Formation
Notre expertise est votre avenir

TCP/IP

TCPIP-ESS – Rév 2

Sommaire

I. INTRODUCTION A TCP/IP ____ *7*

I.1. Suite de protocoles TCP/IP ____ **8**

I.2. Historique ____ **8**

I.2.1 Modèle OSI et modèle TCP/IP ____ 9
I.2.2 Couche physique (Physical layer) ____ 9
I.2.3 Couche liaison de données (Data link layer) ____ 9
I.2.4 Couche réseau (Network layer) ____ 10
I.2.4.1 Protocole IP ____ 10
I.2.4.2 Protocole ICMP ____ 11
I.2.4.3 Protocole IGMP ____ 13
I.2.4.4 Protocole ARP ____ 13
I.2.4.5 Cache ARP ____ 13
I.2.5 Couche transport (Transport layer) ____ 14
I.2.5.1 Protocole TCP ____ 14
I.2.5.2 Session TCP ____ 14
I.2.5.3 Alternatives à TCP ____ 15
I.2.5.4 Protocole UDP ____ 15
I.2.5.5 Protocole SCTP ____ 15
I.2.5.6 Protocole MPTCP ____ 16
I.2.6 Couche application (Application layer) ____ 17

I.3. Concepts fondamentaux d'interconnexion ____ **18**

I.3.1 Typologie des réseaux ____ 18
I.3.1.1 Body Area Network (BAN) et BSN (Body Sensor Network) ____ 18
I.3.1.2 Building Area Network (BAN) ____ 19
I.3.1.3 Personal Area Network (PAN) et Home Area Network (HAN) ____ 19
I.3.1.4 Campus Area Network (CAN) ____ 19
I.3.1.5 Metropolitan Area Network (MAN) ____ 19
I.3.1.6 Wide Area Network (WAN) ____ 19
I.3.2 Principaux équipements et leurs rôles ____ 20
I.3.2.1 Concentrateur (hub) ____ 20
I.3.2.2 Domaine de diffusion (broadcast domain) ____ 20
I.3.2.3 Domaine de collision (collision domain) ____ 21
I.3.2.4 Protocoles d'accès à un média ____ 21
I.3.2.5 CSMA/CD (Carrier Sense Multiple Access/Collision Detection) ____ 22
I.3.2.6 CSMA/CA (Collision Avoidance) ____ 23
I.3.2.7 CSMA/CR (Collision Resolution) ____ 23
I.3.2.8 Commutateur (switch) niveau 2 ____ 24
I.3.2.9 Commutateur (switch) niveau 3 ____ 25
I.3.2.10 Routeur (passerelle ou gateway) ____ 25

I.4. Normes ____ **26**

I.4.1 Organismes ____ 26
I.4.2 Request for comments (RFC) ____ 28

II. ADRESSES IP ____ *31*

II.1. Adressage IP classful ____ **32**

II.1.1 Identificateur réseau et identificateur d'hôte ____ 32
II.1.2 Classes et plage d'identificateurs de réseau ____ 32
II.1.2.1 Classe A ____ 33
II.1.2.2 Classe B ____ 33

II.1.2.3 Classe C 34
II.1.3 Directives d'adressage 34
II.1.3.1 Adresse de loopback (rebouclage) 34
II.1.3.2 Adresse de broadcast local 34
II.1.3.3 Adresse de broadcast du réseau 34
II.1.3.4 Adresse de réseau 34
II.1.3.5 Adresse IP d'hôte 35
II.1.3.6 Masques de sous-réseau 35
II.1.3.7 Adresses IP privées 39
II.1.3.8 Résolution d'une adresse IP en une adresse MAC 40
II.1.3.9 Résolution d'une adresse IP locale 40
II.1.3.10 Résolution d'une adresse MAC en une adresse IP 41

II.2. Adressage IP classless 43
II.2.1 Abolition des classes 43
II.2.2 CIDR (Classless Inter-Domain Routing) 43
II.2.3 Subnetting 44
II.2.4 VLSM (Variable Length Subnet Mask) 46
II.2.5 Supernetting 48

III. INTERCONNEXION DU RESEAU IP 49

III.1. Routage IP 50
III.1.1 Principe du routage 50
III.1.2 Routage IP classful 51
III.1.3 Routage IP classless 51
III.1.4 Catégories de routage 52

III.2. Protocoles IGP et EGP 53
III.2.1 Définition 53
III.2.2 Classes de protocoles IGP 54
III.2.3 Protocoles IGP classful 54
III.2.4 Protocoles IGP classless 55

III.3. Routage inter-LAN 56
III.3.1 VLAN (Virtual LAN) 56
III.3.2 Configuration du routage InterVLAN et de la jonction ISL/802.1Q sur un switch L2 57
III.3.3 Configuration du routage InterVLAN sur un switch L3 59

III.4. Routage inter-LAN 60
III.4.1 Routage statique 60
III.4.2 Routage dynamique 61
III.4.3 RIP 63
III.4.4 OSPF 64

III.5. Network Address Translation 66
III.5.1 Principe 66
III.5.2 Les termes Cisco 67
III.5.3 NAT statique 68
III.5.4 NAT dynamique 69
III.5.5 PAT 70

IV. DEPANNAGE 71

IV.1. Problème au niveau de la couche physique 72
IV.1.1 Symptômes des problèmes 72
IV.1.2 Causes des problèmes 74
IV.1.3 Dépannage de la couche physique (couche 1) 74

IV.2. Problème au niveau de la couche liaison 75
IV.2.1 Symptômes des problèmes 75
IV.2.2 Causes des problèmes 75
IV.2.3 Dépannage de la couche liaison (couche 2) 76

IV.3. Problème au niveau de la couche réseau ____ 76
IV.3.1 Configuration IP dynamique ____ 77
IV.3.2 Configuration IP statique ____ 77
IV.3.3 Dépannage de la couche réseau (couche 3) ____ 77

IV.4. Capturer et analyser un trafic réseau avec Wireshark ____ 78
IV.4.1 C'est quoi une capture réseau ? ____ 78
IV.4.2 Une capture réseau pour quoi faire ? ____ 78
IV.4.3 Où effectuer la capture réseau ? ____ 78
IV.4.4 Comment faire la capture réseau ? ____ 79
IV.4.5 Comment analyser la capture réseau ? ____ 80
IV.4.6 Appliquer un filtre ____ 81

V. TCP/IP : COUCHE TRANSPORT ____ 82

V.1. TCP et UDP ____ 83
V.1.1 Introduction ____ 83

V.2. TCP (Transmission Control Protocol) ____ 84
V.2.1 Introduction ____ 84
V.2.2 Établissement d'une connexion ____ 84
V.2.3 Structure d'un segment TCP ____ 85
V.2.4 Transferts de données ____ 87
V.2.5 Numéros de séquence ____ 88
V.2.6 Somme de contrôle ____ 88
V.2.7 Temporisation ____ 89
V.2.8 Contrôle de flux ____ 90
V.2.9 Contrôle de congestion ____ 90
V.2.10 Pour conclure ____ 90
V.2.11 Terminaison d'une connexion ____ 91
V.2.12 Ports TCP ____ 91

V.3. UDP (User Datagram Protocol) ____ 92
V.3.1 Introduction ____ 92

V.4. Structure d'un datagramme UDP ____ 93
V.4.1 Ports UDP ____ 93

V.5. Ports TCP et UDP ____ 94
V.5.1 Liste des ports ____ 94

VI. TCP/IP : APPLICATIONS ____ 105

VI.1. Adressage IP dynamique ____ 106
VI.1.1 Introduction ____ 106
VI.1.2 Fonctionnement ____ 107
VI.1.3 Compatibilité ____ 107
VI.1.4 Renouvellement du bail ____ 108
VI.1.5 Client et serveur sur des segments différents ____ 108
VI.1.6 Configuration du serveur DHCP ____ 109

VI.2. Résoudre les noms d'hôtes ____ 109
VI.2.1 Introduction ____ 110
VI.2.2 Un système hiérarchique et distribué ____ 110
VI.2.2.1 Hiérarchie du DNS ____ 110
VI.2.2.2 Résolution du nom par un hôte ____ 111
VI.2.2.3 Résolution inverse ____ 112
VI.2.2.4 Résolution inverse CIDR ____ 113
VI.2.3 Serveurs DNS racine ____ 114
VI.2.4 Fully Qualified Domain Name ____ 114
VI.2.5 Nom de domaine internationalisé ____ 114
VI.2.6 Les techniques du DNS Round-Robin pour la distribution de la charge ____ 114
VI.2.7 Principaux enregistrements DNS ____ 115

VI.2.7.1 NS record 116
VI.2.7.2 PTR record 116
VI.2.7.3 MX record 117
VI.2.7.4 CNAME record 118
VI.2.7.5 NAPTR record 119
VI.2.7.6 SOA record 120
VI.2.7.7 Time to live 120
VI.2.7.8 Glue records 121
VI.2.8 Mise à jour dynamique 121
VI.2.8.1 Considérations opérationnelles 121
VI.2.8.2 Mise à jour du DNS 121
VI.2.8.3 Cohérence du DNS 122
VI.2.8.4 Robustesse du DNS 122
VI.2.9 Sécurité du DNS 123
VI.2.9.1 Interception des paquets 123
VI.2.9.2 Fabrication d'une réponse 123
VI.2.9.3 Corruption des données 123
VI.2.9.4 Empoisonnement du cache DNS 123
VI.2.9.5 Déni de service 123
VI.2.10 DNSSEC 124
VI.2.11 Détails du protocole 124
VI.2.12 Outils de diagnostics 125
VI.3. Administration à distance avec le service SSH 126
VI.3.1 Introduction 126
VI.3.2 Protocole SSH 126
VI.3.3 SSH avec authentification par clés 127
VI.3.4 Implémentations logicielles 127
VI.4. Transfert ou copie de fichiers 128
VI.4.1 Introduction 128
VI.4.2 FTP 128
VI.4.2.1 Utilisation 129
VI.4.2.2 Mise en œuvre 129
VI.4.2.3 Interopérabilité 130
VI.4.2.4 Le protocole 130
VI.4.3 TFTP 131
VI.4.4 SFTP 132
VI.4.5 SCP 132
VI.5. Transférer des pages hypertextes 133
VI.5.1 Protocole HTTP 133
VI.5.2 Serveur HTTP 134
VI.5.3 Navigateur Web 134
VI.6. Étudier les protocoles associés à la messagerie 135
VI.6.1 Introduction 135
VI.6.2 Envoi 135
VI.6.3 Livraison 136
VI.6.3.1 Protocole POP 136
VI.6.3.2 Protocole IMAP 136
VI.7. Administration des réseaux IP 137
VI.7.1 Introduction 137
VI.7.2 SNMP 138
VI.7.3 MIB 138
VI.7.4 Logiciels 139
VII. VERS IPV6 140
VII.1. IPv6 141
VII.1.1 Introduction 141

VII.1.2 Fonctionnement d'IPv6 141
VII.1.2.1 Adresse IPv6 142
VII.1.2.2 Structure de l'adresse IPv6 unicast globale 143
VII.1.2.3 Scope 144
VII.1.2.4 Indice de zone 144
VII.1.3 Attribution des blocs d'adresses IPv6 144
VII.1.4 En-tête IPv6 145
VII.1.4.1 Fragmentation et option jumbo 146
VII.1.4.2 En-têtes d'extension 146
VII.1.5 Neighbor Discovery Protocol 147
VII.1.6 Attribution des adresses IPv6 147
VII.1.6.1 Configuration manuelle 147
VII.1.6.2 Configuration automatique 147
VII.1.7 Multicast 148
VII.1.8 DNS 149
VII.1.9 Traduction d'adresse 149
VII.1.10 IPv6 et mobilité 149
VII.1.11 Technologies de transition pour l'accès à l'Internet IPv6 149
VII.1.12 Tunnels statiques 150
VII.1.13 Tunnels automatiques 150
VII.1.14 Passerelles applicatives 150
VII.1.15 Multihoming 151

I. INTRODUCTION A TCP/IP

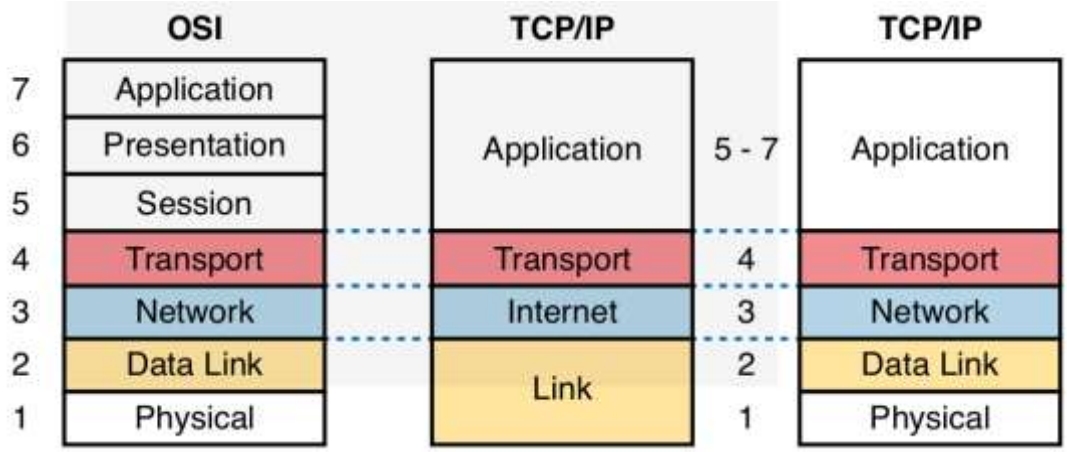
Objectifs

- Présenter les différents protocoles TCP/IP et leurs rôles respectifs.

Références

- https://www.iso.org
- https://www.ietf.org
- https://www.Internetsociety.org
- https://www.icann.org
- https://www.ietf.org/standards/rfcs/

I.1. SUITE DE PROTOCOLES TCP/IP



I.2. HISTORIQUE

L'origine de TCP/IP est intimement liée à celle d'Internet. Tout a commencé par le projet du réseau ARPAnet (acronyme anglais de « Advanced Research Projects Agency Network ») qui est initié en 1966. Il faut attendre 1969 pour qu'il voie le jour. C'est le premier réseau à transfert de paquets développé aux USA par une agence du département de la Défense appelée DARPA (Defense Advanced Research Projects Agency). Elle est chargée de rechercher et de développer de nouvelles technologies destinées à un usage militaire. La première démonstration officielle s'est déroulée en octobre 1972.

