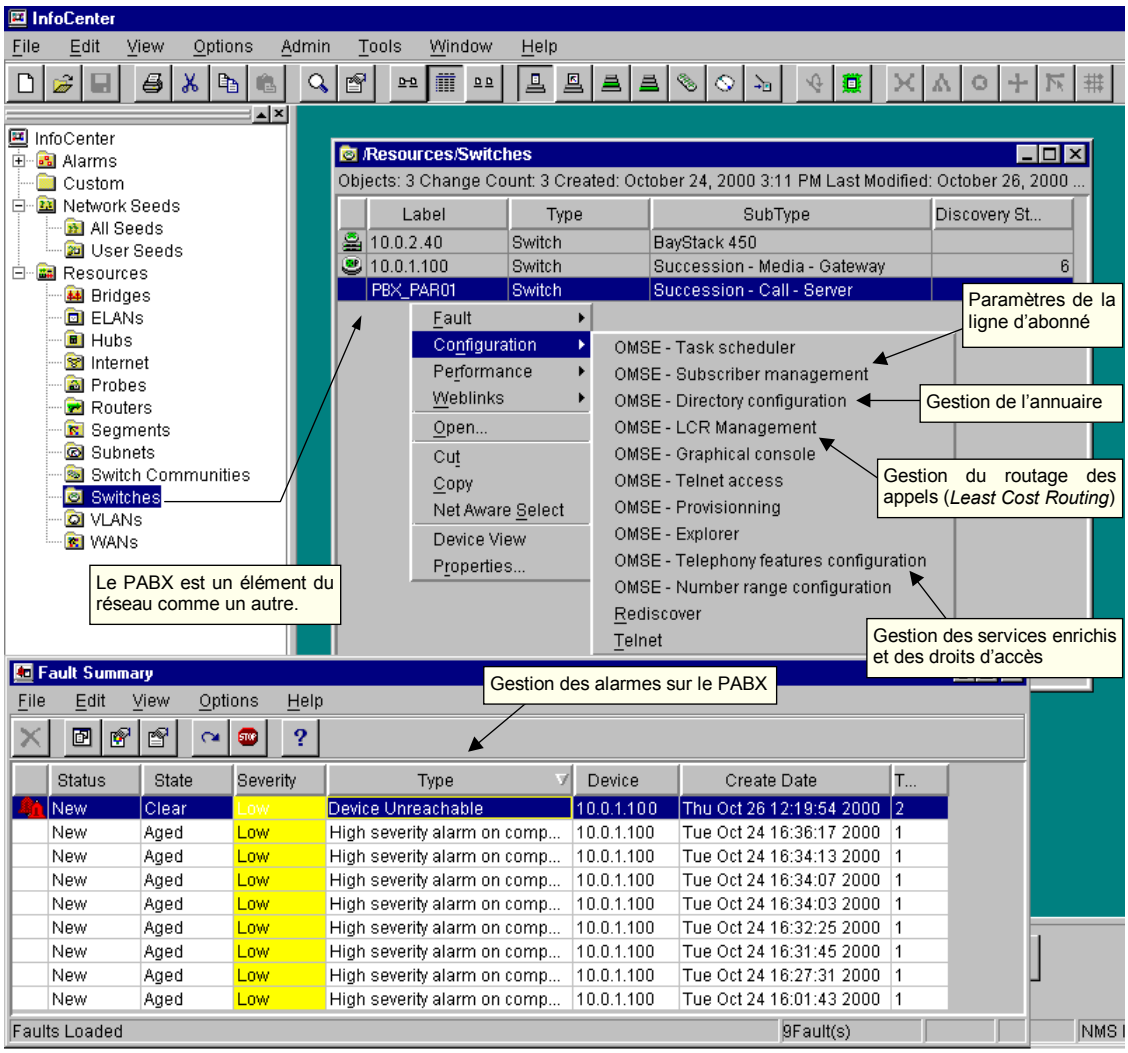
De son côté, la passerelle accueille les terminaux non IP : PABX classiques, téléphones analogiques et numériques Matra, bornes DECT, etc. L'accès au monde extérieur (RTC, RNIS) est réalisé soit *via* le PABX traditionnel, soit directement par la passerelle. Cette dernière peut également accueillir les terminaux IP et peut, de ce fait, coupler le terminal VoIP à un poste DECT.

Le PABX accueille les terminaux H.323 natifs ainsi que ceux de Matra *via* un protocole propriétaire. Ces derniers se comportent comme des terminaux passifs qui ne font qu'afficher les informations envoyées par le PABX (fonction équivalente à celle d'un serveur de terminaux, tel que MetaFrame).

Configurer le PABX et la passerelle VoIP

Pour configurer tous ces éléments, nous nous retrouvons dans un environnement familier d'une interface graphique sous Windows. Désormais, ces systèmes sont, en effet, considérés comme des services à part entière au même titre que les bases de données, les serveurs de fichiers, la messagerie, etc.

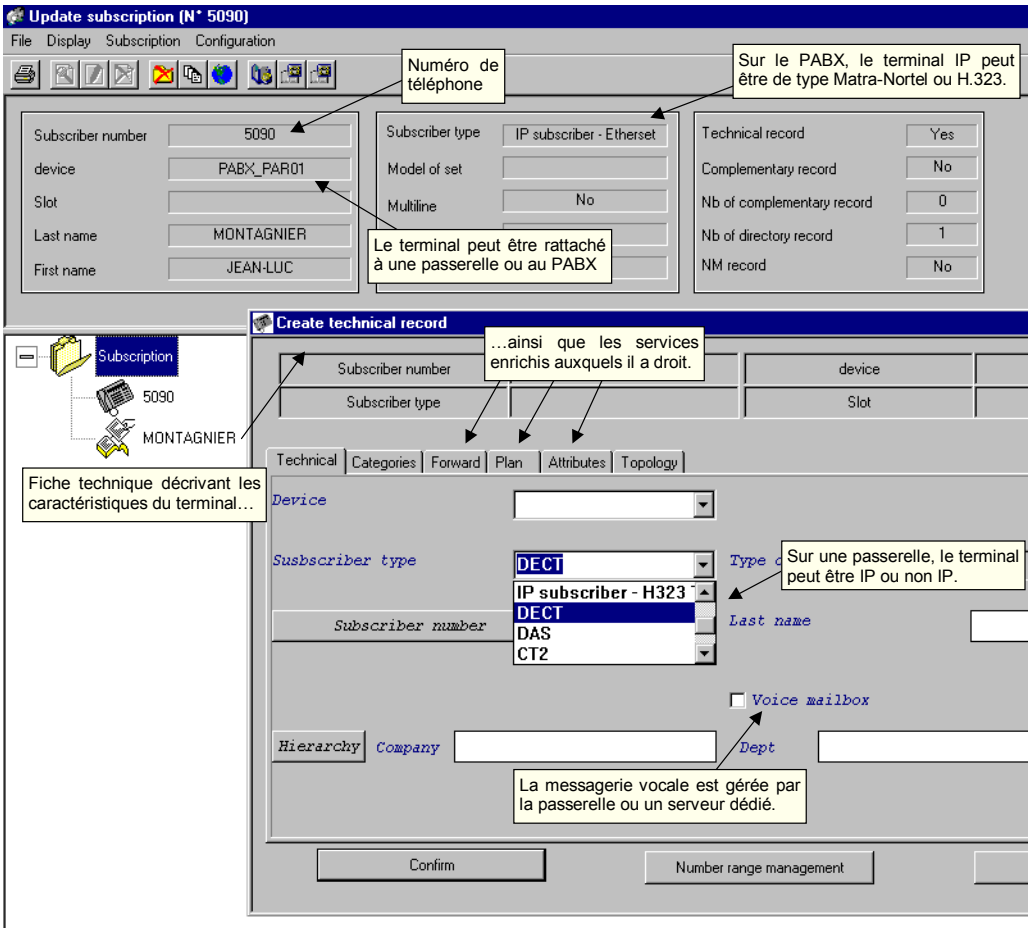
L'écran principal du gestionnaire fourni par Matra présente les éléments de notre réseau VoIP : commutateurs du réseau local (ici un *BayStack 450*), passerelles (la *Media Gateway* de Nortel) et PABX (le *Call Server* de Nortel).



Déclarer les terminaux téléphoniques

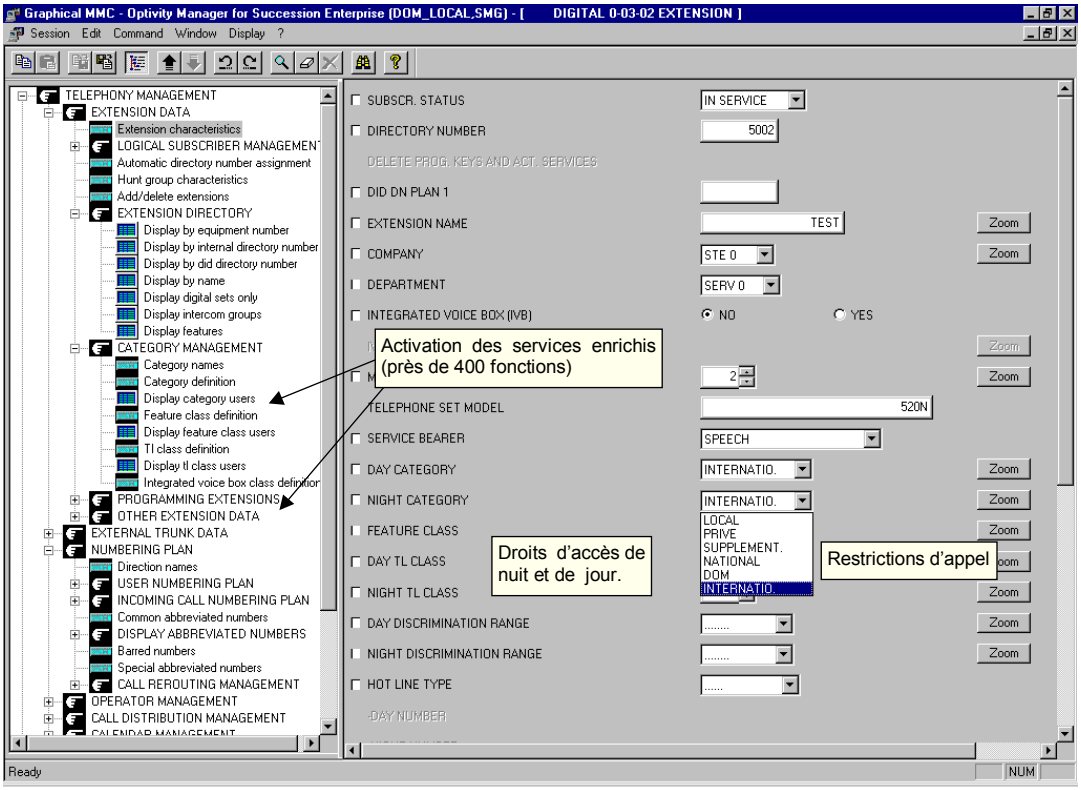
Ensuite, il convient de déclarer les terminaux rattachés à notre PABX : il peut s'agir de téléphones VoIP Matra (i2004), de PC équipés de Netmeeting ou encore de téléphones classiques raccordés à une passerelle.

Nous entrons alors dans le monde des téléphonistes : il faut créer un **abonné** auquel on associe une entrée dans l'**annuaire** ainsi qu'une **fiche technique** qui décrit les caractéristiques du poste et les services auxquels l'utilisateur a droit.



Sous l'arborescence *Subscription* (abonnement), nous avons créé l'abonné Montagnier, puis nous lui avons affecté un numéro de téléphone (le 5090) et spécifié le type de terminal qu'il utilise : dans son cas, il s'agit d'un téléphone IP (*IP subscriber – Etherset*). Nous lui avons ensuite associé une fiche technique décrivant les caractéristiques du terminal (ici, *IP subs-*

criber – H323) et les services auxquels il a droit, par exemple une messagerie vocale (*Voice mailbox*).



Parmi les services enrichis que le PABX peut gérer pour le terminal on trouve :

- l'identification de l'appelant ;
- la restriction d'appel (local, province, étranger, etc.), restrictions horaires ;
- le renvoi, transfert, filtrage ;
- le rappel sur occupation ;
- les numéros programmés (urgence, SVP) ;
- la gestion du second appel : parcage, conférence, va-et-vient ;
- la messagerie vocale ;
- etc.

Assurer la qualité de service

Nous devons également penser à gérer la qualité de service sur l'ensemble de nos routeurs, en activant, par exemple, RSVP et les files d'attente appropriées.

La réservation de la bande passante doit être effectuée à la source des flux VoIP, c'est-à-dire au niveau des passerelles :

```
req-qos guaranteed-delay
```

Demande une garantie sur la bande passante et sur le délai de transit

Ensuite, RSVP doit être activé sur toutes les interfaces de nos routeurs (comme indiqué au chapitre 14) :

45 Kbit/s dédiés à RSVP, dont 15 Kbit/s par flux (8 Kbit/s pour le codec + overhead + signalisation)

```
ip rsvp bandwidth 45 15
fair-queue 64 256 3
```

3 files d'attentes pour RSVP pour 3 communications simultanées

Les débits à réserver dépendent des codec choisis pour coder la voix (voir chapitre 13). Celui en question est le g729 qui nécessite 8 Kbit/s de bande passante.

Rappelons que les commutateurs sur nos LAN et les files d'attente WFQ sur les routeurs prennent en compte la priorité des paquets IP. Il est donc intéressant d'affecter une valeur au champ « IP precedence », ce qui doit être fait à la source des flux VoIP, c'est-à-dire au niveau des passerelles :

```
#gwpar1
dial-peer voice 44 voip
ip precedence 5
```

Priorité 5 = Critical (voir chapitre 14).

Par ailleurs, nos commutateurs doivent assurer la correspondance entre la priorité du paquet IP et celle affectée à la trame Ethernet, ce qui est réalisé comme suit :

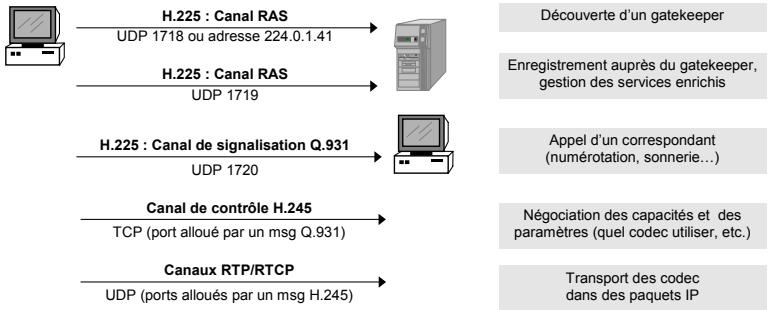
```
set qos acl ip politique_voip trust-ipprec
set qos ipprec-dscp-map 0 8 16 24 32 40 48 56
set qos dscp-cos-map 40:6
```

Transporter les flux multimédias

Un terminal H.323 utilise les services de la couche d'adaptation H.225 pour envoyer et recevoir des messages. En plus des signalisations Q.931 et RAS, cette couche réalise l'interface entre la gestion des canaux logiques par le terminal et la gestion des paquets sur un réseau IP. En fonction de leur nature, H.225 sélectionne ainsi le type de paquet dans lequel envoyer les données : TCP, UDP, RTP ou RTCP.

Canal logique	Protocole de transport / Numéro de port
Canal RAS : découverte d'un gatekeeper	UDP 1718 ou adresse multicast 224.0.1.41
Canal RAS : enregistrement auprès d'un gatekeeper	UDP 1719
Canal de signalisation Q.931	TCP 1720
Canal de contrôle H.245	Port TCP alloué dynamiquement par Q.931 (messages *Connect*, *Alerting* et *Call proceeding*)
Sessions RTP/RTCP	Port UDP alloué dynamiquement par H.245 (message *Open Logical Channel*)
T.120	TCP 1503

Figure 16-7. *Interaction des protocoles H.323.*



Le transport des flux multimédias est donc assuré par H.225 qui reprend les spécifications des RFC suivantes.

RFC	Sujet traité
1889	Spécifications de RTP et de RTCP
1890	Définition des profils pour les conférences audio et vidéo
2032	Transport des codec vidéo H.261
2190	Transport des codec vidéo H.263
2198	Transport des codec audio

Le transport des flux audio et vidéo *via* RTP et RTCP

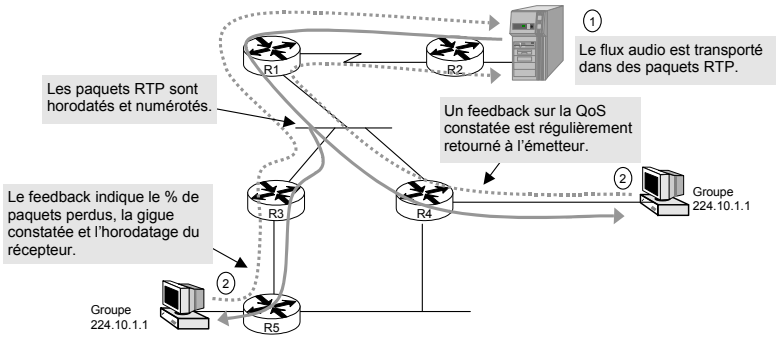
Une conversation téléphonique génère un flux audio qui est découpé en tranches représentant quelques millisecondes d'un son. De même, une image vidéo est découpée en tranches représentant un groupe de pixels. Une tranche correspond ainsi à un échantillonnage (*sampling*) d'un son ou d'une image vidéo numérisés et donc représentés par des octets (voir chapitre 13).

Ces octets sont placés dans des paquets **RTP** (*Real-time Transport Protocol*). Ce standard ne décrit aucun mécanisme de contrôle ni de récupération d'erreur : il se contente de définir le format des paquets et des données transportés. Ces paquets RTP sont ensuite encapsulés dans des paquets UDP, eux-mêmes transportés dans des paquets IP.

RTP est associé à **RTCP** (*RTP Control Protocol*), également transporté dans des paquets UDP, qui renvoie à l'émetteur un retour sur la qualité de service perçue par les récepteurs d'un flux.

Deux couples de ports UDP (source et destination) sont alloués pour un flux audio ou vidéo : l'un pour RTP, l'autre pour RTCP. Le protocole UDP étant orienté sans connexion, les sessions sont gérées par H.245 qui fait l'association avec ses numéros de canaux logiques.

Figure 16-8.
Rôles de RTP et de RTCP.



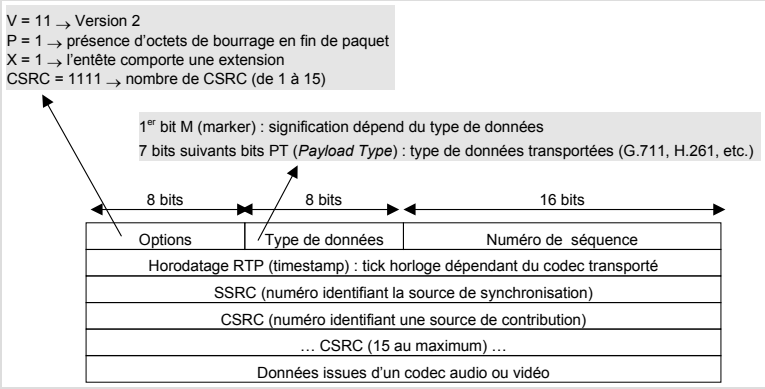
Un flux ne transporte jamais de données mixtes audio et vidéo. Par exemple, pour une visioconférence, les paquets audio et vidéo sont transportés dans deux sessions RTP différentes, chacune étant associée à une session RTCP.

Le codage peut être modifié au cours d'une session pour s'adapter à un changement de la qualité de transmission (débit, erreurs, etc.). L'émetteur peut également modifier la nature du flux au cours d'une session et basculer, par exemple, de la vidéo à l'audio.

LE POINT SUR RTP (RFC 1889 ET 1890)

Le protocole RTP (*Real-time Transport Protocol*) décrit le format des paquets transportant des flux audio ou vidéo.

Utilisant les services de la couche UDP, RTP n'assure pas pour autant un quelconque contrôle de flux, de reprise sur erreur ou même de contrôle d'intégrité du paquet. Son but est simplement de transporter quelques millisecondes de voix ou une portion d'image en y incluant des informations relatives au temps (synchronisation) et permettant d'identifier les émetteurs qui génèrent le signal audio ou vidéo.



Le numéro de port UDP affecté à RTP est alloué aléatoirement par l'application. Il doit simplement être pair. Le numéro suivant (forcément impair) est alors affecté à la session RTCP associée (voir l'encart suivant).

L'**horodatage**, qui dépend de l'horloge de l'émetteur, représente l'instant où le premier octet des données transportées a été échantillonné.

Le champ **SSRC** (*Synchronization Source*) est un numéro unique, généré de manière aléatoire, qui identifie l'émetteur du flux. Un mécanisme permet de gérer les problèmes de collision (deux SSRC identiques).

La RFC prévoit l'utilisation de **mixers** dont le rôle est de mélanger les flux pour n'en faire qu'un qui sera redistribué aux participants d'une conférence audio, par exemple. Il synchronise les flux combinés issus des différentes sources n'utilisant pas le même codage et ne disposant pas de la même source d'horloge ni de la même fréquence d'échantillonnage. Le paquet RTP résultant contient la liste des SSRC qui ont contribué à la formation du flux mélangé (champs **CSRC**, C*ontributing Source*), par exemple, la liste de ceux qui ont parlé en même temps.

Le mixer s'apparente aux fonctions du MCU H.323, mais au niveau RTP, alors que ce dernier assure des fonctions plus évoluées liées à la signalisation. En pratique, un serveur de conférence implémente les fonctions de MCU et de mixer.

La RFC prévoit également l'utilisation de **translators** dont le rôle est de convertir des flux sans les mélanger.

Ce logiciel peut :

- convertir un codec en un autre ;
- convertir un flux multicast en un flux unicast ;
- convertir les ports UDP aléatoires en ports UDP fixes pouvant être filtrés par un firewall ;
- créer un tunnel sécurisé en chiffrant les données.

Dans ces deux derniers cas, un translator s'utilise par paire.

LE POINT SUR RTCP (RFC 1889 ET 1890)

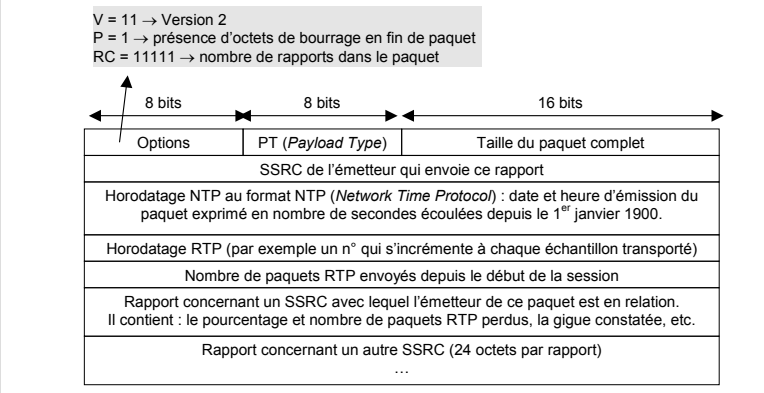
RTCP (*RTP Control Protocol*) décrit le format des informations remontées par le ou les récepteurs d'un flux. Il est utilisé concurremment avec RTP pour remplir trois fonctions :

- fournir un retour sur la qualité de service perçue par les récepteurs ;
- permettre d'identifier tous les participants à la conférence par un nom unique (un nom DNS, tel que machine.laposte.fr). Cela permet de retrouver le participant si le SSRC change en cours de session (suite à une réinitialisation, ou à la détection d'un conflit d'attribution de SSRC) ;
- permettre à chaque participant d'évaluer le nombre total de participants et d'adapter les flux en conséquence (fréquence d'émission des paquets).

Le paquet RTCP contient ainsi des informations relatives à l'**horodatage** des paquets reçus, la **gigue** (*jitter*) constatée par les récepteurs, le nombre de **paquets perdus** par période et en cumulé. Les informations recueillies par les émetteurs et les récepteurs permettent de vérifier que la qualité de service correspond à celle demandée *via* **RSVP**.

Il existe plusieurs types de paquets RTCP : rapport de récepteur (type RR – *Receiver Report*), rapport d'émetteur (type SR – *Sender Report*) si le récepteur est également un émetteur, description de la source (type SDES – *Source Description*), fin d'une participation (BYE) et fonctions spécifiques à une application (type APP), par exemple audio, vidéo, etc.

Ainsi, les rapports RR (PT = 201) et SR (PT = 200) prennent la forme suivante :



Le **rapport SEDS** contient, quant à lui, toutes les informations concernant l'émetteur : nom, adresse e-mail, numéro de téléphone, localisation géographique, nom du logiciel utilisé, etc.

Pour consommer le moins de bande passante possible :

- Les différents types de paquets RTCP sont regroupés dans un seul paquet UDP.
- La périodicité d'émission des paquets RTCP est aléatoire (entre 2 et 5 minutes).
- L'envoi des paquets RTCP par tous les participants doit être espacé de 2 à 7 secondes afin d'éviter les bursts.
- Le flux RTCP ne doit pas dépasser 5 % de la bande passante utilisée par RTP.
- Un participant ne doit pas générer plus de 20 % du total des flux RTCP.

Optimiser les flux multimédias

Un paquet RTP transporte 2 à 30 ms de voix, ce qui représente 20 à 150 octets selon le codec utilisé. La RFC 2298 montre l'exemple d'un paquet RTP transportant 20 ms de voix échantillonnée à 8 KHz et dont les données utiles ne représentent que 98 octets.

Codec	Données numériques
G.722	1 ms par trame
G.723.1	30 ms par trame
G.728	2,5 ms par trame
G.729	10 ms par trame

Afin de réduire l'impact de l'overhead, on pourrait augmenter la taille des données transportées en envoyant plusieurs trames dans un seul paquet. L'inconvénient de cette approche est qu'elle engendre un délai de transit supplémentaire dans une passerelle ou au départ du PC émetteur. Par exemple, au lieu d'attendre 30 ms pour envoyer une trame G.723, on attendra 60 ms pour envoyer deux trames. Résultat : le délai d'attente de la première trame du paquet est allongé de 30 ms, et ce pour une trame sur deux dans le flux.

Les techniques suivantes sont pour cela privilégiées.

Compression des en-têtes

La taille minimale d'un paquet RTP (sans la liste des contributeurs) est de 12 octets, auxquels il faut ajouter 8 octets pour UDP et 20 pour IP, soit 40 octets en tout.

Afin de diminuer cet overhead, les routeurs permettent de compresser ces en-têtes en réduisant leur taille cumulée à 2 ou 5 octets sur les interfaces WAN. Par exemple, les commandes suivantes permettent de compresser les en-têtes de 16 sessions RTP simultanées :

```
int s 0
encapsulation ppp
ip rtp header-compression
ip rtp compression connections 16
```

Le principe repose sur le fait que l'en-tête varie très peu d'un paquet à l'autre : seuls les numéros de séquence et de contrôle d'erreur changent. L'activation de la compression consiste alors à n'envoyer que les données qui ont changé d'un en-tête à l'autre (RTP/UDP/IP), ce qui représente entre 2 et 5 octets la plupart du temps. Cette fonction est surtout efficace pour les liaisons inférieures à 2 Mbit/s.

Utilisation des mixers

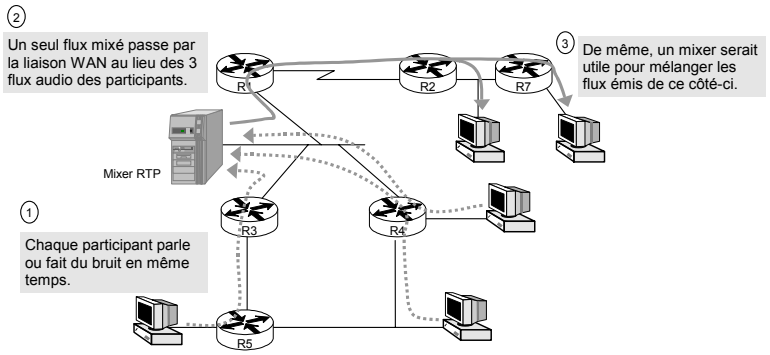
Le son perçu par un participant est la résultante des voix et des bruits émis par les autres. Pourtant, celui-ci reçoit *N* flux audio, provenant des *N* autres participants. Son PC se contente de restituer un son composite.

Afin de diminuer le débit généré par ces flux sur notre réseau, et surtout sur les liaisons WAN limitées en bande passante, il serait intéressant de créer un son composite en amont.

Le **mixer RTP** répond à ce besoin puisqu'il permet de fusionner les différents flux audio en un seul. Rappelons, par ailleurs, qu'un mixer permet également de convertir les codec audio entre participants ne disposant pas des mêmes équipements.

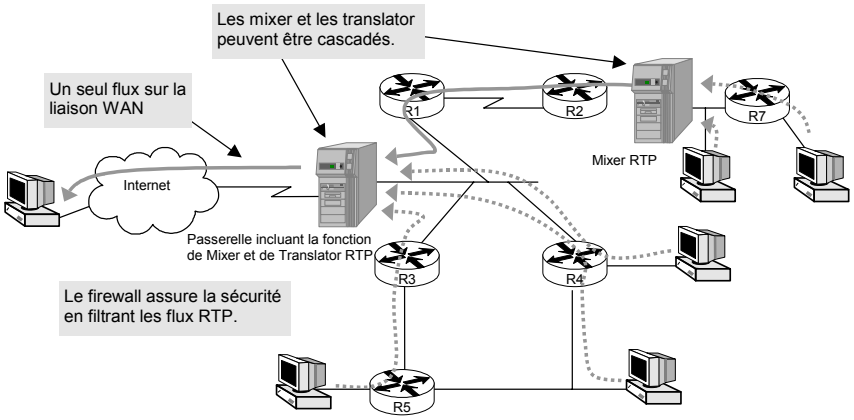
Dans notre réseau, l'installation d'un mixer de chaque côté de la liaison WAN permet de diminuer considérablement les flux générés par une conférence audio. De plus, alors que les participants peuvent envoyer leur flux dans des paquets unicasts, le flux envoyé par le mixer pourra utiliser des paquets multicasts afin de réduire encore davantage la charge globale du réseau.

Figure 16-9.
Utilisation d'un mixer pour optimiser les flux.



L'intranet est connecté à Internet par une liaison (spécialisée, Frame Relay ou autre) dont le débit est souvent limité. Si plusieurs participants à une même conférence sont répartis entre notre intranet et Internet, là encore, l'utilisation d'un mixer permet de réduire la charge de la liaison WAN.

Figure 16-10.
Intégration d'un mixer et d'un translator dans un firewall.

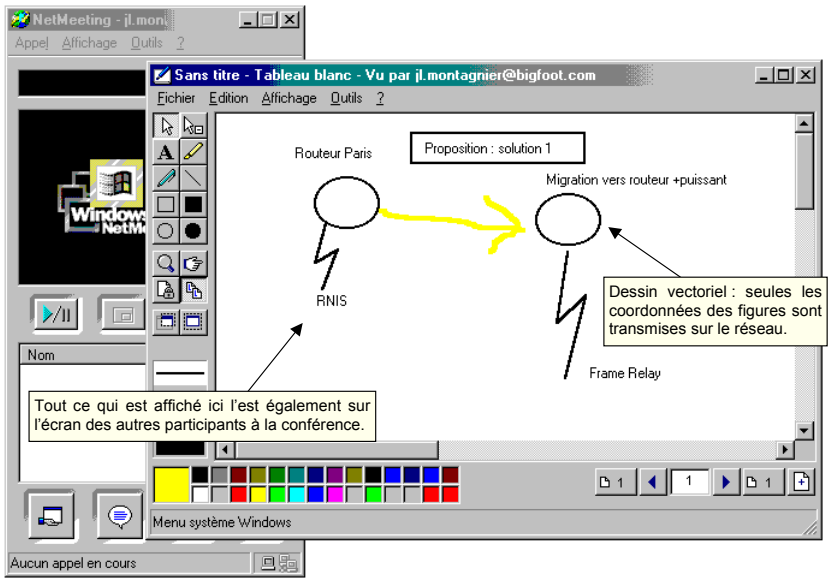


Échanger des données multimédias

Les participants à une visioconférence peuvent vouloir recevoir un document que leur remet l'animateur, par exemple. Le standard T.120 de l'ITU-T offre à ces participants la possibilité d'utiliser quatre applications :

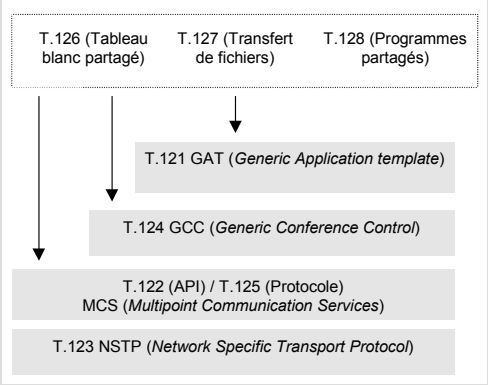
- transfert de fichiers multipoint ;
- partage d'un presse-papiers commun ;
- partage d'un tableau blanc virtuel (*whiteboard sharing*) qui permet à chaque participant de dessiner sur le même schéma ;
- déport écran/clavier pour une prise de contrôle à distance et une démonstration à tous les participants.