ARPAnet a utilisé au début le protocole NCP (Network Control Program). Au début des années 80, il migre sur TCP/IP (Transmission Control Protocol / Internet Protocol). Un second réseau, qui utilise les lignes téléphoniques, est connecté à ARPAnet. Il s'appelle « Internet ».

I.2.1 Modèle OSI et modèle TCP/IP

TCP/IP n'est pas un protocole monolithique mais un ensemble de protocoles. Nous parlons ainsi d'une suite de protocoles ou bien d'une pile de protocoles. Il s'agit de protocoles de communication et d'application pour connecter des systèmes hétérogènes, indépendamment de la couche physique.

Le modèle OSI (Open Systems Interconnection) est un standard de communication pour les réseaux informatiques proposé par l'ISO (International Organization for Standardization). Il décrit en 7 couches les fonctionnalités nécessaires à la communication et l'organisation de ces fonctions.

La suite de protocoles Internet a son propre modèle de communication en 4 couches. Le modèle TCP/IP (appelé aussi modèle Internet ou DARPA) date de 1976 et a été stabilisé bien avant la publication du modèle OSI en 1984. Ce modèle évoluera en 5 couches pour assurer une compatibilité avec les couches 1,2,3 et 4 du modèle OSI.

I.2.2 Couche physique (Physical layer)

La couche 1 représente la boucle locale c'est-à-dire ce qui relie un utilisateur d'un réseau au premier niveau d'équipement du réseau auquel il est connecté. Cet élément peut être physique comme une paire torsadée ou une fibre optique dans le cas du réseau téléphonique, ou immatériel comme une onde électromagnétique dans le cas d'une boucle locale radio ou d'un réseau de téléphonie mobile.

I.2.3 Couche liaison de données (Data link layer)

La couche 2 transfère des données entre les nœuds adjacents d'un réseau étendu (WAN) ou entre des nœuds sur le même segment d'un réseau local (LAN). La couche de liaison de données fournit les moyens fonctionnels et procéduraux pour le transfert de données entre des entités d'un réseau. Dans certains cas, elle fournit également les moyens de détecter et de corriger les erreurs qui peuvent survenir au niveau de la couche physique.

Ethernet pour les réseaux locaux (multi-nœuds), le protocole point à point (PPP), HDLC et ADCCP pour des connexions point à points (double nœud) sont des exemples de protocoles de liaison de données.

I.2.4 Couche réseau (Network layer)

La couche 3 IP construit une voie de communication de bout à bout à partir de voies de communication avec ses voisins directs. Ses apports fonctionnels principaux sont donc :

- Le *routage* : détermination d'un chemin permettant de relier les 2 machines distantes ;
- Le *relayage* qui est la retransmission d'un PDU (Protocol Data Unit ou Unité de données de protocole) dont la destination n'est pas locale pour le rapprocher de sa destination finale. Le PDU est souvent appelé « paquet ». La fonction de « relayage » est parfois appelée « acheminement ».

La couche réseau est donc la seule à être directement concernée par la topologie du réseau. C'est aussi la dernière couche supportée par toutes les machines du réseau pour le transport des données utilisateurs : les couches supérieures sont réalisées uniquement dans les machines d'extrémité.

I.2.4.1 Protocole IP

Lors d'une communication entre deux postes sur un réseau local (LAN), le flux de données provenant de la couche transport (couche 4) sont, par exemple, des segments TCP. Ils sont ensuite encapsulés dans des paquets par le protocole IP lors de leur passage au niveau de la couche réseau (couche 3). Ces paquets sont ensuite transmis à la couche de liaison de données (couche 2) afin d'y être encapsulés dans des trames Ethernet.

Le protocole IP (Internet Protocol) assure l'acheminement au mieux des paquets. Il ne se préoccupe pas du contenu des paquets mais fournit une méthode pour les mener à destination. Aucun chemin pour le transfert des données n'est établi à l'avance. Il est dit que le protocole est « non orienté connexion ».

I.2.4.2 Protocole ICMP

ICMP (Internet Control Message Protocol) est utilisé pour véhiculer des messages de contrôle et d'erreurs, par exemple lorsqu'un service ou un hôte est inaccessible.

Un paquet ICMP est encapsulé dans un datagramme IP :

Bit 0 - 7	Bit 8 - 15	Bit 16 - 23	Bit 24 - 31
Version/IHL	Type de service	Longueur totale	
Identification (fragmentation)		*flags* et *offset* (fragmentation)	
Durée de vie(TTL)	Protocole	Somme de contrôle de l'en-tête	
Adresse IP source			
Adresse IP destination			
Type de message	Code	Somme de contrôle	
Bourrage ou données			
Données (optionnel et de longueur variable)			

Il est composé :

- D'un en-tête IP (en bleu) avec Protocole valant 1 et Type de Service valant 0
- Du type de message ICMP (8 bits)
- Du code de l'erreur (8 bits)
- D'une somme de contrôle (16 bits), calculée sur la partie spécifique à ICMP (sans l'en-tête IP)
- D'une partie aménagée pour des données relatives aux différents types de réponses (32 bits)
- Du message

Les types et codes, codés respectivement sur 8 bits (soit 2 octets), représentent la définition de message d'erreur contenu. Voici la liste des principales combinaisons entre les champs Type et Code :

type=00 et code=00 : Réponse à une demande d'écho

type=03 et code=00 : Réseau inaccessible

type=03 et code=01 : Hôte inaccessible

type=03 et code=02 : Protocole inaccessible

type=03 et code=03 : Port inaccessible

type=03 et code=04 : Fragmentation nécessaire mais interdite

type=03 et code=05 : Echec de routage par la source

type=03 et code=06 : Réseau de destination inconnu

type=03 et code=07 : Hôte de destination inconnue

type=03 et code=08 : Machine source isolée

type=03 et code=09 : Réseau de destination interdit administrativement

type=03 et code=10 : Hôte de destination interdite administrativement

type=03 et code=11 : Réseau inaccessible pour ce type de service

type=03 et code=12 : Hôte inaccessible pour ce type de service

type=03 et code=13 : Communication interdite par un filtre

type=03 et code=14 : Host Precedence Violation

type=03 et code=15 : Precedence cutoff in efect

type=04 et code=00 : Volume de donnée trop importante

type=05 et code=00 : Redirection pour un hôte

type=05 et code=01 : Redirection pour un hôte et pour un service donné

type=05 et code=02 : Redirection pour un réseau

type=05 et code=03 : Redirection pour un réseau et pour un service donné

type=08 et code=00 : Demande d'écho

type=09 et code=00 : Avertissement routeur

type=10 et code=00 : Sollicitation routeur

type=11 et code=00 : Durée de vie écoulée avant d'arrivée à destination

type=11 et code=01 : Temps limite de réassemblage du fragment dépassé

type=12 et code=00 : Entête IP invalide

type=12 et code=01 : Manque d'une option obligatoire

type=12 et code=02 : Mauvaise longueur

type=13 et code=00 : Requête pour un marqueur temporel

type=14 et code=00 : Réponse pour un marqueur temporel

type=15 et code=00 : Demande d'adresse réseau

type=16 et code=00 : Réponse d'adresse réseau

type=17 et code=00 : Demande de masque de sous réseau

type=18 et code=00 : Réponse de masque de sous réseau

I.2.4.3 Protocole IGMP

IGMP (Internet Group Management Protocol) permet de gérer le trafic entre les appareils faisant partie d'un groupe de multidiffusion (multicast) et les routeurs de multidiffusion.

Plusieurs versions d'IGMP existent et diffèrent en fonctionnalité.

Ce protocole sera abordé dans d'autres cours.

I.2.4.4 Protocole ARP

ARP (Address Resolution Protocol) est utilisé pour traduire une adresse de protocole de couche réseau (adresse IPv4), en une adresse de protocole de couche de liaison (adresse MAC). Il se situe à l'interface entre la couche réseau (couche 3) et la couche de liaison (couche 2).

Le protocole ARP est nécessaire au fonctionnement d'IPv4 utilisé au-dessus d'un réseau de type Ethernet.

Un ordinateur connecté à un réseau informatique souhaite émettre une trame Ethernet à destination d'un autre ordinateur dont il connaît l'adresse IP et placé dans le même sous-réseau. Dans ce cas, cet ordinateur va placer son émission en attente et effectuer une requête ARP en broadcast de niveau 2. Cette requête est de type « quelle est l'adresse MAC correspondant à l'adresse IP ? Répondez à mon adresse IP ».

Puisqu'il s'agit d'une diffusion (broadcast), tous les ordinateurs du segment vont recevoir la requête. En observant son contenu, ils pourront déterminer quelle est l'adresse IP sur laquelle porte la recherche. La machine qui possède cette adresse IP sera la seule à répondre en envoyant à la machine émettrice une réponse ARP du type « je suis adresse IP et adresse MAC ».

I.2.4.5 Cache ARP

Pour émettre cette réponse au bon ordinateur, il crée une entrée dans son cache ARP à partir des données contenues dans la requête ARP qu'il vient de recevoir.

La machine à l'origine de la requête ARP reçoit la réponse, met à jour son cache ARP et peut donc envoyer à l'ordinateur concerné le message qu'elle avait mis en attente.

Il suffit donc d'un broadcast et d'un unicast pour créer une entrée dans le cache ARP de deux ordinateurs.

I.2.5 Couche transport (Transport layer)

I.2.5.1 Protocole TCP

TCP (Transmission Control Protocol) est un protocole de transport fiable en mode connecté. Il découpe le flux d'octets en segments dont la taille dépend de la MTU (Maximum Transmission Unit) de la couche 2 (liaison de données).

Le MTU est la taille maximale d'un paquet pouvant être transmis en une seule fois, c'est-à-dire sans fragmentation, sur une interface.

Le path MTU désigne la taille maximale entre une machine source et une machine destination. Il est égal au plus petit MTU des interfaces via lesquelles le paquet est transmis.

La valeur de MTU peut varier selon le type de réseau. Pour IPv4, sa taille minimale est de 68 octets.

La taille par défaut sur un réseau Ethernet est de 1500 octets.

TCP utilise le numéro de port pour identifier les applications. À chaque extrémité (client/serveur) de la connexion TCP est associé un numéro de port sur 16 bits (1 à 65535) assigné à l'application émettrice ou réceptrice.

I.2.5.2 Session TCP

Une session TCP fonctionne en 3 phases :

- L'établissement de la connexion

Il s'effectue par un handshaking en 3 temps :

  - Le client envoie un segment SYN au serveur,
  - Le serveur lui répond par un segment SYN/ACK,
  - Le client confirme par un segment ACK.

- Les transferts de données

Des paramètres comme le numéro de séquence sont initialisés pour assurer la transmission fiable des données afin d'ordonner les segments TCP reçus et de détecter les données perdues, les sommes de contrôle permettent la détection d'erreurs et les acquittements ainsi que les temporisations permettent la détection des segments perdus ou retardés.

- La fin de la connexion

La rupture de connexion, elle, utilise un handshaking en 4 temps. Chaque extrémité de la connexion effectue sa terminaison de manière indépendante. Ainsi, la fin d'une connexion nécessite une paire de segments FIN et ACK pour chaque extrémité.

I.2.5.3 Alternatives à TCP

Certaines applications en temps réel (Real Time) comme les jeux multi-joueurs, la diffusion multimédia, les échanges de fichiers en autres n'ont pas besoin et peuvent même souffrir des mécanismes complexes de transport fiable de TCP.

Dans ce type d'applications, il est souvent préférable de gérer les pertes, erreurs ou congestions, plutôt que d'essayer de les éviter. Pour ces besoins particuliers, d'autres protocoles de transport ont été créés et déployés : UDP, SCTP, MPTCP.

I.2.5.4 Protocole UDP

UDP (User Datagram Protocol) permet la transmission de données sous forme de datagrammes de manière très simple entre deux entités. Chacune est définie par une adresse IP et un numéro de port UDP sur 16 bits (de 1 à 65535).

Aucune communication préalable n'est requise pour établir la connexion, au contraire de TCP qui utilise le procédé de handshaking. UDP utilise un mode de transmission sans connexion.

Une liste de propriétés rend UDP particulièrement adapté à certaines applications.

Il est orienté transaction et donc adapté aux protocoles simples de type requête-réponse tels le DNS (Domain Name System) ou le NTP (Network Time Protocol).

Il fournit des datagrammes utiles pour modéliser d'autres protocoles tel que le tunneling IP ou le RPC (Remote Procedure Call) ainsi que le NFS (Network File System).

Il est très simple, ce qui le rend adapté pour le bootstrapping ou d'autres usages sans pile de protocole complète tels DHCP et le protocole simplifié de transferts de fichiers TFTP (Trivial File Transfer Protocol).

Il est dit sans état, ce qui est utile dans des cas où de nombreux clients sont présents comme les applications de streaming (télévision IP, par exemple).

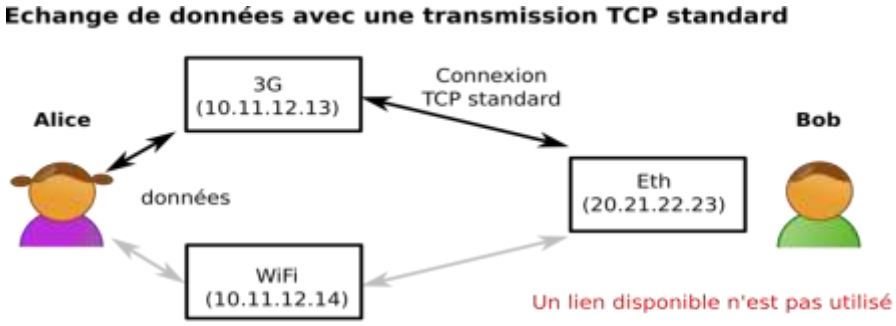
L'absence de délai de retransmission en fait un protocole utile pour les applications en temps réel. Quelques exemples de ces applications sont la VoIP (voix sur IP), les jeux en ligne, et de nombreux protocoles construits sur base du RTSP (Real Time Streaming Protocol).