On retrouve ici des applications classiques, mais sous la forme du partage par plusieurs utilisateurs.



LES DONNÉES AU SEIN D'UNE CONFÉRENCE MULTIMÉDIA (ITU-T T.120)

La norme T.120 décrit un cadre fonctionnel général permettant à des utilisateurs de **partager des données au sein d'une conférence** *via* des sessions TCP gérées par H.245.



Les protocoles utilisés sont les suivants :

- **GAT** (*Generic Application Template*), norme **T.121**. Définit les interfaces d'accès (API) pour des applications T.120 : entrées et sorties des conférences, négociations des fonctionnalités, etc.
- **MCS** (*Multipoint Communication Service*), normes **T.122** pour l'interface d'accès et **T.125** pour le protocole. Assure la diffusion des données au sein d'une conférence selon trois topologies : en étoile autour d'un gestionnaire central (*top provider*), en cascade autour de plusieurs gestionnaires et d'un *top provider*, ou chaînage (*daisy chain*).
- **NSTP** (*Network Specific Transport Protocol*), norme **T.123**. Assure l'interface entre les applications T.120 et le réseau. Il réalise l'adaptation à différents supports, tels que RNIS et TCP/IP, tout en gérant les erreurs de transmission.
- **GCC** (*Generic Conference Control*), norme **T.124**. Il s'agit du logiciel qui organise la conférence. Il gère la liste des participants, les accepte ou les refuse selon les mots de passe et les droits d'accès, décide à qui passer la main, assure la cohérence des informations échangées (mise à jour en temps réel et simultanément pour tous les participants) et surveille les ressources utilisées par le MCS.

CINQUIÈME PARTIE

La sécurisation des réseaux IP

17

Protéger son réseau et ses données

Dès lors qu'il permet à des centaines, voire à des milliers d'utilisateurs de communiquer et d'échanger des informations, le réseau pose inévitablement le problème de la sécurité : les informations qu'il véhicule possèdent de la valeur et ne doivent pas être accessibles à tout le monde.

La sécurité est un vaste sujet, qui dépasse le cadre du réseau tant sur les plans technique que méthodologique. Ainsi, ce chapitre ne traite pas des étapes essentielles que sont l'analyse de la valeur (qu'est-ce qu'on protège) et l'analyse de risque (contre quoi l'on se protège). Il ne traite pas, non plus, des protections appliquées aux postes de travail (anti-virus, par exemple) et aux applications (messagerie, serveurs Web, etc.).

Ce chapitre ne s'attache qu'à une petite partie de la sécurité, celle liée à la protection du réseau IP que nous venons de construire. Dans ce chapitre, vous apprendrez ainsi à connaître :

- les vulnérabilités des protocoles de la famille IP ;
- les différents types d'attaques ;
- les techniques de filtrage utilisées par les firewalls ;
- les architectures à contrôle de flux ;
- les principes des algorithmes de chiffrement ;
- les mécanismes de chiffrement, de contrôle d'intégrité, d'authentification et de signature ;
- les protocoles SSL et IPsec.

L'importance d'une politique de sécurité

Plus que partout ailleurs, définir une politique de sécurité dans le domaine des réseaux est avant tout un gage de cohérence et donc, d'une réelle protection. Sans cela, le développement anarchique des moyens de sécurité, ajoutés comme des surcouches au-dessus des réseaux, peut conduire à une complexité telle, qu'elle entrave la circulation de l'information.

Sur un petit réseau, couvrant un ou quelques sites, la complexité peut encore être maîtrisée, mais dans les grandes entreprises, le réseau permet l'échange de données de natures très diverses entre des populations aussi diverses que géographiquement dispersées. Dans tous les cas, la problématique reste la même ainsi que les moyens mis en œuvre pour la résoudre.

Trop souvent, les responsables sécurité imposent sans discernement des contraintes drastiques : un firewall par-ci, un autre point de coupure par-là, parce que tel type d'attaque est toujours possible. Et, ne sait-on jamais, tel autre cas de figure peut toujours se produire : il faut donc ajouter un autre mécanisme de protection.

Ces réflexes sécuritaires proviennent à la fois d'une responsabilisation excessive des personnes en charge de la sécurité et d'une méconnaissance technique des réseaux qu'ont ces mêmes personnes. Le réseau est alors perçu comme un organisme tentaculaire échappant à tout contrôle. Il est le vecteur de tous les maux réels ou imaginaires.

Ne connaissant pas ou pire, croyant savoir, ces responsables brandissent des menaces qui conduisent à une surenchère permanente, attisée en cela par les sociétés de conseil et les éditeurs qui vivent de ce commerce. Ne voulant prendre aucun risque, ils entretiennent la culture du secret et empilent des barricades.

Afin d'éviter ce cauchemar, il convient donc de ne pas faire de la sécurité pour la sécurité, mais de connaître la valeur de ce que l'on protège et contre quoi l'on se protège. À la suite de quoi, une réflexion technique permet de mettre en place les moyens de protection appropriés. En fin de compte, c'est le bon sens qui nous amène à ne pas utiliser le même type de clé pour la porte de son coffre-fort, pour la porte d'entrée de sa maison et pour celles des autres pièces.

La définition d'une politique de sécurité passe donc par :

- une analyse de la valeur des informations à protéger et une analyse des menaces ;
- l'application de règles et de procédures par les utilisateurs internes à l'entreprise ;
- la mise à jour permanente des logiciels de protection (firewalls, anti-virus...) et des correctifs pour les systèmes d'exploitation et les applications tels que les navigateurs Web ;
- la surveillance du réseau (enregistrement des événements, audits, outils de détection) ;
- la définition d'une architecture permettant de limiter les possibilités d'attaques et la portée d'éventuels dégâts causés par les pirates.

Les sections qui suivent traitent de ce dernier point en commençant par exposer les vulnérabilités, puis en présentant les différents types d'attaques.

Les vulnérabilités des protocoles IP

La famille des protocoles IP présente des failles de sécurité intrinsèques, car ils n'ont pas été conçus dans une optique de sécurité, mais plutôt dans une optique de résilience et de performance.

Il n'y a notamment :

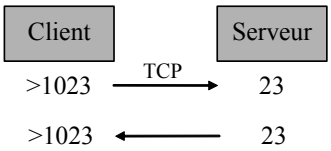
- aucun chiffrement des données (les données et souvent les mots de passe circulent en clair) ;
- aucun contrôle d'intégrité (les données peuvent être modifiées par un tiers) ;
- aucune authentification de l'émetteur (n'importe qui peut émettre des données en se faisant passer pour un autre).

Les sections suivantes décrivent quelques-uns des protocoles les plus courants qui cumulent ces lacunes.

Telnet

L'identifiant et le mot de passe Telnet circulant en clair sur le réseau, il faut demander aux utilisateurs d'employer des mots de passe différents de ceux utilisés pour se connecter à des applications utilisant des protocoles mieux sécurisés. Interceptés, l'identifiant et le mot de passe pourront, en effet, être utilisés par un intrus pour se connecter à des applications dont les mots de passe sont, en ce qui les concerne, chiffrés.

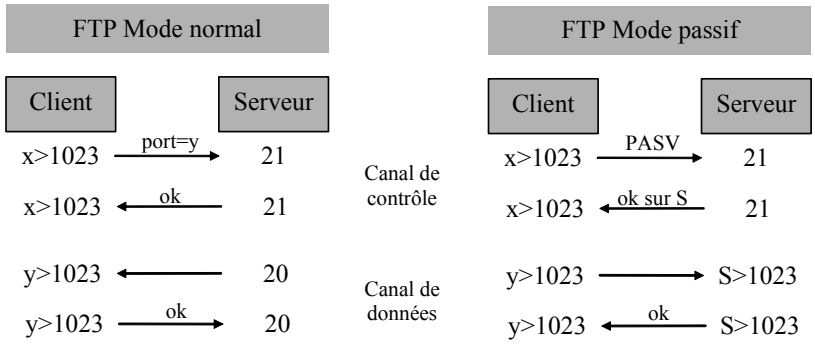
De plus, les données sont véhiculées sous une forme non structurée, tâche qui est laissée à l'initiative de l'application. Il est donc très facile de transférer toutes sortes de données en profitant d'une session Telnet.



Ces deux caractéristiques font de Telnet un protocole très dangereux à utiliser entre une entreprise et Internet. Il doit donc être interdit. Sur le plan interne, son usage peut être autorisé, réglementé ou interdit, selon l'architecture de votre réseau, les mécanismes de sécurité mis en œuvre et la valeur des informations à protéger.

FTP

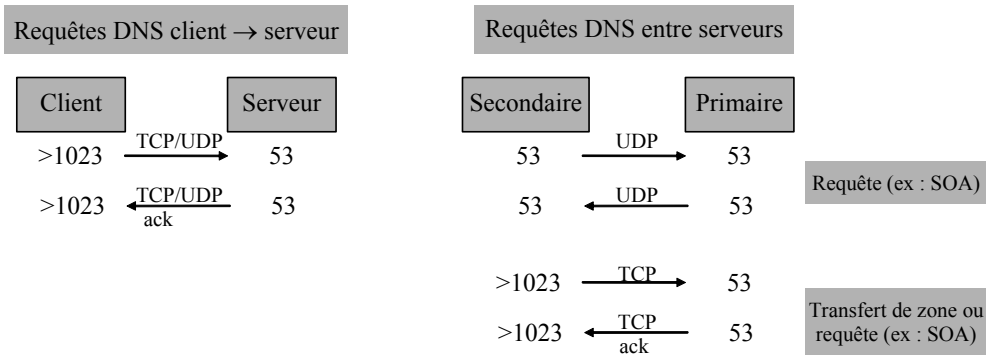
L'identifiant et le mot de passe circulant en clair, les recommandations énoncées pour Telnet s'appliquent également à FTP.



Le mode passif est beaucoup plus sûr que le mode normal, car il évite d'autoriser les connexions entrantes. Cependant, toutes les implémentations ne sont pas compatibles avec ce mode. Si tous vos clients et serveurs FTP supportent le mode PASV, il est recommandé d'interdire le mode normal.

DNS

La recherche de noms et le transfert de zone (cf. chapitre 7) utilisent généralement les ports UDP (53↔53), mais si ces requêtes échouent, elles sont relancées *via* TCP (>1023↔53). Certaines implémentations, ce qui est le cas avec les machines AIX, utilisent exclusivement TCP.

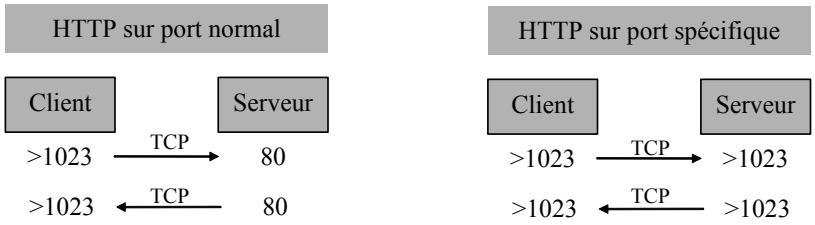


Aucun mot de passe n'est requis pour ces échanges, effectués automatiquement entre machines et ce, de manière transparente pour l'utilisateur. Par ailleurs, la commande nslookup (cf. chapitre 7) permet à tout utilisateur de consulter la base de données.

La base de données du DNS est particulièrement sensible, car elle contient l'ensemble des adresses IP de vos serveurs et équipements réseau, voire plus si vous utilisez Active Directory de Microsoft. Elle doit donc être protégée et en particulier être inaccessible depuis l'extérieur.

HTTP

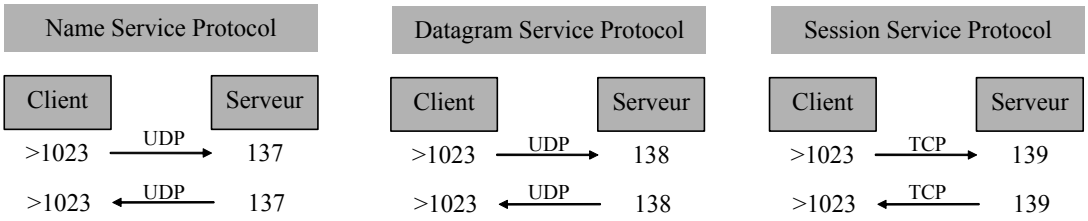
Les serveurs Web peuvent répondre sur des ports spécifiques, autres que 80, tels les 8000, 8080, 8010 ou encore 81, pour ne citer que les plus courants. Ces ports non-standards ne présentent aucun intérêt en termes de sécurité, car les scanners permettent de les trouver rapidement.



En revanche, il est recommandé de séparer les données accessibles en FTP et en HTTP, soit sur deux serveurs différents, soit sur deux répertoires différents. Il est en effet facile de déposer un cheval de Troie sur un répertoire FTP, puis de le faire exécuter par le serveur HTTP. Sous Unix, il est en plus recommandé d'utiliser la fonction *chroot*.

Netbios

Le protocole Netbios (*Network Basic Input Output System*) est difficile à contrôler, car il transporte de nombreux protocoles propres à Microsoft : il s'agit de SMB (*Server Message Block*), de NCB (*Network Control Bloc*) et de toute une série de RPC (*Remote Procedure Calls*), qui sont utilisés par les environnements Windows.



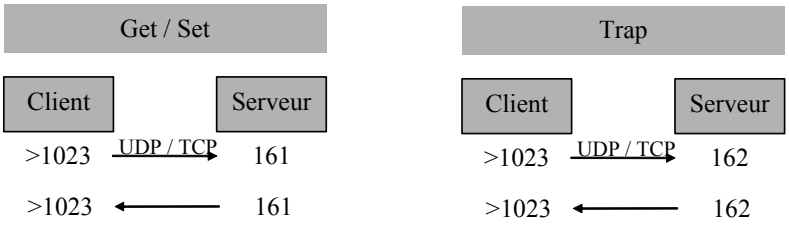
Par exemple, les relations entre contrôleurs de domaines et entre clients et serveurs WINS emploient des relations complexes qu'il est difficile de filtrer. Pour ces raisons, l'utilisation de Netbios doit être interdite entre l'entreprise et Internet.

SNMP

Pour les requêtes « get » et « set » (cf. chapitre 6), le client est la plate-forme d'administration tandis que pour les messages « traps », le client est le matériel (réseau, serveur, etc.). Il est donc recommandé de cantonner ce protocole entre les stations d'administration et les équipements à surveiller :

- La plupart des implémentations utilisent UDP, et il faut l'autoriser dans les deux sens.
- Les noms de communauté circulent en clair.

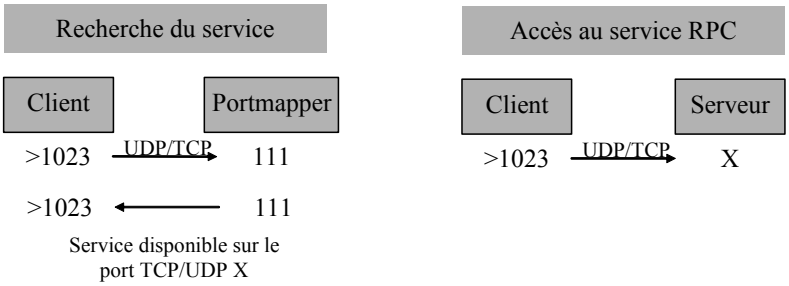
- Les possibilités d'action offertes sont importantes (la commande « get » permet de faire un dump complet d'une configuration et la commande « set » permet de modifier la configuration).



On peut donc envisager de laisser passer les messages SNMP en filtrant les adresses source et destination. La politique de filtrage peut cependant être plus souple au sein d'un même domaine de confiance ou dans le cas de réseaux entièrement commutés reposant sur des VLAN. Pour ce qui concerne Internet, le protocole SNMP doit être interdit.

RPC (*Remote Procedure Calls*)

De nombreuses applications, comme les logiciels de sauvegarde, utilisent les services du *Portmapper* qui répond sur les ports UDP et TCP 111 et qui renvoie au client un numéro de port aléatoire indiquant que le service est disponible sur sa machine.



Ce protocole ne peut pas être efficacement filtré, car de nombreux ports doivent être laissés ouverts de par la nature aléatoire de l'allocation. De plus, les firewalls génèrent de nombreux dysfonctionnements pour les applications qui utilisent les RPC. Les RPC doivent donc être interdits pour Internet.

NFS

Le protocole NFS utilise *a priori* le port UDP 2049, mais la RFC 1094 indique que ce n'est pas une obligation. Certaines implémentations utilisent en fait le *Portmapper*.

Un problème de sécurité important posé par NFS est celui lié au numéro de handle :

- Il est prédictible, car reposant sur la date création du système de fichier.
- Il est valable même lorsque le système de fichiers est démonté.
- Il est utilisable par quiconque l'ayant obtenu (par écoute réseau ou par calcul).

NFS est également vulnérable à la dissimulation d'adresse (spoofing), car le serveur se fie à l'adresse IP pour authentifier la machine cliente. Par ailleurs, les mécanismes *setuid* et *nobody* positionnés par l'administrateur du serveur demeurent des failles de sécurité, s'ils sont mal configurés.

ICMP

De nombreux services ICMP (cf. chapitre 6) peuvent renseigner un intrus, comme par exemple « destination unreachable », qui comporte des informations indiquant la cause du problème ou encore « echo request » et « echo reply », qui indiquent si un nœud est actif.

Les seuls paquets ICMP, qu'il est envisageable de laisser passer vers un autre domaine de confiance, sont les suivants :

- type 4 « source quench », qui permet de contrôler le flux entre un client et un serveur ;
- type 11 « time to live exceeded », qui évite le bouclage des paquets IP ;
- type 12 « parameter problem », qui indique une erreur dans un en-tête ;
- type 8 « echo request » et type 0 « echo reply », utilisés pour vérifier l'activité des nœuds pour des opérations de supervision et de diagnostic.

Vis-à-vis d'Internet, tous les services ICMP doivent être interdits : ils ne sont pas nécessaires et peuvent donner trop d'informations aux intrus.

Les différents types d'attaques

En plus de leur vulnérabilité, la famille des protocoles IP bénéficie d'une grande accessibilité : leurs spécifications sont rendues publiques avec les RFC, et Internet est un réseau ouvert à tous, aux particuliers comme aux professionnels. Les compétences techniques sont donc très répandues, et tout le monde est susceptible de trouver des failles et de lancer des attaques aussi bien sur le réseau public que sur notre réseau interne. Il est à noter que, selon la plupart des statistiques, environ 70 % des problèmes de sécurité ont des origines internes à l'entreprise.

Les attaques de type refus de service (denial of service)

Ce terme désigne un groupe d'attaques destiné à bloquer le service offert par un serveur (Web ou autre), un routeur ou un firewall. Le principe consiste à inonder de requêtes la machine cible de manière à ce qu'elle soit de plus en plus sollicitée, et ce, jusqu'à ce qu'elle ne puisse plus traiter les vraies requêtes des utilisateurs. Le pirate espère, de plus, qu'un débordement (saturation des capacités du logiciel) conduira à l'arrêt brutal de la machine, profitant ainsi d'une situation limite, non prévue par le programmeur (bogue).

Les parades consistent, la plupart du temps, à appliquer des correctifs (*patches*) qui permettent de traiter les cas de figure limites.

Quelques exemples d'attaques par refus de service

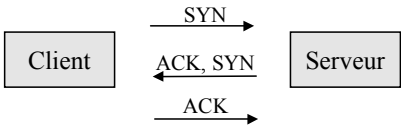
L'attaque la plus classique, le ping de la mort (*ping of death*), consiste à bombarder la machine cible de paquets ICMP de type « echo_request ».

Une variante, appelée *teardrop*, consiste à envoyer des paquets ICMP de taille importante (plusieurs dizaines de Ko) de manière à activer les mécanismes de fragmentation IP. La plupart des machines s'arrêtent de fonctionner lorsqu'elles rencontrent ce cas de figure (fragmentation de paquets ICMP), à moins d'être équipées du correctif adéquat.

L'attaque *land* (littéralement atterrissage), consiste à générer un paquet ayant la même adresse IP source et destination que celle de la machine visée, et avec des ports (TCP ou UDP) source et destination identiques. La machine visée est de préférence un routeur, qui route le paquet indéfiniment pour lui-même. La parade consiste, là encore, à appliquer un correctif qui traite ce cas limite.

Les attaques par inondation SYN (syn flooding)

Dès qu'un client envoie une demande d'ouverture de connexion TCP à un serveur, l'échange normal est le suivant (cf. chapitre 5) :



Si le serveur ne reçoit pas le paquet ACK du client, il reste en attente de la réponse jusqu'à l'expiration d'un timeout.

L'attaque consiste donc, pour le pirate, à écrire d'abord un logiciel qui n'envoie jamais de paquets ACK en réception d'un paquet SYN du serveur, puis à inonder le serveur de demandes de connexions de ce type. Ceci entraîne le serveur à allouer de plus en plus de ressources pour les demandes de connexions TCP qui restent en attente, jusqu'à dépassement de ses capacités.

Le firewall protège contre ce type d'attaque en détectant puis en bloquant la requête après N paquets SYN consécutifs issus du même client ou à destination du même serveur.

La dissimulation d'adresses (spoofing)

Cette technique consiste, pour le pirate, à utiliser une adresse IP interne, c'est-à-dire celle d'un réseau de l'entreprise protégé par le firewall. Le but est de faire croire aux machines (serveurs, firewalls) qu'il s'agit d'un paquet issu du réseau interne, de manière à bénéficier des mêmes droits d'accès que nos utilisateurs comme, par exemple, l'équivalence de droits pour les *remote commands* Unix ou les règles de filtrages du firewall basées sur l'adresse source.

Les routeurs ne se soucient pas de savoir par quelle interface proviennent les adresses sources, car les algorithmes de routage doivent prendre en compte le cas de routes multiples et de secours.

En revanche, si le réseau interne est connecté à Internet, aucun paquet, ayant comme adresse source celle d'un réseau interne, ne doit en provenir. Le mécanisme d'antispoofing, utilisé par les firewalls, consiste donc à contrôler l'origine des paquets sur la base du couple « interface réseau physique » / « adresse IP source ».

Il est à noter que le même principe s'utilise avec les noms DNS et SMTP ou encore les mots de passe des procédures d'authentification.

Les attaques par tunnel

Le principe consiste à utiliser un port TCP ou UDP dédié à un service (Telnet, par exemple), pour faire passer des flux autres que ceux prévus.

Le cas le plus classique est celui du serveur SMTP qui permet de se connecter en Telnet, non pas sur le port 23 dédié habituellement aux serveurs Telnet mais sur le port 25 dédié aux serveurs SMTP. L'utilisateur peut ainsi dialoguer manuellement avec le serveur *via* les commandes SMTP.

Pour les cas de figure les plus évolués, les pirates développent des logiciels spécifiques capables de dialoguer avec un protocole donné sur n'importe quel port TCP ou UDP.

Le vol de session (session stealing, splicing ou hijacking)

Cette technique consiste à repérer une session TCP ouverte depuis l'intérieur sur un serveur situé sur Internet, puis à se substituer à ce serveur. La session ouverte par l'utilisateur interne devient ainsi un tunnel permettant d'opérer en toute quiétude.

Cette technique est très sophistiquée puisqu'elle nécessite de sonder le réseau Internet ou interne à des endroits bien précis, à trouver les numéros aléatoires qui identifient les paquets TCP au sein d'une session, puis à se substituer au serveur.

La seule parade réellement efficace consiste à chiffrer les données et à contrôler l'intégrité des paquets IP à l'aide du protocole IPsec.

Le rebond

Cette technique est la plus simple puisqu'elle consiste à profiter d'une faille de sécurité pour investir un serveur, puis, à partir de là, à se connecter à d'autres machines en utilisant les droits accordés à ce serveur.

La protection contre ce type d'attaque relève de la sécurisation des serveurs (mots de passe et correctifs) et de l'architecture (voir plus loin). En positionnant les serveurs les plus exposés sur un segment dédié, différent de celui hébergeant les machines les plus sensibles, un firewall peut contrôler les flux.

Les chevaux de Troie

Un cheval de Troie est un programme, importé sur une machine à l'insu de ses utilisateurs, qui se substitue à un programme normal, par exemple, au client FTP. Quand l'utilisateur lancera la commande FTP, au lieu d'utiliser son programme habituel, il lancera le faux FTP, dont l'interface utilisateur ressemble au vrai programme, à la différence qu'il ouvrira une session parallèle sur un serveur appartenant au pirate.

L'importation de tels programmes s'opère *via* la messagerie, une connexion directe à un serveur en profitant de ses failles de sécurité (mot de passe ou bogue), *via* la technique du tunnel ou, plus simplement, en installant le logiciel avec les droits d'accès de l'administrateur.

La première parade consiste à réaliser un contrôle d'intégrité sur les programmes binaires et les scripts, puis à vérifier qu'il n'y a pas eu de changement dans leur taille, leur date, etc. La seconde parade consiste à repérer en temps réel l'introduction d'un tel cheval de Troie en l'identifiant par sa signature (une série de codes binaires caractérisant les instructions qui le composent). Encore faut-il que ce cette signature soit connue et qu'elle ait été intégrée dans l'anti-virus chargé de cette détection, d'où l'importance des mises à jour fréquentes (hebdomadaires, voire journalières).

Les vers

Ce type de programme utilise le réseau pour se propager : installé sur une machine, non seulement il l'infecte, mais il cherche également d'autres machines sur le réseau pour essayer de s'y installer. Le ver désigne donc un mode de propagation qui peut intégrer les fonctions de virus, de scanner et de cheval de Troie.

À l'heure actuelle, tous ces programmes ne peuvent se propager qu'entre machines du même type, par exemple, entre PC Windows, entre Macintosh ou entre machines Unix. S'il s'agit d'un exécutable, les processeurs doivent, de plus, être de la même famille (pentium, power PC, etc.). S'il s'agit d'un script, le programme peut s'exécuter sur différents types de machines disposant du même système d'exploitation. Là encore, cette restriction est susceptible d'être levée dans le futur.

Le contrôle de flux

Les parades à ces vulnérabilités et à ces attaques sont de deux ordres : le contrôle de flux (qui accède à quoi et comment) et la confidentialité des données. Commençons par la première.

Le contrôle de flux consiste à filtrer les paquets IP selon les adresses source et destination, les ports TCP et UDP, les types de protocoles (ICMP, OSPF, TCP, UDP, etc.), et éventuellement, selon des informations issues des couches applicatives. Il peut être réalisé par les routeurs ou par des équipements dédiés appelés **firewalls** (pare-feu en français).

Les technologies utilisées par les firewalls

Quatre technologies sont utilisées pour contrôler les flux :

- le filtrage de paquet (*packet filtering*) ;
- le filtrage par état des sessions (*stateful inspection*) ;
- le relais applicatif (*application gateway*), encore appelé mandataire, ou, abusivement, *application proxy*, *proxy gateway ou proxy* tout court ;
- le relais de circuit (*circuit relay* ou *circuit-level gateway*).

La confusion quant à l'emploi du terme de « proxy » est historique : le NCSA (*National Computer Security Association*) a, le premier, utilisé le terme de « *proxy service* » fonctionnant sur un firewall appelé « *application gateway* ». Les journalistes ont ensuite cédé à la facilité en prenant pour raccourci le mot « *proxy* » pour désigner ce type de firewall. Par amalgame, les termes « *application proxy* » et « *proxy gateway* » ont ensuite été abusivement employés pour désigner le firewall alors qu'il désigne en réalité le processus. Le vrai terme pour désigner ce type de firewall est donc « relais applicatif » (*application gateway* dans la littérature anglo-saxonne).

Un *proxy*, ou serveur *proxy*, désigne en réalité un serveur cache, c'est-à-dire une machine qui héberge les URL les plus récemment demandées dans le but d'économiser de la bande passante sur la connexion Internet (64 Kbit/s à 2 Mbit/s). Le navigateur Web est configuré de manière à interroger le serveur cache. Si l'URL demandée n'est pas dans le cache, le serveur relaie la requête vers le serveur cible. En toute rigueur, le proxy relaie la requête de la même manière qu'un firewall de type relais applicatif, mais on ne peut pas le considérer comme un firewall à part entière, car il ne dispose pas des mécanismes de protection utilisés par cette classe de produit.

Pour ajouter à la confusion, le paramètre « proxy », qui est défini au niveau des navigateurs, indique à ces derniers de rediriger les requêtes Web à destination de l'extérieur vers un serveur spécialisé (à la manière du paramètre « *default gateway* » qui indique au protocole IP de rediriger les paquets à destination de l'extérieur vers un routeur). À l'origine, ce serveur était bien un serveur proxy (un cache), mais par la suite, l'utilisation de cette fonction a été étendue pour désigner les firewalls qui contrôlent les droits d'accès à Internet.

Il est à noter qu'à cette fonction de cache, vient s'ajouter celle déjà gérée au niveau de chaque navigateur Web.

Le filtrage de paquet

Le filtrage de paquet est historiquement parlant la première technique, encore largement utilisée par les routeurs. Elle consiste à filtrer chaque paquet individuellement au niveau de la couche réseau en fonction des adresses source et destination, du port TCP ou UDP, du type de protocole (ICMP, RIP, etc.), et du sens du flux. Très peu d'informations (alertes, statistiques) sont par ailleurs enregistrées et aucun mécanisme de protection n'est implémenté contre les attaques.

Le « stateful inspection »

La technologie stateful inspection reprend les principes du filtrage de paquet mais conserve un état (historique) de toutes les sessions. Les paquets ne sont plus filtrés individuellement, mais en fonction des précédents qui appartiennent à un même échange.

Pour ce faire, le firewall examine les données du paquet et remonte jusqu'à la couche transport, voire applicative. Le filtrage est bien réalisé au niveau 3 (couche réseau), mais en fonction d'informations issues des couches supérieures. De ce fait, il est également possible de filtrer au niveau applicatif (commandes FTP, SMTP, DNS, etc.).

Pour s'adapter aux protocoles qui utilisent des ports aléatoires (les ports client pour tous les protocoles et les RPC), les règles sont générées dynamiquement pour le temps de la session à partir des règles de base définies par l'administrateur.

En outre, seul le premier paquet d'une session est comparé aux règles de filtrage, les suivants étant seulement comparés à l'état de la session, d'où des gains de performance par rapport aux routeurs filtrants. C'est la technologie la plus rapide.

Le relais applicatif

Cette technologie repose sur le principe du mandat : le firewall intercepte la requête du client et se substitue à ce dernier. En réalité, c'est donc le firewall qui envoie une requête au serveur. Il le fait de la part du client. Ce dernier croit dialoguer directement avec le serveur, alors que le serveur dialogue en fait avec un client qui est le firewall (le relais).

Cette technique implique de développer un client spécifique pour chaque application supportée (Telnet, FTP, HTTP, etc.), ce qui en fait la technologie la moins évolutive. Le support d'une application dépend des capacités de l'éditeur du produit.

La sécurité repose sur le fait qu'aucune connexion directe n'est réalisée entre le client et le serveur et que le client du firewall (le relais) est une version spéciale sans bogue, ne comportant aucune faille de sécurité, et susceptible d'intercepter les attaques.

Le relais de circuit

Le relais de circuit reprend la technologie de relayage sans permettre de filtrage au niveau applicatif et, surtout, sans tenir compte des spécificités liées à chaque protocole (échanges entre le client et le serveur, connexions ouvertes au sein d'une même session, etc.). Les paquets sont relayés individuellement. Cette technique offre donc un moins bon niveau de protection que le relayage applicatif.

Le premier produit à avoir introduit cette technologie est le freeware Socks. La mise en place d'un firewall Socks (on parle souvent de serveur Socks) requiert l'utilisation de clients (FTP, HTTP, etc.) spécifiques. Le client émet une requête au serveur Socks qui comprend l'adresse du serveur cible, le type de service demandé (port TCP ou port UDP) et le nom de l'utilisateur. Le serveur authentifie l'utilisateur, puis ouvre une connexion vers le serveur cible. Les flux de données sont ensuite relayés sans aucun contrôle particulier.

Cette technique est souvent utilisée en complément des relais applicatifs, plus particulièrement pour les protocoles non supportés par les relais, mais n'est jamais employée seule.

Comparaison des technologies

Les techniques de filtrage de paquet et de relais de circuit sont aujourd'hui obsolètes. Avec les techniques de piratage actuelles, il n'est pas envisageable de les utiliser sur un firewall. Les technologies stateful inspection et relais applicatif dominent actuellement le marché des firewalls et sont, de plus en plus souvent, combinées. C'est, par exemple, le cas de Netwall, de Cyberguard, de Watchguard et, dans une moindre mesure, de Firewall-1.

	Stateful Inspection	Relais applicatif
Évolutivité	+ Supporte tous les protocoles.	- Nécessite de développer un client pour tout nouveau protocole, à moins d'utiliser un relais de circuit (relais générique).
Performances	+ Filtrage optimisé et opéré au niveau du driver.	- Un processus par session.Les paquets sont traités par le relais et deux fois par l'OS.
Impact sur le système d'exploitation	+ Aucun, puisque l'OS est court-circuité	- L'OS doit être sécurisé puisqu'il traite les paquets.
Niveau de sécurité	+ L'état des sessions est conservé et les paquets (en-tête et données) analysés. Le filtrage applicatif élimine les commandes dangereuses. + Limite les possibilités de vol de session. - Les attaques par tunneling restent possibles.	+ Le relayage évite les attaques directes, et le filtrage applicatif élimine les commandes dangereuses. - Le relayage doit être associé à un filtrage (ports et adresses). - Les attaques par vol de session restent possibles.

La puissance de traitement des processeurs actuels permet de masquer les différences de performances entre les deux technologies pour des débits réseau compris entre 64 Kbit/s et 2 Mbit/s. La différence devient significative dans un réseau d'entreprise où le firewall doit supporter plusieurs interfaces Ethernet et de nombreuses règles de filtrage.

Choix d'un firewall

Il est recommandé d'utiliser deux firewalls de marque différente, un pour les connexions externes, dont Internet, et un autre pour le contrôle des flux internes :

- Si un pirate connaît bien un produit (et donc ses faiblesses), les chances qu'il en connaisse également un autre sont moindres.
- Un bogue présent sur un firewall a peu de chance de se retrouver également sur un autre.

Il est conseillé d'utiliser un firewall français pour les accès Internet :

- Leur origine offre plus de garanties qu'un produit étranger. Un firewall peut, en effet, comporter des trappes logicielles, sorte de portes dérobées (*back-doors*), qui contourne tous les mécanismes de protection du firewall, généralement au profit d'un service de renseignement. C'est tellement simple à réaliser et cela peut rapporter gros.
- Ils offrent les mêmes qualités que leurs homologues étrangers.
- Ils sont moins connus des pirates.

Un firewall plus répandu peut, en revanche, être utilisé pour la sécurisation du réseau interne. Sa fonction est en effet plus liée au partitionnement du réseau.

	Firewall-1 de Check Point	**Cyberguard de Cyberguard Corp**	**Netwall de Bull**
Remarques générales	Le plus répandu Facile à paramétrer	Un des meilleurs à résister aux attaques	Développements français
Plate-forme	Unix, NT	Unix	Unix
Technologie	Filtrage à état de lien	Filtrage à état de lien et relais applicatif	Filtrage à état de lien et relais applicatif
Authentification supportée	Radius, Secure ID	Radius, Secure ID, SSL	CP8, Radius
Partage de charge et redondance	Oui	Oui	Oui
Translation d'adresse	Oui	Oui	Oui
Protections contre les attaques	Oui	Oui	Oui
Risque de trappe logicielle	Oui	Oui	Oui, mais français !
Génération de filtres pour d'autres routeurs	Cisco, Nortel	Non	Non
Filtrage des commandes FTP	Oui	Oui	Oui
Filtrage des commandes SMTP	Non	Oui	Oui

Comparaison entre les routeurs et les firewalls

On peut se poser la question de l'intérêt d'un firewall par rapport à un routeur. Rappelons cependant les avantages du premier sur le deuxième :

- meilleure protection aux attaques par refus de service et par dissimulation d'adresse ;
- analyse de l'état des connexions TCP et UDP pour chaque session (à la place d'un simple filtrage paquet par paquet ne tenant pas compte des sessions) ;
- plus grande richesse de filtrage ;
- visualisation en temps réel des sessions ouvertes ;
- journalisation des tentatives d'intrusion et remontée d'alerte ;
- ergonomie de l'interface d'administration.

Les routeurs sont bien dévolus à l'acheminement des paquets selon des contraintes autres que sécuritaires : calcul des routes, reroutage en cas de panne d'un lien, gestion de la qualité de service, diffusion multicast, ou gestion des interfaces et protocoles WAN.

Les firewalls sont, quant à eux, dévolus au filtrage des sessions et des applications selon des règles complexes. Alors que les routeurs agissent entre les couches 2 et 3 (liaisons et IP), les firewalls agissent à partir de la couche 3.

Sous la pression du marché, ces différences tendent cependant à s'estomper, et les routeurs embarquent désormais des firewalls. Cette intégration présente l'avantage de réduire les coûts et d'enlever un étage de complexité à la chaîne de liaison LAN - routeur - firewall - LAN - routeur - WAN.

Définition d'une architecture à contrôle de flux

Un firewall seul ne permet pas d'interconnecter des réseaux de manière sûre. Il ne suffit pas d'installer un firewall et de programmer quelques règles de filtrage. Cette approche vous procure plus un sentiment de sécurité qu'une réelle sécurité. Il convient plutôt de déterminer une architecture globale dont le firewall n'est qu'un des composants.

Les principes de base devant présider à la politique de filtrage sont les suivants :

- Les filtres sont positionnés sur les firewalls.
- Tout ce qui n'est pas expressément permis est interdit.
- Toute demande de connexion depuis Internet est interdite.
- Les paquets TCP entrants doivent tous avoir le bit ACK positionné à « 1 » et le bit SYN à « 0 » (réponse à un paquet issu de l'extérieur).

En outre, plus le réseau de l'entreprise est grand, plus il peut être nécessaire de le partitionner et d'appliquer le concept de **défense en profondeur**. Ce principe consiste à définir plusieurs niveaux de protection de manière à ce que si le premier est pénétré (le firewall Internet, par exemple), les autres protègent des zones plus sensibles. Les dégâts causés par un pirate, un virus ou un ver, sont ainsi limités à la zone initiale d'attaque.

Ce principe, que nous retenons pour notre réseau, nous amène à définir la notion de **domaine de confiance**.

Une approche macroscopique permet de distinguer globalement deux domaines de confiance : le réseau interne couvrant les sites de notre entreprise et les réseaux externes regroupant les partenaires et sous-traitants. Une analyse plus détaillée montre que notre réseau se compose de trois grands ensembles :

- les sites parisiens formant notre réseau interne ;
- le reste du Groupe et ses filiales formant le réseau de notre entreprise ;
- les utilisateurs en accès distant.

L'objet de l'étude est donc de décrire l'architecture à mettre en place sur notre réseau parisien. La sécurité est abordée du point de vue de notre réseau : ce sont les ressources situées sur notre réseau qu'il faut protéger.

Il est à noter que le choix du firewall est indépendant de l'architecture et relève d'une autre étude.

Acteurs et matrice de confiance

Un certain nombre d'acteurs et de ressources auxquels il nous faut accéder ont été identifiés et classés en fonction de critères propres à notre site parisien et faisant suite à une étude d'analyse de la valeur et d'analyse de risque. Ainsi, les niveaux de protection et de confiance indiqués ci-dessous sont ceux vus de notre site parisien, site sur lequel le firewall est mis en place. Le point de vue d'autres sites, mettant en place leur propre firewall, serait sans doute différent. Il s'agit ici de déterminer les protections à mettre en place sur le firewall de Paris.

Les valeurs des niveaux de confiance ont les significations suivantes :

0 Aucune confiance n'est faite à cet acteur (domaine de non-confiance).

1 Il est fait moyennement confiance à cet acteur, par exemple, parce qu'il fait partie de la même organisation, ou parce qu'il s'engage sur des niveaux de protection.

2 Pleine confiance est faite à cet acteur, car nous maîtrisons son réseau et ses ressources.

Les valeurs des niveaux de protection ont les significations suivantes :

0 L'architecture firewall n'a pas de ressources à protéger.

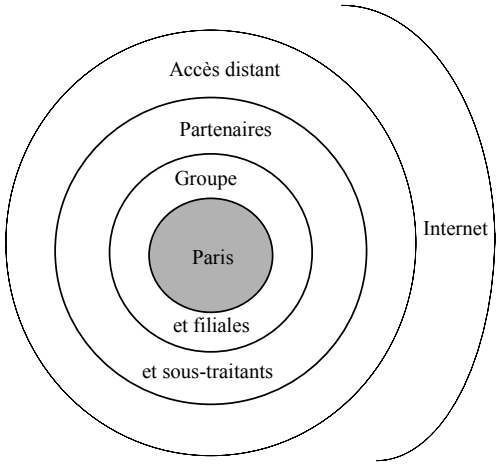
1 L'architecture firewall offre un premier niveau de protection, notamment par rapport à des acteurs externes à l'entreprise.

2 L'architecture firewall doit protéger absolument les ressources ; c'est ce pourquoi elle est mise en place.

La matrice de confiance est élaborée selon des critères qui dépendent amplement de considérations organisationnelles et techniques propres à chaque entreprise.

Acteurs	Niveau de confiance	Ressources de l'acteur	Niveau de protection
Site parisien	2	Sites intranet, fichiers, SGBD, bureautique	2
Groupe / Filiales	1	Sites intranet, SGBD, bureautique	1
Partenaire / Sous-traitants	0	Sites intranet, fichiers	0
Acteurs en accès distants	0		0

On peut donc dresser une cible de confiance, le centre étant le domaine à sécuriser, et les domaines devenant de moins en moins sûrs au fur et à mesure que l'on s'en éloigne. Les pointillés indiquent des limites non maîtrisées par notre entreprise et donc, un domaine de non-confiance.



Matrice de communication

Le tableau suivant précise quel acteur doit accéder aux ressources détenues par tel autre acteur.

L'acteur X ↓ accède aux ressources de l'acteur Y →	Paris	Groupe / Filiales	Partenaires / Sous-traitants	Accès distants
Paris	---	Oui	Oui	---
Groupe / Filiales	Oui	---	Oui	---
Partenaires / Sous-traitants	Oui	Non	---	---
Accès distants	Oui	Oui	Non	---

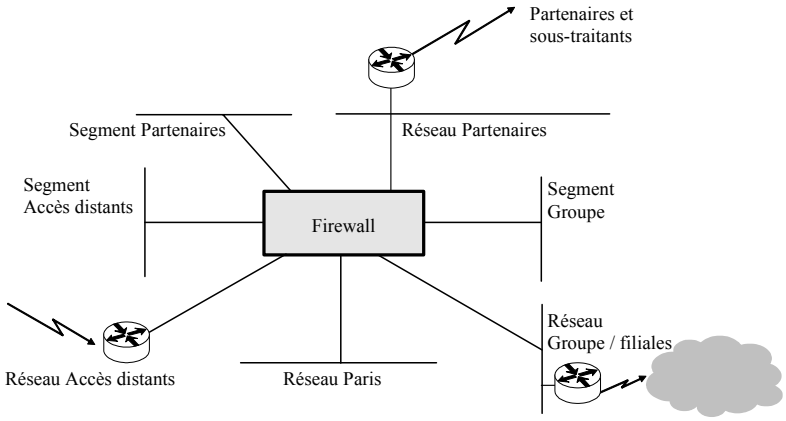
En grisé, apparaissent les acteurs ayant un niveau de confiance inférieur à ceux détenant des ressources cibles. Il nous faut donc mettre en place des ressources partagées et augmenter la sécurité en isolant l'acteur ou la ressource sur un réseau dédié.

Le cas de l'acteur « Accès distants » est particulier, car il doit accéder directement aux ressources de notre entreprise, alors qu'il est situé dans un domaine de non-confiance. Ce qui signifie qu'il faut mettre en place des mécanismes de sécurité supplémentaires permettant de renforcer la sécurité.

On peut, dès lors, identifier trois segments dédiés hébergeant des ressources partagées :

- un segment « Accès distants » hébergeant des ressources dédiées aux accès distants ;
- un segment « Partenaires » partagé par les partenaires et notre site parisien ;
- un segment « Groupe » partagé par le Groupe et notre site parisien.

Les segments, communément appelés DMZ (*DeMilitarized Zones*), permettent ainsi d'éviter tout flux direct entre acteurs ayant des niveaux de confiance différents. De la même manière, les acteurs de niveaux de confiance différents seront placés sur des réseaux distincts, c'est-à-dire connectés à différentes interfaces du firewall.



Mécanismes de sécurité supplémentaires

Les protections contre les attaques et le contrôle de flux offerts par un firewall ne suffisant pas à se protéger des domaines de non-confiance, les mécanismes suivants permettent d'ajouter un niveau supplémentaire de sécurité vis-à-vis des réseaux externes tels ceux des partenaires :

- masquage des adresses internes par translation ;
- coupure des flux avec des relais applicatifs, de préférence intégrés au firewall ;
- interdiction des sessions Netbios et RPC, qui ne sont pas maîtrisables par les firewalls ;
- sécurisation des serveurs situés sur les segments dédiés par durcissement du système d'exploitation, restriction des droits d'accès et désactivation des services système et réseau non utilisés.

LES TRANSLATIONS D'ADRESSES ET DE PORTS (RFC 3022)

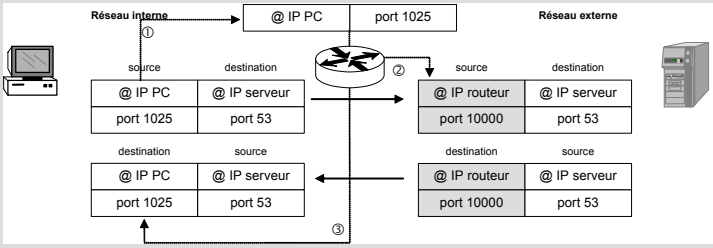
La translation d'adresse, appelée **NAT** (*Network Address Translation*) consiste à changer l'adresse source et/ou destination des paquets IP traversant un routeur ou un firewall. Ce type de mécanisme peut être utilisé à plusieurs desseins :

- masquage du plan d'adressage interne vis-à-vis de réseaux externes tels qu'Internet pour des questions de sécurité ;
- résolution des conflits d'adresses : lorsque deux sociétés fusionnent, la probabilité est non négligeable que leurs plans d'adressage soient identiques, surtout si elles respectent la RFC 1918 (cf. chapitre 5).

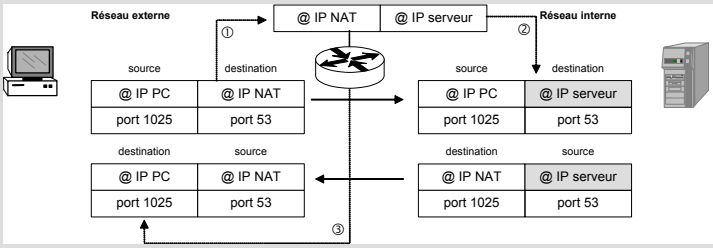
Pour Internet, la translation d'adresse a été la solution privilégiée pour palier à la pénurie d'adresses publiques, puisqu'à l'origine toutes les entreprises voulant se raccorder à Internet demandaient ce type d'adresses. Depuis, les entreprises optent majoritairement pour un plan d'adressage RFC 1918 ou privé hors RFC 1918 (c'est-à-dire des adresses non réservées auprès des organismes d'Internet), avec une translation d'adresse vis-à-vis de l'Internet et de leurs partenaires externes.

Il existe trois types de translations d'adresses :

- **Dynamique N pour 1** : N adresses sources sont translatées en 1 adresse source. Vis-à-vis de l'extérieur, les paquets sont donc tous issus de la même source. Afin de distinguer les destinataires des paquets reçus en retour, le port TCP ou UDP source des paquets émis est également modifié et associé à l'adresse IP source originelle. Pour des protocoles comme ICMP, le discriminant est généralement un numéro de requête ou de session contenu dans le segment de données transporté dans IP.



- **Dynamique N pour M** (avec M ≤N), qui est une variante du précédent type. Si M est inférieur à N, le mécanisme de discriminant précédemment décrit est utilisé.
- **Statique 1 pour 1** : chaque adresse est translatée en une autre adresse. Ce mode est surtout utilisé pour les adresses de destination correspondant à des serveurs à atteindre.



La translation de port **PAT** (*Port Address Translation*) consiste à modifier les ports destination dans le but d'ajouter un niveau de sécurité supplémentaire en conjonction avec un relais applicatif (cf. section traitant des firewalls).

Le firewall traitera des translations d'adresses pour masquer les adresses internes vis-à-vis des réseaux des partenaires. Cette traduction doit être statique pour les machines cibles auxquelles accèdent les partenaires et sous-traitants. Les clients sortants feront quant à eux l'objet d'une translation dynamique N pour 1.

Quant aux utilisateurs en accès distant, qui doivent néanmoins accéder aux ressources de notre réseau, il est difficile de les orienter vers des ressources partagées sur un segment dédié, car les mécanismes de réplication avec les serveurs internes seraient trop complexes, voire impossibles à mettre en œuvre. De ce fait, les mécanismes devant renforcer la sécurité doivent avoir trait à la confidentialité :

- authentification forte permettant de s'assurer de l'identité de l'utilisateur ;
- chiffrement des données afin d'assurer la confidentialité face aux réseaux externes traversés (réseau téléphonique, ADSL, etc.) ;
- contrôle d'intégrité permettant de s'assurer que les paquets IP (en-tête et données) ne sont pas modifiés.

Cas du DNS

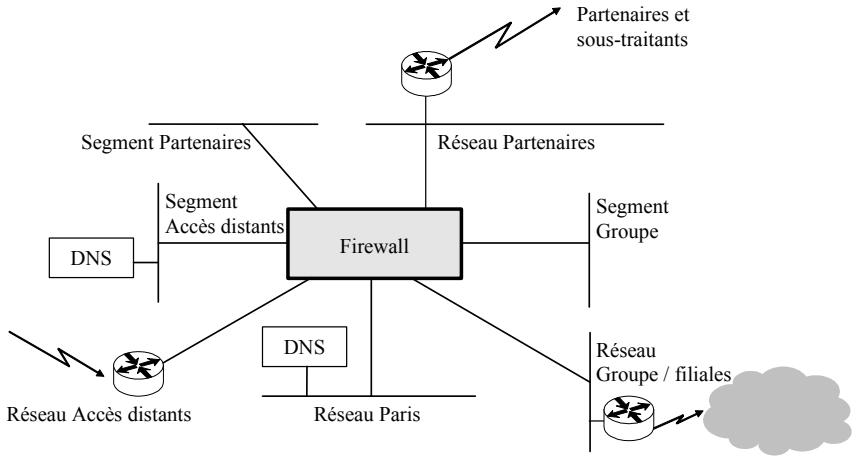
Les partenaires disposent de leur propre DNS, qu'ils renseignent avec les adresses que nous leur donnons et auxquelles ils associent les noms souhaités. De cette manière, aucun échange n'est nécessaire entre leur DNS et le nôtre.

Quant à nos utilisateurs en accès distant qui peuvent être localisés n'importe où dans le monde, il est préférable de leur dédier un DNS qui contient des informations partielles.

Nous devons, en conséquence, considérer deux serveurs DNS :

- un vrai serveur DNS sur le réseau parisien à usage exclusif de nos utilisateurs ;
- un serveur DNS contenant des informations partielles à l'usage de nos utilisateurs en accès distant.

Chacun de ces serveurs réside sur son réseau ou segment respectif et appartient à un système DNS distinct, c'est-à-dire disposant de sa propre racine sans aucun échange avec l'autre.



Le DNS du réseau parisien fournit les informations aux clients du réseau parisien :

- Il gère le SOA *siege.societe* contenant les ressources situées sur notre réseau interne.
- Il gère le SOA *partenaires.societe* contenant les ressources des partenaires auxquelles nos utilisateurs doivent accéder. Ce SOA contient les adresses qui nous sont présentées par les partenaires et les noms connus par nos utilisateurs. Cette configuration présente l'avantage d'éviter tous flux DNS entre les réseaux Partenaires et notre réseau, ce qui renforce la sécurité. On suppose ici que le nombre de machines est raisonnablement limité de manière à ne pas compliquer les tâches d'exploitation.
- Il gère le SOA *part-ext.siege.societe* contenant les ressources situées sur le segment Partenaires. Ce SOA n'est connu que de nos utilisateurs.
- Les requêtes concernant d'autres SOA sont redirigées vers les serveurs compétents du Groupe *via* le serveur racine, utilisant en cela les mécanismes classiques du DNS.

Le DNS du segment Accès distants fournit les informations à nos utilisateurs en accès distant. Il gère une copie partielle du SOA *siege.societe*, qui contient uniquement les adresses et les noms des machines situées sur notre réseau interne et qui doivent être accessibles depuis l'extérieur. Ces adresses doivent être les mêmes que celles configurées dans les règles de filtrage du firewall.

Cas de Netbios

La résolution des noms Netbios utilise les serveurs WINS (*Windows Internet Name Service*) de Microsoft. Pour les mêmes raisons que DNS, le segment Accès distants doit disposer de son propre serveur, qu'il est préférable de paramétrer avec des adresses statiques. De cette manière, seules les machines explicitement autorisées seront visibles par les utilisateurs en accès distant, et le protocole WINS sera bloqué.

Un serveur WINS exporte ses ressources vers un autre serveur WINS en mode « push » *via* le protocole Nameserver sur le port TCP 42. Le montage de ressources disques et la messagerie Outlook utilisent, quant à elles, le protocole Netbios sur TCP 139, que nous devons laisser passer.

Les utilisateurs en accès distant doivent disposer de leur propre domaine de compte leur permettant d'accéder, par le biais des relations de confiance (au sens Windows du terme), au domaine de ressources du siège. Un PDC (*Primary Domain Controler*) doit donc être installé sur le segment Accès distants.

Il est toutefois possible de contourner les mécanismes de relations de confiance en utilisant la procédure « connect as ». Il faut donc filtrer les messages SMB correspondants. En revanche, les connexions entre les domaines Paris et Groupe utiliseront la procédure « connect as » afin d'éviter de diffuser tous les comptes (plusieurs milliers).

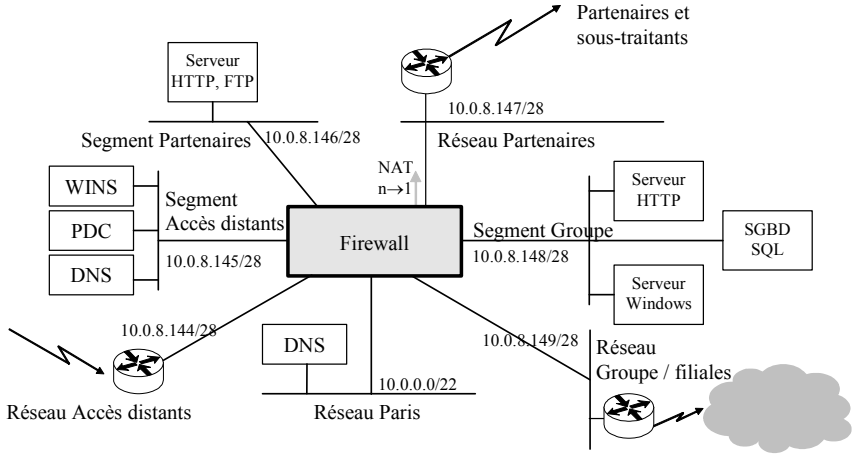
Services IP et matrice de flux

Chacune des ressources est accessible par un protocole et un numéro de port UDP ou TCP bien identifiés, qui vont servir de base au filtrage à appliquer sur le firewall. La matrice de flux ci-dessous présente les sens de connexion, ce qui n'interdit pas les échanges dans les deux sens. Ainsi, les connexions ne peuvent être initiées qu'à partir d'un réseau situé dans un domaine de confiance supérieur à un autre (exception faite de nos utilisateurs en accès distant), et aucune connexion ne peut être initiée depuis un segment.

Vers → / De ↓	Réseau Paris	Réseau Groupe	Réseau Partenaires	Réseau Accès distants	Segment Partenai-res	Segment Groupe
Réseau Paris	---	DNS	HTTP, FTP		HTTP, FTP	HTTP, SQL, Netbios
Réseau Groupe	DNS	---			HTTP, FTP	HTTP, SQL, Netbios
Réseau Partenaires			---		HTTP, FTP	
Réseaux Accès distants	Telnet(1), HTTP, SQL, DNS, Netbios	HTTP, SQL, DNS, Netbios		---		
Segment Public					---	
Segment Groupe						---

Case vide = aucun service accessible. (1) Pour la population des exploitants uniquement.

Il est désormais possible de préciser l'architecture à mettre en place, de déterminer le plan d'adressage et les mécanismes de translation d'adresses.



Dans la mesure du possible, les serveurs utilisant des protocoles différents doivent être isolés sur des segments distincts, ce qui est le cas de Netbios, protocole difficilement maîtrisable. Cela permet d'appliquer une politique de filtrage plus restrictive, par adresse plutôt que par subnet. Si l'acteur Groupe avait eu un niveau de confiance égal à zéro, les serveurs accessibles *via* Netbios auraient même dû être installés sur des segments dédiés, de manière à éviter les rebonds entre serveurs Web et Windows. Même si les serveurs ne sont pas accessibles *via* ces deux protocoles, un cheval de Troie peut toujours l'utiliser.

Conformément à notre plan d'adressage défini au chapitre 5, les réseaux et segments dédiés autour du firewall partagent un réseau de classe C subneté sur 28 bits, 10.0.8.144/28, dans le but d'économiser des adresses. On obtient ainsi 16 sous-réseaux de 14 adresses chacun. Cinq sous-réseaux, un par interface du firewall hors notre réseau interne, sont ici utilisés.

Le firewall ne fonctionne avec aucun protocole de routage dynamique tels qu'OSPF ou RIP, là encore pour des questions de sécurité. Le routage est donc statique entre les routeurs et le firewall.

La mise en place de relais ne se justifie pas sur un réseau interne, car d'une part nous exerçons un contrôle sur les acteurs avec qui nous sommes en relation (nos utilisateurs, les partenaires, le reste du Groupe), et d'autre part la diversité et la complexité des protocoles à filtrer ne permettent pas aux relais de fournir un niveau de sécurité supplémentaire. Le raisonnement vis-à-vis d'Internet serait, en revanche, différent, puisque nous aurions affaire à un réseau public, donc par définition non maîtrisé, et parce que le nombre de protocoles, nécessairement restreints, (HTTP, SMTP, etc.) seraient, de plus, facilement filtrables par des relais.

Matrice de filtrage

La matrice de flux permet d'élaborer la matrice de filtrage, c'est-à-dire les règles à implémenter sur le firewall. Celle-ci doit être construite sur le principe de l'ouverture de flux minimale : dans la mesure du possible, les flux doivent être ouverts entre couples d'adresses plutôt qu'entre couples de subnets. Ainsi :

- Les flux impliquant un serveur doivent être ouverts sur la base de l'adresse du serveur.
- Les flux impliquant des utilisateurs authentifiés doivent être ouverts sur la base des noms des utilisateurs (qui seront associés à une adresse lorsqu'ils seront authentifiés).
- Les flux impliquant des utilisateurs non authentifiés doivent être ouverts sur la base de leur subnet. Il est envisageable d'appliquer une politique de filtrage plus restrictive vis-à-vis des partenaires et de filtrer par couple d'adresses.

Une fois établie, la matrice doit pouvoir être directement transposée dans les règles de filtrage du firewall.

Rôle et fonctionnement des firewalls

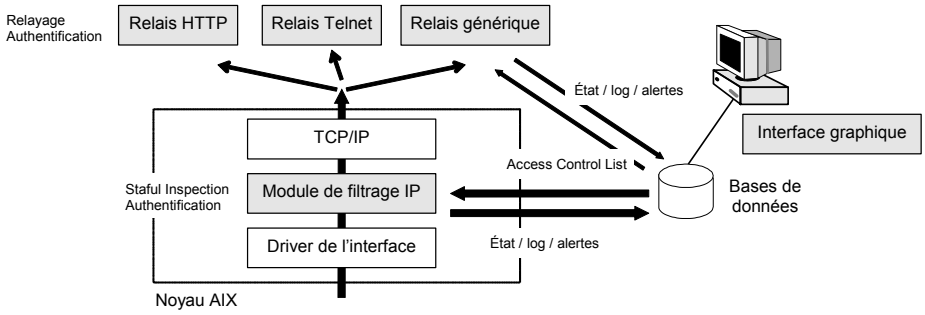
L'objet de cette section est de montrer, avec l'exemple de Netwall, le fonctionnement d'un firewall et les fonctions qu'il remplit :

- protection active contre les attaques ;
- détection d'intrusion ;
- filtrage des paquets IP sur la base des adresses et des ports source et destination ;
- translation d'adresses IP (NAT) et des ports TCP/UDP (PAT) ;
- filtrage des commandes applicatives (HTTP, FTP, DNS, etc.) ;
- authentification des utilisateurs, permettant ainsi de filtrer les sessions par utilisateur et non plus sur la base des adresses IP sources ;
- chiffrement et intégrité des données à l'aide du protocole IPsec ;
- décontamination anti-virus, le plus souvent en liaison avec un serveur dédié à cette tâche ;
- filtrage des URL, soit directement, soit en liaison avec un serveur dédié ;
- filtrage des composants actifs (ActiveX, applet Java, Java scripts, etc.).

Il appartient à l'étude d'architecture de déterminer si toutes ces fonctions doivent être ou non remplies par un seul équipement et à quel endroit du réseau.

Architecture

Netwall fonctionne sur un serveur Estrella ou Escala avec le système d'exploitation AIX de Bull ou avec un AIX sécurisé, appelé BEST-X, qui est certifié F-B1/E3 par le CERT. Une version sous Windows reprend l'interface graphique, le module de filtrage IP et uniquement le relais SMTP. Netwall utilise simultanément deux technologies : le firewall est composé d'un module de filtrage IP, qui repose sur le principe de stateful inspection et de plusieurs relais applicatifs.



Le module de filtrage IP, qui réside au niveau du noyau Unix, est intercalé entre le driver de la carte et les couches IP. Tous les paquets entrants et sortants passent obligatoirement par le

module de filtrage IP et, en fonction du protocole, sont ensuite transmis au relais applicatif correspondant.

La partie relayage applicatif de Netwall est constituée de plusieurs relais (démon Unix ou service NT), un par type d'application à relayer (Telnet, SMTP, etc.) dérivés des sources du kit TIS (*Trusted Information Systems*) et de Netscape Proxy Server pour le relais HTTP.

Les relais sont lancés dans un environnement restreint (*chroot*) et sans les privilèges superviseur (*root*). Ils ne reçoivent jamais les flux directement.

Fonctionnement

Tous les flux pour un service particulier (Telnet, SMTP, etc..) sont obligatoirement traités par le relais applicatif correspondant. Si un relais s'arrête de fonctionner (arrêt manuel, disque plein ou bogue), le module de filtrage IP bloque le service associé.

Le module de filtrage IP offre toutes les caractéristiques présentées par la technologie stateful inspection :

- Il garde une trace de toutes les sessions (TCP, UDP, RPC, etc.).
- Il génère dynamiquement les règles de filtrage.
- Il analyse au niveau applicatif (FTP, RPC, etc.).
- Il gère la fragmentation IP.

Les règles sont définies dans une base appelée ACL (*Acces Control List*). Pour chacune d'entre elles, on définit l'action à entreprendre (accepter, rejeter, authentifier), le niveau d'information à enregistrer pour les événements (normal, détaillé, debug) et la manière de remonter une alerte (trap SNMP, e-mail, *pager* ou déclenchement d'un script personnalisé). Les règles peuvent être activées pendant des périodes données (heure, jour de la semaine).

Les protections contre le *spoofing*, les attaques par inondation *syn*, les *ping of death*, etc., sont activées au niveau du module de filtrage IP, entre le driver réseau et la couche IP.