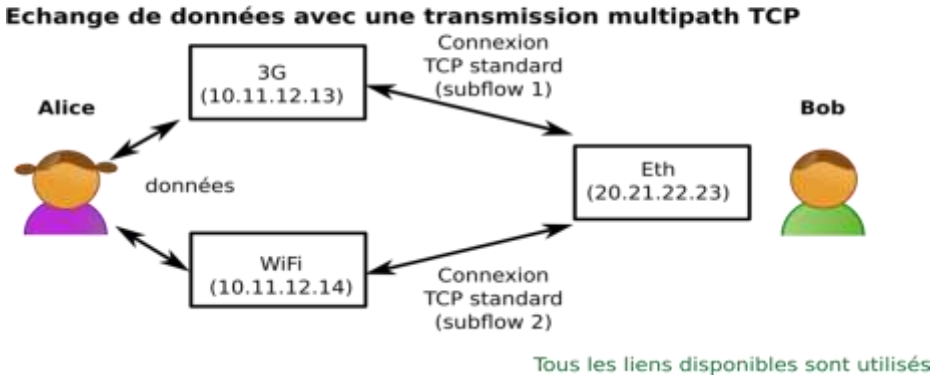
I.2.5.5 Protocole SCTP

Alors que TCP est orienté flux, SCTP (Stream Control Transfert Protocol) est, comme UDP, orienté message. Il gère plusieurs flux au sein d'une communication. Il sait aussi répartir des flux entre plusieurs adresses IP cibles afin de faire de la répartition de charge.

I.2.5.6 Protocole MPTCP

MPTCP (Multipath TCP) permet d'utiliser plusieurs chemins d'accès pour maximiser l'utilisation des ressources et l'augmentation de la redondance tout en restant compatible avec les équipements actuels tels que les firewalls, les routeurs NAT (Network Address Translation).





I.2.6 Couche application (Application layer)

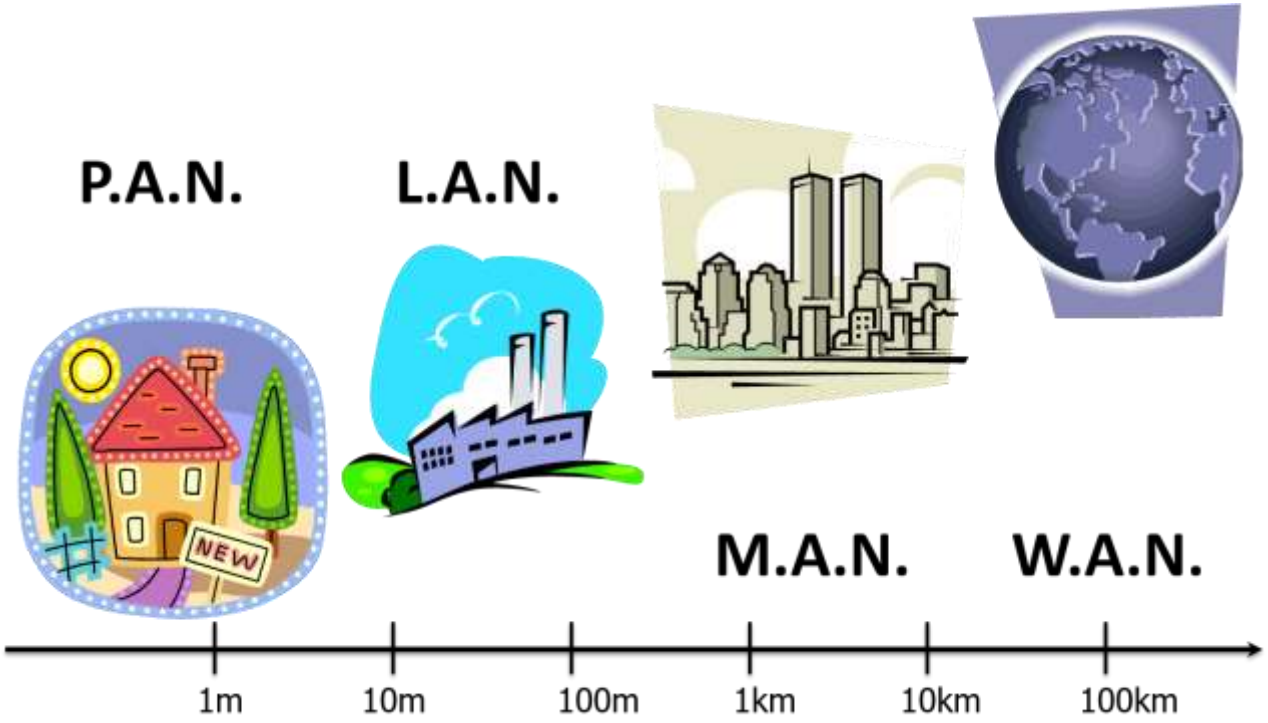
La couche 5 application est le point d'accès aux services réseaux. Les modèles de communication TCP/IP et OSI n'ont pas pour rôle de spécifier les applications, ils ne spécifient pas de service à ce niveau.

La couche d'application représente des données pour l'utilisateur ainsi que du codage et un contrôle du dialogue.

En se limitant au monde TCP/IP, voici quelques exemples de protocoles :

Protocoles		Description
FTP	File Transfer Protocol	Protocoles orientés transfert de fichiers
NFS	Network File System	
SMB/CIFS	Server Message Block Common Internet File System	
SMTP	Simple Mail Transfer Protocol	Protocoles orientés messageries
POP	Post Office Protocol	
IMAP	Internet Message Access Protocol	
NNTP	Network News Transfer Protocol	Protocole Usenet (news)
TELNET	terminal network ou telecommunication network, ou encore teletype network)	Protocoles de type « session distance »
SSH	Secure Shell	
HTTP	Hypertext Transfer Protocol	Transfert de pages HTML
DNS	Domain Name System	Traduction des noms de domaine Internet en adresse IP ou autres enregistrements
LDAP	Lightweight Directory Access Protocol	Interrogation et modification des services d'annuaire
SNMP	Simple Network Management Protocol	Supervision et de diagnostic des problèmes réseaux et matériels à distance
NTP	Network Time Protocol	Synchronisation de l'heure

I.3. CONCEPTS FONDAMENTAUX D'INTERCONNEXION



I.3.1 Typologie des réseaux

La classification des réseaux peut s'effectuer selon l'étendue ou la distance couverte.

I.3.1.1 Body Area Network (BAN) et BSN (Body Sensor Network)

Le BAN ou le réseau corporel est une technologie de réseau sans fil défini par IEEE 802.15.6 qui est basée sur les radiofréquences. Elle consiste à interconnecter sur, autour ou dans le corps humain de minuscules dispositifs pouvant effectuer des mesures avec des capteurs ou agir de façon active avec des actionneurs.

Ces capteurs très miniaturisés, disposant d'une grande autonomie et utilisant des courants de très faible puissance, peuvent être capables de dialoguer avec un centre de service distant, pour alerter un service d'urgences hospitalières par exemple.

Les principales applications se trouvent dans les domaines de la santé, des premiers secours, du militaire, du sport, de l'Intelligence ambiante, etc.

L'architecture de communication d'un réseau BAN peut se décliner en 3 niveaux :

1. Le premier niveau représente les communications Intra-BAN. Les communications Intra-BAN font référence aux échanges radio qui ont lieu à proximité immédiate du corps humain. On distinguera les communications entre les capteurs et les communications des capteurs vers un PDA (assistant personnel).

2. Le deuxième niveau représente les communications Inter-BAN. Un réseau BAN fonctionne rarement de manière autonome. L'Inter-BAN recouvre les communications entre le PDA et un point d'accès au réseau.
3. Le troisième niveau représente les communications hors BAN. C'est à ce niveau que le service est fourni. Dans le domaine de la santé, ce niveau représentera le service de télémédecine d'un hôpital.

Cette technologie a émergé début des années 2000. Le terme BAN peut être ambigu avec le Building Area Network.

I.3.1.2 Building Area Network (BAN)

Le BAN est un réseau local dont l'envergure couvre un bâtiment.

I.3.1.3 Personal Area Network (PAN) et Home Area Network (HAN)

Le PAN définit le réseau d'une personne avec son ordinateur, PDA, Smartphone, tablette, imprimantes... Le HAN détermine un réseau domestique d'une ou plusieurs personnes.

I.3.1.4 Campus Area Network (CAN)

Le CAN désigne l'espace rassemblant les bâtiments et l'infrastructure d'une université, d'une entreprise...

I.3.1.5 Metropolitan Area Network (MAN)

Le MAN désigne un réseau qui interconnecte un ensemble de réseaux LAN dans l'agglomération d'une ville. Le réseau utilise généralement des fibres optiques pour relier les différents sites.

I.3.1.6 Wide Area Network (WAN)

Un réseau étendu, souvent désigné par son acronyme anglais WAN, est un réseau informatique ou de télécommunications qui couvre une grande zone géographique, typiquement à l'échelle d'un pays, d'un continent, ou de la planète. Le plus grand WAN est le réseau Internet.

I.3.2 Principaux équipements et leurs rôles

Les principaux équipements réseaux, que l'on retrouve dans un LAN, sont :

- Les commutateurs (switches),
- Les routeurs,
- Les concentrateurs (hubs).

Application	
Transport	
Réseau	routeur, switch
Liaison de données	switch
Physique	hub

I.3.2.1 Concentrateur (hub)

Un hub est un appareil informatique permettant de concentrer les transmissions Ethernet de plusieurs équipements sur un même support dans un LAN.

Chaque équipement attaché au hub partage le même domaine de diffusion ainsi que le même domaine de collision.

I.3.2.2 Domaine de diffusion (broadcast domain)

C'est une aire logique d'un réseau informatique ou n'importe quel équipement connecté à celui-ci peut directement transmettre à tous les autres machines de ce domaine, sans devoir passer par un routeur ou une passerelle.

I.3.2.3 Domaine de collision (collision domain)

C'est une zone logique d'un réseau informatique où les paquets de données peuvent entrer en collision entre eux, en particulier avec le protocole de communication Ethernet. Il peut être un seul segment de câble Ethernet, un seul concentrateur ou même un réseau complet de concentrateurs.

Lorsque l'Ethernet est utilisé en mode full-duplex, il n'y a plus de domaine de collision, car aucune collision n'est possible.

Comme dans tout segment de réseau Ethernet, une seule des machines connectées peut y transmettre à la fois. Dans le cas contraire, une collision se produit, les machines concernées doivent retransmettre leurs trames après avoir attendu un temps calculé aléatoirement par chaque émetteur.

I.3.2.4 Protocoles d'accès à un média

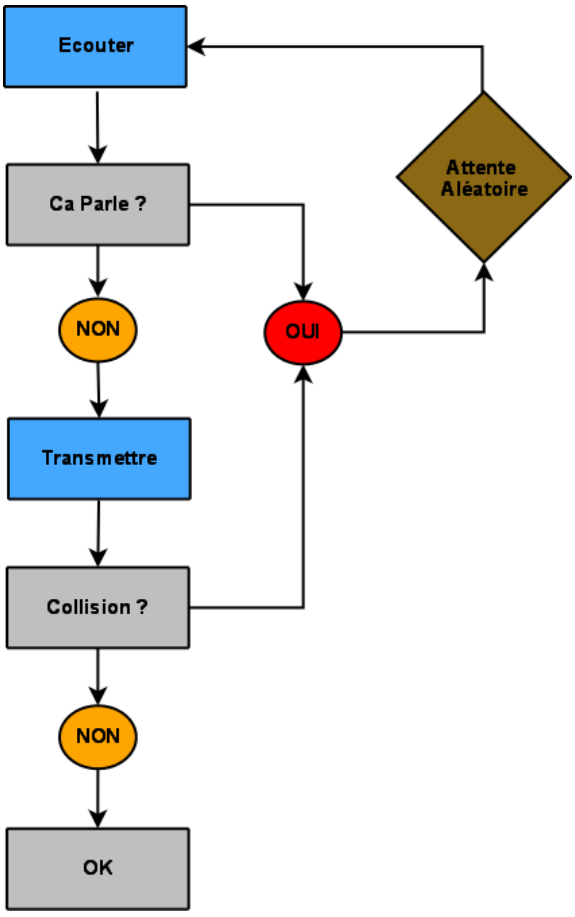
CSMA (Carrier Sense Multiple Access) représente un ensemble de protocoles d'accès à un média, lesquels vérifient si le support est disponible avant de commencer l'envoi d'une trame. Ils permettent aussi de détecter ou bien éviter les collisions de messages dans les transmissions.

Il existe 3 méthodes principales employées dans les réseaux :

- CSMA/CD : Collision Detection
- CSMA/CA : Collision Avoidance
- CSMA/CR : Collision Resolution

  (Aussi appelé CSMA/BA pour "Bitwise Arbitration" ou CSMA / AMP pour "Arbitration on Message Priority")

I.3.2.5 CSMA/CD (Carrier Sense Multiple Access/Collision Detection)



CSMA/CD est un protocole qui gère le partage de l'accès physique au réseau Ethernet, selon la norme IEEE 802.3. Cette méthode permet à une station d'écouter le support physique de liaison pour déterminer si une autre station transmet une trame de données.

Si tel n'est pas le cas, elle suppose qu'elle peut émettre. Si la distance entre les périphériques est telle que la latence des signaux d'un périphérique implique que les signaux ne sont pas détectés par un deuxième périphérique, ce dernier peut, lui aussi, commencer à transmettre son message. Les supports disposent désormais de deux périphériques qui transmettent simultanément des signaux.

Les messages sont propagés sur les supports jusqu'à ce qu'ils se rencontrent, à ce stade les signaux se mélangent et les messages sont détruits: une collision se produit. Bien que les messages soient endommagés, les signaux restants désordonnés continuent à se propager sur les supports.

L'accès multiple implique que plusieurs stations peuvent émettre au même moment ce qui provoque une collision (donc une perte de données). Comme les stations écoutent aussi les collisions elles savent qu'elles doivent réémettre après avoir attendu pendant un délai aléatoire.

Ce type de protocole est dit « probabiliste », c'est-à-dire qu'il n'est pas possible de déterminer avec certitude le délai d'envoi d'un message. Rappelons que dans un réseau Ethernet les stations se partagent le même média de communication, qu'il n'y a pas de jeton ni de priorité d'émission.