La configuration des interfaces est particulière. Trois types d'interfaces sont prédéfinis à la base : interne, externe et DMZ (*DeMilitarized Zone*). Lorsque le firewall est configuré avec plus de trois interfaces, l'administrateur doit définir des domaines de sécurité basés sur les réseaux IP puis affecter les interfaces à un domaine. La définition des règles est ensuite réalisée à partir de ces domaines.

Relais applicatif

Les applications relayées sont HTTP, FTP, SMTP et Telnet. Chaque relais gère le contrôle de flux, le filtrage des commandes (HTTP, FTP, SMTP), l'authentification, l'enregistrement des événements et la connexion vers la machine cible.

Deux processus SMTP séparés fonctionnent, un pour les connexions avec Internet et l'autre pour les connexions internes. Le relais SMTP empêche toute connexion directe à partir d'un client Telnet sur port 25.

Les autres protocoles peuvent être traités par un relais générique fonctionnant sur le mode relais de circuit : la demande de connexion est interceptée pour authentifier la session, puis la connexion est autorisée vers le serveur cible. Les flux sont ensuite relayés sans contrôle particulier, autre que celui réalisé par le module de filtrage IP.

Le relais HTTP offre les fonctions suivantes :

- gestion des changements de mots de passe utilisateur depuis un navigateur ;
- filtrage des URL ;
- filtrage des applets Java et des scripts Java.

Les composants ActiveX peuvent également être filtrés.

Cas des protocoles non relayés

Tous les protocoles, qu'ils soient relayés ou non, sont traités par le module de filtrage IP sur le mode stateful inspection. Les protocoles non relayés sont donc traités au niveau du noyau.

Les paquets peuvent être filtrés en fonction du protocole (UDP, TCP, ICMP), des adresses source et destination, du port (UDP ou TCP) ou sur le champ « option » situé dans l'en-tête du paquet IP.

Authentification

L'authentification est gérée individuellement par chaque relais applicatif et globalement par le module de filtrage IP. Les niveaux de fonctionnalités sont cependant différents.

Tous les protocoles peuvent être authentifiés par le module de filtrage IP *via* la carte à puce CP8 de Bull. Cette solution nécessite un lecteur de carte sur le poste de l'utilisateur et un serveur CP8LAN. L'authentification est transparente pour l'utilisateur : le firewall intercepte une demande de connexion et demande l'autorisation au serveur CP8LAN qui interroge la carte de l'utilisateur. Dans le cas d'une réponse positive, le firewall permet aux paquets de passer. Si la carte CP8 est retirée de son lecteur, le serveur CP8LAN la détecte et en informe le firewall qui ferme toutes les sessions de l'utilisateur en question. Les communications entre les clients, le firewall et le serveur CP8LAN sont chiffrées *via* DES.

Les proxy FTP et Telnet supportent, en plus, les authentifications *via* S/Key, SecurID, les mots de passe Unix classiques et également ISM/Access Master. En revanche, le relais HTTP supporte uniquement le mode classique (mot de passe non chiffré) et la variante chiffrée *via* SSL.

Deux procédures de connexion sont proposées aux utilisateurs :

- mode non transparent : l'utilisateur se connecte d'abord sur le firewall pour s'authentifier, puis ouvre une session sur la machine cible et, éventuellement, s'authentifie à nouveau ;
- mode transparent : l'utilisateur se connecte directement sur la machine cible, mais cette demande est interceptée par le firewall, qui demande une authentification préalable. La session est alors relayée normalement par le relais.

Translation d'adresses

Deux modes de translation sont proposés :

- Le mode statique offre une gestion au niveau du module de filtrage IP : chaque adresse est translatée en une adresse (mode 1→1) qui sera seule visible de l'extérieur (par rapport à l'interface d'où elle provient).
- Le mode dynamique offre une gestion au niveau des relais applicatifs : les adresses d'un réseau ou d'un sous-réseau sont translatées en une seule (mode n→1 dit de masquage).

Netwall prend en compte les protocoles DNS et Netbios, qui présentent la particularité de transporter les adresses IP dans la partie « données » des paquets.

Redondance

Un module complémentaire, appelé SafeNetwall, permet d'offrir le partage de charge et la redondance entre deux firewalls. Les deux machines dialoguent *via* une interface Ethernet de préférence dédiée pour des questions de sécurité. Les sessions en cours ne sont pas coupées.

Administration

L'administration du firewall peut être réalisée de différentes manières :

- à partir d'une interface graphique X/Windows ;
- à partir d'un navigateur Web *via* HTTP ;
- *via* des menus en mode texte (interface SMIT fournie en standard avec AIX).

Si la station est déportée, les sessions avec Netwall sont authentifiées *via* DES et l'intégrité des données est contrôlée par une signature DES. Les paquets en eux-mêmes ne sont pas chiffrés. Les accès au firewall peuvent être contrôlés *via* les mécanismes propres à ISM/Access Master.

Les statistiques affichées en temps réel comprennent le nombre de paquets traités et rejetés ainsi que les ressources système utilisées. L'administrateur peut, à travers l'interface graphique, modifier des paramètres système tels que la taille des caches, les files d'attente et les tampons utilisés pour la fragmentation.

Un produit distinct, appelé Netwall Partitioner, permet de générer des règles de filtrage pour différentes plates-formes : les routeurs Cisco et NCS ainsi que les firewalls Netwall et PIX (Cisco).

Log et audit

Comme pour les autres fonctionnalités, les événements sont générés à deux niveaux : relais applicatifs et module de filtrage IP. Ils sont enregistrés dans deux types de fichiers : audit et alertes.

Le module de filtrage IP enregistre dans le fichier d'audit les adresses IP source et destination, les ports, l'en-tête du paquet (mode détaillé), le contenu du paquet (mode trace) et la date de l'interception. Le fichier des alertes contient les adresses IP source et destination, les ports, les protocoles et la date de l'événement.

Les relais enregistrent des informations propres à leur application (FTP, Telnet, HTPP, etc.) : heure de début et durée de la session, nombre d'octets échangés, adresses source et destination, nom de l'utilisateur et informations sur les sessions d'authentification. En mode trace, le contenu du paquet est également enregistré.

Les alertes peuvent être déclenchées à partir d'un événement simple (rejet d'un paquet, refus d'authentification) ou sur des conditions spécifiques fondées sur le nombre d'occurrences d'un événement ou sa fréquence (nombre d'occurrences de l'événement pendant une période donnée).

Mécanismes de protection

La fragmentation des paquets est gérée de la manière suivante : lorsqu'un fragment arrive et qu'il n'est pas le premier d'une série, il est mémorisé jusqu'à la réception du premier fragment qui contient toutes les informations. Cette approche est à double tranchant : elle offre plus de sécurité puisque tous les fragments sont lus avant le traitement complet du paquet, mais elle est potentiellement vulnérable aux attaques de type *teardrop*, puisqu'il faut allouer de la mémoire pour stocker les fragments.

Intégrité

Le module de filtrage IP bloque par défaut tout le trafic et contrôle l'activité des autres processus. Si l'un d'eux s'arrête, le module de filtrage IP bloque le trafic correspondant. Les fonctionnalités liées à la TCB (*Trusted Computing Base*) peuvent être étendues à Netwall et vérifier l'intégrité des fichiers (checksum sur le contenu, date de modification, propriétaire).

Chiffrement des données

Le firewall est compatible avec les VPN (*Virtual Private Network*) IPsec. Pour des débits supérieurs à quelques Mbit/s, mieux vaut utiliser un boîtier dédié appelé SecureWare.

Remarques sur l'administration des firewalls

Il faut souligner que, dans tous les cas de figure, la sécurité réclame une administration continue. Une équipe d'exploitation doit être formée spécifiquement aux techniques liées à la sécurisation des réseaux IP et au firewall lui-même. Les tâches sont notamment les suivantes :

- positionnement des règles de filtrage ;
- analyse de toutes les règles de filtrage (une règle peut en effet mettre en défaut toutes les autres) ;
- gestion des autorisations d'accès ;
- analyse et purge des log.

La confidentialité

La sécurité, exprimée en termes de confidentialité, recouvre plusieurs fonctionnalités :

- le chiffrement des données ;
- l'intégrité des données, qui consiste à vérifier que les données reçues sont bien celles qui ont été originellement émises ;
- la signature, qui consiste à s'assurer qu'un objet a bien été émis par celui qui prétend l'avoir fait ;
- le certificat, qui consiste à s'assurer que la clé publique du destinataire est bien la sienne.

Les mécanismes mis en œuvre pour ces fonctionnalités reposent sur les algorithmes de chiffrement. Des protocoles intègrent ensuite ces algorithmes dans des applications telles que les échanges réseau au sens large, les cartes bancaires, le paiement électronique, etc.

Les algorithmes de chiffrement

Il existe trois catégories d'algorithmes de chiffrement :

- les algorithmes à clé secrète, dits **symétriques** : l'émetteur doit partager une clé secrète avec chacun des destinataires ;
- les algorithmes à clé publique, dits **asymétriques** : deux clés (une privée et une publique) peuvent chiffrer et déchiffrer un message ;
- Les algorithmes à clé secrète **non réversibles**.

Les algorithmes **symétriques** nécessitent qu'émetteur et destinataire se soient, préalablement à tout échange, mis d'accord sur un mot de passe connu d'eux seuls. Chaque émetteur doit donc gérer autant de clés qu'il a de correspondants. Les algorithmes de ce type les plus répandus sont DES (*Data Encryption Standard*), une version améliorée appelée Triple DES, RC4, RC5, RC6, IDEA (*International Data Encryption Algorithm*) et AES (*Advanced Encryption System*), le successeur désigné de DES.

Le principe des algorithmes **asymétriques** est le suivant :

- La clé publique de B est diffusée auprès de tous ses correspondants.
- Une personne A, désirant envoyer un message à B, chiffre le message avec la clé publique de B.
- B est le seul à pouvoir déchiffrer le message avec sa clé privée.

Il est à noter que les clés sont commutatives :

- Il est possible de chiffrer avec la clé publique et de déchiffrer avec la clé privée.
- Il est possible de chiffrer avec la clé privée et de déchiffrer avec la clé publique.

Cette propriété est utilisée pour la signature numérique (voir plus loin).

L'avantage d'un algorithme asymétrique est qu'un destinataire B n'a besoin que d'une clé pour tous ses correspondants. Avec un algorithme symétrique, il faut en effet une clé secrète

à partager avec chaque correspondant (à moins qu'on accepte l'idée que chaque correspondant puisse déchiffrer les communications entre B et quiconque partageant la même clé).

Les algorithmes à clé publique les plus en pointe actuellement sont RSA (*Rivest, Shamir et Adleman*), et ECC (*Elliptic Curve Cryptosystem*).

Les algorithmes **non réversibles** consistent à transformer un message en un mot de 128 bits ou plus. L'algorithme de hachage utilisé doit être à collision presque nulle, c'est-à-dire que la probabilité que deux messages différents produisent le même mot de 128 bits, doit être la plus faible possible. Les algorithmes de ce type les plus répandus sont MD5 (*Message Digest* – RFC 1319 et 1321) et SHA (*Secure Hash Algorithm*).

Algorithme	Norme ou propriété	Type	Longueur de la clé
ECC	Libre	Asymétrique	128 bits
RSA	Propriété RSA	Asymétrique	40, 56, 128 ou 1024 bits
IDEA	Propriété MediaCrypt	Symétrique	128 bits
CAST	RFC 2144 et 2612	Symétrique	128 ou 256 bits
AES	FIPS 197	Symétrique	128, 256 bits ou plus
DES	FIPS 42-2	Symétrique	40 ou 56 bits
Triple DES	FIPS 42-3	Symétrique	3 fois 40 bits
RC4	Propriété RSA	Non réversible	40 bits
RC5	Propriété RSA	Non réversible	40, 56 ou 1024 bits
SHA	FIPS 180-1	Non réversible	160 bits

FIPS = *Federal Information Processing Standard* : normes édictées par le NIST.

Efficacité des algorithmes de chiffrement

La performance d'un algorithme à clé (secrète ou publique) se caractérise par son inviolabilité qui repose sur deux facteurs :

- un problème mathématique, réputé insoluble (un « *hard problem* »), souvent basé sur la factorisation des grands nombres premiers en utilisant des modulos dans des groupes générateurs Zp ;
- la longueur de la clé : 40, 56, 128 et 1024 étant des valeurs courantes.

Le tableau suivant présente le temps mis pour trouver une clé DES. Il est extrait d'un rapport rédigé en 1996 par sept spécialistes parmi lesquels Rivest (de RSA), Diffie (de Diffie-Hellman) et Wiener (auteur d'une méthode permettant de casser l'algorithme DES).

Budget en $	Matériel utilisé	Temps mis pour trouver la clé DES	
		40 bits	56 bits
Presque rien	Ordinateurs	1 semaine	infaisable
400 $	Ship programmable	5 heures	38 ans
10 000 $	Ship ou ASIC	12 minutes	556 jours
300 000 $	ASIC	18 secondes	3 heures
10 M$	ASIC	0,005 secondes	6 minutes

Depuis, les choses ont bien évolué : un étudiant de l'École polytechnique a réussi à casser la clé 40 bits de l'algorithme RC4 en 3 jours, rien qu'en utilisant les temps CPU libres de super-calculateurs. Le record est détenu par des étudiants de l'université de Berkeley qui, en février 1997, ont cassé la clé en 3 heures 30 à l'aide de 250 stations de travail en réseau.

La société RSA, qui commercialise l'algorithme du même nom, a organisé un concours dont le but était de casser la clé 56 bits de l'algorithme DES, épreuve remportée par un consultant indépendant.

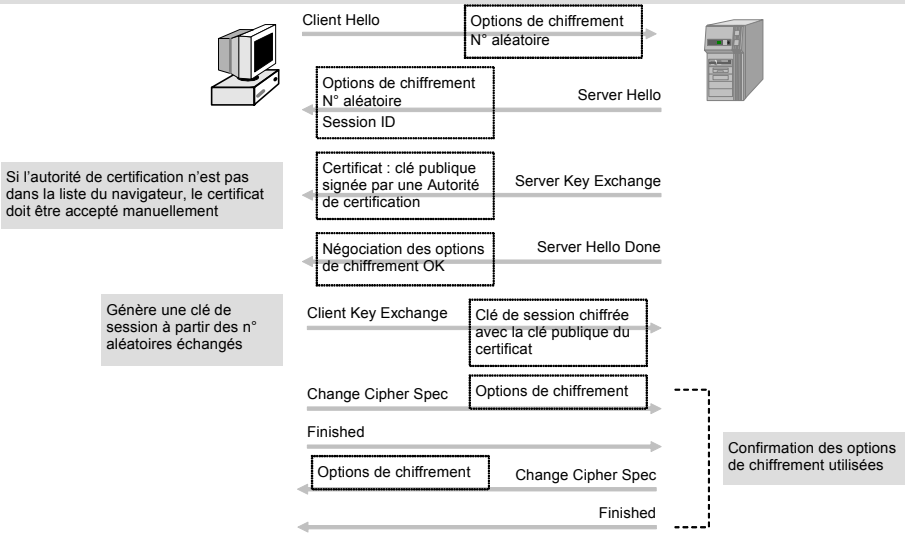
L'infaisabilité de la propriété mathématique utilisée par l'algorithme est également un critère important : par exemple, ECC à 128 bits est aussi sûr que RSA à 1024 bits, et RSA à 40 bits est aussi sûr que DES à 56 bits.

Tout organisme privé ou public peut investir dans la production d'un ASIC spécialisé. La hauteur de son investissement dépend simplement du gain qu'il peut en tirer, d'où l'intérêt de bien évaluer la valeur commerciale de l'information à protéger.

En conclusion, les clés à 40 bits sont aujourd'hui considérées comme non sûres et l'algorithme DES comme obsolète.

LE POINT SUR SSL / TLS (RFC 2246 ET 3546)

Le protocole **TLS** 1.0 (*Transport Layer Security*) est la version standardisée du plus connu **SSL** 3.0 (*Secure Socket Layer*). Il permet d'**authentifier** le serveur et, éventuellement, le client, ainsi que de **chiffrer** et de contrôler l'**intégrité** des données de bout en bout entre un PC et un serveur. Le principe repose sur l'utilisation de clés publiques diffusées sous forme de **certificats X.509**, c'est-à-dire signées par une **autorité de certification**. Dans le cas du paiement électronique, c'est le serveur, sur lequel est effectué le paiement, qui envoie son certificat, mais certaines applications peuvent requérir un certificat de la part du client afin d'authentifier ce dernier.



La clé publique (1024 bits ou plus) sert à chiffrer la clé de session à l'aide d'un algorithme asymétrique tel que **RSA**. La clé de session (128 bits au minimum, 40 bits étant courant mais pas suffisant) est utilisée pour chiffrer les données avec un algorithme symétrique tel que 3DES ou **IDEA** (RC4 et DES étant courants mais trop faibles). Un algorithme non réversible, tel que MD5 ou, mieux, SHA, est utilisé pour contrôler l'intégrité des données échangées.

De nombreux protocoles reposant sur TCP (pas UDP) existent en une version SSL :

Protocole	Port TCP	Protocole	Port TCP	Protocole	Port TCP
SMTPS	465	FTPS	989/990	HTTPS	443
IMAPS	993	Telnets	992		
POP3S	995	LDAPS	636		

Le chiffrement est très consommateur de ressources CPU, d'autant qu'une page Web comportant 4 images nécessite la création de 5 sessions SSL (une pour chaque requête GET), ce qui signifie 5 négociations séparées. Des boîtiers dédiés sont donc souvent positionnés devant les serveurs devant traiter plusieurs sessions SSL en parallèle. Enfin, bien que la plupart des navigateurs supportent SSL 2.0 et 3.0, il est fortement recommandé de ne pas utiliser la version 2.0, qui comporte des failles de sécurité connues.

Les protocoles de sécurité

L'intégration des algorithmes à des produits commerciaux est rendue possible grâce aux protocoles de sécurité. Ceux-ci comprennent, au minimum, le chiffrement, mais peuvent également prendre en compte l'intégrité des données, la signature, ou encore la gestion des certificats.

Protocole	Origine / Norme	Algorithmes utilisés
IPsec	RFC 2401 à 2406	DES, Triple DES, MD5, etc.
SSL (*Secure Socket Layer*)	Netscape RFC 2246 et 3546	Authentification et chiffrement RC4 40 ou 128 bits
PGP (*Pretty Good Privacy*)	RFC 1991	Signature MD5, SHA Chiffrement IDEA, CAST et RSA jusqu'à 2048 bits
S-HTTP (*Secure HTTP*)	Enterprise Integration Technologies	RC4 40 ou 128 bits Certificat X.509
S/Mime (*Secure Multi Purpose Mail Extensions*)	Draft IETF conforme à PKCS	DES ou RC2 40 bits et RSA 40 bits
PGP/Mime	RFC 2015	Idem PGP appliqué à Mime
PEM (*Privacy Enhanced Mail*)	RFC 1421 et 1424	Intégrité Authentification Chiffrement
PKCS (*Public Key Cryptography Standards*)	RSA	Chiffrement DES et RSA Signature MD5 et RSA Échange de clé Diffie-Hellman

Ces protocoles définissent les échanges, le format des données, les interfaces de programmation et l'intégration dans un produit. Par exemple, SSL (*Secure Socket Layer*) et SHTTP sont destinés aux navigateurs Web alors que S/Mime et PGP/mime s'intègrent à la messagerie Internet.

L'intégration de ces algorithmes est décrite par le standard PKCS sous la forme de profils décrits dans le tableau suivant.

	Profil PKCS #								
Standards	**1**	**3**	**5**	**6**	**7**	**8**	**9**	**10**	**Autres standards liés**
Syntaxe indépendante des algorithmes									
Signature numérique des messages					x		x		
Enveloppe numérique des messages					x				
Demande de certificat							x	x	
Certificats									X.509, RFC 1422
Certificats étendus				x			x		
Liste de certificats révoqués									
Chiffrement par clé privée						x	x		
Échange de clés									
Syntaxe dépendant des algorithmes									
RSA à clé publique	x								
RSA à clé privée	x								
Algorithmes									
Message digest : MD2, MD5									RFC 1319, 1321
Chiffrement à clé privée DES									RFC 1423
Chiffrement à clé publique RSA	x								
Signature: MD2, MD4, MD5 w/RSA	x								
Chiffrement des mots de passe			x						
Échange de clés Diffie-Hellman		x							

Ainsi, S/MIME repose sur les standards PKCS #1, #3, #7 et #10. De même, le GIE Mastercard/Visa utilise PKCS #7 comme base des spécifications SET (*Secure Electronic Transaction*). Les algorithmes retenus sont RSA 1024 bits pour la signature et l'enveloppe, DES 56 bits pour le chiffrement (uniquement pour des faibles montants) et Triple DES.

Les protocoles d'échange de clés

Le problème des algorithmes symétriques réside en la diffusion préalable de la clé entre les deux parties désirant échanger des messages : le destinataire doit connaître la clé utilisée par l'émetteur, la clé devant demeurer secrète. Le moyen le plus simple est l'échange direct ou *via* un support autre qu'informatique. Outre les problèmes de confidentialité, des problèmes pratiques peuvent survenir, surtout s'il faut changer souvent de clé.

Afin d'éviter cela, d'autres algorithmes ont été développés. Il s'agit de :

- Diffie-Hellman ;
- SKIP de SUN ;
- PSKMP (*Photuris Secret Key Management Protocol*) ;

- ECSVAP1 et ECSVAP2 (*Elliptic Curve Shared Value Agreement Protocol*) ;
- ISAKMP (*Internet Security Association and Key Management Protocol*), Oakley et Skeme ;
- IKE (*Internet Key Exchange*) utilisé par IPsec.

L'algorithme de Diffie-Hellman est le plus connu. Son principe est le suivant :

- Soit un générateur g du groupe Zp où p est un grand nombre premier.
- A choisit une clé a, calcule g^a et l'envoie à B.
- A choisit une clé b, calcule g^b et l'envoie à A.
- A calcule $(g^b)^a$ et B calcule $(g^a)^b$.
- On obtient la clé secrète $(g^b)^a = (g^a)^b$.

Les valeurs a et b ne sont connues que de A et de B respectivement et ne circulent pas sur le réseau. Même si l'échange est intercepté, il sera très difficile de calculer $(g^a)^b$ à partir de g^b ou de g^a. L'avantage est que cette opération peut se répéter automatiquement et plusieurs fois au cours d'un échange, de manière à changer de clé avant qu'elle ne soit éventuellement cassée.

Les autres algorithmes sont moins répandus, soit parce qu'ils sont moins efficaces (exemple de SKIP), soit parce qu'ils sont encore trop complexes à mettre en œuvre ou encore parce que trop récents, ce qui est le cas de ECSVAP1.

La gestion des certificats

Les algorithmes symétriques posent le problème de l'échange préalable des clés secrètes. Les algorithmes asymétriques n'ont pas ce type de problème, car la clé publique est connue de tous. Cela soulève cependant un autre point : est-on certain que la clé publique est bien celle de la personne à qui l'on veut envoyer un message ?

Les clés publiques sont, en effet, hébergées sur des serveurs, qui sont donc susceptibles d'être piratés. Pour contourner l'obstacle, le destinataire peut communiquer sa clé publique à qui la demande, mais on retombe dans les difficultés liées à la diffusion des clés symétriques.

L'autre solution consiste à certifier la clé auprès d'une autorité au-dessus de tout soupçon, appelée CA (*Certificate Authority*). L'émetteur A désirant envoyer un message à B demande au serveur de certificats de confirmer que la clé publique dont il dispose, est bien celle de B :

- A demande un certificat pour la clé publique de B.
- Le CA utilise sa clé privée pour signer la clé publique de B : cette signature constitue le certificat.
- A vérifie la signature avec la clé publique du CA.

Mais comment être certain que la clé publique du CA est bien la sienne ? Le problème subsiste en effet. On peut alors désigner une autorité qui certifierait les CA, un gouvernement

ou un organisme indépendant, par exemple. Aux États-Unis, les CA habilités à émettre des certificats sont VeriSign, Entrust, RSA. En France, la DCSSI (Direction centrale de la sécurité des systèmes d'information) a habilité des sociétés telles que CertPlus, CertEurope ou Certinomis ainsi que des filiales spécialisées créées par des banques françaises (cf. http://www.minefi.gouv.fr/dematerialisation_icp/dematerialisation_declar.htm).

Le standard en matière de certificats est X.509 de l'ISO repris dans le RFC 1422. Il est utilisé par les navigateurs Web *via* SSL, les annuaires LDAP (*Lightweight Directory Access Protocol*) ou encore par le SET (*Secure Electronic Transaction*) pour le paiement électronique.

Les serveurs qui gèrent, distribuent et valident les certificats sont appelés PKI (*Public Key Infrastructure*). Pour plus de détails, le lecteur pourra se référer au site www.pki-page.org.

La signature numérique

La signature numérique permet de prouver que l'émetteur du message est bien celui qu'il prétend être. Une fonction de hachage est opérée sur le message, et la valeur obtenue (le *digest*) est chiffrée avec la clé privée de l'émetteur. Tous les destinataires peuvent vérifier que l'émetteur est bien celui qu'il prétend être en déchiffrant la valeur de hachage avec la clé publique de l'émetteur et en la comparant avec la valeur calculée sur le message reçu, à l'aide de la même fonction de hachage.

Émetteur A	Destinataire B
1. Applique MD5 au message et obtient la signature σ	4. Décrypte Σ avec clé publique de A et obtient σ
2. Chiffre σ avec sa clé privée et obtient Σ	5. Applique MD5 au message déchiffré et obtient X
3. Envoie Σ et le message chiffré	6. si $X=\sigma$, alors le message est bien issu de A

Pour ce type d'algorithme, on trouve :

- DSA (*Digital Signature Algorithm*) qui repose sur le standard DSS (*Digital Signature Standard*) du NIST (*National Institute of Standard and Technologies*), et utilise une clé privée à 1024 bits et un algorithme basé sur le log de Zp analogue à Diffie-Hellman.
- ECDSA (*Elliptic Curve de DSA*) qui est analogue à DSA mais se base sur un algorithme ECC.
- Rabin Digital Signature qui repose sur la racine carrée des nombres modulo Zn où n est le produit de 2 grands nombres premiers : la racine carrée de Zn est difficile à trouver (la racine carrée de X modulo Zn est très difficile à trouver quand on ne connaît pas n).

L'enveloppe numérique

L'enveloppe numérique est une technique de plus en plus utilisée. On la trouve dans PGP, Mime, PEM (*Privacy Enhanced Mail*). Le principe est le suivant :

- Le message est chiffré à l'aide d'un algorithme à clé secrète tel IDEA ou CAST.
- La clé secrète est à son tour chiffrée à l'aide d'un algorithme à clé publique tel RSA.
- Les destinataires déchiffrent la clé secrète à l'aide de leur clé privée.
- La clé secrète ainsi déchiffrée est utilisée pour déchiffrer le message.

Par exemple, PGP peut utiliser IDEA 128 bits pour chiffrer le message et RSA, 1 024 bits pour chiffrer la clé IDEA.

Les algorithmes à clé publique sont plus puissants que les algorithmes à clé symétrique : ils sont, de ce fait, plus lents et soumis à de plus grandes restrictions d'utilisation et d'exportation. L'intérêt de la technique de l'enveloppe est qu'elle permet de protéger la clé de chiffrement avec un algorithme réputé inviolable, et d'utiliser un algorithme moins puissant mais plus rapide pour chiffrer le message.

Les contraintes législatives

En France, la DCSSI (Direction centrale de la sécurité des systèmes d'information) autorise une utilisation libre de tous les algorithmes d'authentification, de signature et de contrôle d'intégrité, quelle que soit la longueur de leur clé.

En ce qui concerne le chiffrement, l'utilisation est libre dès lors que la longueur de clé n'excède pas 128 bits, et que le logiciel ait été déclaré à la DCSSI par le fournisseur. Au-delà de cette longueur, le chiffrement est soumis à autorisation, à moins que la clé soit gérée par un tiers de confiance agréé par la DCSSI et susceptible de remettre les clés aux autorités judiciaires et de sécurité, à leur demande.

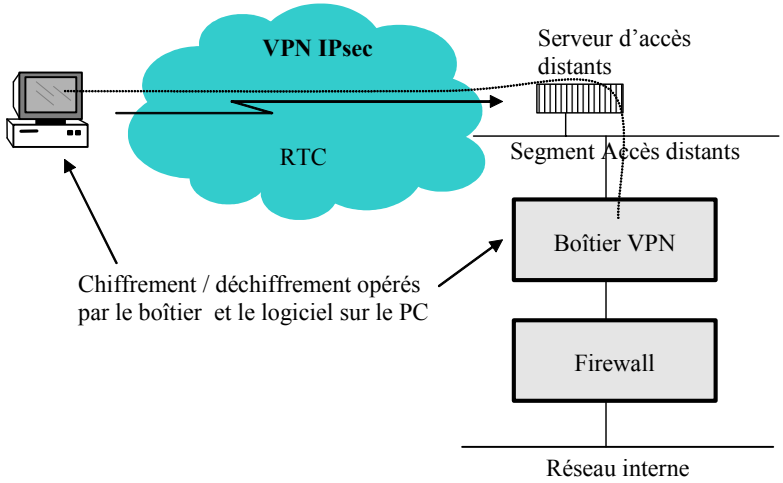
Toute exportation d'un moyen de chiffrement doit faire l'objet d'une déclaration à la DCSSI. Le site www.ssi.gouv.fr permet de consulter les textes les plus à jour.

Le chiffrement des données

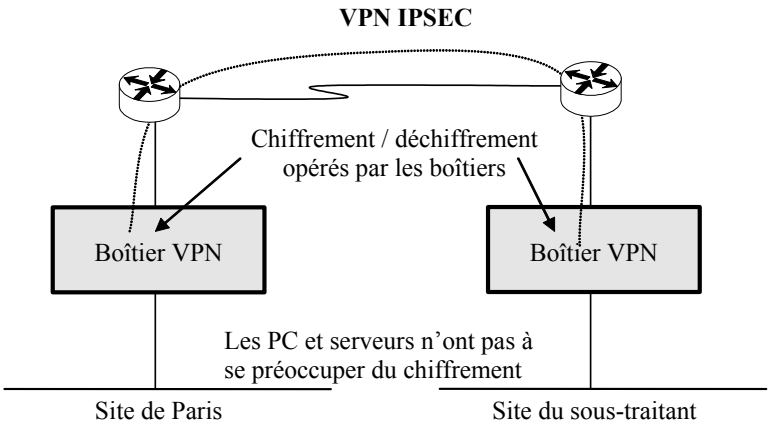
Les VPN IPsec (*Virtual Private Network – IP security*), cités au chapitre 10, consistent à créer un tunnel IP chiffré entre un PC isolé et un site (mode *Client-to-LAN*) ou entre deux sites (mode *LAN-to-LAN*). Le terme de VPN est donc (un peu) usurpé, car le réseau privé virtuel ne repose pas sur un partitionnement logique du réseau d'un opérateur, mais sur le chiffrement des données. Il est d'ailleurs possible d'utiliser IPsec au-dessus d'un VPN de niveau 2 ou de niveau 3 (cf. chapitre 10) et de toute liaison WAN ou LAN en général. La notion de VPN appliquée à IPsec ne vient qu'avec le chiffrement des données ; il est donc possible de créer un réseau privé au-dessus d'Internet, qui est un réseau public résultant de la concaténation de réseaux opérateur (cf. annexes).

D'une manière générale, le chiffrement permet de renforcer la sécurité pour les données sensibles circulant dans un domaine de non-confiance. Si le niveau de sécurité l'exige, il peut s'appliquer :

- aux accès distants qui empruntent le réseau téléphonique d'un opérateur (mode PC-to-LAN) ;



- ou entre deux sites interconnectés par un réseau opérateur, que ce soit une LS, un VPN de niveau 2 ou 3 ou encore Internet (mode LAN-to-LAN) ;

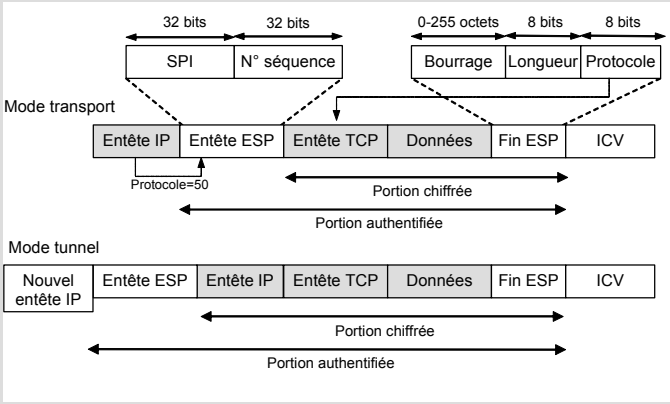


- et, éventuellement, entre deux PC ou serveurs de notre réseau, qu'il soit de type LAN, MAN, WAN ou WLAN.