I.3.2.6 CSMA/CA (Collision Avoidance)

La méthode CSMA/CA s'utilise dans les réseaux sans-fil (802.11). Contrairement aux réseaux filaires, deux stations peuvent émettre vers une troisième sans se détecter. Par exemple, la première étant hors de portée de la seconde.

Pour éviter cela, une station est considérée comme le maître des transmissions qui autorise une station à communiquer lorsque celle-ci le demande. Pour cela, la station doit émettre une courte trame RTS (Ready To Send) contenant quelques informations sur la communication (débit, longueur de la trame...).

Si la station maîtresse accepte cette communication, elle renvoie alors une trame CTS (Clear To Send) et la station peut transmettre son message. En revanche, si la station ne reçoit pas de message elle doit attendre à nouveau avant de redemander une autorisation d'émettre.

C'est la méthode utilisée dans les réseaux Wi-Fi (802.11) et la station maîtresse est généralement le point d'accès (AP - Access Point).

I.3.2.7 CSMA/CR (Collision Resolution)

Cette méthode est légèrement plus évoluée que la méthode CSMA/CD.

Si plusieurs stations transmettent un message, elles appliquent un ET logique entre le signal reçu et le signal émis. Dans le cas d'une inégalité, la station s'arrête de transmettre. Comme le 0 est une valeur dominante, elle écrase donc le 1 (état récessif). Cela signifie que la communication de l'une des stations n'est pas modifiée et permet ainsi de terminer cette communication sans délai d'attente ou de retransmission.

Un réseau utilisant cette méthode peut alors être déterministe. C'est la méthode employée dans les réseaux CAN (Campus Area Network).

I.3.2.8 Commutateur (switch) niveau 2

Un commutateur est un équipement qui relie plusieurs segments (câbles ou fibres) dans un réseau informatique et de télécommunication. Ce qui permet de créer des circuits virtuels.

Dans les réseaux locaux (LAN), il s'agit le plus souvent d'un boîtier disposant de plusieurs ports RJ45 (entre 4 et plusieurs centaines), il a donc la même apparence qu'un concentrateur (hub).

La commutation est un des deux modes de transport de trame au sein des réseaux informatiques et de communication, l'autre étant le routage.

Contrairement à un concentrateur, un commutateur ne reproduit pas sur tous les ports chaque trame qu'il reçoit : il sait déterminer sur quel port il doit envoyer une trame, en fonction de l'adresse de destination de cette trame. Les commutateurs sont souvent utilisés pour remplacer des concentrateurs car ils encombrent moins le réseau.

Dans le cas d'un réseau IP/Ethernet, un commutateur, étant en couche 2 du modèle TCP/IP, utilise les adresses MAC (Media Access Control) pour diriger les données. Il établit et met à jour dynamiquement une table de commutation qui associe des adresses MAC avec les ports correspondants.

Lorsqu'il reçoit une trame destinée à une adresse présente dans cette table, le commutateur renvoie la trame sur le port correspondant. Si le port de destination est le même que celui de l'émetteur, la trame n'est pas transmise. Si l'adresse du destinataire est inconnue dans la table, alors la trame est traitée comme un broadcast, c'est-à-dire qu'elle est transmise à tous les ports du commutateur à l'exception du port de réception.

Un commutateur de niveau 2 est similaire à un concentrateur dans le sens où il fournit un seul domaine de diffusion. En revanche, chaque port a son propre domaine de collision. Le commutateur utilise la micro-segmentation pour diviser les domaines de collision, un par segment connecté. Ainsi, seules les interfaces réseau directement connectées par un lien point à point sollicitent le medium. Si le commutateur auquel il est connecté prend en charge le full-duplex, le domaine de collision est éliminé.

I.3.2.9 Commutateur (switch) niveau 3

Les switches de niveau 2 ou 3 fournissent les mêmes fonctionnalités en couche 2. Ils permettent, aussi, de segmenter également le réseau en VLAN (Virtual LAN).

Les VLAN présentent les intérêts suivants :

- Améliorer la gestion du réseau.
- Optimiser la bande passante.
- Séparer les flux.
- Segmentation : réduire la taille d'un domaine de broadcast.
- Sécurité : permet de créer un ensemble logique isolé des uns et des autres pour améliorer la sécurité.

Le seul moyen pour communiquer entre des machines appartenant à des VLAN différents est alors de passer par un routeur ou par un switch niveau 3 qui a la fonction de routage.

I.3.2.10 Routeur (passerelle ou gateway)

Le rôle d'un routeur est d'interconnecter 2 réseaux.

Les routes sont définies manuellement (routes statiques) ou bien les routes sont élaborées par des protocoles de routage dynamiques tels que RIP, EIGRP, OSPF...

La fonction de routage traite les adresses IP et les dirige selon l'algorithme de routage. Sa table associée, contient la correspondance des adresses réseau avec les interfaces physiques du routeur où sont connectés les autres réseaux.

I.4. NORMES

I.4.1 Organismes

L'IETF (Internet Engineering Task Force) est un groupe informel, international, sans statut, ouvert à tout individu qui participe à l'élaboration des standards Internet. Les groupes de travail (task forces) sont répartis dans une dizaine de domaines d'intérêts, chaque domaine ayant un ou deux directeurs, lesquels font partie de l'IESG (Internet Engineering Steering Group).

L'IESG est un groupe de l'IETF chargé du pilotage de l'activité de production des standards Internet. Il examine tous les projets, sert de chambre d'appel contre les projets contestés, valide les processus de normalisation et donne l'accord final de l'IETF pour que la RFC (Request For Comments) soit publiée comme standard Internet.

L'IAB (Internet Architecture Board) se charge de l'orientation à long terme d'Internet et de ce fait des activités données à l'IETF. Il a, en autre, pour rôle de :

- Approuver les nominations proposées par le Comité de nomination de l'IESG.
- Examiner en appel les requêtes contre certaines décisions de l'IESG.
- Approuver la nomination de l'IANA (aujourd'hui sous le contrôle de l'ICANN).
- Conseiller l'ISOC.
- Encadrer les relations de l'IETF avec les autres organismes de standardisation.

L'IESG et l'IAB sont chapeautés par l'ISOC (Internet Society) qui est une association de droit américain à vocation internationale créée en janvier 1992 par les pionniers de l'Internet pour promouvoir et coordonner le développement des réseaux informatiques dans le monde. Elle est devenue, en 2005, l'autorité morale et technique la plus influente dans l'univers du réseau Internet.

L’ICANN (Internet Corporation for Assigned Names and Numbers) est une autorité de régulation de l'Internet. C'est une société de droit californien à but non lucratif ayant pour principales missions d'administrer les ressources numériques d'Internet telles que l'adressage IP et les noms de domaines de premier niveau (TLD - Top Level Domain) et de coordonner les acteurs techniques.

Ces services étaient initialement assurés dans le cadre d'un contrat avec le gouvernement fédéral américain par l’IANA (Internet Assigned Numbers Authority) et d'autres organismes. L'ICANN assume, à présent, les fonctions de l'IANA. Le 1er octobre 2016, le contrat liant l’ICANN aux Etats-Unis est arrivé à expiration.

Par le contrôle qu'elle exerce sur l'affectation des noms de domaines de premier niveau, l'ICANN délivre en pratique un droit de délégation sur la vente des noms de domaines à différentes organisations comme VeriSign pour les domaines génériques .com, .org et .net ou la société Afilias qui gère .info. Quant aux domaines de niveau national, chaque pays a une délégation : l'AFNIC pour les domaines .fr (France), .re (Réunion), .tf (les Terres australes et antarctiques françaises), .wf (Wallis-et-Futuna), .yt (Mayotte) et .pm (Saint-Pierre-et-Miquelon) ou Switch pour .ch (Suisse)…

I.4.2 Request for comments (RFC)

Les RFC sont une série numérotée de documents officiels décrivant les aspects techniques d'Internet ou de différents matériels informatiques. Peu de RFC sont des standards, mais tous les documents publiés par l'IETF sont des RFC.

Elles sont rédigées sur l'initiative d'experts techniques puis sont revues par la communauté Internet dans son ensemble. La majorité des RFC utilisent les termes MUST, MUST NOT, SHOULD, MAY... tels que définis dans la RFC 2119 présentée ci-dessous :

Network Working Group — *S. Bradner*
Request for Comments: 2119 — *Harvard University*
BCP: 14 — *March 1997*
Category: Best Current Practice

Key words for use in RFCs to Indicate Requirement Levels

Status of this Memo

This document specifies an Internet Best Current Practices for the Internet Community, and requests discussion and suggestions for improvements. Distribution of this memo is unlimited.

Abstract

In many standards track documents several words are used to signify the requirements in the specification. These words are often capitalized. This document defines these words as they should be interpreted in IETF documents. Authors who follow these guidelines should incorporate this phrase near the beginning of their document:

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119.

Note that the force of these words is modified by the requirement level of the document in which they are used.

1. MUST This word, or the terms "REQUIRED" or "SHALL", mean that the definition is an absolute requirement of the specification.

2. MUST NOT This phrase, or the phrase "SHALL NOT", mean that the definition is an absolute prohibition of the specification.

3. SHOULD This word, or the adjective "RECOMMENDED", mean that there may exist valid reasons in particular circumstances to ignore a particular item, but the full implications must be understood and carefully weighed before choosing a different course.

4. SHOULD NOT This phrase, or the phrase "NOT RECOMMENDED" mean that there may exist valid reasons in particular circumstances when the particular behavior is acceptable or even useful, but the full implications should be understood and the case carefully weighed before implementing any behavior described with this label.

5. MAY This word, or the adjective "OPTIONAL", mean that an item is truly optional. One vendor may choose to include the item because a particular marketplace requires it or because the vendor feels that it enhances the product while another vendor may omit the same item. An implementation which does not include a particular option MUST be prepared to interoperate with another implementation which does include the option, though perhaps with reduced functionality. In the same vein an implementation which does include a particular option MUST be prepared to interoperate with another implementation which does not include the option (except, of course, for the feature the option provides.)

6. Guidance in the use of these Imperatives

Imperatives of the type defined in this memo must be used with care and sparingly. In particular, they MUST only be used where it is actually required for interoperation or to limit behavior which has potential for causing harm (e.g., limiting retransmisssions) For example, they must not be used to try to impose a particular method on implementors where the method is not required for interoperability.

7. Security Considerations

These terms are frequently used to specify behavior with security implications. The effects on security of not implementing a MUST or SHOULD, or doing something the specification says MUST NOT or SHOULD NOT be done may be very subtle. Document authors should take the time to elaborate the security implications of not following recommendations or requirements as most implementors will not have had the benefit of the experience and discussion that produced the specification.

8. Acknowledgments

The definitions of these terms are an amalgam of definitions taken from a number of RFCs. In addition, suggestions have been incorporated from a number of people including Robert Ullmann, Thomas Narten, Neal McBurnett, and Robert Elz.

9. Author's Address

Scott Bradner
Harvard University
1350 Mass. Ave.
Cambridge, MA 02138

phone - +1 617 495 3864

email - sob@harvard.edu

Les termes "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" permettent donc de définir leurs exigences (obligation, interdiction, recommandation...). Pour plus d'informations à propos des RFC et les procédures associées, voyez la RFC 2026 « Procédures Standards d'Internet. Révision 3 ».

Les RFC font d'abord l'objet d'un draft (brouillon). Tout le monde soumet un draft à l'IETF en le transmettant à rfc.editor@rfc.editor.org. Les drafts soumis qui ne sont pas dignes d'intérêt ont une date de péremption. Sinon, un groupe de travail peut être créé pour la rédaction d'une RFC. La RFC 2223 donne les instructions pour les futurs auteurs.

Quelques RFC finissent par devenir des standards d'Internet. La procédure complète pour la transcription d'une RFC en standard est la suivante :

RFC → Proposed Standard → Draft Standard → Internet Standard

Malgré leur nom, les RFC sont le plus souvent stables. Toute modification apportée à une RFC entraîne l'écriture d'une nouvelle RFC qui rend la précédente obsolète.

Voici 2 exemples :

- Le protocole ICMP a été défini par la RFC 792 et mise à jour par les RFC successives 950, 4884, 6633 et 6918.
- La RFC de la suite de protocoles TCP/IP est définie dans la RFC 1122. Les mises à jour sont effectuées par les RFC 1349, 4379, 5884, 6093, 6298, 6633, 6864 et 8029.

II. ADRESSES IP

Objectifs

- Étudier l'adressage IP classful et classless, la résolution d'une adresse MAC et l'adressage IP statique ou dynamique
- Créer un réseau et des sous-réseaux (subnetting)
- Agréger plusieurs réseaux IP en un seul (supernetting)
- Accéder à un réseau public avec une adresse IP privée (NAT)

II.1. ADRESSAGE IP CLASSFUL

Classe	Identificateurs de réseau (1er octet)	ID réseau	ID hôtes	Nombre de réseaux	Nombre d'hôtes par réseau	Masque de sous-réseau
A	1 - 126	1 octet	3 octets	126	16 777 214	255.0.0.0
B	128 - 191	2 octets	2 octets	16 384	65 534	255.255.0.0
C	192 - 223	3 octets	1 octet	2 097 152	254	255.255.255.0
D	224 - 239	Adresses uniques				Non défini
E	240 - 255	Adresses uniques				Non défini

II.1.1 Identificateur réseau et identificateur d'hôte

Une adresse IPv4 est composée de 4 octets. Ce qui donne 4x8=32 bits.

La RFC 790 prévoit que l'adresse est divisée en deux parties :

- L'identificateur de réseau (net id) ;
- L'identificateur d'un hôte sur ce réseau (host id).

II.1.2 Classes et plage d'identificateurs de réseau

Les réseaux classful utilisent des classes pour répertorier les adresses IP. Les classes A, B, C fournissent des adresses unicast pour des tailles de réseaux différents. La classe D est destinée au multicast tandis que la classe E est réservée pour des expériences futures.

Nous allons donc nous intéresser aux classes A, B et C, qui sont destinées à être affectées aux équipements réseaux :

II.1.2.1 Classe A

Une adresse IP de classe A dispose d'un seul octet pour identifier le réseau et de trois octets pour identifier les hôtes (hosts) sur ce réseau. Donc un réseau de classe A peut comporter jusqu'à 23×8-2 hôtes, soit 16 777 214 adresses IP.

Pourquoi faut-il « -2 » pour calculer le nombre d'adresses IP ? Parce que la 1ère et la dernière adresse sont réservées à un usage précis.

Le premier octet d'une adresse IP de classe « A » commence toujours par le bit 0. La plage d'adresse en binaire s'étend de 00000000 à 01111111, soit de 0.X.X.X à 127.X.X.X.

Ceci étant, les réseaux 0.0.0.0 et 127.0.0.0 sont réservés pour un usage spécifique. Ils n'existent donc pas sur Internet ou dans un LAN. La plage réelle commence à 00000001 à 01111110, soit 1.0.0.0 jusqu'à 126.0.0.0.

Voici la représentation d'une adresse IPv4 pour une classe A :

Octet 1	Octet 2	Octet 3	Octet 4
0xxxxxxx	xxxxxxxx	xxxxxxxx	xxxxxxxx
Partie réseau	Partie hôte		

II.1.2.2 Classe B

Une adresse IP de classe B dispose de deux octets pour identifier le réseau et de deux octets pour identifier les hôtes sur ce réseau. Un réseau de classe B peut comporter jusqu'à 22 × 8-2 postes, soit 65 534 adresses IP.

Le premier octet d'une adresse IP de classe B commence toujours par la séquence de bit 10, il est donc compris entre 128 et 191.

Voici la représentation d'une adresse IPv4 pour une classe B :

Octet 1	Octet 2	Octet 3	Octet 4
10xxxxxx	xxxxxxxx	xxxxxxxx	xxxxxxxx
Partie réseau		Partie hôte	

II.1.2.3 Classe C

Une adresse IP de classe C dispose de trois octets pour identifier le réseau et d'un seul octet pour identifier les hôtes sur ce réseau. Un réseau de classe C peut comporter jusqu'à 28-2 postes, soit 254 adresses IP.

Le premier octet d'une adresse IP de classe C commence toujours par la séquence de bits 110, il est donc compris entre 192 et 223.

Voici la représentation d'une adresse IPv4 pour une classe C :

Octet 1	Octet 2	Octet 3	Octet 4
110xxxxx	xxxxxxxx	xxxxxxxx	xxxxxxxx
Partie réseau			Partie hôte

II.1.3 Directives d'adressage

II.1.3.1 Adresse de loopback (rebouclage)

L'identificateur de réseau ne peut pas être 127 puisqu'il est réservé aux fonctions de loopback.

Un loopback est un système matériel ou logiciel en informatique, réseaux ou télécommunications, destiné à renvoyer un signal reçu vers son envoyeur sans modification ni traitement et qui peut, par exemple, être utilisé à des fins de test.

L'adresse de loopback en TCP/IP version 4 est 127.0.0.1. On l'appelle également « localhost ».

II.1.3.2 Adresse de broadcast local

L'identificateur de réseau et l'identificateur d'hôte ne peuvent être 255, soit tous les bits à 1. L'adresse 255.255.255.255 est une adresse de diffusion locale.

II.1.3.3 Adresse de broadcast du réseau

Souvenez-vous pourquoi on utilise « -2 » pour calculer le nombre d'adresses IP ? La dernière adresse IP d'une plage réseau est réservée pour effectuer un broadcast au sein de ce réseau.

II.1.3.4 Adresse de réseau

L'identificateur de réseau et l'identificateur d'hôte ne peuvent être 0, soit tous les bits à 0 car la valeur 0 signifie « ce réseau uniquement ».

Souvenez-vous pourquoi on utilise « -2 » pour calculer le nombre d'adresses IP ? La première adresse IP d'une plage réseau est réservée pour identifier ce réseau.

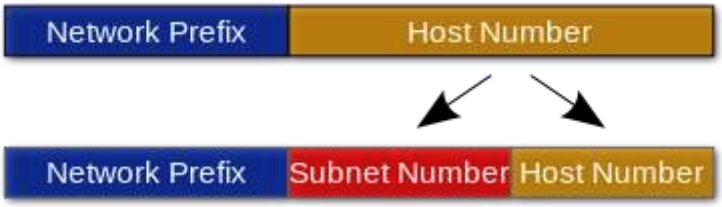
II.1.3.5 Adresse IP d'hôte

L'identificateur d'hôte doit être unique dans le réseau sinon cela provoque un conflit d'adresse IP.

II.1.3.6 Masques de sous-réseau

Historiquement, on appelle « sous-réseau » chacun des réseaux connectés à Internet.

En 1984, devant la limitation du modèle de classes, la RFC 917 « Internet subnets » crée le concept de sous-réseau qui introduit un niveau hiérarchique supplémentaire entre le numéro de réseau et le numéro d'hôte. Le masque de sous-réseau permet de distinguer la partie de l'adresse utilisée pour le routage et celle utilisable pour numéroter des interfaces. Un sous-réseau correspond typiquement à un réseau local sous-jacent.



Un masque de sous-réseau (subnet mask, netmask ou address mask) est un masque distinguant les bits d'une adresse IPv4 utilisés pour identifier le sous-réseau de ceux utilisés pour identifier l'hôte.

L'adresse du sous-réseau est obtenue en appliquant l'opérateur **et** binaire entre l'adresse IPv4 et le masque de sous-réseau. L'adresse de l'hôte à l'intérieur du sous-réseau est quant à elle obtenue en appliquant l'opérateur et entre l'adresse IPv4 et le complément à un du masque.

Les masques de sous-réseau utilisent la même représentation que celles des adresses IPv4. Une adresse IP est codée sur 4 octets, soit 32 bits qui sont représentés en notation décimale à point. Un masque de sous-réseau possède lui aussi 4 octets.

On utilise en pratique des masques constitués sous leur forme binaire d'une suite de 1 suivi d'une suite de 0, il y a donc 32 masques réseaux possibles.

Le nombre de bits à 1 dans le masque de réseau est représenté par r et s pour le nombre de bits à 1 dans le masque de sous-réseau. Le nombre de sous-réseaux possibles est donné par : 2^s-r et le nombre d'hôtes par sous-réseau est : 2^32-s-2

Nous ôtons de 2 car, comme nous l'avons déjà dit, la première adresse de ce sous-réseau est réservée à l'identité de ce sous-réseau et la dernière au broadcast de ce sous-réseau. Ils ne peuvent pas être utilisés pour numéroter une interface de communication.

Liste des masques de sous-réseaux :

Bits à 1	Bits à 0	Masque de sous-réseau	Nombre d'hôtes par sous-réseau
1	31	128.0.0.0	231-2 = 2147483646
2	30	192.0.0.0	230-2 = 1073741822
3	29	224.0.0.0	229-2 = 536870910
4	28	240.0.0.0	228-2 = 268435454
5	27	248.0.0.0	227-2 = 134217726
6	26	252.0.0.0	226-2 = 67108862
7	25	254.0.0.0	225-2 = 33554430
8	24	255.0.0.0	224-2 = 16777214
9	23	255.128.0.0	223-2 = 8388606
10	22	255.192.0.0	222-2 = 4194302
11	21	255.224.0.0	221-2 = 2097150
12	20	255.240.0.0	220-2 = 1048574
13	19	255.248.0.0	219-2 = 524286
14	18	255.252.0.0	218-2 = 262142
15	17	255.254.0.0	217-2 = 131070
16	16	255.255.0.0	216-2 = 65534
17	15	255.255.128.0	215-2 = 32766
18	14	255.255.192.0	214-2 = 16382
19	13	255.255.224.0	213-2 = 8190
20	12	255.255.240.0	212-2 = 4094
21	11	255.255.248.0	211-2 = 2046
22	10	255.255.252.0	210-2 = 1022
23	9	255.255.254.0	29-2 = 510
24	8	255.255.255.0	28-2 = 254
25	7	255.255.255.128	27-2 = 126
26	6	255.255.255.192	26-2 = 62
27	5	255.255.255.224	25-2 = 30
28	4	255.255.255.240	24-2 = 14
29	3	255.255.255.248	23-2 = 6
30	2	255.255.255.252	22-2 = 2
31	1	255.255.255.254	21=2
32	0	255.255.255.255	20=1

Un sous-réseau dans un réseau local (LAN) devient un VLAN dans un switch.

Exemple : comment créer des sous-réseaux ?

Un LAN a une adresse IP de classe B 172.16.0.0 avec le masque de réseau 255.255.0.0, soit 172.16.0.0/255.255.0.0

Nous voulons diviser ce dernier en 6 sous-réseaux. Combien de bits à 1 dans le 3ème octet du masque de sous-réseau nous faut-il ?

Nous allons utiliser la formule 2^ (nombre de bits à 1)-2. Nous ne pouvons pas utiliser le 1er et le dernier sous-réseaux. Nous reviendrons sur ce point dans la seconde étape.

Si nous prenons 3 bits, nous avons 2^3-2=6. Nous pouvons donc créer les 6 sous-réseaux souhaités.

2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0
128	64	32	16	8	4	2	1
1	1	1	0	0	0	0	0

La valeur décimale de cet octet est 128+64+32=224.

Le masque de sous-réseau est : 255.255.224.0

1. Nous devons maintenant définir les identificateurs de sous-réseau. Le tableau ci-dessous représente toutes les combinaisons du 3ème octet sur 3 bits, soit 2^3 qui fait 8 sous-réseaux.

La formule est 2^3-2. Du fait que le 1er sous-réseau a les 3 bits à 0 et que le dernier sous-réseau a les 3 bits à 1, le protocole IP ne sait pas les gérer.

255 11111111	255 11111111	224 11100000	0 00000000
Inutilisable →		~~00000000~~	~~= 0~~
		00100000	= 32
		01000000	= 64
		01100000	= 96
		10000000	= 128
		10100000	= 160
		11000000	= 192
Inutilisable →		~~11100000~~	~~= 224~~

Le pas d'incrémentation est de 32 pour changer de sous-réseaux.

Identificateur de sous-réseau	Début de la plage d'identificateurs d'hôte	Fin de la plage d'identificateurs d'hôte	Adresse de Broadcast
172.16.32.0	172.16.32.1	172.16.63.254	172.16.63.255
172.16.64.0	172.16.64.1	172.16.95.254	172.16.95.255
172.16.96.0	172.16.96.1	172.16.127.254	172.16.127.255
172.16.128.0	172.16.128.1	172.16.159.254	172.16.159.255
172.16.160.0	172.16.160.1	172.16.191.254	172.16.191.255
172.16.192.0	172.16.192.1	172.16.223.254	172.16.223.255

Pour communiquer entre ces sous-réseaux, il faut mettre en place un système de routage IP. Ce sujet sera abordé dans le chapitre 3.

2. Pour terminer, nous pouvons calculer le nombre d'adresses IP disponibles dans chaque sous-réseau. La formule est 2^(nombre de bits à 0 dans le masque)-2 adresses IP (la première et la dernière).

Le masque 255.255.224.0 à 19 bits à 1 et 13 bits à 0.

2^13-2=8190.

II.1.3.7 *Adresses IP privées*

- **RFC 1918 « Address Allocation for Private Internets »**

Cette RFC, définie en février 1996, spécifie les plages d'adresses IP pour les réseaux privés qui ne sont pas routées sur Internet :

Identificateur de réseau	Masque de sous-réseau	Plage IP	Nombre d'adresses IP
10.0.0.0	255.0.0.0	10.0.0.0 – 10.255.255.255	232-8 = 16 777 216
172.16.0.0	255.255.0.0	172.16.0.0 – 172.31.255.255	232-12 = 1 048 576
192.168.0.0	255.255.0.0	192.168.0.0 – 192.168.255.255	232-16 = 65 536

L'utilisation d'adresses publiques dans un LAN privé poserait des problèmes à des services réseaux. Comment distinguer 2 réseaux ayant la même identité c'est-à-dire le vrai réseau public et le vôtre qui est privé ? Comment aller vers la bonne destination ?

Vous devez configurer votre LAN avec des adresses privées. Mais, comment des adresses privées peuvent accéder à des ressources sur Internet puisque ces dernières ne sont pas routées sur celui-ci ? La solution est la traduction d'adresse réseau ou NAT (Network Address Translation). Cette fonctionnalité doit être configurée dans le routeur qui vous interconnecte avec Internet.

- **RFC 3927 « Dynamic Configuration of IPv4 Link-Local Addresses »**

L'APIPA (Automatic Private Internet Protocol Addressing) ou connu aussi sous le nom de IPv4LL (IPv4 Link-Local) est un processus qui permet à un système d'exploitation de s'attribuer automatiquement une adresse IP, lorsque le serveur DHCP (Dynamic Host Configuration Protocol) est hors service ou injoignable.

Le serveur DHCP est un service réseau qui permet de distribuer une configuration TCP/IP à un hôte qui l'a demandé. Ce sujet sera abordé plus loin dans ce chapitre.

APIPA utilise la plage d'adresses IP 169.254.0.0/255.255.0.0 c'est-à-dire que la plage des adresses va de 169.254.0.0 à 169.254.255.255. Cette plage est réservée à cet usage auprès de l'ICANN.

II.1.3.8 Résolution d'une adresse IP en une adresse MAC

- **Protocole ARP**

Le protocole ARP (Address Resolution Protocol), défini dans la RFC 826, a un rôle important parmi la pile de protocoles Internet. Il permet de connaître l'adresse physique d'une carte réseau (adresse MAC - Media Access Control) correspondant à une adresse logique (adresse IP).

Chaque interface de communication sur le réseau a une adresse matériel ou physique de 48 bits. Elle est aussi parfois appelée : Adresse Ethernet, UAA (Universally Administered Address), BIA (Burned-In Address), MAC-48 ou bien EUI-48. L'IEEE a défini un format similaire à 64 bits appelé EUI-64.

L'adresse MAC, attribuées par l'IEEE (Institute of Electrical and Electronics Engineers), est normalement un numéro unique en format hexadécimal qui est fixé dès la fabrication de la carte en usine. Certain fabriquant ne respecte malheureusement pas les normes et affecte n'importe quelle adresse MAC à leurs équipements. Si deux machines ont la même adresses MAC sur le réseau, elles ne pourront pas communiquer car vous aurez un conflit d'adresses. La solution est de changer de carte réseau ou de changer l'adresse MAC si vous en avez la possibilité.

- **Table ARP**

Chaque hôte maintient un cache des adresses les plus récemment utilisées dans une table d'association d'adresses appelées « table ARP » ou parfois « cache ARP ».