La fonction VPN peut être intégrée ou non dans le firewall.

LE POINT SUR IPSEC (RFC 2401, 2402, 2406)

Le protocole IPsec permet de **chiffrer** et de contrôler l'**intégrité** des données, soit de bout en bout (mode transport), soit sur une portion du réseau (mode tunnel), ainsi que d'**authentifier** l'émetteur. Pour ce faire, le paquet IP est encapsulé dans un paquet **ESP** (*Encapsulation Security Payload* – RFC 2406).



SPI (*Security Parameters Index*) : n° incrémental qui, en combinaison de l'adresse IP destination, identifie de manière unique la session.

N° séquence du paquet au sein de la session.

Bourrage : permet d'aligner le champ ICV sur un mot de 4 octets. Peut également être requis par l'algorithme de chiffrement.

Longueur : nombre d'octets du bourrage.

Protocole : protocole originellement encapsulé dans IP (tcp, udp…).

ICV (*Integrity Check Value*) : valeur permettant de contrôler l'intégrité des données. La longueur de ce champ dépend de l'algorithme retenu lors de la négociation.

Le **mode transport** est utilisé entre un client et un serveur pour protéger les données transportées dans le paquet IP originel, mais pas l'entête IP lui-même. Le **mode tunnel** est, quant à lui, utilisé si l'une des deux extrémités au moins est un firewall ou un boîtier IPsec : il permet de chiffrer entièrement le paquet IP originel, l'entête et ses données. Il est à noter qu'il est possible d'encapsuler le mode transport dans le mode tunnel afin, par exemple, de chiffrer entièrement et de bout en bout le paquet IP original, tout en traversant des firewalls. L'inverse est également possible afin, par exemple, de masquer à un réseau tiers tel qu'Internet le plan d'adressage interne à l'entreprise.

Un autre type d'encapsulation, appelé **AH** (*Authentication Header* - RFC 2402), offre les mêmes fonctionnalités hormis le chiffrement des données. Il est très peu utilisé.

La taille maximale des données transportées dans un paquet IP est définie par le MTU (*Maximum Transfer Unit*) de la couche sous-jacente (par exemple, 1 500 octets pour Ethernet). Si le paquet IP est déjà au maximum de son MTU, l'overhead ajouté par IPsec provoque systématiquement l'activation du mécanisme de fragmentation IP (cf. chapitre 5).

Les sessions IPsec, appelées ***security associations***, sont initiées et gérées à l'aide du protocole **IKE** (*Internet Key Exchange* – RFC 2409), qui s'appuie sur les procédures **ISAKMP** (*Internet Security Association and Key Management Protocol* – RFC 2408), Oakley et Skeme. Les algorithmes d'authentification, de chiffrement et de contrôle d'intégrité, ainsi que les longueurs de clés, sont à cette occasion négociés. Selon le **sélecteur** retenu, deux nœuds peuvent établir une session IPsec, soit pour l'ensemble de leurs échanges, soit pour chaque couple d'adresses IP et/ou de port TCP/UDP.

L'authentification

L'authentification apporte une protection supplémentaire par rapport à la connexion par mot de passe classique, car celui-ci, même s'il est chiffré, circule entre le client et le serveur. Intercepté et déchiffré, il peut donc être réutilisé par un éventuel pirate. De plus, comme il ne change pas d'une session à l'autre, la session interceptée peut être rejouée sans avoir besoin de déchiffrer le mot de passe.

L'authentification consiste donc à s'assurer que l'émetteur est bien celui qu'il prétend être.

Ce type de protection peut aussi bien s'appliquer à des utilisateurs qu'à des protocoles réseau tels qu'OSPF, afin de s'assurer que seuls les routeurs dûment identifiés soient habilités à échanger des informations de routage.

Les mécanismes d'authentification

La connexion d'un utilisateur sur un ordinateur implique la saisie d'un nom de compte et d'un mot de passe. Deux mécanismes de protection sont mis en œuvre à cette occasion :

- le chiffrement du mot de passe à l'aide d'un algorithme quelconque ;
- la génération d'un mot de passe différent à chaque tentative de connexion.

Le premier mécanisme utilise les algorithmes classiques symétriques et asymétriques. Le deuxième mécanisme, appelé **One Time Password**, peut être réalisé de deux manières :

- par défi/réponse (*challenge*/*response*) ;
- par mot de passe variant dans le temps à l'aide de calculettes appelées *token cards*.

Il existe beaucoup de produits qui utilisent des variantes. Parmi les plus représentatifs, on trouve :

- CHAP (*Challenge Handshake Authentication Protocol* – RFC 1994), mécanisme logiciel ;
- S/Key (RFC 1938), mécanisme logiciel ;
- Secure ID (Security Dynamics), calculette ;
- ActivCard (société française), calculette, clé USB ou logiciel ;
- Access Master de Bull, carte à puce CP8.

Le protocole **CHAP** (*Challenge Handshake Authentication Protocol*) est utilisé en conjonction avec PPP (cf. chapitre 9) pour les accès distants. C'est un mécanisme d'authentification qui offre un premier niveau de sécurité :

- Le serveur envoie un challenge aléatoire.
- Le client répond avec un hachage MD5 opéré sur la combinaison d'un identifiant de session, d'un secret et d'un challenge.

Le secret est un mot de passe connu du logiciel client et du serveur.

L'authentification de CHAP repose donc sur une clé secrète et un algorithme de hachage. Le mot de passe ne circule pas sur le réseau mais doit être codé en dur sur le logiciel client et le serveur. Sur un serveur Radius, ce mot de passe peut même être inscrit en clair dans la base de données.

Le mécanisme **S/Key** de Bellcore repose également sur l'algorithme de hachage MD5 appliqué à une liste de mots de passe valables pour une seule tentative de connexion :

- La liste des clés est générée à partir d'une clé initiale : la clé N s'obtient en appliquant MD5 à la clé N-1 et ainsi de suite.
- Le serveur qui connaît la clé N+1 demande la clé N (codée à partir de la clé initiale).
- Soit l'utilisateur possède la liste des clés codées, soit il utilise un programme qui la génère à la demande à partir de la clé initiale.
- Le serveur applique MD5 à n et le résultat est comparé avec la clé N+1.
- Le serveur enregistre la clé n et demandera la prochaine fois la clé N-1.
- Arrivé à la clé 1, il faudra régénérer une nouvelle liste.

Le problème de S/Key est la gestion de cette liste. De plus, il est relativement facile de se faire passer pour le serveur (technique du *spoofing*) :

- Le vrai serveur demande la clé N à l'utilisateur.
- Le faux serveur détourne la communication et demande la clé N-10 à l'utilisateur.
- L'utilisateur consulte sa liste (ou génère la clé) et renvoie la réponse au faux serveur.
- Disposant de la clé N-10, il suffit alors d'appliquer MD5 à N-10 pour obtenir la clé N-9, et ainsi de suite.
- Il suffit alors de lancer MD5 de 2 à 10 fois sur cette clé pour obtenir les réponses pour les clés N-1 à N-10 que le vrai serveur demandera aux prochaines demandes de connexion.

La calculette **SecureID** repose cette fois sur un microprocesseur qui génère un mot de passe temporel toutes les 60 secondes. Celui-ci dérive de l'horloge et d'une clé secrète partagée par la carte et le serveur :

- L'utilisateur saisit son PIN (*Personnal Identification Number*) sur sa carte.
- La carte génère un mot de passe en fonction du temps et de la clé secrète (*seed*) contenue dans la carte.
- L'utilisateur envoie au serveur le mot de passe affiché par la carte.
- Le serveur génère en principe le même mot de passe et compare la réponse avec son propre calcul.

Le serveur doit cependant tenir compte du décalage de son horloge avec celle de la carte.

Fonctionnement d'une calculette d'authentification

La clé 3DES utilisée par la calculette **ActiveCard** est une combinaison de plusieurs clés :

- une clé secrète « organisation », commune à plusieurs utilisateurs ;
- une clé secrète « ressource » propre à une application, à une machine sur laquelle on veut se connecter, à une utilisation particulière de la calculette ;
- une clé « utilisateur » qui est le nom de l'utilisateur, le nom d'un compte, ou un identifiant quelconque propre au propriétaire de la calculette.

La clé ressource désigne une machine cible sur laquelle on veut se connecter ou toute autre application (par exemple, « Accès distants »). Dix ressources sont ainsi programmables sur la calculette, ce qui veut dire que l'on peut utiliser la carte pour se connecter à dix machines différentes. Pour la version logicielle, ce nombre est illimité.

La calculette génère des clés sur la base d'informations prédéfinies (les trois clés précédemment citées, l'identifiant de l'utilisateur et le numéro de série de la calculette), et sur la base d'une information variable qui est un compteur. Les informations statiques sont stockées dans la calculette et dans la base de données du serveur.

À chaque connexion, le compteur du serveur d'authentification s'incrémente de N. De même, à chaque fois que l'utilisateur appuie sur la touche de fonction ressource, la calculette incrémente son compteur interne de N. Les deux compteurs, ceux de la calculette et du serveur, doivent être synchronisés. Si l'utilisateur appuie trop de fois sur la touche de fonction par inadvertance, les deux compteurs peuvent être désynchronisés, ce qui empêche l'utilisateur de s'authentifier.

Dès qu'il reçoit le mot de passe, le serveur procède au même calcul et compare les résultats. S'ils sont identiques, il renvoie une autorisation. Afin de prendre en compte les éventuels écarts avec le compteur de la calculette, il procède à ce calcul avec les N/2 prochaines valeurs de son compteur et avec les N/2 précédentes.

Ce principe de fonctionnement est le même que celui utilisé par les commandes d'ouverture à distance des portes de voiture.

Les serveurs d'authentification

Pour sécuriser la connexion à un ordinateur, il existe, on vient de le voir, une multitude de techniques et de produits différents. Ces produits ne sont pas disponibles sur toutes les plates-formes, car le travail d'intégration est important.

Afin de pallier à cette difficulté, il est possible d'ajouter une couche supplémentaire en intercalant un serveur entre le client et la machine cible. Les éditeurs de systèmes d'authentification n'ont alors plus qu'à développer une seule interface avec le serveur de sécurité, et les machines cibles n'ont qu'à supporter un seul protocole d'authentification avec le serveur d'authentification.

Il existe trois grands standards sur le marché.

Produit	Protocoles et produits supportés
Radius (RFC 2138 et 2139)	SecureID, CHAP (qui utilise MD5)
Tacacs (RFC 1492) et Tacacs+	SecureID, CHAP, MD5
Kerberos (RFC 1510)	DES

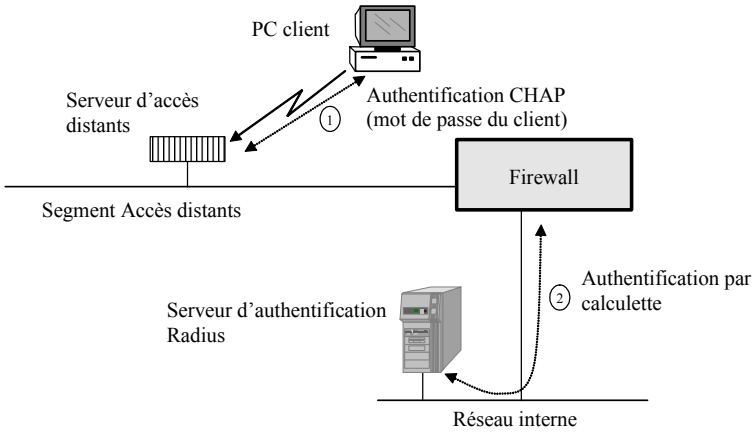
Radius est le plus utilisé. Kerberos a été le premier standard en la matière, mais sa mise en œuvre trop complexe a freiné son utilisation. Le support de Kerberos par Windows 2000 pourrait annoncer cependant la résurgence de ce standard qui n'a jamais réellement percé. Tacacs est plus simple mais présente trop de lacunes. Tacacs+ a été développé par Cisco mais n'est pas réellement un standard.

Exemple d'authentification forte pour les accès distants

Les accès des utilisateurs depuis l'extérieur de l'entreprise, qu'ils soient nomades ou sédentaires, sont particulièrement dangereux, car ce type de besoin nécessite l'accès aux ressources de l'entreprise depuis un domaine de non-confiance. En plus du contrôle de flux réalisé par un firewall, il est nécessaire d'authentifier les utilisateurs, un peu à la manière d'une banque avec les cartes à puce.

Une première amélioration a été apportée par le système d'authentification CHAP lors de la connexion PPP entre le client et le serveur d'accès distants. Ce mécanisme ne suffit cependant pas, car il repose sur le partage d'un secret entre l'utilisateur et le serveur. Il est préférable d'utiliser un mécanisme d'authentification « fort » fondé sur des calculettes. Chaque calculette appartient à son détenteur, et tout comme une carte bancaire, son utilisation peut être associée à la saisie d'un code personnel.

Ce type de système repose sur un serveur Radius et, au choix, de calculettes hardwares et logicielles installées sur les postes de travail. Ce serveur est situé sur le réseau interne ou sur un segment dédié, et dialogue avec le firewall ou le serveur d'accès distants, installé sur un autre segment dédié. Il est également utilisé comme station d'administration permettant de configurer et d'activer les calculettes.



La solution mise en œuvre est indépendante du serveur d'accès distants et du mécanisme d'authentification. Il y a en fait deux niveaux d'authentification :

1. Au niveau PPP, entre le PC client et le serveur d'accès distants (*via* CHAP), permettant d'avoir accès au service réseau. Cette étape est transparente pour l'utilisateur à partir du moment où le mot de passe est stocké dans le poste de travail client.

2. Au niveau du firewall, entre le PC client et le serveur d'authentification (à l'aide des calculettes), permettant d'accéder à notre réseau intranet. Cette étape nécessite que l'utilisateur saisisse le mot de passe affiché sur la calculette.

La procédure de connexion qui en résulte est la suivante :

- Lors de la connexion, le PC envoie le mot de passe au serveur d'accès distants *via* le protocole CHAP. Vous utilisez le même mécanisme pour vous connecter à Internet avec votre modem RTC ou ADSL.
- Le firewall intercepte le premier paquet issu du PC. Il demande alors le nom de l'utilisateur. Celui-ci le saisit.
- Le firewall demande ensuite le mot de passe. L'utilisateur doit alors saisir son code d'identification sur la calculette ou sur son PC, s'il s'agit d'une clé USB ou d'une carte à puce connectée au PC *via* un lecteur.
- La calculette affiche alors un code constitué de chiffres et de lettres, que l'utilisateur saisit tel quel comme mot de passe. S'il s'agit d'une clé USB ou d'une carte à puce, l'utilisateur n'a rien à faire.
- Le firewall transmet les informations (nom de l'utilisateur et mot de passe) au serveur d'authentification *via* le protocole Radius.
- Le serveur d'authentification vérifie les informations par rapport à sa base de données et donne ou non l'autorisation au firewall (toujours *via* Radius).

Le mot de passe a été transmis du PC client au serveur d'authentification. Mais cela n'est pas grave puisqu'il n'est pas rejouable : celui qui l'intercepterait ne pourrait pas le réutiliser, car il n'est valable qu'une fois.

L'intérêt d'opérer l'authentification au niveau du firewall plutôt qu'au niveau du serveur d'accès distants est multiple :

- L'authentification permet de créer des règles de filtrage par nom utilisateur, indépendamment de leurs adresses IP.
- Le firewall enregistre dans son journal toutes les sessions des utilisateurs identifiés par leur nom.
- La politique de sécurité est implémentée et exploitée en un point unique, à partir de la console d'administration du firewall.

Avec ce principe, le serveur d'accès distants gère les connexions et les modems RTC ou ADSL, tandis que le firewall implémente la politique de sécurité de notre entreprise.

La gestion des calculettes (configuration, génération des codes secrets, etc.) est réalisée à partir d'une station d'administration, qui héberge également le serveur d'authentification Radius.

Annexes

Le câblage

Normes CEN relatives au câblage

Référence	Objet
EN 55022	Compatibilité électromagnétique pour les équipements informatiques
EN 55024	Protection contre les champs électromagnétiques (dérivé de IEC 801)
EN 50081-1 EN 50082-1	Compatibilité électromagnétique des composants électroniques et des câbles
EN 50167 EN 50168 EN 50169	Caractéristiques et performances des câbles FTP
EN 50173	Spécifications des performances des systèmes de câblage (dérivé de l'ISO/IEC DIS 11801)
EN 186000-1	Spécifications génériques pour les câbles et connecteurs fibres optiques
EN187000	Spécifications générales des câbles fibres optiques
EN 188000	Spécifications générales des fibres optiques

Normes EIA/TIA relatives au câblage

Référence	Objet
EIA/TIA-569	Spécifications générales des infrastructures d'immeuble (espace alloué, chemins de câbles...)
EIA/TIA-568	Spécifications générales des systèmes de câblage
EIA/TIA-586	Définition des 5 catégories de câbles UTP (*Unshielded Twisted Pair*)
EIA/TIA TSB 36	Spécifications des câbles de catégorie 5
EIA/TIA TSB 40	Spécifications des connecteurs de catégorie 5
EIA-455-48	Spécifications du cœur des fibres optiques 62.5/125
EIA-455-27 EIA-455-48	Spécifications des caractéristiques mécaniques du revêtement des fibres optiques 62.5/125
EIA-455-57	Spécifications de l'ouverture numérique des fibres optiques 62.5/125

Normes ITU-T relatives au câblage

Référence	Objet
ITU-T G651	Caractéristiques des câbles fibres optiques à gradient d'indice 50/125µm
ITU-T G652	Caractéristiques des câbles monomode

Les interfaces physiques

Avis de l'ITU-T relatifs aux interfaces physiques

Avis	Contenu de la spécification
V.6	Standardisation des vitesses de transmission synchrone sur les lignes spécialisées
V.11	Caractéristiques électriques des circuits d'échange à double courant symétrique jusqu'à 10 Mbit/s
V.24	Définition des échanges entre ETTD et ETCD et du format de la prise
V.28	Caractéristiques électriques pour les circuits d'échange double courant asymétrique
V.33	Modem 14,4 Kbit/s *full duplex* sur 4 fils
V.34	Modem 33,6 Kbit/s *full duplex* (V.FAST)
V.42	Correction d'erreurs LAPM (*Link Access Procedure for Modem*)
V.42 bis	Compressions de données
V.54	Équipement de test de modem en boucle
V.90	Modem à 56,4 Kbit/s

Normes EIA/TIA relatives aux interfaces physiques

Référence	Objet
EIA/TIA-232	Spécifications de l'interface RS-232 proches de celles de la norme V.24. Débit maximal de 64 Kbit/s.
EIA/TIA-449	Spécifications de l'interface RS-449. Amélioration de la norme RS-232. Débit maximal de 2 Mbit/s.
EIA-530	Spécifications de deux implémentations d'échange de la norme EIA-TIA-449 : mode balancé (RS-422) et mode non balancé (RS-423).

Avis de l'ITU-T relatifs aux échanges ETTD-ETCD

Avis	Contenu de la spécification
X.20	Interface entre un ETTD et un ETCD pour les services de transmission asynchrone
X.21	Interface entre un ETTD et un ETCD pour fonctionnement synchrone
X.24	Définitions des circuits d'échanges entre un ETTD et un ETCD

Les réseaux locaux

Normes IEEE relatives aux réseaux locaux

Référence	Désignation	Objet
802.1	*High Level Interface*	Traite des architectures (802.1 a), des ponts et du *spanning tree* (802.1d) et du *System Load Protocol* (802.1 e).
802.1p	*Traffic Class Expediting and Dynamic Multicast Filtering*	Gestion du flux et des priorités sur Ethernet
802.1q	VLAN *Virtual Bridged Local Area Networks*	Ajoute un entête de 4 octets (une étiquette ou encore *tag*) aux trames Ethernet définissant le numéro de réseau virtuel
802.2	LLC *Logical Link Control*	Spécifications de la sous-couche LLC du niveau 2 du modèle OSI (802.2c, f et h)
802.3	Ethernet CSMA/CD	Spécifications des réseaux Ethernet
802.3u	Ethernet 100bT	Spécifications du Fast Ethernet. Couche MII (*Media Independant Interface*), 100bTX, 100bT4...
802.3x	Full Duplex et contrôle de flux	Signal intercommuteurs émis pour arrêter le trafic lorsque la mémoire est saturée
802.3ab	Ethernet 1000bT	Spécifications du Gigabit Ethernet sur cuivre en paires torsadées
802.3af	DTE power via MDI	Alimentation 48v via les câbles en paires torsadées pour les postes de téléphonie sur IP connectés au réseau local Ethernet
802.3ah	EFM (*Ethernet in the First Mile*)	Réseau Ethernet entre les points d'accès opérateurs et les particuliers ou les entreprises (livraison d'une prise Ethernet à domicile)
802.3z	Ethernet 1000bX	Spécifications du Gigabit Ethernet sur fibre optique
802.4	Réseaux Token-Bus	Spécifications des réseaux Token-Bus
802.5	Réseaux Token-Ring	Spécifications des réseaux Token-Ring
802.6	Réseaux MAN DQDB	Spécifications des réseaux métropolitains
802.7	Réseaux large bande	Groupe de travail BBTAG (*Broadband Technical Advisory Group*). Norme *Slotted Ring*.
802.8	Réseaux fibre optique	Groupe de travail FOTAG (*Fibre Optics Technical Advisory Group*)
802.9	Réseaux voix/données	IS LAN (*Integrated Services LAN*) Ethernet Isochrone - IsoEnet
802.10	Sécurité des réseaux	Méthodes d'accès entre les couches MAC et LLC (niveau 2) ainsi que pour la couche application (niveau 7) pour les données confidentielles
802.11	WLAN (*Wireless LAN*)	Réseaux locaux sans fil

Référence	Désignation	Objet
802.12	100bVG-AnyLAN	Spécifications des réseaux locaux à 100 Mbit/s avec DPMA (*Demand-Priority Access Method*)
802.14	CATV (*Cable-TV*)	Réseaux sur les câbles télévision CATV
802.15	WPAN *Wireless Personnal Area Network*	Spécifications des Réseaux personnels sans fils tels que Bluetooth
802.16	WMAN (*Wireless MAN*)	Spécifications des boucles locales radios
802.17	RPR (*Resilient Packet Ring*)	Boucles optiques adaptée au transport des paquets

La famille des protocoles TCP/IP

RFC relatives aux protocoles TCP/IP

Référence	Objet
791	Spécifications de **IP** (*Internet Protocol*) - MIL-STD-1777
792	Spécifications de **ICMP** (*Internet Control Message Protocol*)
793, 761, 675	Spécifications de **TCP** (*Transmission Control Protocol*) - MIL-STD-1778
768	Spécifications de **UDP** (*User Datagramm Protocol*)
813	Algorithmes d'acquittement de TCP
815	Algorithmes de réassemblage des paquets TCP/IP
816	Mécanisme de **dead gateway**
919, 922	Diffusion des datagrammes Internet (*broadcast*)
917, 932, 936, 950	Spécifications et description du ***subnetting*** des adresses IP
1219	Subnetting variable
826	**ARP** (*Address Resolution Protocol*) sur Ethernet
903	**RARP** (*Reverse Address Resolution Protocol*)
1293	**Inverse ARP**
1027	**Proxy ARP** (pour les stations ne supportant pas les *subnets*)
1011	Description officielle des protocoles Internet
1108	Spécifications de IPSO (*IP Security Option*)
894	Encapsulation de IP dans une trame Ethernet V2
1042	Encapsulation de IP dans une trame Ethernet 802.3
1078	Multiplexage des ports des services TCP
1144	Compression des en-têtes TCP
1505	En-tête des messages Internet
1918	Subnets IP réservés à l'adressage privé
1001 – 1002	**Netbios** sur TCP/IP : concepts et spécifications
1356	Encapsulation dans X25
1434	**DLSw** - Encapsulation des trames SNA dans des paquets TCP/IP
3168	**ECN** (*Explicit Congestion Notification*) pour IP

*Standards originaux du DOD (*Department Of Defense*)*

Référence	Objet
MIL-STD-1777	Spécifications de **IP** (*Internet Protocol*)
MIL-STD-1778	Spécifications de **TCP** (*Transport Control Protocol*)
MIL-STD-1780	Spécifications de **FTP** (*File Transfer Protocol*)
MIL-STD-1781	Spécifications de **SMTP** (*Simple Mail Transfer Protocol*)
MIL-STD-1782	Spécifications de **Telnet**

RFC relatives *aux protocoles de routage IP*

Référence	Objet
1256	Messages **ICMP** de découverte de routes
2328	**OSPF** (*Open Shortest Path First*) version 2
1245, 1246	Analyse du fonctionnement de OSPF
1771, 1772	**BGP** (*Border Gateway Protocol*)
827, 904, 911	**EGP** (*Exterior Gateway Protocol*)
1058, 1723	**RIP** (*Routing Information Protocol*)
950	Procédures de *subnetting*

RFC relatives aux applications utilisant TCP/IP

Référence	Objet
821, 822, 974	Spécifications de **SMTP** (*Simple Mail Transfer Protocol*)
1869	Extended **SMTP**
1777, 1778	Spécifications de **LDAP** v3 (*Lightweight Directory Access Protocol*)
1344, 1437	Spécifications de Mime (*Multipurpose Internet Mail Extensions*)
2045, 2046	Structure des messages Mime et des formats supportés
1939	Spécifications de **POP** 3 (*Post Office Protocol*)
2060	Spécifications de **IMAP** 4 (*Internet Message Access Protocol*)
854	Spécifications de **Telnet**
1205	Spécifications de l'émulation 5250 sur Telnet
1184, 1091, 1372	Spécifications de diverses options Telnet
1282	Spécifications de Rlogin (*remote login*)
959	Spécifications de **FTP** (*File Transfert Protocol*)
1350	Spécifications de **TFTP** (*Trivial File Transfer Protocol*)
1094, 1813	Spécifications de **NFS** (*Network File System*)

Référence	Objet
1945	Spécifications de **HTTP** 1.0 (*Hyper Text Transfert Protocol*)
2068, 2069, 2109, 2145	Spécifications de **HTTP** 1.1 (*Hyper Text Transfert Protocol*)
1630, 1738	Spécifications des **URL** (*Uniform Resource Locators*)
1034, 1035	Concept, spécifications et implémentation du **DNS** (*Domain Name System*)
1982, 1995, 1996, 2136, 2137, 2181, 2308	Mises à jour de DNS
1183, 1706, 1712, 2052, 2230	Extensions DNS expérimentales
2535	Extensions DNS sécurisé
1591, 1912	Implémentations et bonnes pratiques du DNS
906, 951, 1542	Spécifications de **BootP** (*Bootstrap Protocol*)
1534	Interopérabilité entre DHCP et BootP
2131	Spécifications de **DHCP** (*Dynamic Host Configuration Protocol*)
2132, 2224	Options **DHCP**

RFC relatives à IP sur Frame-Relay

Référence	Objet
2427	Transport de IP dans Frame-Relay (**NLPID** (*Network Level Protocol ID*) / **SNAP** (*Sub Network Access Protocol*)

RFC relatives à IP sur ATM

Référence	Objet
1332	Base de travail pour IP sur ATM (*Asynchronous Transfert Mode*)
1483	Utilisation de la couche AAL-5 pour l'encapsulation des protocoles IP dans un réseau ATM. Spécifications **DXI** (*Data eXchange Interface*)
2225	Spécifications du routage IP et de la résolution d'adresse ARP sur les réseaux ATM (***Classical IP***)
1626	Taille des paquets IP pour ATM **AAL-5**
1629	Guide pour l'allocation des adresses **NSAP** (*Network Service Access Point*) au sein de l'Internet
1680	Support des Services ATM pour IPv6
1755	Signalisation pour IP sur ATM
2022	Multicast sur UNI 3.0/3.1
2149	Spécifications de **MARS** (*Multicast Server Architectures for MARS-based ATM multicasting*)

RFC relatives à PPP

Référence	Objet
1661, 1662, 1663	Spécifications du protocole **PPP**
1332	Spécifications de la couche IPCP (*IP Control Protocol*)
1552	Spécifications de la couche IPXCP (*IPX Control Protocol*)
1570	Spécifications de la couche LCP (*Link Control Protocol*)
1989	Spécifications de la couche LQM (*Link Quality Monitoring*)
1990	Spécifications de PPP MP (*PPP Multilink Protocol*)
1321	Spécifications de l'algorithme MD5 (*Message Digest 5*)
1993	Algorithme de compression FZA pour PPP
1994	Spécifications des protocoles **CHAP** (*Challenge-Handshake Authentication Protocol*)

RFC relatives à SNMP

Référence	Objet
1089, 1157	Spécifications de **SNMP** (*Simple Network Management Protocol*)
1303	Description conventionnelle des **agents SNMP**
1352	Protocoles de sécurité SNMP
2571	Architecture de l'administration avec SNMP
2572	Traitement des messages SNMP
1418, 1419, 1420	SNMP sur OSI, AppleTalk et IPX
1115	Structure de la **MIB** pour les protocoles TCP/IP
1212	Spécifications génériques de la **MIB** (*Management Information Base*)
1213	Spécifications de la **MIB-II** pour TCP/IP
2863	Le groupe **« Interface »** de la MIB
2011, 2012, 2013	MIB-II pour SNMP V2 (mises à jour du RFC 1213)
1231	Spécifications de la MIB Token Ring
1381	Spécifications de la MIB X25
1382	Spécifications de la MIB HDLC LAP-B
1398	Spécifications de la MIB Ethernet
1512, 1285	Spécifications de la MIB FDDI
1513, 1271	Spécifications de la MIB RMON (*Remote Monitoring*)
1559	Définition de la MIB Decnet Phase IV

Normes ISO et équivalents ITU-T relatifs à la syntaxe ASN.1

Référence ISO	Référence ITU-T	Objet
8824	X.208	Spécifications du langage ASN.1 (*Abstract Syntax Notation 1*)
8825	X.209	Spécifications des règles de codage BER (*Basic Encoding Rules*)

RFC relatives à IPv6

Référence	Objet
2460 et 2675	Spécifications de IPv6 (Ipng - *IP Next Generation*)
1924, 2462, 2471, 2373, 1881	Adressage IPv6 (configuration, allocations)
2374, 1887	Format des adresses *unicast*
1809	Utilisation du champ de contrôle de flux
2461	Protocole de découverte des nœuds IPv6
2472	IPv6 sur **PPP**
2464	IPv6 sur **Ethernet**
2463	**ICMPv6** (*Internet Control Message Protocol*)
1886, 2874, 3152	Extension du **DNS** pour le support de IPv6
2740	**OSPF** pour IPv6
2080, 2081	**RIPng** (*Routing Information Protocol next generation*) pour IPv6
2553	Extension des sockets pour IPv6
1888	NSAP OSI pour IPv6
2893	Transition de IPv4 vers IPv6
de 1667 à 1688, 1550, 1705, etc.	Discussions sur IPng (*Internet Protocol next generation*)

Le multimédia sur IP

RFC relatives à la voix sur IP

Référence	Objet
2543	**SIP** (*Session Initiation Protocol*)
2976	Extension de SIP : méthode INFO
1889	**RTP** (*Real-time Transport Protocol*) et **RTCP** (*Real-time Transport Control Protocol*)
1890	Définition des profils pour les conférences audio et vidéo
2032	Transport des codec vidéo **H.261**
2190	Transport des codec vidéo **H.263**
2198	Transport des codec audio

Avis de l'ITU-T relatifs à la voix sur IP

Référence	Objet
H.323	Systèmes de communications Multimedia basé sur des paquets
H.323 – Annexe A	Messages **H.245** utilisés par H.323
H.323 – Annexe B	Procédures pour les codec vidéo
H.323 – Annexe C	H.323 sur ATM AAL5
H.323 – Annexe D	Fax temps réel sur H.323
H.323 – Annexe E	Transport de la signalisation d'appel
H.323 – Annexe G	Spécifications des terminaux H.323
H.323 – Annexe H	Mobilité
H.323 – Annexe I	Opération sur des réseaux avec une basse qualité de service
H.323 – Annexe J	Terminaux sécurisés
H.323 – Annexe K	Interface avec http
H.323 – Annexe L	Gestion des communications un serveur (*feature server*)
H.323 – Annexe M	Tunneling Q.SIG
H.323 – Annexe N	Qualité de service
H.225	Protocole de signalisation d'appel (**RAS** et **Q.931**)
H.235	Sécurité
H.245	Protocole de contrôle pour les communications multimédias
Q.931	UNI niveau 3 pour le contrôle d'appel
H.450	Protocole générique pour les services complémentaires
H.261	Codec vidéo aux débits multiples de 64Kbit/s
H.263	Codec vidéo aux débits inférieurs à 64Kbit/s

Référence	Objet
G.711, G722, G.723	Codec audio
G.728, G.729	Codec audio
T.120	Protocole de données pour les conférences multimédias

RFC relatives à la qualité de service

Référence	Objet
791, 1349	Champs TOS (*Type Of Service*) et Precedence du paquet IP
2474, 2475, 3260	DiffServ (*Differentiated Service*)
2873	Compatibilité du champ TOS avec le Code Point de DiffServ
1633	IntServ (*Integrated Service*)
2205	RSVP (*Resource reSerVation Protocol*)
2206-2209	Extensions MIB et IP-SEC et mise en œuvre
2210	Utilisation de RSVP dans IntServ
2211	IntServ – Classe de service par contrôle de charge
2212	IntServ – Classe de service par garantie de service
2213	IntServ – MIB de l'intégration de service
2214	IntServ – MIB de l'intégration de service - Extension à la garantie de service
2215	IntServ – Définition des paramètres permettant de calculer une QoS
2216	IntServ – Définitions de la QoS pour les éléments du réseau (définit un cadre général sur lequel s'appuient les RFC 2211 et 2212)

RFC relatives au routage multicast

Référence	Objet
1112	Support **multicast** par les piles TCP/IP
2236	**IGMP** (*Internet Group Membership Protocol*)
1075	**DVMRP** (*Distance Vector Multicast Routing Protocol*)
1584	**MOSPF** (*Multicast Open Shortest Path First*)
2362	**PIM-SM** (*Protocol Independent Multicast*)

Les réseaux RNIS

Organisation et nomenclature des normes

Série I.600 Principes d'administration du RNIS	
Série I.400 Interfaces usager-réseau	**Série I.500** Interfaces d'interconnexion
Série I.200 Services assurés par le RNIS	**Série I.300** Aspects réseaux du RNIS
Série I.100 Concepts généraux du RNIS Structure des recommandations, terminologie, méthodes générales	

Le terme RNIS (réseau numérique à intégration de services) ou ISDN (*Integrated Services Digital Network*) désigne le réseau connu des utilisateurs (c'est-à-dire l'accès de base T0 et l'accès primaire T2) et, plus généralement, l'ensemble des réseaux numériques comme l'ATM, le relais de trames, SDH, etc. Tous ces protocoles sont fédérés par un ensemble de normes communes. L'ATM est ainsi référencé sous le nom de RNIS large bande (*Broad-band ISDN*).