Cette table va permettre de fluidifier et d'accélérer les prochains échanges avec les émetteurs enregistrés en évitant de reproduire une requête ARP à chaque échange.

II.1.3.9 Résolution d'une adresse IP locale

1. Lors d'une tentative de résolution d'adresse, l'hôte consulte préalablement la table ARP. Si une entrée correspond à l'adresse recherchée, on l'utilise.
2. Sinon l'hôte diffuse un message de demande de résolution (broadcast) : « Qui a cette adresse IP ? »

Si la réponse ne parvient pas dans un délai imparti alors Time-Out (délai d'attente dépassée) et réexpédition de la requête.

3. La station recevant un message de demande de résolution répondant à cette demande va mettre à jour sa table ARP avec les informations concernant l'hôte qui a initié cette demande de résolution.

Les messages de demande et de réponse de résolution d'adresses comportent une paire de couples {adresse MAC ; adresse IP} de l'émetteur et du récepteur.

- **Proxy ARP (promiscuous ARP ou ARP hack)**

Le broadcast ARP ne peut s'appliquer qu'entre deux hôtes se trouvant sur le même réseau. Le Proxy ARP offre la possibilité de constituer un seul réseau IP (partie adresse Réseau Commune) de deux réseaux physiques séparés par un routeur.

Le Proxy ARP est une technique permettant à un routeur de faire croire que les hôtes des deux réseaux se trouvent sur le même réseau. Le routeur répond à une requête ARP provenant d'un hôte à destination d'un hôte se trouvant sur l'autre réseau. Le routeur ment en trompant la source qui associera l'adresse IP de la machine destinatrice avec l'adresse physique Ethernet du routeur. Dans cette configuration, le routeur a connaissance des différentes machines se trouvant sur chaque réseau.

Cette technique fera qu'une adresse Ethernet unique (routeur) sera associée à plusieurs adresses IP (celles des machines du réseau opposé). Cette tolérance n'est pas acceptée par certaines implémentations d'ARP. Ce type de correspondances peut être considéré comme une éventuelle mystification et donc une faille potentielle dans la sécurité.

- **ARP gratuit**

Dans certaines implémentations, au démarrage d'un hôte, une requête ARP concernant l'adresse IP de la machine émettrice est envoyée. Normalement, il ne devrait y avoir aucune réponse.

Cette possibilité offre deux avantages :

- Faire connaître aux autres stations de l'introduction d'une nouvelle station donc d'une nouvelle adresse MAC. Ainsi, une mise à jour des entrées dans les tables ARP sera effectuée.
- Détecter si une autre station a la même adresse IP. Dans le cas d'une réponse à la requête, il s'avère qu'il y a duplication d'adresses IP et un message d'erreur est affiché indiquant l'adresse de la carte Ethernet d'où provient la réponse ARP.

II.1.3.10 Résolution d'une adresse MAC en une adresse IP

- **Protocole RARP**

Le protocole RARP (Reverse Address Resolution Protocol), défini dans la RFC 903, est beaucoup moins utilisé. Il s'agit donc d'une sorte d'annuaire inversé des adresses logiques et physiques.

Il permet à une station de connaître son adresse IP à partir d'une table de correspondance entre adresse MAC (adresse physique) et adresses IP hébergée par une passerelle (routeur) située sur le même réseau local (LAN).

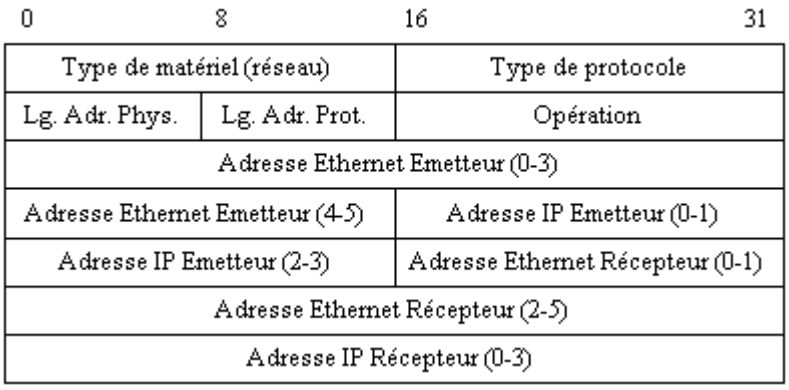
Pour cela il faut que l'administrateur paramètre le routeur avec la table de correspondance des adresses MAC/IP. En effet, à la différence de ARP ce protocole est statique. Il faut donc que la table de correspondance soit toujours à jour pour permettre la connexion de nouvelles cartes réseau.

- **Format des paquets ARP et RARP**

Un paquet ARP ou RARP est encapsulé dans une trame Ethernet. Dans l'en-tête de la trame, le champ type de protocole permet de préciser quel est le protocole encapsulé :

- Type de protocole = 0x0806 pour ARP
- Type de protocole = 0x8035 pour RARP

Les paquets ARP et RARP ont le format suivant :



- Type de matériel = 1 pour Ethernet
- Type de protocole = 0x0800 pour IP
- Lg. Adr. Phys. = 6 octets pour Ethernet
- Lg. Adr. Prot. = 4 octets pour IP
- Opération :
  1. (requête ARP),
  2. (réponse ARP),
  3. (requête RARP),
  4. (réponse RARP)
- Adresses MAC sur 48 bits et adresses IP sur 32 bits de l'émetteur et du récepteur remplis selon l'opération en cours.

II.2. ADRESSAGE IP CLASSLESS

- Abolition des classes
- CIDR (Classless Inter-Domain Routing)
- Subnetting
- VLSM (Variable Length Subnet Mask)

II.2.1 Abolition des classes

En 1992, la RFC 1338 « Supernetting: an Address Assignment and Aggregation Strategy » propose d'abolir la notion de classe qui n'était plus adaptée à la taille d'Internet.

La distinction entre les adresses de classe A, B ou C a été ainsi rendue obsolète de sorte que la totalité de l'espace d'adressage unicast puisse être gérée comme une collection unique de sous-réseaux indépendamment de la notion de classe.

Le masque de sous-réseau ne peut plus être déduit de l'adresse IP elle-même, les protocoles de routage compatibles avec CIDR, dits « classless », doivent donc accompagner les adresses du masque correspondant. Il écrit sous la forme « /nombre de bits à 1 dans le masque ».

Par exemple : 255.128.0.0 devient /9.

II.2.2 CIDR (Classless Inter-Domain Routing)

Le CIDR est mis au point en 1993 dans le but de diminuer la taille de la table de routage contenue dans les routeurs. Cet objectif est atteint en agrégeant plusieurs entrées de cette table en une seule.

Les protocoles de routage dynamique ont été modifiés pour être conformes à ce changement :

- BGP (Border Gateway Protocol) version 4
- OSPF (Open Shortest Path First),
- EIGRP (Enhanced Interior Gateway Routing Protocol),
- RIPv2 (Routing Information Protocol)

II.2.3 Subnetting

Le subnetting est une technique qui consiste à diviser un réseau plus large en plusieurs sous-réseaux. Ce que nous avions déjà fait.

Reprenons l'exemple que nous avons vu dans IP classful :

1. Un LAN a une adresse IP 172.16.0.0/16. Nous voulons diviser ce dernier en 6 sous-réseaux. Combien de bits à 1 dans le 3ème octet du masque de sous-réseau nous faut-il ?

   Nous allons utiliser la formule 2^(nombre de bits à 1). Nous n'ôtons plus le 1er et le dernier sous-réseaux dans la formule car nous pouvons, désormais, d'exploiter tous les sous-réseaux.

   Si nous prenons 2 bits, 2^2=4. Nous ne pouvons pas faire 6 sous-réseaux dans 4. Alors nous allons prendre 3 bits et nous avons 2^3=8. Là, nous pouvons donc créer 6 sous-réseaux dans 8. Certes, il en restera 2.

2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0
128	64	32	16	8	4	2	1
1	1	1	0	0	0	0	0

La valeur décimale de cet octet est 128+64+32=224 et le masque de sous-réseau est : 255.255.224.0 ou /19.

2. Nous devons maintenant définir les identificateurs de sous-réseau. Le tableau ci-dessous représente toutes les combinaisons du 3ème octet sur 3 bits, soit 2^3 qui fait 8 sous-réseaux.

255	255	224	0
11111111	11111111	11100000	00000000

00000000	= 0
00100000	= 32
01000000	= 64
01100000	= 96
10000000	= 128
10100000	= 160
11000000	= 192
11100000	= 224

Le pas d'incrémentation est de 32 pour changer de sous-réseaux.

Identificateur de sous-réseau	Début de la plage d'identificateurs d'hôte	Fin de la plage d'identificateurs d'hôte	Adresse de Broadcast
172.16.32.0/19	172.16.32.1	172.16.63.254	172.16.63.255
172.16.64.0/19	172.16.64.1	172.16.95.254	172.16.95.255
172.16.96.0/19	172.16.96.1	172.16.127.254	172.16.127.255
172.16.128.0/19	172.16.128.1	172.16.159.254	172.16.159.255
172.16.160.0/19	172.16.160.1	172.16.191.254	172.16.191.255
172.16.192.0/19	172.16.192.1	172.16.223.254	172.16.223.255

Pour communiquer entre ces sous-réseaux, il faut mettre en place un système de routage IP. Ce sujet sera abordé dans le chapitre 3.

3. Pour terminer, nous pouvons calculer le nombre d'adresses IP disponibles dans chaque sous-réseau. La formule est 2^(nombre de bits à 0 dans le masque)-2 adresses IP (la première et la dernière).

Le masque 255.255.224.0 à 19 bits à 1 et 13 bits à 0.

2^13-2=8190.

4. En Conclusion.

Si nous voulons diviser le réseau 172.16.0.0/16 en 7 sous-réseaux.

En IP classful, 4 bits à 1 supplémentaires dans le masque sont nécessaires, soit 2^4-2=14 car pour rappel 2^3-2=6.

En IP classless, 3 bits à 1 suffisent : 2^3=8. Il nous reste 1 sous-réseau.

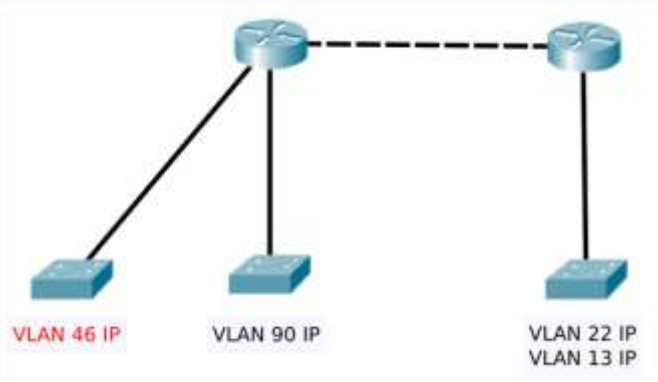
II.2.4 VLSM (Variable Length Subnet Mask)

Le concept VLSM est né avec la volonté d'économiser des adresses IP. En effet, si une entreprise décide d'utiliser un masque de sous-réseau /24 (255.255.255.0) pour tous ses sous-réseaux, cela veut donc dire que chaque sous réseau a 254 adresses IP respectivement.

L'utilisation de masques de longueur variable découpe l'espace d'adressage en blocs de taille variable permettant une utilisation plus efficace de l'espace d'adressage.

Exemple :

1. Vous avez 4 sous-réseaux à concevoir dans un réseau 192.168.0.0/24 :



Pour rappel, un VLAN est un sous-réseaux.

Sous-réseaux	Nombre d'adresses IP nécessaires
1	90+1
2	46+1
3	22+1
4	13+1
5	2
TOTAL	177

Nous avons 171 (90+46+22+13) adresses IP pour les 4 VLAN. Nous devons aussi attribuer 4 adresses IP aux interfaces de VLAN 4 et 1 adresse IP à chaque routeur qui nous permet d'accéder au réseau distant, soit 2 IP.

Au total, cela fait 6 adresses IP supplémentaires. Nous avons un total de 177. Ces adresses IP seront réparties sur 5 sous-réseaux : 4 VLAN + 1 sous-réseau pour l'interconnexion des 2 routeurs.

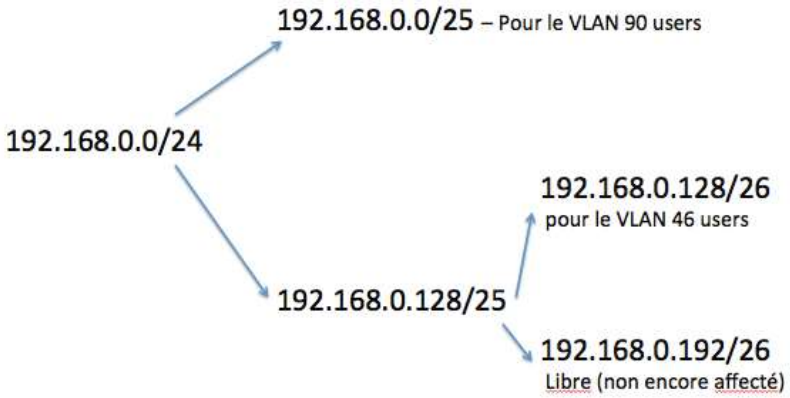
2. Découpons en 5 sous-réseaux. Cependant, le nombre de machines sur chaque sous-réseau est variable. Cela va de 91 (90 IP + 1 pour l'interface de VLAN) à 2.

Commençons par le sous-réseau le plus important en termes d'adresses IP (91 IP). Un masque /26 donnerait 2^(32-26)-2=62 IP. Je n'ai pas assez d'adresses IP. Il nous faut alors un masque /25. 2^(32-25)-2=126 IP. Le 1er sous-réseau sera 192.168.0.0/25

Avec /25, nous avons pris un bit supplémentaire donc le pas d'incrémentation est de 128.

Les autres réseaux sont créés en découpant la deuxième plage restante, soit 192.168.0.128/25. Nous allons créer un second sous-réseau de 46 IP + 1 IP pour l'interface de VLAN, soit 47 IP. Un masque de /26 est nécessaire. 2^(32-26)-2=62 IP Le VLAN sera 192.168.0.128/26.

- Faisons un schéma pour situer où nous en sommes :



Pour le VLAN 22 IP + 1 IP, nous allons utiliser la plage laissée libre qui est 192.168.0.192/26. Quelle taille de masque a-t-on besoin ? Un masque de sous-réseaux /27. 32-27=5 bits pour les hôtes, soit : 2^5-2=30 IP.

Le masque du troisième sous-réseau sera /27, soit 192.168.0.192/27. La plage laissée libre pour la suite est donc la plage 192.168.0.224/27 que nous utiliserons pour le prochain sous-réseau.

Le quatrième sous-réseau a besoin de 13 IP + 1 IP, soit 14 IP. Un masque /28 permet d'avoir 14 adresses IP, ce qui est exactement le nombre requis.

Nous allons donc attribuer au quatrième VLAN la plage 192.168.0.224/28. La plage 192.168.0.240/28 reste libre.

Le sous-réseau d'interconnexion entre 2 routeurs nécessite 2 adresses IP. Il faut donc un masque /30. 2^(32-30)-2=2 IP. Il est possible d'utiliser la plage 192.168.0.240/30.

Des plages restent libres et pourront être utilisées pour un usage futur.

Voici un tableau pour résumer le plan d'adressage :

Sous-réseaux	Nombre d'adresses IP nécessaires	Identificateur de sous-réseaux	Nombre d'IP
1	90+1	192.168.0.0/25	2^(32-25)-2=126
2	46+1	192.168.0.128/26	2^(32-26)-2=62
3	22+1	192.168.0.192/27	2^(32-27)-2=30
4	13+1	192.168.0.224/28	2^(32-28)-2=14
5	2	192.168.0.240/30	2^(32-30)-2=2
TOTAL	177		

II.2.5 Supernetting

Un sur-réseau ou supernet est une normalisation RFC 1519 mais aussi une technique de CIDR qui permet de définir un préfixe réseau englobant plusieurs sous-réseaux.

La technique du sur-réseau consiste à agréger plusieurs réseaux IP en un seul. Cette technique est utilisée dans les tables de routage ou pour simplifier les règles de filtrage dans un pare-feu ou un routeur, par exemple. Elle permet de réduire le nombre de lignes tout en conservant la totalité des destinations.

Cette technique n'est pas sans risque car elle peut autoriser des destinations qui ne sont pas prévues initialement.

- *Exemple de sur-réseau* :

Les routes des sous-réseaux 128.203.1.128/26 et 128.203.1.0/25 peuvent être collectivement agrégés avec un sur-réseau 128.203.1.0. Le préfixe réseau est sur 3 octets donc 24. Le supernet est 128.203.1.0/24.

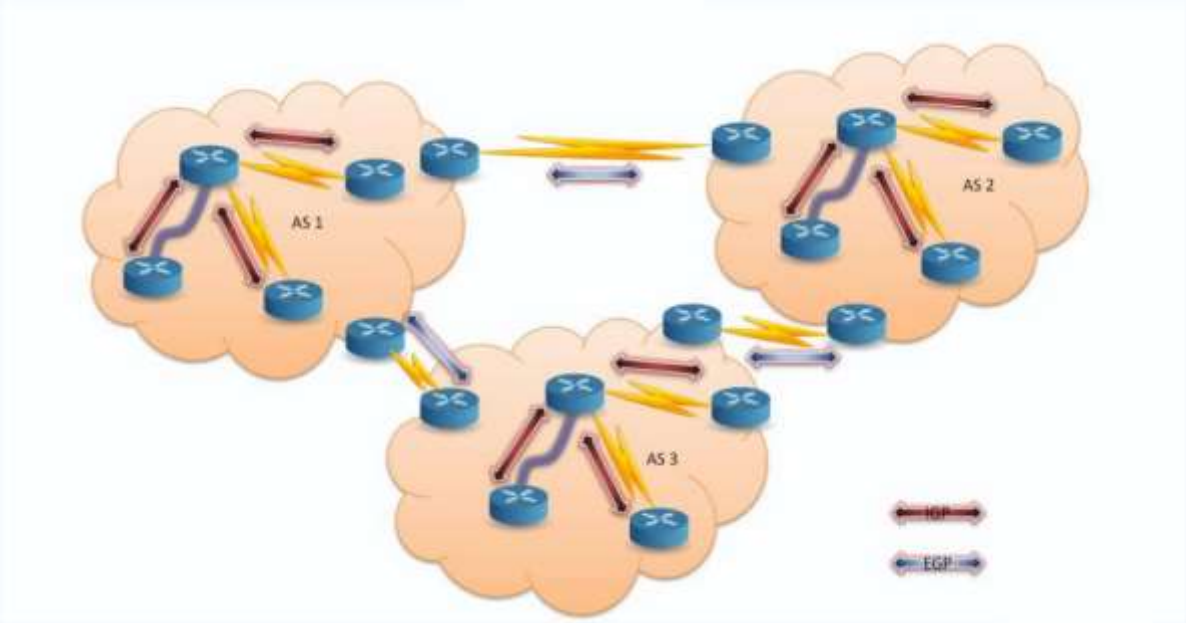
III. INTERCONNEXION DU RESEAU IP

Objectifs

- Étudier le routage inter-VLAN et le routage inter-LAN
- Configurer les équipements de niveau 3
- Mettre en œuvre le routage IP statique
- Mettre en œuvre le routage IP dynamique avec RIPv2 et OSPF
- Étudier les tables de routages
- Mettre en œuvre le NAT

Références

III.1. ROUTAGE IP



III.1.1 Principe du routage

Lorsqu'un routeur reçoit un paquet IP, il examine l'adresse IP de destination.

Si celle-ci appartient au réseau 127.0.0.0/8 : le paquet ne sera pas transmis mais traité localement.

Si l'adresse de destination correspond exactement à l'adresse de l'une des interfaces du routeur, le paquet est à destination de celui-ci. Le paquet ne sera donc pas transmis mais traité localement.

Si aucune des conditions précédentes n'est respectée alors le paquet n'est pas à destination du routeur. Celui-ci va tenter de déterminer si la destination est directement joignable.

Pour chaque interface sans adresse IP (interface série de type unnumbered), l'adresse IP de destination est comparée au « router ID » du routeur à l'autre bout de la ligne. Si les deux correspondent exactement, le paquet est envoyé au routeur distant.

Sinon :

- Pour chaque interface, prendre le préfixe réseau qui lui est associé (pour cela un ET logique est effectué entre l'adresse IP de l'interface et le masque associé),
- Un ET logique est calculé entre l'adresse IP destination et le masque de l'interface,
- Si le préfixe de l'interface et celui de la destination correspondent, la destination est directement joignable par l'interface.
- Répéter cela pour chaque interface.

Si aucune interface n'a été sélectionnée durant l'étape précédente, cela signifie que la destination n'est pas directement joignable. Le routeur doit alors consulter sa table de routage afin de découvrir le routeur intermédiaire vers lequel doit être transmis le paquet.

III.1.2 Routage IP classful

Ce type de routage est le comportement par défaut sur les routeurs Cisco avant iOS 12.0 ou si la commande « no ip classless » apparaît dans la configuration.

L'algorithme utilisé est le suivant :

1. Vérifier l'adresse réseau de la destination (ici l'adresse destination appartient au réseau de classe B 172.16.0.0). Si aucune entrée pour ce réseau n'est disponible dans la table de routage, aller à l'étape 2.
   a. Si le réseau n'est pas divisé en sous-réseaux : transmettre le paquet au routeur assigné à ce réseau.
   b. Si le réseau est divisé en sous-réseaux, vérifier si le préfixe de l'adresse destination correspond à l'un de ces sous-réseaux. Si tel est le cas, transmettre le paquet au routeur assigné à ce sous-réseau, sinon envoyer un message « ICMP unreachable » à l'expéditeur.
2. Si une route par défaut est présente, il faut l'utiliser sinon envoyer un message « ICMP unreachable » à l'expéditeur.

III.1.3 Routage IP classless

Ce type de routage est le comportement par défaut sur les routeurs depuis IOS 12.0 ou si la commande « ip classless » apparaît dans la configuration.

L'algorithme est le suivant :

1. Pour chaque entrée dans la table de routage : comparer le préfixe de l'adresse de destination apparaissant dans le paquet et celle de la route. Pour cela, effectuer un ET logique entre l'adresse IP destination et le masque associé à la table de routage.
   a. S'il y a correspondance, placer cette route dans la liste des routes potentielles.
   b. S'il n'y a pas correspondance, ne pas tenir compte de la route.
2. Parmi toutes les routes potentielles, sélectionner celle qui a le préfixe correspondant le plus long.

Exemple :

Pour l'adresse IP de destination 172.16.3.206, après consultation de la table de routage, nous obtenons ceci :

Préfixe	Longueur préfixe	IP destination / Masque	Prochain Routeur
172.16.3.128	25	172.16.3.128	R2
172.16.3.192	26	172.16.3.192	R3
172.16.4.0	24	172.16.3.0	R4
172.16.3.0	24	172.16.3.0	R4
0.0.0.0	0	0.0.0.0	R2

Après mise l'étape 1, il reste 4 routes potentielles sur les 5 du départ. De ces 4 routes, celle qui offre la plus grande correspondance est la route qui passe par le routeur R3 et à destination de 172.16.3.192/26.

III.1.4 Catégories de routage

Nous disposons de 2 catégories de routage :

- **Routage statique**

L'ajout des routes est manuel dans les tables de routage des routeurs. Cette technique est utile sur des réseaux simples, dans lesquels il n'existe pas plusieurs routes possibles pour une même destination.

Ce type de routage ne permet pas une détection automatique des nouvelles routes ou des routes devenues indisponibles.

- **Routage dynamique**

Toutes routes sont apprises automatiquement par les routeurs. Pour cela, il est nécessaire d'utiliser des protocoles de routage permettant l'échange d'informations entre les routeurs. Tous les réseaux un tant soit peu complexes utilisent ce type de routage.

III.2. PROTOCOLES IGP ET EGP

III.2.1 Définition

Il existe 2 types de protocoles de routage dynamique :

- IGP (Interior Gateway Protocol),
- EGP (Exterior Gateway Protocol).

Afin de bien comprendre la différence entre un IGP et un EGP, il est nécessaire de définir la notion de Système Autonome (AS - Autonomous System).

Un Système Autonome est un ensemble de réseaux appartenant au même domaine d'administration. Le réseau Internet peut être représenté comme un ensemble de Système Autonome.

Les Systèmes Autonomes s'échangent des informations de routage à l'aide d'un protocole de type EGP qui est BGP (Border Gateway Protocol) version 4.

Au sein d'un Système Autonome, les routeurs s'échangent des informations de routage en utilisant des IGP tels que :

- RIP (Routing Information Protocol),
- IGRP (Interior Gateway Routing Protocol),
- EIGRP (Enhanced Interior Gateway Routing Protocol),
- OSPF (Open Shortest Path First).

NOTE : Dans ce document, nous n'aborderons pas le protocole BGP.

III.2.2 Classes de protocoles IGP

Il existe principalement 3 classes de protocoles :

- **Les protocoles dits à vecteurs de distance**

Les routeurs annoncent régulièrement leur table de routage. Cette annonce n'est faite qu'à leurs voisins directs. Il est à noter que des annonces sont effectuées même si aucune modification n'est intervenue dans la topologie du réseau. Dans cette catégorie peuvent être classés : RIP et IGRP.

- **Les protocoles à état de liens**

Les routeurs utilisant ce type de protocole annoncent uniquement l'état de leurs liens (interfaces). Cependant, cette annonce est à destination de tous les routeurs du domaine de routage (Aire). Ces protocoles sont peu bavards et ne nécessitent pas d'annonces fréquentes. Généralement celles-ci surviennent lors de la modification de topologie. Donc tous les routeurs connaissent les liens des autres. Nous pouvons classer dans cette catégorie : OSPF.

- **Les protocoles hybrides**

EIGRP est un protocole à vecteur de distance. Cependant, il agit comme les protocoles à état de liens il n'effectue pas d'annonces périodiques mais uniquement lors d'une modification de la topologie.

III.2.3 Protocoles IGP classful

Les protocoles de routage de type classful (RIPv1, IGRP) n'annoncent pas le masque de réseau avec la route.

Les routeurs pour déterminer le masque vont effectuer les vérifications suivantes :

- Si une interface du routeur dispose d'une adresse se trouvant dans le même réseau principal que la route, utiliser le masque de l'interface pour la route.
- Sinon, utiliser le masque par défaut de la classe (A, B ou C) correspondant au réseau.

Les annonces effectuées par un protocole de routage classful entre deux réseaux différents sont des résumés. C'est-à-dire qu'ils annoncent le réseau complet (classe A, B ou C).

Si nous prenons le premier exemple ci-dessus, nous pouvons constater que R2 disposera de deux routes pour se rendre sur le réseau 10.0.0.0. Il y a donc une chance sur deux pour que le message « ICMP ECHO REQUEST » de l'utilisateur parvienne à destination.

Lorsqu'au sein d'un même réseau des masques de longueurs différentes sont utilisés, les routeurs n'ont aucun moyen de déterminer le masque correct.

Ainsi, dans l'exemple suivant, il est peu probable que le routage fonctionne.

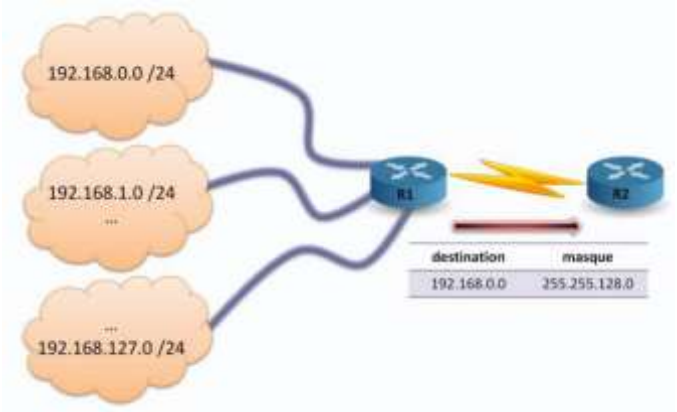
En recevant les annonces de R1 et R3, R2 utilisera le masque associé à ses interfaces car celles-ci appartiennent au réseau 10.0.0.0. Cela ne correspond pas aux masques associés aux sous-réseaux 10.47.1.0 et 10.47.2.0.