	RNIS		ATM		Relais de Trames	
Couche 3	I.450 - Q.930 I.451 - Q.931 I.452 - Q.932	UNI Signalisation Procédures	Q.933 - Q.2931	UNI	Q.933 - Q.2931	UNI
Couche 2	I.440 - Q.920 I.441 - Q.921	UNI Lap-D	I.362, I.363 I.361 I.371	AAL ATM Congestion	Q.921, Q.922 I.370	Couche liaison Congestion
Couche 1	I.420, I.430 I.421, I.431	Accès de base Accès primaire	I.432	PHY	I.430 I.431	Accès de base Accès primaire

Série I.100 : Concepts généraux du RNIS

Avis	Contenu des spécifications
I.110	Structure générale des recommandations de la série I
I.111	Relations avec les autres recommandations
I.112	Vocabulaire RNIS (Réseau Numérique à Intégration de service)
I.120	Description du RNIS (ISDN - *Integrated Services Digital Network*)
I.130 et I.140	Services du RNIS et fonctionnalités du réseau (méthode et technique)

Série I.200 : Services assurés par le RNIS

Avis	Contenu des spécifications
I.200	Description de la série I.200
I.210	Généralités sur les services offerts par le RNIS
I.211	Définition des services réseaux (au niveau du SSTM)
I.212	Description des téléservices
I.220	Description dynamique commune des services de télécommunications
I.221	Caractéristiques des services spécifiques courants
I.230	Définition des catégories des services support
I.231	Description des catégories des services support en mode circuit
I.232	Description des catégories des services support en mode paquet
I.240 et I.241	Définition et description des téléservices offerts par le RNIS
I.250	Définition des compléments de service
I.251 à I.257	Description des compléments de service (identification de l'appelant, conférence, facturation, transfert d'appel...)

Série I.300 : Aspects réseaux du RNIS

Avis	Contenu de la spécification
I.310	Principes fonctionnels du réseau
I.320	Modèle de référence des protocoles
I.324	Architecture du RNIS
I.325	Configurations de référence pour les types de connexion du RNIS
I.326	Configurations de référence pour les besoins de ressources
I.330	Principes de la numérotation et de l'adressage
I.331	Plan de numérotation E.164
I.332	Principes d'adressage pour l'interconnexion de plusieurs RNIS

Avis	Contenu de la spécification
I.333	Sélection du terminal dans le RNIS
I.334	Principes d'adressage et de sous-adressage par rapport au modèle OSI
I.335	Principes de routage
I.340	Interconnexion des RNIS (attributs et topologies)
I.350	Objectifs de performance pour le réseau téléphonique et le RNIS
I.351	Objectifs de performance au point de référence T
I.352	Objectifs de performance en termes de délais de connexion

Série I.400 - Interfaces usager-réseau

Avis	Contenu de la spécification
I.410	Généralités et principes de la UNI (*User Network Interface*) : indique les types d'interfaces, le paramétrage (débit, codage...), la vérification de compatibilités entre terminaux, etc.
I.411	Configurations de référence : définitions de la TNR, TNA, etc.
Couche physique du RNIS (couche 1)	
I.412	Structure d'interface et possibilités d'accès Définitions des types de canaux : 2B+D Canal H0 à 384 Kbits/s Canal H11 à 1 536 Kbits/s (1 544 Kbits/s en accès primaire) Canal H12 à 1 920 Kbits/s (2 040 Kbits/s en accès primaire)
I.420	Structure des canaux et interface d'accès de base
I.421	Structure des canaux et interface d'accès primaire
I.430	Spécifications de la couche physique de l'accès de base. Interface usager-réseau. Points de référence S et T
I.431	Spécifications de la couche physique de l'accès primaire
Couche liaison du RNIS (couche 2)	
I.440 - Q.920	Généralités sur la signalisation UNI (*User Network Interface*)
I.441 - Q.921	Spécifications de la couche 2 : LAP-D, signalisation DASS (*Digital Access Signaling System*)
Couche réseau du RNIS (niveau 3)	
I.450 - Q.930	Généralités sur la signalisation UNI
I.451 - Q.931	Spécifications de la couche réseau signalisation des services
I.452 - Q.932	Procédures pour le contrôle des services complémentaires
Support d'autres interfaces par le RNIS	
I.460	Adaptation de débit au multiplexage et support des interfaces existantes
I.461 - X.30	Support des interfaces X21, X21 bis et X20 bis

Avis	Contenu de la spécification
I.462 - X.31	Support des terminaux en mode paquet
I.463 - V.110	Support des ETTD série V
I.464	Multiplexage, adaptation de débit et support des interfaces existantes pour des possibilités réduites de transfert à 64 Kbit/s
I.465 - V.120	Support des ETTD dotés d'interfaces de la série V et multiplexage statistique
I.470	Relations avec les fonctions du terminal RNIS

Série I.500 : Interfaces d'interconnexion du RNIS

Avis	Contenu de la spécification
I.500	Structure générale de l'interconnexion de réseaux RNIS
I.510	Définition et principes généraux de l'interconnexion
I.511	Interface physique de l'interconnexion
I.515	Échanges de paramètres
I.520	Spécifications générales de l'interconnexion des réseaux RNIS
I.530	Interconnexion avec un réseau public
I.540 - X.321	Spécifications générales pour l'interconnexion avec un réseau public à commutation de circuits
I.550 - X.325	Spécifications générales pour l'interconnexion avec un réseau public à commutation de paquets
I.560 - U.202	Support du service télex par le RNIS

Série I.600 : Administration du RNIS

Avis	Contenu de la spécification
I.601	Principes généraux de l'administration
I.602	Application des principes à l'installation d'un abonné
I.603	Application des principes à l'accès de base
I.604	Application des principes à l'accès primaire
I.605	Application des principes aux accès multiplexés

Avis de l'ITU-T relatifs aux réseaux ATM

Référence	Objet
I.113	Vocabulaire relatif au B-ISDN (*Broadband-Integrated Services Digital Network*)
I.121	Généralités sur les aspects large bande du B-ISDN
I.150 et I.351	Caractéristiques fonctionnelles ATM du B-ISDN
I.211	Présentation générale des services du B-ISDN
I.311 - I.312	Présentation générale du réseau B-ISDN
I.321	Modèle de référence du protocole B-ISDN et ses applications
I.327	Architecture fonctionnelle du B-ISDN
I.353	Définition des événements et paramètres permettant de mesurer les performances de la commutation de cellules
I.356	Performances de la commutation de cellules
I.361	Spécifications de la couche ATM : format des cellules
I.362	Description fonctionnelle de la couche AAL (*ATM Adaptation Layer*)
I.363	Spécifications de la couche AAL
I.371	Contrôle de flux et gestion des congestions
I.413	Spécifications de l'UNI (*User Network Interface*)
I.432	Spécifications de la couche physique (niveau 1) de l'UNI, de l'en-tête HEC et du mécanisme cell delineation
I.600	Application des principes de gestion des abonnés B-ISDN
I.610	Principes des gestion du B-ISDN (couche OAM - *Operations, Administration and Maintenance*)
Q.93B - G.771	Spécifications du contrôle des services et des appels (UNI dérivée de la norme Q.931).
Q.2931 - Q.933	Couche 3 UNI (version 3.0) de l'ATM Forum

Avis de l'ITU-T et équivalents ANSI relatifs au Frame Relay

Référence ITU-T	Référence ANSI	Objet
I.233	T1.606	Description des services du Frame relay
I.370	T1.606	Gestion des congestions
I.372		Besoins pour les interfaces NNI
I.555		Frame relay sur ATM
Q.921		Description de la couche liaison pour la signalisation
Q.922	T1.618	Principes du protocole, spécifications de la couche liaison
Q.933	T1.617	Spécifications de la signalisation UNI de niveau 3

Avis de l'ITU-T relatifs aux systèmes de transmission numérique MIC

Référence	Objet
G.703	Caractéristiques physiques et électriques des interfaces
G.704	Caractéristiques fonctionnelles des jonctions
G.711	Modulation par impulsions codées (MIC) des fréquences vocales
G.721	ADPCM (*Activity Detection Pulse Code Modulation*) : codage MIC différentiel adaptatif (MIC-DA) à 32 Kbit/s
G.722	Codage des fréquences audio à 7 KHz sur 64 Kbit/s
G.725	Aspects relatifs à l'utilisation du codage audio à 7KHz sur 64 Kbit/s
G.726	ADPCM avec compression à 16k, 24k ou 32k
G.728	LDCELP : compression de la voix à 16k
G.729	CS-ACELP : compression de la voix de 4,8k à 16k
G.732	Caractéristiques des équipements de multiplexage MIC primaires à 2 048 Kbit/s.

Avis de l'ITU-T relatifs aux réseaux SDH

Référence	Objet
G.702	Spécifications des débits de la hiérarchie numérique
G.703	Caractéristiques physiques et électriques des interfaces MIC
G.704	Structure des trames des niveaux primaires et secondaires
G.706	Spécifications du CRC (*Cyclic Redundancy Check*) et de l'alignement des trames
G.707	Spécifications des débits de la hiérarchie numérique synchrone
G.708	Spécifications des interfaces des nœuds du réseau
G.709	Structure de multiplexage synchrone
G.733	Caractéristiques physiques et électriques des interfaces PCM (D2)
G.771 - Q.93B	Spécifications du contrôle des services et des appels (UNI)
G.774	Modèle d'administration pour les éléments du réseau
G.781	Aspects généraux des systèmes de transmission (multiplexeurs)
G.782	Caractéristiques générales des équipements de multiplexage
G.783	Caractéristiques générales des blocs fonctionnels des multiplexeurs
G.784	Aspects généraux des équipements d'administration
G.804	Format d'encodage des cellules ATM sur le support physique
G.821	Performance pour un réseau international
G.826	Paramètres et objectifs de performance pour les classes de trafic constant

Organisation de l'Internet

L'Internet est la concaténation de différents réseaux appartenant à des opérateurs privés ou publics (par exemple, France Télécom en France). La plupart des opérateurs ne disposent pas de leur propre infrastructure (les câbles) ; ils louent tout ou partie des liaisons et y connectent leurs équipements (la plupart du temps de marque Cisco et Nortel).

Les réseaux des opérateurs se superposent, se dédoublent et se rejoignent par moments au niveau de points de concentration.

Figure A-1.
L'Internet.

Les points de concentration reposent sur des réseaux à hauts débits (ATM, FDDI, voire Ethernet commuté) qui utilisent la fibre optique, parcourant une ville importante. Il en existe de trois types :

- Les **MAE** (*Metropolitain Area Exchange*) qui sont détenus par Worldcom (*via* sa filiale MFS). Il en existe sept aux États-Unis (dont les plus gros sont situés à San Jose, dans la Silicon Valey, et à Washington) et un à Paris.
- Les **NAP** (*Network Access Point*) qui sont détenus par le NSF (*National Science Foundation*) et gérés par des opérateurs privés. Il en existe actuellement quatre aux États-Unis (un à San Francisco géré par Pacific Bell, un à Chicago géré par Bellcore, un à Washington DC géré par Ameritech, et un quatrième près de New York géré par Sprint).
- Les **PPP** (*Private Peering Point*) qui sont gérés directement par des ISP qui se sont associés afin de contourner les MAE et les NAP.

Les ISP (*Internet Service Provider*) connectent leurs routeurs aux MAE, NAP et PPP (moyennant une redevance) interconnectant ainsi leurs réseaux. L'ensemble de ces interconnexions forme l'Internet.

Les liaisons entre les villes sont constituées de câbles en fibre optique ou en cuivre au bout desquels on retrouve les routeurs, équipements réseau de base qui permettent d'acheminer toutes les communications sur l'Internet. Les routeurs utilisent les protocoles Frame Relay et, de plus en plus, ATM.

Les routeurs des ISP ainsi que les commutateurs des MAE, NAP et PPP sont installés dans des locaux techniques, sans doute dans un immeuble devant lequel vous passez tous les jours sans vous en rendre compte. Certains NAP sont même perdus au fond d'un parking souterrain…

L'Internet, c'est donc cela : une collection de réseaux détenus et gérés par des opérateurs privés ou gouvernementaux (essentiellement le NSF) autour desquels gravitent les ISP qui revendent leurs services aux consommateurs que nous sommes.

La conséquence de cette situation est qu'il n'y a aucune garantie de service : les temps de réponse dépendent de la charge réseau, et les pannes sur l'Internet ne sont pas des légendes, certaines sont mêmes célèbres :

- En 1996, une erreur de configuration sur un routeur a créé une nouvelle route redirigeant ainsi 25 % du trafic Internet sur une seule liaison à 1,5 Mbit/s. L'engorgement résultant a provoqué un arrêt quasi total de l'Internet pendant deux heures.
- En 1997, la panne d'un serveur a abîmé un fichier DNS contenant près d'un million de noms. Le technicien a ignoré l'alarme (normal, un bip de plus ou de moins…) et le fichier s'est répliqué sur d'autres serveurs DNS. Le temps de tout remettre en ordre, des dizaines de milliers de sites web n'ont plus été accessibles pendant plusieurs heures.
- Le 13 avril 1998, le backbone Frame Relay d'ATT s'est complètement arrêté suite à une erreur de configuration sur un commutateur doublée d'un bogue dudit équipement. Celui-ci a généré des messages d'alerte vers les autres commutateurs qui se sont engorgés et qui ont eux mêmes généré d'autres messages d'alerte, et ainsi de suite. En moins d'une demi-heure, le réseau était par terre. L'impact sur l'Internet a été faible, car il existait des routes de secours (celles d'autres opérateurs !), mais les clients d'ATT ont été privés de réseau pendant 12 à 24 heures.

Cependant, depuis 1985, l'Internet n'a cessé d'évoluer afin de faire face à l'augmentation constante du nombre d'utilisateurs, et l'on peut espérer que, tant que les opérateurs gagneront de l'argent, ils assureront un minimum de qualité de service afin d'éviter de perdre leurs clients.

Quelques chiffres

L'Internet est en perpétuelle évolution, ce qui fait que les statistiques sont difficiles à obtenir. Les chiffres diffèrent selon les sources ; de plus, ils ne s'appuient pas toujours sur les mêmes critères. Même le NIC (*Network Information Center*), qui gère le plan d'adressage et de nommage, publie des informations provenant de sociétés privées qui scrutent l'Internet (dans le but de vendre leurs études de marchés). Il est vrai qu'à sa décharge, le NIC a un fonctionnement très décentralisé.

On ne peut donc que se fonder sur une photographie prise à un instant donné, et encore, assez floue. Début 1999, il y aurait ainsi eu :

- plus d'un million de domaines DNS dont 75 % dans " .com " ;
- plus d'un million et demi de serveurs web (nommés www) dont 44 % de marque Apache (un freeware) et 20 % de marque IIS (*Internet Information Server*, de Microsoft) ;
- plus de 160 000 réseaux de classe C et près de 10 000 réseaux de classe B ;
- plus de 40 millions d'ordinateurs connectés ;
- et près de 100 millions d'internautes (dont 50 aux États-Unis, 11 au Japon, 7 en Allemagne et 2 en France).

Pour vous donner une idée de la croissance phénoménale de l'Internet, il y aurait eu, début 2000, 80 millions d'ordinateurs connectés dont 6 millions d'internautes en France…

Vous pouvez prendre connaissance des statistiques les plus récentes en consultant, par exemple, le site http://www.isc.org/ds.

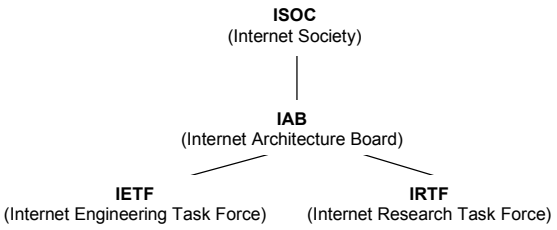
La gestion de l'Internet

L'Internet, c'est d'abord un réseau. C'est aussi un ensemble de protocoles (plusieurs centaines) couramment appelés protocoles Internet ou protocoles TCP/IP. Ils couvrent aussi bien les couches réseaux que les applications, telles que la messagerie ou le web. Tous ces protocoles respectent des standards définis dans des documents techniques appelés **RFC** (*Request For Comments*).

L'Internet, c'est également une communauté regroupant des organismes de recherche, des universités, des constructeurs de matériels, des éditeurs de logiciels, des opérateurs et, de plus en plus, des sociétés qui veulent simplement gagner de l'argent.

Pour organiser tout cela, l'Internet est structuré en plusieurs organisations, chacune ayant un rôle bien défini.

Figure A-2.
Les organismes chargés de développer les protocoles Internet..



L'**ISOC** (*Internet Society*) est une association ayant pour but de promouvoir l'Internet et d'en financer le développement. Son rôle est donc de coordonner et de financer les organismes qui régulent l'Internet, tels que l'IETF. L'ISOC entretient également des relations avec l'ITU (*International Telecommunication Union*, organisation dépendant de l'ONU), les acteurs de l'industrie (constructeurs et éditeurs de matériels et de logiciels réseaux) et les gouvernements, afin d'officialiser les relations et de collecter des fonds.

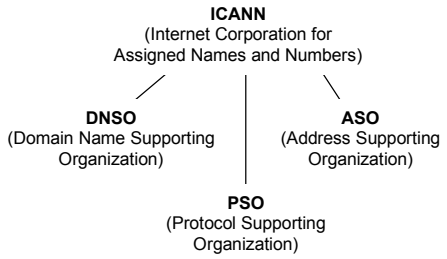
Le résultat concret de ces échanges est l'apparition des RFC dans les recommandations de l'ITU-T (ITU, secteur des télécommunications) et une plus grande diffusion des protocoles édictés par cet organisme, dont les documents sont longtemps restés inaccessibles au public.

L'**IAB** (*Internet Architecture Board*) est un comité de quelques personnes qui décide des évolutions de l'Internet, telles que l'adressage, la mise en chantier d'IPv6, l'évolution de l'architecture du réseau et du DNS ou encore la sécurité. Parmi la quinzaine de membres que compte ce comité on trouve des représentant de Cisco, 3com, Microsoft, Netscape, Sun, MCI, ATT, IBM, plus quelques autres représentants d'universités américaines. La plupart d'entre eux viennent de l'IETF, les autres de l'IESG, du IANA et de l'IRTF.

L'**IETF** (*Internet Engineering Task Force*) regroupe des ingénieurs de divers horizons (instituts de recherche, universités, constructeurs et éditeurs, etc.) qui travaillent à l'élaboration des protocoles utilisés sur l'Internet. Les résultats de ses travaux aboutissent à la rédaction des **RFC** (*Request For Comments*) approuvés par l'IESG, puis validés par l'IAB, et enfin estampillés et diffusés par l'ISOC. L'IESG (*Internet Engineering Steering Group*) est le comité de validation technique de l'IETF.

L'**IRTF** (*Internet Research Task Force*) est structuré en groupes de recherche dont les objectifs sont de travailler aux protocoles de demain. L'IRTF réalise des travaux analogues à l'IETF mais sur le long terme. Le RFC 2014 décrit le fonctionnement de cette organisation. L'IRSG (*Internet Research Steering Group*) est le comité de validation technique de l'IRTF.

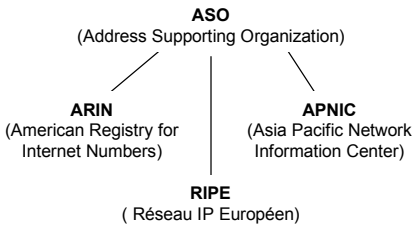
Figure A-3.
Les organismes chargés de réguler l'Internet.



L'**ICANN** (*Internet Corporation for Assigned Names and Numbers*) a en charge la gestion des adresses IP et des noms DNS (et des serveurs racines) ainsi que l'affectation des paramètres aux protocoles IP. Cet organisme est pour cela structuré en trois SO (*Supporting Organization*) : l'ASO, le DNSO et le PSO.

L'**ASO** (*Address Supporting Organization*) est chargé de gérer le plan d'adressage de l'Internet. Il affecte à ce titre les adresses IP et s'appuie sur des délégations régionales.

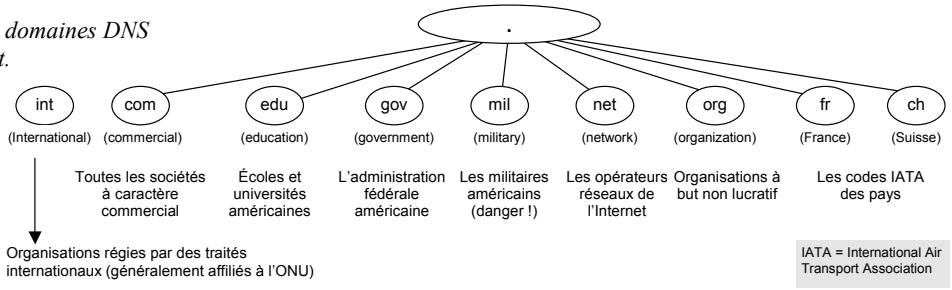
Figure A-4.
La gestion des adresses IP sur l'Internet.



Le **DNSO** (*Domain Name Supporting Organization*) est chargé de gérer le plan de nommage de l'Internet. Il accrédite les sociétés auprès desquelles vous pouvez demander l'enregistrement d'un nom de domaine et celles qui peuvent gérer des serveurs de noms.

Les domaines de plus haut niveau sont appelés *Top Level Domains* (TLD). Parmi eux, on distingue les gTLD (*general TLD*) qui regroupent " .com ", " .edu ", etc., et les ccTLD (*country-code TLD*) qui désignent chacun un pays, par exemple " .fr ", " .uk ", etc.

Figure A-5.
Les noms de domaines DNS sur l'Internet.



Par exemple, l'AFNIC, financé par l'INRIA et les ISP opérant en France, gère les serveurs DNS du domaine *fr* et enregistre les noms de domaines demandés par les sociétés (délai moyen : 48 heures). La liste de ces domaines est disponible à l'adresse : ftp://ftp.nic.fr/pub/annuaire/Liste-Des-Domaines-Français.

Le **PSO** (*Protocol Supporting Organization*) s'occupe d'enregistrer toutes sortes de valeurs utilisées par les protocoles Internet. Cela concerne, entre autres :

- les numéros des ports TCP et UDP (les *Well Known Port Numbers*) ;
- les numéros des protocoles utilisant IP (TCP=6, UDP=17, etc.) ;
- tous les codes utilisés par tous les protocoles (par exemple, les types de messages DNS, OSPF et PPP, la signification de tel octet t de tel bit, etc.) ;
- les variables des MIB ;
- les codes vendeurs des adresses MAC (repris de l'IEEE) ;
- etc.

Dans les faits, le **IANA** (*Internet Assigned Numbers Authority*) est actuellement le seul PSO. La liste actualisée de toutes les valeurs peut être consultée sur le site web de l'IANA, à www.iana.org.

Quelques autres organismes d'intérêt général

L'**IEPG** (*Internet Engineering Planning Group*) a pour objectif de coordonner les activités des opérateurs dont les réseaux constituent l'Internet. La charte de l'IEPG fait l'objet du RFC 1690. Au sein de ce groupe, les opérateurs décident de la manière dont ils interconnectent leurs réseaux, diffusent leurs tables de routage, etc.

Dans le domaine de la sécurité, le **CERT/CC** (*Computer Emergency Response Teams/Coordination Center*) coordonne l'activité d'une dizaine de groupes de surveillance à travers le monde. Il diffuse en permanence des informations ou des bulletins d'alerte relatifs à des problèmes de sécurité (virus, bogues logiciels, attaques répertoriées, statistiques, etc.). Le **FIRST** (*Forum of Incident Response and Security Teams*) regroupe également divers groupes de surveillance, dont le CERT/CC, liés à des organismes gouvernementaux (NASA, NIH), à des universités (Oxford, Israeli Academic Network) ou à des sociétés privées.

Les anciens organismes de régulation

La réorganisation de 1999 résulte de la volonté du gouvernement américain de privatiser la gestion de l'Internet auparavant exercée par des organismes qu'il subventionnait.

Le **NSF** (*National Science Foundation*), l'équivalent américain du CNRS, a créé le premier backbone de l'Internet dans les années 80 (il n'existe plus et a été remplacé par ceux des opérateurs). Jusqu'à fin 98, le NSF conservait la maîtrise de l'exploitation de la partie commune de l'Internet, c'est-à-dire l'adressage IP et le nommage DNS.

Le **NIC** (*Network Information Center*) était financièrement soutenu par le NSF. Cet organisme était chargé de gérer le plan d'adressage de l'Internet, ainsi que les domaines *com*, *edu*, *net* et *org*. Les ccTLD étaient gérés par les délégations nationales du NIC, par exemple, l'INRIA, en France. Les domaines *gov* et *mil* sont toujours gérés par leurs organismes respectifs.

Où les contacter

Organisme	Site web
Internet Society	www.isoc.org
IAB	www.iab.org
IETF	www.ietf.org
IRTF	www.irtf.org
Éditeur des RFC	www.rfc-editor.org
ICANN, ASO, DNSO et PSO	www.icann.org
IANA	www.iana.org

Organisme	Site web
NIC en France chez les militaires américains zone Asie-Pacifique	rs.internic.net www.nic.fr ou afnic.asso.fr www.nic.mil www.apnic.net
ARIN	www.arin.net
RIPE NCC	www.ripe.net
IEPG	www.iepg.org
FIRST	www.first.org
CERT	www.cert.org

Glossaire

A

ACR (*Attenuation Crosstalk Ratio*) : mesure, exprimée en décibels, du rapport signal/bruit d'un câble.

ADSL (*Asymmetric Digital Subscriber Line*) : technique de transmission à haut débit sur cuivre (de 1,5 à 8 Mbit/s dans un sens et 16 à 640 Kbit/s dans l'autre) sur une portée de 3,7 à 5,4 km. Permet de connecter les particuliers à l'Internet de façon permanente et à haut débit.

AMRC (*Accès multiples à répartition en code*) : désigne les techniques de multiplexage par attribution d'un code d'identification. Utilisé dans les réseaux cellulaires UMTS.

AMRF (*Accès multiples à répartition de fréquence*) : désigne les techniques de multiplexage fréquentiel. Utilisé dans les réseaux cellulaires analogiques.

AMRT (*Accès multiples à répartition dans le temps*) : désigne les techniques de multiplexage temporel. Utilisé dans les réseaux cellulaires numériques de type GSM.

APPN (*Advanced Peer to Peer Network*) : évolution des réseaux SNA vers un réseau non hiérarchique.

Arcnet (*Attached Ressource Computer Network*) : réseau local de type bus à jeton sur câble coaxial en étoile. Le débit est de 2,5 Mbit/s.

ARP (*Address Resolution Protocol*) : protocole basé sur un *broadcast* permettant d'obtenir l'adresse MAC (niveau 2) à partir d'une adresse réseau (niveau 3).

Arpanet (*Advanced Research Projects Agency Network*) : historiquement le premier réseau de type TCP/IP développé pour le DoD (*Department Of Defense*) américain.

ART (*Autorité de régulation des télécommunications*) : organisme de régulation dans le domaine des télécommunications en France (tarifs, concurrence, obligations des opérateurs, etc.).

ASIC (*Application-Specific Integrated Circuit*) : circuit intégré développé spécifiquement pour une application donnée et à la demande. Il est constitué d'une matrice de transistors qui forment des circuits logiques (AND, OR, XOR...) agissant sur des signaux en entrée et générant des signaux en sortie.

ASCII (*American Standard Code for Information Interchange*) : code sur 7 ou 8 bits, utilisé dans l'informatique pour représenter un caractère alphanumérique (128 ou 256 combinaisons sont possibles). Ce code est également normalisé ITU-T n° 5.

ASN.1 (*Abstract Syntax Notation 1*) : langage normalisé ISO permettant de décrire diverses structures comme les bases de données et les paquets des protocoles.

ATM (*Asynchronous Transfert Mode*) : protocole haut débit reposant sur la commutation de cellules de 53 octets.

AUI (*Attachment Unit Interface*) : type de connecteur à 15 broches utilisé pour les matériels Ethernet.

Autocom (abréviation de *autocommutateur*) : désigne un ordinateur spécialisé dans la commutation de circuits. Établit automatiquement une communication téléphonique en fonction d'un numéro de téléphone.

B

Backbone (*épine dorsale*) : désigne un réseau fédérateur à hauts débits permettant d'interconnecter des réseaux secondaires.

Balun (*Balanced Unbalanced*) : connecteur permettant d'adapter l'impédance entre deux câbles de nature différente.

Baud (du nom de *Émile Baudot*, inventeur du code télégraphique) : unité exprimant le nombre de modulations par seconde. Elle équivaut au bit par seconde si un signal représente une valeur binaire.

Bluetooth (du nom d'un roi norvégien) : technologie de réseau sans fil sur courte de distances dans le but connecter des périphériques à un ordinateur.

BBS (*Bulletin Board Systems*) : messagerie pour les PC permettant de télécharger des fichiers.

BERT (*Bit Error Rate Tester*) : test de la qualité d'une ligne consistant à générer des trames et à mesurer le taux d'erreur.

BISDN (*Broadband Integrated Services Digital Network*) : désigne le réseau numérique à haut débit reposant sur ATM et un support de transmission SDH (ou Sonet).

Bit (*BInary digiT*) : représente le plus petit élément d'information. Il prend les valeurs binaires « 1 » et « 0 ».

BNC (*Basic Network Connector*) : connecteur propre aux câbles coaxiaux utilisés par les réseaux Ethernet.

BOC (*Bell Operating Company*) : compagnies de téléphone américaines régionales au nombre de sept issues du démantèlement d'ATT (*American Telephone and Telecommunication*).

Bit/s (*bits par seconde*) : nombre de bits transmis par seconde.

BRI (*Basic Rate Interface*) : désigne l'accès de base RNIS (2B+D).

BSC (*Binary Synchronous Communications*) : protocole synchrone utilisé par certains équipements d'IBM.

BUS (*Broadcast and unknown Server*) : couche logicielle permettant de gérer les *broadcasts* émis par les réseaux locaux sur des chemins virtuels ATM.

C

CAP (*Carrierless Amplitude Phase*) : méthode de codage en ligne utilisée pour les réseaux hauts débits, les liaisons xDSL. Basé sur une quadruple modulation d'amplitude (QAM). Moins performant que DMT, mais moins cher.

CATV (*Cable Antenna TV*) : désigne la télévision par câble et tous les dispositifs s'y rattachant (câble, modulateur...).

CBDS (*Connectionless Broadband Data Service*) : transmission de données pour les réseaux de télécommunications hauts débits tels que ATM.

CDDI (*Copper Distributed Data Interface*) : version de FDDI sur des câbles cuivre à paires torsadées de catégorie 5.

CDMA (*Code Division Multiple Access*) : désigne les techniques de multiplexage par attribution d'un code d'identification. Utilisé dans les réseaux cellulaires UMTS. (cf. AMRC)

CELP (*Code-Excited Linear Predictive*) : algorithme de codage d'un signal audio consistant à synthétiser des approximation du signal à venir en fonction des signaux précédents.

CICS (*Customer Information Control System*) : moniteur transactionnel d'IBM.

CISC (*Complex Instruction Set Component*) : type de microprocesseur offrant un grand nombre d'instructions (par opposition à RISC).

CMOS (*Complementary Metal Oxyde Semiconductor*) : technologie de fabrication de puces électroniques à base de silicium.

CPU (*Central Processor Unit*) : unité centrale exécutant les instructions d'un programme.

CRC (*Cyclic Redundancy Check*) : mécanisme de contrôle d'erreurs basé sur le calcul d'un polynôme générateur. Permet de contrôler que les données d'un paquet, d'une trame ou d'une cellule n'ont pas été endommagées lors de la transmission sur le réseau.

CSU (*Channel Service Unit*) : désigne un modem numérique raccordant l'équipement terminal (un routeur par exemple) à un nœud du réseau, comme un commutateur ATM (voir également DSU).

CSMA-CD (*Carrier Sense Multiple Access - Collision Detection*) : méthode d'accès au support physique par écoute du réseau et propre au réseau Ethernet.

CSMA-CR (*CSMA - Contention Resolution*) : méthode d'accès au support physique par gestion de priorité et propre à l'accès de base du RNIS.

CT2 (*Cordless Telephone 2nd generation*) : norme de radiotéléphones numériques sans fil. Exemple du Bi-Bop en France.

CV (*Circuit Virtuel*) : désigne un lien établi lors d'une procédure de connexion à travers plusieurs commutateurs d'un réseau. Les CVC (c*ircuits virtuels commutés*) sont établis à la demande. Les CVP (*circuits virtuels permanents*) sont établis une fois pour toutes lors de l'initialisation des équipements d'extrémité.

CWDM (*Coarse Wavelength Division Multiplexing*) : version économique du WDM qui multiplexe jusqu'à 16 longueurs d'onde.

D

DECT (*Digital European Cordless Telecommunication*) : norme européenne de radiotéléphones numériques sans fil.

DES (*Data Encryption Standard*) : norme de cryptage spécifiée par l'ANSI (*American National Standards Institute*) et par le NIST (*National Institute of Standards and Technology)*.

DHCP (*Dynamic Host Configuration Protocol*) : permet, à partir d'un serveur, de télécharger la configuration réseau vers un ordinateur (adresse IP, paramètres TCP/IP, etc.).

DNS (*Domain Name System*) : service de noms reposant sur des serveurs. Permet de convertir un nom en une adresse IP.

DLSw (*Data-Link Switching*) : désigne une méthode d'encapsulation des trames SNA dans des paquets TCP/IP (RFC 1434). Les acquittements des trames sont locaux.

DMT (*Discrete MultiTone*) : méthode de codage en ligne utilisée pour les réseaux hauts débits, les liaisons spécialisées et l'ADSL. Basé sur une quadruple modulation d'amplitude (QAM). Plus performant que CAP.

DMTF (*Desktop Management Task Force*) : groupe de travail ayant défini la norme du même nom qui a pour objet de décrire les protocoles et bases de données dans le domaine de l'administration bureautique.

DNA (*Digital Network Architecture*) : spécification des réseaux de la société DEC (*Digital Equipment Corporation*).

DQDB (*Distributed Queue Dual Bus*) : désigne la norme IEEE 802.6 pour les réseaux MAN (*Metropolitan Area Network*).

DSA (*Distributed System Architecture*) : modèle d'architecture réseau de la société Bull.

DSL (*Digital Subscriber Line*) : ensemble de technologies de transmission numérique à haut débit (quelques Mbit/s) fonctionnant sur des paires de cuivre comme celles du téléphone.

DSU (*Data Service Unit*) : désigne un adaptateur entre l'interface physique d'un ETTD et un support de transmission numérique tel qu'une LS (voir également CSU). Le protocole

d'échange entre les deux est référencé sous le nom de DXI (*Data eXchange Interface*, RFC 1483).

DTMF (*Dual Tone MultiFrequency*) : système de signalisation utilisé pour la numérotation sur les réseaux téléphoniques analogiques.

DVD (*Digital Versatile Disk*) : disque optique numérique permettant de stocker des films au format MPEG-2 (133 minutes) ou des données informatiques (4,7 Giga octets). Les données peuvent être stockées sur deux couches, doublant ainsi la capacité.

DWDM (Dense *Wavelength Division Multiplexing*) : version très haut débit du WDM qui combine jusqu'à 64 longueurs d'onde (> 100 Gbit/s).

E

EBCDIC (*Extended Binary Coded Decimal Interchange Code*) : code sur 8 bits permettant de coder les caractères alphanumériques.

Edge device (*Équipement de bordure*) : commutateur réseau local (Ethernet ou Token-Ring) disposant d'une interface ATM. Selon les cas, supporte les protocoles LANE, MPOA et PNNI. Dans certains cas, également appelé commutateur multiniveau.

ECMA (*European Computer Manufacturer Association*) : organisme de normalisation agissant dans les domaines des ordinateurs et des réseaux locaux.

EDI (*Electronic Data Interchange* ou *Échange de Données Informatisées*) : procédure d'échange de documents sous forme électronique. Spécifie la structure des données informatiques.

ELAN (*Emulated Local Area Network*) : réseau virtuel sur ATM.

EMC (*Electromagnetic Compatibility*) : mesure qui caractérise l'aptitude des différents composants informatiques, dont les câbles, à ne pas perturber d'autres composants.

ETEBAC (*Échanges télématiques entre les banques et leurs clients*) : ensemble de protocoles spécifiés par les organismes bancaires français pour l'échange de données informatiques.

ETSI (*European Telecommunication Standards Institute*) : organisme de normalisation participant aux travaux de l'ITU-T.

F

FAX (*Fac-similé*) : document télécopié.

FAI (*Fournisseur d'accès à Internet*) : désigne un opérateur réseau qui permet aux entreprises et aux particuliers de se connecter à l'Internet (*voir* ISP).

FCC (*Federal Communications Commission*) : organisme de régulation dans le domaine des télécommunications. Équivalent de l'ART en France.

FCS (*Frame Check Sequence*) : code de contrôle d'erreur équivalent au CRC.

FDDI (*Fiber Distributed Data Interface*) : réseau local haut débit (100 Mbit/s) en anneau à jeton spécifié par l'ANSI. FDDI-II est une version améliorée offrant le support de la voix et de la vidéo.

FEXT (*Far End Cross Talk*) : mesure, exprimée en décibels, du taux de réflexion dans un câble. Également appelé télédiaphonie.

FIFO (*First In, First Out*) : mode d'empilement de données dans une pile ou une file d'attente, dans laquelle la première donnée entrée est la première à ressortir. Voir également LIFO.

Frame Relay (*Relais de trames*) : protocole couvrant les couches 2 et 3 du modèle OSI utilisé sur les liaisons longues distances. Considéré comme un protocole X25 allégé.

G

Gbit/s (*Gigabits par seconde*) : nombre de milliards de bits transmis par seconde.

GFA (*Groupe fermé d'abonnés*) : ensemble de nœuds interconnectés formant virtuellement un seul réseau inaccessible de l'extérieur bien que partageant la même infrastructure que d'autres clients au sein du réseau d'un opérateur.

Go (gigaoctet) : équivalent à 1 073 741 824 octets ($1\ 024^3$).

GPRS (*General Packet Radio Service*) : évolution de la norme GSM permettant la transmission de données rapide.

GSM (*Global System for Mobile*) : norme européenne définissant les réseaux téléphoniques numériques sans fil (réseaux cellulaires).

H

H.323 (*Protocole 323, Série H de l'ITU-T*) : architecture et protocole permettant d'établir des conversations téléphoniques sur un réseau IP (cf. VoIP et SIP).

HDB3 (*High-Density Bi-polar modulus 3*) : méthode de codage utilisée dans les liaisons MIC (avis ITU-T G703).

HDLC (*High-level Data Link Control*) : désigne une méthode d'encapsulation des données sur un support de transmission synchrone comme une ligne spécialisée.

HDSL (*High-bit-rate Digital Subscriber Line*) : technique de transmission à haut débit sur cuivre (1,544 et 2,048 Mbit/s sur une portée de 4,6 km) en *full duplex* à la différence de l'ADSL.

HDTV (*High Definition Television*) : standard de la télévision numérique 16/9 offrant deux résolutions, 1280x720 et 1920x1080, et codée avec MPEG-2.

Hz (*Hertz*) : unité de fréquence correspondant à un cycle par seconde.

HPPI (*High Performance Parallel Interface*) : norme d'interface parallèle fonctionnant à 800 Mbit/s en transmettant 32 bits en parallèle ou fonctionnant à 2,6 Gbit/s sur 64 bits.

HTML (*HyperText Markup Language*) : langage de description d'une page web. La navigateur (browser web) exécute ces commandes pour afficher la page.

HTTP (*HyperText Transfer Protocol*) : protocole permettant d'envoyer une page web d'un serveur web vers un ordinateur équipé d'un navigateur. Protocole à la base du web.

I

IEEE (*Institute of Electrical and Electronics Engineers*) : organisme de standardisation des réseaux LAN et MAN, essentiellement de la famille Ethernet.

IETF (*Internet Engineering Task Force*) : organisme de standardisation des protocoles TCP/IP regroupant les constructeurs et opérateurs réseaux. Émet les RFC (*Request for Comments*).

INRIA (*Institut national de la recherche en informatique et en automatique*) : organisme français notamment en charge de la gestion de l'Internet en France.

Internet (*Inter Network*) : interconnexion de réseaux IP formant le plus grand réseau public à commutation de paquets du monde.

Intranet (*Intra Network*) : version privée de l'Internet dans une entreprise.

IP (*Internet Protocol*) : protocole de niveau 3 (couche réseau) à commutation de paquet. Protocole à la base de l'Internet.

IPX (*Internetwork Packet Exchange*) : protocole réseau utilisé par les serveurs Netware. IPX/SPX est l'équivalent de TCP/IP.

IPNS (*ISDN PABX Networking Specification*) : norme spécifiant la signalisation entre autocommutateurs privés. Permet d'interconnecter des PABX de marques différentes.

ISDN (*Integrated Services Digital Network*) : désigne le réseau téléphonique numérique (voir RNIS).

IPsec (*IP Security*) : protocole IP sécurité. Les paquets sont cryptés et authentifiés. Standard à la base des VPN-IP.

ISP (*Internet Service Provider*) : désigne un opérateur réseau qui permet aux entreprises et aux particuliers de se connecter à l'Internet (cf. FAI).

ITSP (*Internet Telephony Service Provider*) : opérateur proposant des services VoIP (voix sur IP) sur Internet.

ITU-T (*International Telecommunication Union – Telecommunication standardization sector*) : organisme affilié à l'ONU qui définit les normes et la réglementation en matière de réseaux de télécommunications.

J

Jitter (*gigue*) : désigne le déphasage des signaux d'horloge dû à la distorsion des signaux sur la ligne.

JPEG (*Joint Photographic Experts Group*) : groupe de travail et normes relatives à la compression d'images fixes.

K

Kbit/s (*Kilobits par seconde*) : nombre de milliers de bits transmis par seconde.

Ko (*Kilo-octet*) : équivalent à 1 024 octets.

KHz (*Kilohertz*) : désigne le nombre de milliers de cycles par seconde (fréquence) d'une onde ou d'une horloge.

L

L2TP (*Layer Two Tunneling Protocol*) : protocole permettant d'étendre une session PPP au-delà du point de terminaison, au dessus d'un réseau Frame-Relay, ATM ou IP.

LAN (*Local Area Network*) : désigne les techniques de réseau local.

LANE (*LAN Emulation Protocol*) : protocole permettant d'émuler un réseau local sur ATM. Utilise les composants LEC, LECS, LES et BUS.

LAT (*Local Area Transport*) : protocole, développé par Digital Equipement, permettant de relier des terminaux et imprimantes aux serveurs VAX en environnement Decnet.

LDAP (*Lightweight Directory Access Protocol*) : version allégée de l'annuaire X.500 et adaptée aux réseaux intranet et Internet.

LEC (*Lan Emulation Client*) : couche logicielle permettant d'adapter les protocoles de réseaux locaux sur ATM.

LECS (*Lan Emulation Configuration Server*) : couche logicielle contrôlant les ELAN. Gère les LES et indique aux *edge device* et stations ATM de quel LES ils dépendent en fonction de leur ELAN.

LES (*Lan Emulation Server*) : couche logicielle contrôlant un ELAN. Gère tous les *edge device* ou stations ATM appartenant à son ELAN ainsi que toutes les adresses MAC qui y sont rattachées côté réseau local.

LIFO (*Last In, First Out*) : mode d'empilement de données dans une pile ou une file d'attente, dans laquelle, la dernière donnée entrée est la première à ressortir. Voir également FIFO.

LLC (*Logical Link Control*) : couche logicielle qui a pour objet d'assurer le transport des trames entre deux stations. Elle se situe au niveau 2 du modèle OSI et est fonctionnellement proche du protocole HDLC.

LS (*ligne spécialisée*) : ensemble de liaisons permanentes et vues du client comme étant point à point.

LU (*Logical Unit*) : dans les réseaux SNA, désigne une entité définissant des droits d'accès et des règles de communication avec d'autres entités. La LU 6.2 décrit par exemple les fonctions de communication entre les programmes.

M

MAC (*Medium Access Control*) : couche logicielle qui a pour rôle de structurer les bits d'information en trames adaptées au support physique et de gérer les adresses physiques des cartes réseaux (on parle d'adresse MAC).

MAN (*Metropolitain Area Network*) : désigne un réseau étendu, généralement en fibre optique, à l'échelle d'un campus ou d'une ville.

MAP (*Manufacturing Automation Protocol*) : réseau local spécifié par General Motors pour les environnements industriels. La méthode d'accès est celle du jeton sur un câble CATV. Proche de la norme IEEE 802.4.

Mbit/s (*Mégabits par seconde*) : nombre de millions de bits transmis par seconde.

MHS (*Message Handling System*) : sous-ensemble de la norme X400 spécifiant les mécanismes de gestion des messages.

MHz (*Mégahertz*) : désigne le nombre de millions de cycles par seconde (fréquence) d'une onde ou d'une horloge.

MIC (*Modulation par impulsions codées*) : technique de transmission utilisée pour véhiculer des signaux analogiques sous forme numérique par échantillonnage des signaux. Par extension, désigne ce type de ligne utilisé en France et comportant 32 canaux de 64 Kbit/s.

MIB (*Management Information Base*) : base de données structurée selon la syntaxe ASN.1 décrivant les objets d'un équipement réseau.

MIPS (millions d'instructions par secondes) : nombre d'instructions qu'un microprocesseur peut exécuter en une seconde. Les instructions sont exécutées en un ou plusieurs cycle horloge. Le nombre de cycles par seconde est égal à l'inverse de la fréquence du processeur exprimée en MHz.

Modem (*modulateur-démodulateur*) : appareil transmettant des signaux analogiques sur le réseau téléphonique. Offre les fonctions de numérotation, de connexion, et éventuellement de compression et de correction d'erreur.

Mo (*Mégaoctet*) : équivalent à 1 048 576 octets ($1\ 024^2$).

MPEG (*Moving Pictures Expert Group*) : groupe de travail et normes relatives à la compression d'images animées.

MTA (*Message Transfer Agent*) : élément logiciel qui achemine les messages entre les différents nœuds d'un système de messagerie.

MTBF (*Mean Time Between Failure*) : mesure statistique, exprimée en nombre d'heures, du temps de fonctionnement d'un équipement avant une panne.

MTS (*Message Transfer System*) : désigne dans la norme X400 l'ensemble des MTA d'un même domaine.

MTTR (*Mean Time To Repair*) : exprime le temps de rétablissement suite à une panne.

MTU (*Maximum Transmission Unit*) : longueur (ou taille) maximale d'une trame ou d'un paquet d'un protocole réseau.

N

NDIS (*Network Driver Interface Specification*) : spécification par Microsoft d'une interface logicielle universelle d'accès aux cartes réseau (NIC) d'un PC. Permet au logiciel situé au niveau de la couche 2 du modèle OSI d'utiliser n'importe quelle carte réseau (voir ODI).

Netbios (*Network Basic Input/Output System*) : protocole de niveau session permettant de partager des ressources entre postes de travail et serveurs en environnement Windows (95, 98 ou NT).

NEXT (*Near-End Cross(X) Talk*) : mesure de la paradiaphonie d'un câble cuivre (aptitude d'un câble à ne pas être perturbé par les parasites).

NFS (*Network File System*) : système de gestion de fichiers réseau sur TCP/IP. Permet de partager des fichiers en donnant à l'utilisateur l'impression qu'ils sont locaux.

NIC (*Network Interface Card*) : désigne une carte réseau dans un ordinateur.

NOS (*Network Operating System*) : désigne les systèmes d'exploitation des serveurs de fichiers tels que Netware ou Windows NT.

NRZ (*Non Return to Zero*) : méthode de codage des signaux numériques.

NTI (*Nœud de Transit International*) : commutateur assurant l'interconnexion des réseaux X25 entre les différents pays.

O

ODI (*Open Data-link Interface*) : spécification par Novell d'une interface logicielle universelle d'accès aux cartes réseau (NIC) d'un PC. Permet au logiciel situé au niveau de la couche 2 du modèle OSI d'utiliser n'importe quelle carte réseau (voir NDIS).

OEM (*Original Equipement Manufacturer*) : contrat de fabrication sous licence.

OLE (*Object Linking and Embedding*) : protocole d'origine Microsoft pour la communication des applications réparties.

OLTP (*On Line Transaction Processing*) : désigne un moniteur transactionnel en ligne.

OSPF (*Open Shortest Path First*) : protocole de routage TCP/IP.

P

PABX (*Private Automatic Branch eXchange*) : appelé autocommutateur ou autocom, il assure la concentration des postes téléphoniques et la commutation des circuits.

PAD (*Paquet Assembler / Desassembler*) : équipement permettant aux terminaux asynchrones (travaillant caractère par caractère) d'accéder à un réseau de paquets X25.

PAV (*Point d'accès Vidéotex*) : variante du PAD pour l'accès aux services Minitel reposant sur le réseau X25 de Transpac.

PC (*Personal Computer*) : micro-ordinateur. Désigne un ordinateur compatible avec les spécifications des sociétés IBM et Intel en matière d'ordinateur personnel.

PCM (*Pulse Code Modulation*) : équivalent américain du MIC (Modulations par Impulsions Codées) comportant 24 canaux à 64 Kbit/s.

PESIT (*Protocole d'Échange pour le Système Interbancaire de Télécompensation*) : ensemble de procédures spécifiées par les banques françaises pour les échanges de données électroniques entre les banques.

Ping (*Packet Internet Groper*) : petit logiciel permettant de mesurer les temps de réponse entre deux nœuds réseau. Le programme attend un paquet en écho de chacun qu'il a envoyé.

Pixel (*Picture Element*) : désigne un point élémentaire dans une image.

PNNI (*Private to Private Network*) : protocole de routage ATM. Permet aux commutateurs ATM de déterminer les meilleurs chemins virtuels.

POP (*Point Of Presence*) : salle informatique dans laquelle un opérateur héberge les équipements télécom lui permettant de connecter ses clients à son réseau.

PPP (*Point-to-Point Protocol*) : protocole de la couche liaison utilisé sur les lignes série téléphoniques ou spécialisées.

PRI (*Primary Rate Interface*) : désigne l'accès primaire RNIS.

PSDN (*Packet Switching Data Network*) : désigne un réseau à commutation de paquets.

PSTN (*Public Switched Telephone Network*) : désigne le réseau téléphonique. Équivalent au RTC français (réseau téléphonique commuté).

Q

Q-SIG (*Q signalisation*) : norme basée sur la signalisation CITT Q.931 définissant les échanges entre les systèmes de signalisation publics et privés (PABX).

R

RADSL (*Rate-adaptive Asymmetric Digital Subscriber Line*) : technique de transmission à haut débit analogue à ADSL mais avec modification du débit en fonction de la qualité de la ligne.

RARP (*Reverse Address Resolution Protocol*) : protocole basé sur un *broadcast* permettant d'obtenir l'adresse réseau (niveau 3) à partir d'une adresse MAC (niveau 2).

RFC (*Request For Comments*) : documents issus de l'IETF (*Internet Engineering Task Force*) spécifiant tous les standards en matière de protocoles Internet.

RISC (*Reduced Instruction Set Components*) : type de microprocesseur caractérisé par un jeu d'instructions réduit au minimum (opposé à CISC).

RIP (*Routing Information Protocol*) : protocole de routage TCP/IP et IPX/SPX.

RNIS (*réseau numérique à intégration de services*) : désigne le réseau téléphonique numérique censé remplacer le RTC petit à petit (cf. ISDN).

RPIS (*réseau privé à intégration de services*) : interconnexion de PABX et de lignes pour former un réseau RNIS privé.

RPV (*réseau privé virtuel*) : voir VPN.

RSA (*Rivest Shamir et Adelman*) : initiales des inventeurs de l'algorithme de chiffrement du même nom.

RSVP (*Resource Reservation Protocol*) : protocole de signalisation permettant de garantir une qualité de service sur les réseaux TCP/IP pour les applications temps réel comme le transport de la voix.

RTC (*réseau téléphonique commuté*) : désigne le réseau téléphonique analogique classique.

RVA (*réseau à valeur ajoutée*) : désigne un réseau associé à des services tels que la conversion de protocoles, la facturation, l'accès à des bases de données, etc.

S

SAP (*Service Access Point*) : mécanisme logiciel de pointeurs permettant à un logiciel réseau d'utiliser les services d'une couche inférieure. Le SAP est un numéro unique permettant d'identifier le logiciel qui a envoyé une trame ou un paquet.

SAN (*Storage Area Network*) : réseau haut débit sur fibre optique, reposant sur la technologie Fibre Channel, et dédié aux sauvegarde des données.

SDH (*Synchronous Digital Hierarchy*) : mode de transmission numérique pour les réseaux de télécommunications hauts débits.

SDSL (*Single-line Digital Subscriber Line*) : technique de transmission à haut débit sur cuivre (1,544 et 2,048 Mbit/s sur une portée de 3 km) similaire à HDSL.

SIP (*Session Initiation Protocol*) : architecture et protocole de l'IETF permettant d'établir des conversations téléphoniques sur un réseau IP (cf. VoIP et H.323).

SIT (*Système Interbancaire de Télécompensation*) : réseau d'échange de données entre les banques françaises.

SMDS (*Switched Multimegabit Data Service*) : réseau à commutation de paquets à hauts débits pour des liaisons longues distances.

SMTP (*Simple Mail Transfert Protocol*) : protocole utilisé par les systèmes de messagerie dans le monde TCP/IP.

SNA (*Systems Network Architecture*) : spécifications des réseaux de la société IBM (*International Business Machine*).

SNMP (Simple *Network Management Protocol*) : protocole de la famille TCP/IP utilisé pour administrer à distance les équipements réseaux à partir d'une station d'administration.

SPX (*Sequenced Packet Exchange*) : protocole réseau utilisé par les serveurs Netware. IPX/SPX est l'équivalent de TCP/IP.

STP (*Shielded Twisted Pair*) : désigne un câble blindé composé de 2 ou 4 paires en cuivre (voir UTP).

T

TASI (*Time Assignment Speech Interpolation*) : technique de multiplexage temporel statistique adaptée à la transmission de la voix numérisée et utilisée dans les satellites et les câbles sous-marins.

TAXI (*Transparent Asynchronous Transmitter/Receiver Interface*) : spécification d'une interface fibre optique à 100 Mbit/s pour FDDI et ATM. Utilise le codage 4B/5B.

TCP/IP (*Transport Control Protocol / Internetwork Protocol*) : protocole de transport des données sous forme de paquets, universellement utilisé sur les réseaux LAN et WAN. Désigne également toute une famille de protocoles de niveau session ou application (HTTP, FTP, SMTP, SNMP, etc.).

TDM (*Time Division Multiplexing*) : technique de multiplexage dans le temps (*voir* AMRT).

TIC (*Token-Ring Interface Coupler*) : dénomination d'un mode d'attachement des contrôleurs IBM SNA 3174, 3745, etc. Désigne une connexion SNA sur réseau local.

TNR (*Terminaison Numérique de Réseau*) : coffret marquant la séparation entre le réseau RNIS public et la partie privée chez l'utilisateur. Correspond à la prise téléphonique.

TOP (*Technical Office Protocol*) : ensemble de protocoles développés par la société Boeing pour la CAO (*conception assistée par ordinateur*). TOP respecte entièrement la normalisation OSI.

TTL (*Time To Live*) : compteur permettant de déterminer le temps de validité restant pour une donnée ou un paquet réseau.

U

UA (*User Agent*) : partie cliente des systèmes de messagerie.

UHF (*Ultra High Frequency*) : bande de fréquence, située entre 30 MHz et 300 MHz, utilisée pour transmettre des émissions de télévision analogiques.

UDP (*Network Interface Card*) : équivalent de TCP mais en mode non connecté, sans les mécanismes de contrôle de flux, de reprise sur erreur et autres options.

UTP (*Unshielded Twisted Pair*) : désigne un câble non blindé composé de 4 paires en cuivre. La norme EIA-TIA 586 définit 5 catégories de câbles de ce type (voir STP).

V

VCC (*Virtual Channel Connection*) : chemins virtuels ATM prédéfinis ou affectés dynamiquement pour les besoins des protocoles LANE et MPOA. Un VCC héberge une encapsulation LLC telle que définie par la RFC 1483.

VDSL (*Very-high-bit Digital Subscriber Line*) : technique de transmission à haut débit sur cuivre (de 13 à 52 Mbit/s sur une portée de 300 à 1 300 mètres) similaire à ADSL.

VHF (*Very High Frequency*) : bande de fréquence, située entre 300 MHz et 3 GHz, utilisé pour transmettre des émissions de télévision analogiques.

VLAN (*Virtual Local Area Network*) : désigne les réseaux locaux virtuels. L'extension des VLAN sur ATM passe par les ELAN.

VoIP (*Voice Over IP*) : désigne les technologies permettant de transmettre des conversations téléphoniques (et des visioconférences) sur un réseau IP. Repose sur les protocoles H.323 et SIP.

VPN (*Virtual Private Network*) : désigne un réseau WAN (reposant sur ATM, Frame-Relay ou MPLS) dédié à une entreprise mais reposant sur un backbone haut débit que l'opérateur partage avec d'autres clients. L'opérateur garantit que les données ne sont pas mélangées avec celles d'un autre client.

VPN-IP (*Virtual Private Network – Internet Protocol*) : liaison IP protégée par authentification et cryptage des données et reposant sur le protocole IPsec et/ou L2TP.

VSAT (*Very Small Aperture Terminals*) : protocole de transmission satellite utilisant des paraboles de faibles diamètres (inférieurs à 3,7 m).

W

WAN (*Wide Area Network*) : désigne un réseau longue distance (réseau étendu) reposant sur les technologies de transmissions de données en série et des protocoles tel que le Frame Relay et ATM.

WDM (*Wavelength Division Multiplexing*) : mode de transmission numérique sur fibre optique multiplexant différentes longueurs d'onde et autorisant de très hauts débits (> 100 Gbit/s).

Web (*la toile*) : désigne l'ensemble des serveurs de page web de l'Internet.

WLAN (*Wireless LAN*) : réseau local sans fil à 11 Mbit/s, compatible Ethernet, et normalisé IEEE 802.11.

WML (*Wireless Markup Language*) : protocole de description d'une page web (version simplifiée de HTML et de HTTP) et adapté aux faibles débits des téléphones GSM.

WWW (*World Wide Web*) : désigne l'ensemble des serveurs web de l'Internet. Nom DNS généralement donné aux serveurs web.

X

X25 (*norme X.25 de l'ITU-T*) : protocole de niveau 3 spécifiant l'interface d'accès à un réseau à commutation de paquets.

X400 (*norme X.400 de l'ITU-T*) : systèmes de messagerie OSI (définit les protocoles et les logiciels MTA, UA)

X500 (*norme X.500 de l'ITU-T*) : annuaire de messagerie OSI.

XNS (*Xerox Network System*) : l'un des tous premiers protocoles LAN de la société Xerox, équivalent de TCP/IP et IPX/SPX.

xDSL (*x Digital Subscriber Line*) : regroupe les techniques de transmission ADSL, RADSL, HDSL, SDSL et VDSL.

Bibliographie

T. ANTTALAINEN, *Introduction to Telecommunications*, Artech House, 2002.

R. BATES, *Broadband Telecommunications Handbook*, Mc Graw Hill, 2002.

U. BLACK, *Advanced Internet Technologies*, Prentice Hall, 2000.

U. BLACK, *MPLS and Label Switching Networks*, Prentice Hall, 2002.

U. BLACK, *Voice Over IP*, Prentice Hall, 2002.

JA. CANAVAN, *Fundamentals of network security*, Artech House, 2000.

D. COMER, *Internetworking with TCP/IP*, Prentice Hall, 2000.

B. FOROUZAN, *Local Area Networks*, Mc Graw Hill, 2002.

M.S. GAST, *802.11 Wireless Networks – The Definitive Guide*, O'Reilly, 2002.

B. GI LEE, W-J KIM, *Integrated Broadband Networks*, Artech House, 2002.

O. HERSENT, DAVID GURLE, J-P PETIT, *IP Telephony*, Addison Wesley, 2000.

S. JHA, M. HASSAN, *Engineering Internet QoS*, Artech House, 2002.

J-P. LAUDE, *DWDM Fundamentals, Components, and Applications*, Artech House, 2002.

B. KHASNABISH, *Implementing voice over IP*, Wiley Interscience, 2003.

D. MCDYSAN, D. PAW, *ATM Theory & Application*, Mc Graw Hill, 2002.

D. MINOLI, *Voice over MPLS*, Mc Graw Hill, 2002.

D. MINOLI, P JOHNSON, *SONET-based Metro Area Network*, Mc Graw Hill, 2002.

J-L. MONTAGNIER, *Pratique des réseaux d'entreprise*, Eyrolles, 1999.

G. PUJOLLE, *Les Réseaux*, Eyrolles, 2002.

W. STALLINGS, *Local and Metropolitan Area Networks,* Prentice Hall, 2000.

A. TANENBAUM, *Computer networks,* Pearson, 2002.

Sites web

Accès aux RFC

www.rfc-editor.org/rfcsearch.html
- *Recherche dans la base de données des RFC*

www.ietf.org/rfc
- *Téléchargement rapide des RFC*

www.iana.org/numbers.html
- *Recherche des valeurs réservées par les protocoles*

Câblage

www.eia.org
- *Electronic Industries Alliance*

www.tiaonline.org
- *Telecommunications Industry Association*

www.broadband-guide.com

www.scope.com/standards
- *Testeurs de câblage Wirescope*

www.cablingstandards.com

www.necanet.org
- *National Electrical Contractors Association*

www.nema.org
- *National Electrical Manufacturers Association*

www.rs232.net
- *Portail spécialisé sur le câblage*

Internet

www.afnic.asso.fr
- *Association française pour le nommage Internet encoopération*

www.apnic.net
- *Asia Pacific Network Information Center*

www.arin.net
- *American Registry for Internet Numbers*

www.darpa.mil
- *Defence Advanced Research Projects Agency*

www.iab.org
- *Internet Architecture Board*

www.iana.org
- *Internet Assigned Numbers Authority*

www.icann.org
- *Internet Corporation for Assigned Names and Numbers*

www.ietf.org
- *Internet Engineering Task Force*

www.isoc.org □ *Internet Society : portail d'accès à tous les organismes régulant l'Internet*

www.internic.net □ *Enregistrement des noms de domaines*

www.register.com

www.ripe.net □ *Réseau IP européen*

www.rfc-editor.org □ *Accès à toutes les RFC (Request for Comments)*

www.6bone.net □ *Réseau Internet IPv6*

www.mbone.com □ *Réseau Internet multicast*

www.renater.fr □ *Réseau national de la recherche*

www.internet2.edu □ *L'Internet dédié aux universités*

www.ucaid.edu/abilene □ *Abilene : l'Internet dédié aux universités*

www.cybergeography.org/mapping.html

□ *Cartographie de l'Internet*

www.isc.org/ds □ *Internet Survey*

Modem-câble

www.opencable.com

www.ietf.org/html.charters/ipcdn-charter.html

□ *IP over Cable Data Network*

www.cabledatacomnews.com □ *Portail dédié aux modems câble*

Organismes de normalisation

web.ansi.org □ *American National Standards Institute*

www.etsi.org □ *European Telecommunications Standards Institute*

www.ema.org □ *Electronic Messaging Association*

www.iso.ch □ *International Standard Organization*

www.itu.ch □ *International Telecommunications Union*

www.ietf.org □ *Internet Engineering Task Force*

www.ietf.org/html.charters/wg-dir.html

	□ *Portail d'accès à tous les groupes de travail produisant les RFC*
www.regulate.org	□ *World Regulatory Telecommunications*
standards.ieee.org	□ *Institute of Electrical and Electronics Engineers*
www.w3.org	□ *WWW Consortium*

Organismes de régulation

www.art-telecom.fr	□ *Autorité de Régulation des Télécommunications*
www.fcc.gov	□ *Federal Communication Commission*
www.oftel.gov.uk	□ *Office of Telecommunications*
irgis.icp.pt	□ IRG – *Independent Regulators Group*

Protocoles

www.protocols.com	□ *Description de tous les protocoles*
www.atmforum.com	□ *ATM Forum*
www.frforum.com	□ *Frame-Relay Forum*
www.gigabit-ethernet.org	□ *Gigabit Ethernet Alliance*
www.fibrechannel.com	□ *Fibre Channel Forum*
www.t1.org	□ *Committee T1*
www.adsl.com	□ *ADSL Forum*
www.dslforum.org	□ *Forum des technologies DSL*
www.nmf.org	□ *Network Management Forum*
www.mplsworld.com	□ *Le monde MPLS*
www.mplssrc.com	□ *MPLS Resource Center*

Qualité de service

www.itmcenter.com	□ *Internet Traffic management ressource Center*
diffserv.lcs.mit.edu	□ *Portail d'accès aux informations sur Diffserv*
www.cis.ohio-state.edu/~jain/refs/ipqs_ref.htm	
	□ *Liens vers d'autres sites*

www.qosforum.com □ *Quality Of Service Forum*
www.ietf.org/html.charters/diffserv-charter.html
□ *Le groupe de travail DiffServ*

Réseaux sans fils

www.mobinet.com □ *L'Internet mobile*
www.ieee.org/wireless □ *Les standards WLAN de l'IEEE*
www.wi-fi.org □ *Le site de la WECA*
www.wlana.com □ *Wireless LAN Alliance*
www.bluetooth.org □ *Le site du consortium Bluetooth*
www.bluetooth.com □ *Portail spécialisé sur Bluetooth*
www.irda.org □ *Consortium des liaison infrarouges*
www.homerf.org □ *Consortium du HomeRF*
www.dectweb.com □ *Wireless Telecommunications Internet Services*
www.dect.ch □ *DECT Forum*
www.umts-forum.org □ *UMTS Forum*

Revues de presse

www.reseaux-telecoms.fr
www.01-informatique.com
www.lmi.fr □ *Le Monde informatique*
www.data.com □ *Data Communications*
www.networkcomputing.com
www.telecoms-mag.com
www.cmpnet.com □ *Portail d'accès à de nombreuses revues spécialisées*
www.zdnet.com □ *Toute l'actualité informatique*
www.techguide.com □ *Portail d'accès aux technologies*
www.techweb.com □ *Portail d'accès aux technologies*
www.entmag.com □ *Windows NT & 2000 News*
www.lantimes.com □ *LAN times*

www.samag.com □ *SysAdmin : le journal des administrateurs Unix*

Sécurité

www.2600.com □ *The Hacker Quarterly*
www.cert.org □ *Computer Emergency Response Team*
www.ciac.org □ *Computer Incident Advisory Capability*
www.cerias.purdue.edu/coast □ *Computer Operations, Audit, and Security Technology*
www.pki-page.org □ *Portail PKI* (*Public Key Infrastructure*)
www.ssi.gouv.fr □ *Direction Centrale de la Sécurité des Systèmes d'Information*

VoIP

www.ectf.org □ *Enterprise Computer Telephony Forum*
www.imtc.org □ *International Multimedia Telecommunications Consortium*
itel.mit.edu □ *Internet & Telecoms Convergence Consortium*
standard.pictel.com □ *Picturetel : accès aux standards H.323*
www.telephonyworld.com/iptelep/iptelep.htm □ *Portail VoIP*
www.theipsite.com
www.von.com □ *Voice On The Net*
www.gvcnet.com □ *Portail spécialisé sur la vidéoconférence*
www.openh323.org □ *Informations sur les protocoles H.323*
www.mpeg.org □ *Moving Picture Experts Group*

Index

1

1000bSX, 305
10bT, 100bT, 48, 59, 301

2

2B1Q (codage), 240, 245

8

802.11, 84, 86, 87, 91
802.11i, 104
802.15, 83
802.1d, 72, 303
802.1p, 303, 361
802.1q, 302, 307, 361, 364
802.1s, 304
802.1w, 75
802.3, 49, 70, 123, 303, 361
802.3ab, 52
802.3ad, 56
802.3u, 52
802.3z, 52

A

AAL (*ATM Adaptation Layer*), 283, 287, 288, 335, 411
Accès de base/primaire, 228
Accès distants, 478, 483
Adaptative (switch), 54
Administration réseau, 142
Adressage, 124
  ATM, 284, 290, 295
  Frame Relay, 277, 296
  IP, 17, 112, 127, 220, 380
  IPv4/IPv6, 187
  IPv6, 188, 194
  MAC, 49, 62, 70, 75, 112, 127, 152, 194, 276, 382
  plan, 112, 151
  public/privé, 115, 120
  unicast/multicast, 17, 380
ADSL, 211, 244, 247
AES (*Advanced Encryption System*), 469
Affaiblissement/Atténuation, 27, 30, 33, 35, 43, 55, 210
Agent SNMP, 142, 144
Agrégation de canaux B, 227
Agrégation de liens Ethernet, 56, 65, 67, 304
Aire OSPF, 311, 395
A-law, μ-law, 330, 415
Alias DNS, 175
AMI (codage), 240
Analogique, 210
Analyseur réseau, 140
Anneau (topologie), 4
Annuaire, 256
Anycast, 189
ARP, 129, 224, 225, 276
ART (Autorité de Régulation des Télécommunications), 78, 87
ASN.1, 147
Asynchrone, 222
ATM, 48, 59, 122, 211, 245, 262, 281, 298, 321, 411
Attaques IP, 447
AUI, 8, 50
Authentification, 466, 480
  mécanismes, 480
  calculette, 482
  serveur, 482
Autorité de certification (sécurité), 472, 475
Autosense, 48, 55
AWG, 30, 245

B

Backbone, 59
Baie (câblage), 36
BECN (*Backward Explicit Congestion Notification*), 278
BERT (*Bit Error Tests*), 264
Best effort, 283, 346, 353, 354, 357

BGP, 312, 317, 319, 323, 395
BLR (boucle locale radio), 106
Bluetooth, 83, 85
BNC, 7, 36
Bootp, 151, 154, 157
Boucle locale, 235, 236, 242, 246
BPDU, 72
brassage, 9
Brassage, 26, 37, 57, 217
Bridge, 217
Broadcast, 17, 154
  IP, 112, 276, 380
  MAC, 52, 127, 156, 211, 224, 300, 303
Browser de MIB, 148
Burst, 266
Bus (topologie), 4

C

Câblage (infrastructure), 24
Câble blindé/écranté, 28
Câble cuivre, 6
Cache
  ARP, 129
  DNS, 165, 168, 170
  navigateur, 257
Cahier des charges (câblage), 38
Canal B (RNIS), 228
CAP (codage), 246
Carte réseau, 7, 13, 49, 56, 126, 149, 152, 220, 361
Catégorie 5, 6, 7, 29, 31
CCTP, 22, 38
CDMA, 82
Cellule ATM, 245, 287, 288
CELP (codage), 329
CEM, 33
Certificat (sécurité), 472, 474, 475
Challenge-response (sécurité), 480
CHAP (*Challenge Handshake Authentication Protocol*), 222, 231, 480
Cheval de Troie (attaque), 450
Chiffrement
  algorithmes, 469
  IPsec, 477
  performances, 471
  SSL, 472
CIDR (*Classless Inter-Domain Routing*), 114, 187, 319
CIF (*Common Intermediate Format*), 337
CIR Frame Relay, 266, 278
Circuit virtuel, 235
  Frame Relay, 272
  ATM, 286, 287
  MPLS, 322
Classe D, E, F, 29, 31
Classe de service, 357, 361, 366, 374
Classical IP, 123, 263, 296, 299
Classificateur, 357, 368
Classification, 363
Clé de chiffrement, 469, 474, 477
CLLM (*Consolidated Link Layer Management*), 269, 278
Coaxial (câble), 28, 34
Codage, 82, 93, 240, 245, 328, 432
Codec, 330
  audio, 331, 340, 411
  vidéo, 334, 411
Collapse backbone, 60, 298
Collision, 11, 49, 52, 54, 55
Commutateur
  niveau 2, 52, 54, 62, 68, 72
  niveau 2/3, 300
  niveau 3, 114, 300, 303, 308
Commutation
  de cellule, 284, 410
  de circuit, 295, 339, 341, 410
  de paquet, 341, 410
  Ethernet, 68
  Frame Relay, 274
  IP/MPLS, 320
  matrice, 56, 62, 68
Compilation de MIB, 146
Compression, 328, 335
  Frame Relay, 281
  TCP/IP, 226
  RTP, 435
Concentrateur, 4, 10, 52, 53, 68
Conditionneur, 357
Conflit d'adresses, 459
Congestion, 269, 345, 346, 348
Contrôle d'admission, 368
Contrôle d'erreur
  ATM, 288
  Frame Relay, 269
  TCP, 132
Contrôle d'intégrité, 468
Contrôle de flux
  Frame Relay, 269
  Gigabit, 304
  HDLC, 228

LAP-F, 274
TCP, 132
UDP, 131
COPS, 376
CoS (*Class of Service*), 360, 377
Couches réseaux, 68, 122
CPE, 236, 242
CR-LDP (*Constraint-Based LDP*), 324
CSMA/CA, 87, 90, 91
CSMA/CD, 49, 55
CSMA/CR, 228
CSU/DSU, 211, 241, 264, 281, 285, 286
Cuivre/fibre optique (choix), 27, 63, 66, 301
Cut-through (switch), 54

D

DB9/DB25, 219
DCE/DTE, 211, 219
DCT (codage), 336
Dead gateway, 308
Débit, 48, 61, 213, 258, 266
Débordement RNIS, 232
DECT, 23, 33, 426
Default gateway, 223, 307
Délai de propagation, 210
Délai de transit, 269, 273, 283, 293, 338, 374, 435
DES (*Data Encryption Standard*), 469
DHCP, 149, 157, 226
Diffie-Hellmann (sécurité), 475
Diffserv, 351, 377
DLCI Frame Relay, 268, 276
DMT (codage), 246
DMZ (*DeMilitarized Zone*), 458, 465
DNS (*Domain Name System*), 162, 168, 420
base de données, 170
client, 180
filtrage, 460
serveur, 168
vulnérabilité, 444
Domaine de broadcast, 187, 189
Domaine de confiance (sécurité), 456
Driver, 14, 16
DS0 (canal 64 Kbit/s), 240, 328
DSA (*Digital Signature Algorithm*), 476
DSCP, 362, 364
DSL (*Digital Subscriber Line*), 211, 242, 246
DSP (*Digital Signaling Processing*), 330
DSS1, 270
DSSS, 82
DTMF, 339
DVMRP, 386, 401
DWDM (*Dense Wavelength Division Multiplexing*), 250
DWMT (codage), 246
DXI (*Data eXchange Interface*), 286

E

E.164, 273, 296, 417
E1/T1, 240, 330
EAP (*Extensible Authentication Protocol*), 104
ECC (*Elliptic Curve Cryptosystem*), 470
Échantillonnage, 328, 432
Echo, 338
ECN (*Explicit Congestion Notification*), 348
EGP, 312
EIA/TIA, 30, 31
EIGRP, 312
ELMI (Enhanced LMI), 273, 278
EMC, 33
EN50173, 31
Encapsulation, 122, 340
ATM, 286, 287
Ethernet, 129
Frame Relay, 272
IP, 127
Netbios, 131
Enveloppe numérique, 477
Erlang, 341
Espace d'adressage, 124
Ethernet, 4, 27, 48, 56, 68, 122, 210
Ethernet v2/802.3, 123
Ethertype, 125
Étoile (topologie), 4, 6

F

FDDI, 59
FDMA, 81
Fédérateur (réseau), 59, 305
FHSS, 82
Fibre optique, 27, 32, 43, 52, 67, 245, 305
tests optiques, 43
Fichier DNS
cache, 171
initialisation, 179
Fichier MIB, 145

FIFO (*First In, First Out*), 346
File d'attente, 344, 350
Filtrage (firewall), 455
Filtrage de paquet, 451
Firewall, 118, 450, 453, 464, 478, 484
Flux
client-serveur, 254
conversationnel, 252
transactionnel, 253
FRAD (*Frame Relay Access Device*), 265, 271
Fragmentation, 128, 191
Frame Relay, 211, 262, 265, 321
FTP, 9, 178, 270, 344
vulnérabilité, 443
FTP (câble), 30, 33
FTPS, 472
Full duplex (Ethernet), 55

G

G.709, 249
G.711, 415
G.711, G.722, G.723, etc., 331, 333, 340, 411
GARP, 76
Gatekeeper, 412, 420, 424, 425
Générateur de trafic, 141
Gigabit, 29, 32, 52, 59, 263, 304, 305
Gigue, 293, 339, 340, 357, 376
GMII, 50
GMRP, 76, 408
Groupe multicast, 194, 204, 382
GSM, 33
GTB (gestion technique du bâtiment), 23
GVRP, 76, 303

H

H.245, 414, 424, 431
H.261, H.263, 335, 337
H.323, 34, 411
HDLC, 221, 228
HDSL, 240, 243
Hiperlan / Hiperman, 85
Home RF, 83, 85
Hop limit IPv6, 191, 201
Hosts (fichier), 162
Hot-spot, 96
HSRP (*Hot Standby Router Protocol*), 308
HTTP, 126, 354
vulnérabilité, 445
HTTPS, 472
Hub, 4, 10, 52, 53, 68

I

IANA (*Internet Assigned Numbers Authority*), 131, 150, 194, 380
ICMP, 136, 137
vulnérabilité, 447
ICMPv6, 196
IDEA (*International Data Encryption Algorithm*), 469, 472
IDSL, 242
IETF (*Internet Engineering Task Force*), 191, 324, 377
IGMP (*Internet Group Membership Protocol*), 187, 204, 382, 387, 392, 397
IGMP snooping, 408
IGP, 312
IKE (*Internet Key Exchange*), 475, 479
ILMI (*Integrated Local Management Interface*), 291
IMAPS, 472
Impédance, 9, 28, 30
Intégrité (contrôle), 450, 468, 472, 479
Interface routeur, 114, 216, 220, 230, 265, 276, 280
Interface série, 220, 221, 270
Interframe/Intraframe (compression vidéo), 335
Intersite (réseau), 209, 234, 251, 264, 265, 405
Intranet, 175
Intserv, 351, 367, 377
Inverse ARP, 276, 287, 296
IP, 128
IP precedence, 352, 363, 430
Ipconfig, 158
IPsec, 238, 473, 477
IPv4/IPv6, 187
IPv6, 170, 178
ISAKMP (*Internet Security Association and Key Management*), 479
ITU (*International Telecommunication Union*), 234

J

Jitter, 293, 338, 434
JPEG, 335

K

Karn (algorithme), 132
Kbit/s, 210
Kerberos, 483

L

Label, 322
LAN, 4, 117, 234, 310, 365
LANE (*LAN Emulation*), 263, 284, 291, 299
LAP-D, 228
LAP-F, 272
LDAPS, 472
LDP (*Label Distribution Protocol*), 323
Leaky-bucket, 293, 347
Législation chiffrement, 477
Lien virtuel OSPF, 315
LIFO (*Last In, First Out*), 323
LLC, 76
LLC (*Logical Link Control*), 123, 125, 287
LMDS (*Local Multipoint Distribution Service*), 106
LMI (*Local Management Interface*), 272, 278
Local technique, 25
Longueur d'onde, 35, 43, 250
Loopback, 113, 188
LS (ligne spécialisée), 210, 227, 235, 262, 264
LTE (locaux techniques d'étage), 24, 27, 34, 36, 218

M

MAC, 49, 62, 68, 74, 87, 154, 224, 308
MAN, 305
Manchester (codage), 8, 52
Marquage, 361
Masquage d'adresses, 459
Matrice de flux (firewall), 462, 463
MAU (*Medium Attach Unit*), 4
MBONE, 402, 408
MCU H.323, 412
MD5 (*Message Digest*), 480
Messagerie, 174
Méthode d'accès, 5
MIB (*Management Information Base*), 147, 291
MIC (modulation par impulsions codées), 330
Micro-ondes, 107
Microsegmentation, 62
MII, 50
Mixer RTP, 433, 436
M-JPEG (Motion JPEG), 335
MLQ (codage), 329
MMDS (*Multichannel Multipoint Distribution Service*), 106
Modulation du signal, 80
MOS (*Mean Option Score*), 333
MOSPF, 391, 401
MP3, 336
MPEG, 335
MPLS, 320, 377
MPOA (*Multi Protocol Over ATM*), 299
MSTP (*Multiple Spanning Tree Protocol*), 304
MTA (*Message Transfer Agent*), 174
MTU (*Maximum Transfer Unit*), 123, 128, 191, 201, 204, 374
Multicast
  IP, 112, 369, 380, 405
  IPv6, 190, 194, 204
  MAC, 211, 303
Multimédia, 328, 431, 437
Multimode/monomode, 32
Multipaire (câble), 28
Multiplexage de fréquence, 81
Multiplexage de circuits virtuels ATM, 286, 287
Multiplexage des protocoles, 125
Multiplexage temporel, 245
Multiplexeur, 235, 240, 241
MX (enregistrement DNS), 174

N

Nagle (algorithme), 133, 253
NAT (*Network Address Translation*), 459
Navigateur, 354
NDIS (interface), 14
Netbios, 14, 131, 150
  vulnérabilité, 445
Netstat, 159
NEXT (câblage), 30

NFS (*Network File System*), 446
NLPID (IP dans Frame-Relay), 263, 272, 296
NNI (*Network to Network Interface*), 270
Nom DNS complet/relatif, 181
Nommage, 162
NSAP, 284, 291
Nslookup, 179, 181
Numérique, 210
Numérisation, 328

O

ODI (interface), 14
OFDM, 82, 106
Ohm, 9, 28
Opérateur, 25, 212, 235, 265, 271
Ordonnanceur, 368
OSPF, 114, 311, 314, 391, 395
OTP (*One Time Password*), 480
OUI, 70
Outsourcing, 239
Overhead, 211, 259, 263, 435

P

PABX, 23, 28, 271, 416, 418
PABX IP, 426
Paire torsadée, 6, 30, 34
PAM (codage), 246
Paquet
  IP, 125, 126, 212, 352
  TCP, 125, 132, 133, 253
  UDP, 131
Paradiaphonie, 29, 30, 43
Pas de torsade, 30
Passerelle H.323, 412, 416, 425
PAT (*Port Address Translation*), 459
PCM (*Pulse Code Modulation*), 330
PGP (*Pretty Good Privacy*), 473, 477
PIM, 396, 401
Ping, 136, 226
PKCS (*Public Key Cryptography Standard*), 474
Plan d'adressage, 112
Plan de nommage, 162
Plate-forme d'administration, 142
Plésiochrone, 241
Policing, 354, 355, 362
Pont, 96, 217, 224, 272
POP (opérateur), 236, 265, 271
POP3S, 472
Port (tcp/udp), 125, 130, 132, 363
POTS (*Plain Old Telephone Service*), 417
Powersum Next (câblage), 30, 31
PPP, 122, 211, 221, 231
Préfixe IPv6, 188, 194
Priorité, 344
Promiscus, 140
Proxy
  ARP, 155, 225
  cache, 451
  firewall, 451
  signalisation ATM, 284
PVC
  ATM, 284, 290
  Frame Relay, 268, 273, 275

Q

Q.922, 269
Q.931, 414, 431
QAM (codage), 246
QCIF, 337
QoS (*Quality of Service*), 343, 369
Qualité de service, 239, 278, 282, 292, 341, 351, 374, 424, 430
Quantification, 328, 329, 336

R

Radius, 481, 483, 484
RADSL, 244
Rapport signal/bruit, 30, 31
RARP, 129, 276
RAS (signalisation), 414, 422, 431
Rebond (attaque), 449
Recette (câblage), 22, 40
RED, 347, 349
Redirection IP, 202
Redondance, 65, 227, 308, 313
Réflectométrie, 41
Refus de service, 447
Registry Windows
  dead gateway, 308
  default gateway, 154
  DHCP, 149
  DNS, 169, 179

encapsulation IP dans Ethernet, 126
TTL, 136
Relais applicatif (firewall), 452, 465
Relais de circuit (firewall), 452, 466
Relais DHCP, 155, 156
Requête SNMP, 144
Résolution d'adresse, 129, 198
Résolution d'adresse inverse, 199, 276, 287, 296
Résolution DNS inverse, 178
Resolver DNS, 162, 180
RFC 1918, 115, 120
RIP, 312, 386
RJ45, 6, 9, 27, 28, 36, 57, 217, 219
RLE (compression), 335
RNIS, 210, 217, 242, 411
accès de base, 228
secours, 227
Roaming, 96
Routage, 68, 69, 128, 191, 221, 300, 307, 311
Route (commande windows), 159
Routeur, 114, 156, 212, 216, 220, 234, 270, 276, 300, 308, 383
Routeur (configuration), 219
Routeur vs firewall, 455
RPC (*Remote Procedure Calls*), 446
RS-232, 219
RSA (*Rivest Shamir et Adleman*), 470, 472
RSTP (*Rapid Spanning Tree Protocol*), 75
RSVP, 323, 367, 375, 377, 424, 434
RTC, 122, 210, 235, 339, 411
RTP/RTCP, 424, 431, 432

S

S/Mime (*Secure Multi Purpose Mail Extensions*), 473
SAAL, 284
Salle informatique, 25
SAP, 14, 76, 123
SC (connecteur), 27, 32, 36
SDH (*Synchronous Digital Hierarchy*), 211, 241, 245, 247
SDLC, 125
SDSL/SHDSL, 244
Secours RNIS, 227
Sécurité (politique), 442
Segment Ethernet, 62, 217
Segmentation, 12, 53
Sérialisation, 339
Série (câble/port), 217, 219
Série (interface), 218, 265, 281, 286
Serveur DNS
cache, 177
primaire, 172
racine, 175
secondaire, 177
Service opérateur, 215, 236
SET (*Secure Electronic Transaction*), 476
SFTP, 9
SHA (*Secure Hash Algorithm*), 470, 472
Shannon, 328
Signal radio, 79
Signalisation, 228, 240, 270, 367, 411
Signalisation
ATM, 284, 291
Frame Relay, 273
H.245, 424
Q.931, 414
RAS, 422
Signature numérique, 476
SIP, 409
SLA (*Service Level Agreement*), 239
SLC (codage), 246
SMTP, 174
SMTPS, 472
SNA, 125
SNAP, 123, 125, 263, 272, 287
SNMP, 11, 56, 121, 142, 144, 291
vulnérabilité, 445
Socket, 130
SONET (*Synchronous Optical Network*), 241, 247
Source quench, 138, 346, 349
Spanning tree, 67, 68, 71, 303
Spécificateur, 367
Spoofing (attaque), 448, 465
SSL (*Secure Socket Layer*), 472
Stackable, 12, 50
Stateful inspection (firewall), 452, 465
Station d'administration, 142
STM, 245
Store-and-forward (switch), 54
STP, 9
STP (câble), 29, 30
Stub area, 314
STUN, 125
Subnet, 115, 117, 120, 121, 219, 221, 225, 307
Support de transmission, 214
SVC, 272
ATM, 284, 290
Frame Relay, 268, 273, 277
Switch Ethernet, 52, 54, 62, 72

SYN flooding (attaque), 448
Synchrone, 222, 241, 247
Système autonome (OSPF), 312, 317

T

T.120, 411, 431, 438
T1/E1, 240, 328
Tableau virtuel, 437
TCP, 130, 132, 349
TCP/IP, 13, 222, 270
TCP/IP (configuration), 149, 157, 223
TDM, 245
TDMA, 81
Teardrop (attaque), 448
Telnet, 126, 127, 130, 213, 252, 344, 372
  vulnérabilité, 443
telnetS, 472
Temps de réponse, 53, 59, 63, 132, 136, 137, 140, 213, 226, 253, 260, 268, 344, 346, 351, 372
TFTP, 157
TKIP (*Temporal Key Integrity Protocol*), 104
TLS (*Transport Layer Security*), 472
TMS (*Traffic Management Specification*), 282
TNR (Terminaison numérique de réseau), 230
Token-bucket, 375
Token-Ring, 4, 48
Topologie, 4, 5
TOS (*Type of Service*), 352, 363, 392
Traceroute, 137
Traduction d'adresses, 115
Traffic shaping, 278, 282, 348
Trame, 8
  Ethernet, 123, 125
Transceiver, 8, 50, 65
Translation d'adresses, 459, 467
Trap SNMP, 144
Trunk (Ethernet), 65, 304
TTL
  ARP, 129
  DNS, 170, 177
  Hop limit IPv6, 191, 201
  IP, 128, 136, 137, 150, 407
Tunneling (attaque), 449

U

UDP, 130, 131
UNI (*User Network Interface*), 269, 270, 284
Unicast, 127, 130, 188, 369, 380
Uplink (port/lien), 11, 55, 56, 57, 61, 301
UTP (câble), 9

V

V.24, 218
V.35, 218, 264, 286
VDSL, 245
Ver (attaque), 450
Vidéo, 34, 328
VLAN, 56, 226, 299, 302, 324
VoIP, 411, 417, 426
VoIP, VoFR, VoATM, 341
Vol de session (attaque), 449
Volumétrie, 255
VPI/VCI ATM, 284
VPN (*Virtual Private Network*), 238, 477
VPN-IP, 235
VPN-MPLS, 320
VRRP (*Virtual Router Redundancy Protocol*), 309
Vulnérabilités IP, 443

W

WAN, 120, 210, 234, 310, 365, 416
WDM (*Wavelength Division Multiplexing*), 250
Web, 210, 213
WEP, 87, 103
WFQ (*Weighted Fair Queueing*), 346, 352, 368, 372, 430
Wi-Fi, 84, 85, 86
Winipcfg, 157
WINS, 131, 150, 383
  filtrage, 461
Winsock, 130, 370, 383
WLAN (*Wireless Local Area Network*), 84, 85, 91, 93
WMAN (*Wireless Metropolitan Area Network*), 84, 85, 106
WPA (*Wi-Fi Protected Access*), 104
WPAN (*Wireless Personnnal Area Network*), 84, 85
WRED, 348, 352, 368, 372
www, 167, 175, 181

X

X.121, 273, 296
X.509, 472, 474, 476
X21/V11, 218, 242, 264

Z

Zigbee, 84

Table des encarts

À quoi sert l'adressage ? 17
Adressage IPv6 (RFC 2373, 2374, 2450) 188
Adressage IP v4 (RFC 791) 113
Adressage multicast IPv6 (RFC 2373 et 2375) 194
Adresses MAC (802.3) 70
Analyseur réseau 140
ARP (RFC 826) 129
ATM (I.361) 288
BGP (RFC 1171 à 1174) 319
Câbles cuivre en paires torsadées 9
Canal de contrôle H.245 424
Champ TOS (RFC 791, 1349) 353
Chantier de câblage 41
Circuit virtuel 235
Codec audio 329
Comment fonctionne un commutateur ? 62
Commutateurs Ethernet 300
Commutation Frame Relay 274
Compatibilité électromagnétique 33
Composants d'un système de câblage 26
DHCP (RFC 2131) 154
Diffserv (RFC 2474, 2475) 357
DNS (base de données) 170
DNS (RFC 1034, 1995, 2181) 165
Drivers, NDIS, ODI 14
DVMRP (RFC 1075) 389
Encapsulation dans ATM (RFC 1483) 287
Encapsulation dans Frame-Relay 272
Encapsulation IP dans Ethernet 123
Ethernet (802.3) 49
Ethernet et Token-Ring 5
Ethernet full duplex (802.3x) 55
Ethernet sans fil (802.11) 87
Fibre optique 35
Files d'attente (FIFO, WFQ, RED) 347
Files d'attente (rôle) 345
Frame Relay (Q.922) 269
GARP (802 .1d) 76
ICMP (RFC 792, 950, 1256) 138
ICMPv6 (RFC 2463 et 3122) 196
IGMP (RFC 1112 et 2236) 384
Infrastructure d'un système de câblage 24
Interface DXI (ATM Forum) 286
IntServ (RFC 1633) 368
Inverse ARP (RFC 1293) 276
IPsec (RFC 2401, 2402, 2406) 479
IPv4 (RFC 791) 128
IPv6 (RFC 2460) 191
Locaux techniques 25
MIB (RFC 1212, 1213, 1155 et 2863) 147
MOSPF (RFC 1584) 393
MPEG (ISO CD 11172-2) 336
MPLS (RFC 3031 et 3032) 322
Notification de congestion (RFC 3168) 349
Options DHCP (RFC 2132) 150
OSPF (RFC 2328) 314
Paire torsadée 30
Partage de données (T.120) 438
PIM-SM (RFC 2362) 400
Pourquoi un plan d'adressage ? 112
PPP (RFC 1661, 1662) 222
Proxy ARP (RFC 1027) 225
Qu'est ce qu'un commutateur ? 54
Qu'est ce qu'un routeur ? 114
Qu'est ce qu'un segment ? 11
Qu'est ce que la signalisation ? 270
Qu'est ce que le spanning tree ? 69
Qu'est ce que SNMP ? 11
Qu'est-ce qu'un pont ? 217
Qu'est-ce qu'un protocole de routage ? 312
Qu'est-ce qu'un réseau local ? 4
Qu'est-ce qu'une carte Ethernet ? 8
Qualité de service Ethernet (802.1p) 361
Quels cordons de brassage ? 57
Réseaux Ethernet (802.3) 52
Réseaux locaux, étendus et intersites 234
Réseaux sans fil 85
RNIS (ITU série I) 228
RSVP (RFC 2205 à 2210) 373
RTCP (RFC 1889 et 1890) 434
RTP (RFC 1889 et 1890) 433
SDH (G.709) 249
Signalisation ATM (ATM Forum) 284
Signalisation d'appel (Q.931) 414
Signalisation Frame Relay (Q.933) 273
Signalisation ILMI 4.0 (ATM Forum) 291
Signalisation RAS (H.225) 422
SNMP v1 (RFC 1157, 2571, 2572) 144
Spanning tree (802.1d) 72
SSL/TLS (RFC 2246, 3546) 472
TCP (RFC 793) 132
Translations d'adresses et de ports 459
UDP (RFC 768) 131
VLAN (802 .1q) 303

VRRP (RFC 2338) 309 WLAN (802.11) 91