III.2.4 Protocoles IGP classless

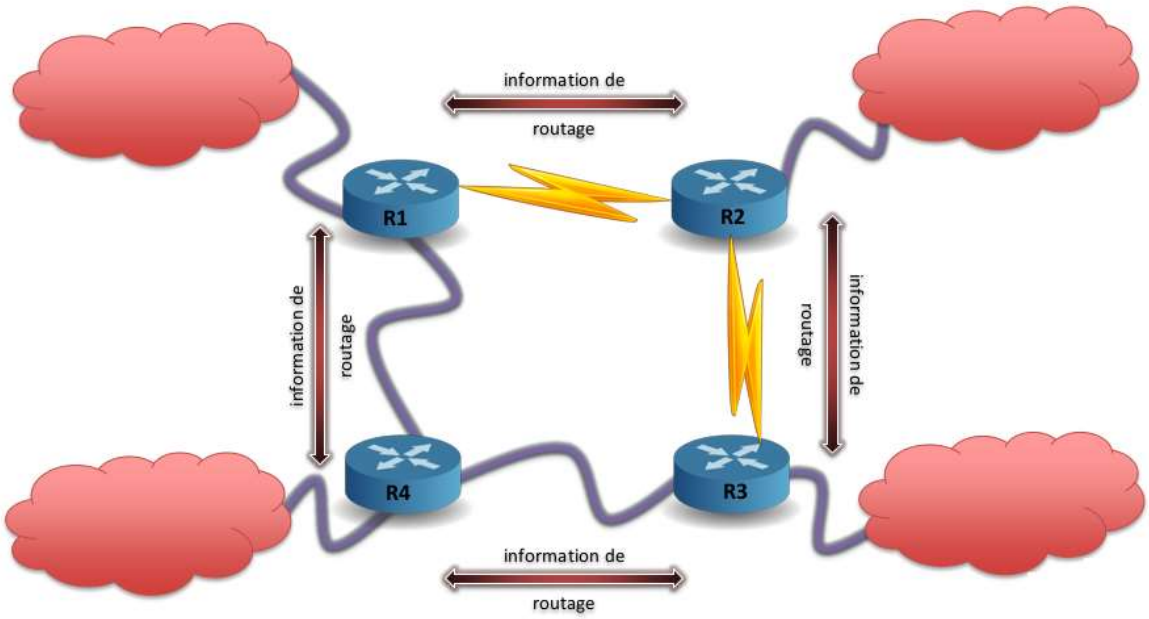
Les protocoles classless (RIPv2, EIGRP, OSPF) permettent de résoudre les problèmes associés aux protocoles classful. Il est ainsi possible d'avoir une longueur de masque différente pour un même réseau.

Le supernetting, que nous avons évoqué dans le chapitre précédent, consiste à pouvoir regrouper plusieurs réseaux dans une annonce en spécifiant un masque de longueur inférieure au masque par défaut de la classe des réseaux.

Dans le cas suivant, il n'est pas nécessaire d'annoncer chacun des 128 réseaux ; une seule annonce suffit. Un masque de 17 bits permet de regrouper les 128 réseaux.



III.3. ROUTAGE INTER-LAN



III.3.1 VLAN (Virtual LAN)

Un réseau local virtuel (VLAN) est un réseau logique indépendant. De nombreux VLAN peuvent coexister sur un même switch (commutateur) réseau.

Les VLAN présentent plusieurs intérêts tels que :

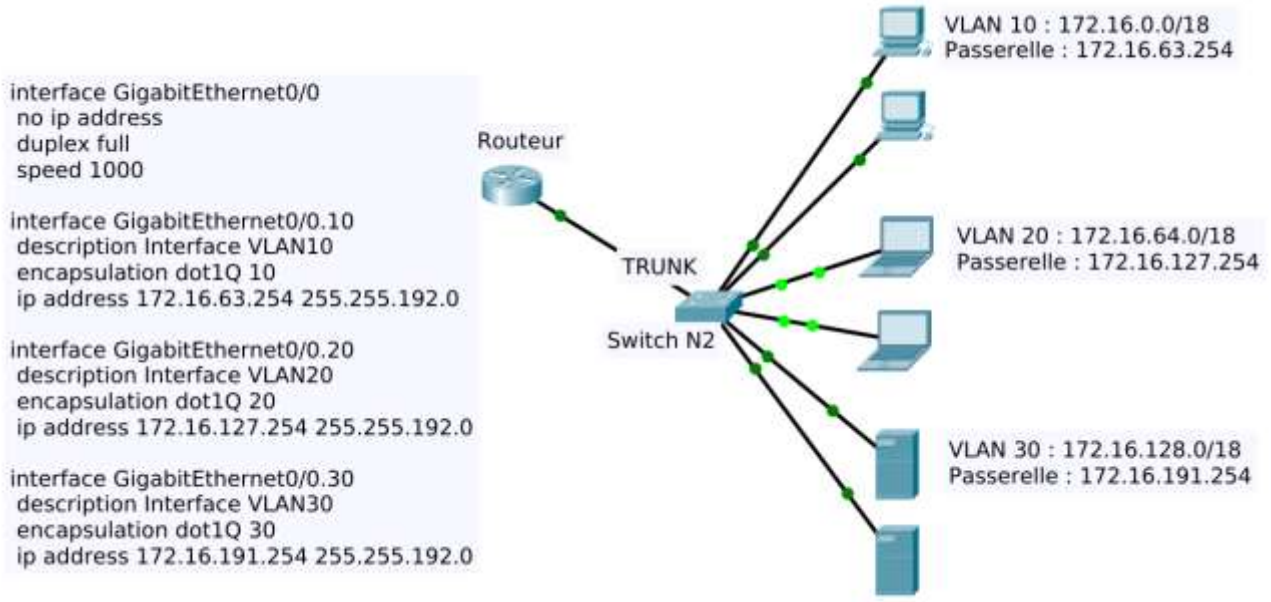
- D'améliorer la gestion du réseau ;
- D'optimiser bande passante ;
- De séparer les flux ;
- De segmenter le réseau donc de réduire la taille d'un domaine de broadcast (diffusion) ;
- De créer un ensemble logique isolé pour améliorer la sécurité.

L'unique moyen pour communiquer entre des hosts (hôtes) appartenant à des VLAN différents est de passer par un équipement qui sait router :

- Un switch de niveau 3,
- Un routeur.

III.3.2 Configuration du routage InterVLAN et de la jonction ISL/802.1Q sur un switch L2

Un switch de niveau 2 (L2 - Layer 2) est capable de gérer des VLAN mais il ne sait pas assurer le routage des paquets entre ces derniers. Normal, puisque c'est un équipement de couche 2 et que le domaine de routage se situe en couche 3. La solution consiste donc à utiliser un routeur.



Vous constatez que le routeur à un seul lien vers le switch dans lequel passent les trames de tous les VLAN. Le port du switch est un port « trunk ».

Pour préserver l'appartenance aux VLAN de chaque trame, elles sont taguées (marquées) pour que les équipements réseaux sachent à quel VLAN elles appartiennent. Ceci est accompli en encapsulant chaque trame de façon à conserver son numéro de VLAN.

Cisco Inter-Switch Link (ISL) est un protocole propriétaire de Cisco qui permet de transférer des trames Ethernet avec leur numéro de VLAN entre deux commutateurs Ethernet ou bien entre un commutateur et un routeur. Ce dernier a précédé le standard IEEE 802.1Q (dot1q).

NOTE :

Cisco favorise désormais le protocole standard IEEE 802.1Q dans ses équipements.

Côté routeur, nous avons configuré une seule interface (GigabitEthernet0/0) du routeur avec des paramètres communs :

- Interface GigabitEthernet0/0
- No ip address
- Duplex full
- Speed 1000

Puis, il faut créer une interface virtuelle par VLAN. Son nom est composé ainsi :

« nom de l'interface physique »	« un point »	« un identifiant numérique »
GigabitEthernet0/0	.	10

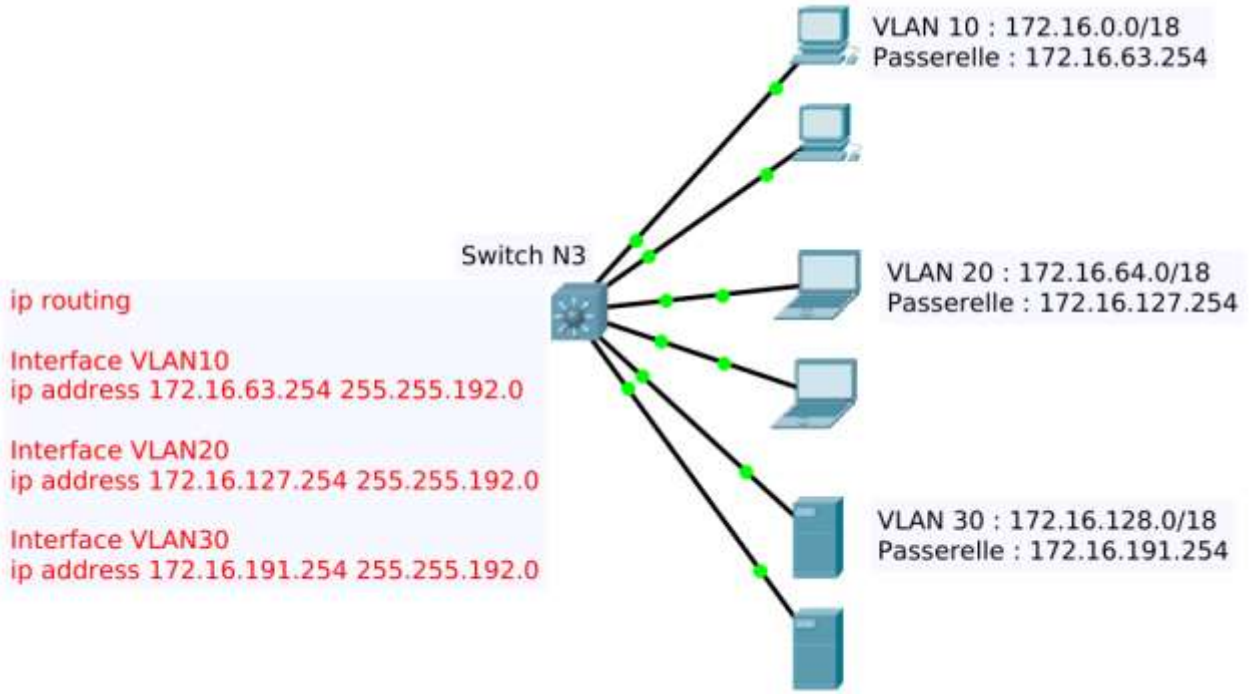
Il suffit de reporter le numéro du VLAN pour l'identifiant numérique pour éviter la confusion.

La directive « description » permet d'écrire un commentaire tandis que la directive « encapsulation dot1Q » permet de marquer les trames.

- Interface GigabitEthernet0/0.10
- Description Interface du VLAN 10
- Encapsulation dot1Q
- IP adresse 172.16.63.254 255.255.192.0

III.3.3 Configuration du routage InterVLAN sur un switch L3

Le switch de niveau 3 intègre le domaine de routage.



Il faut créer une interface de VLAN qui porte le nom de celui-ci et lui affecter une adresse IP. Cette dernière servira de passerelle par défaut.

Contrairement au routeur, le routage doit être activé avec la commande : « ip routing ».

III.4. ROUTAGE INTER-LAN

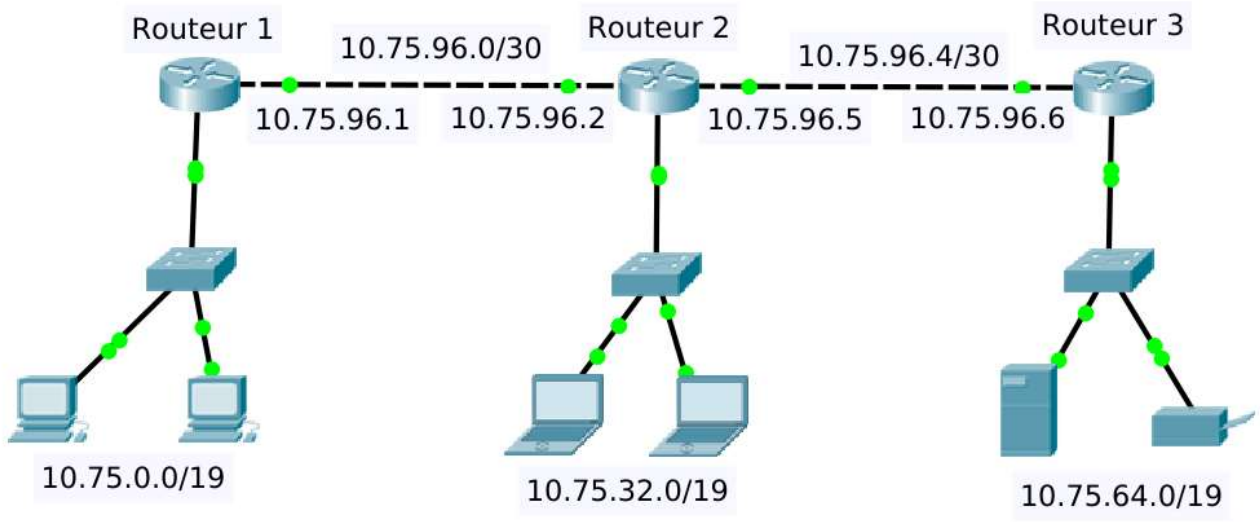
III.4.1 Routage statique

Le routage statique consiste à remplir manuellement les tables de routage. Il est utilisé sur de petits réseaux ou sur des réseaux d'extrémité.

La commande du système IOS de Cisco est :

- ip route « réseau destination » « masque du réseau destination » « IP routeur voisin »
- ip route 192.168.1.0 255.255.255.0 10.255.255.254

Exemple :



Nous avons 5 segments réseaux.

Pour que les machines des 3 segments réseaux 10.75.0.0/19, 10.75.32.0/19 et 10.75.64.0/19 puissent communiquer entre elles, il faut ajouter les routes suivantes :

Routeur 1 :
ip route 10.75.32.0 255.255.224.0 10.75.96.2
ip route 10.75.64.0 255.255.224.0 10.75.96.2

Routeur 2 :
ip route 10.75.0.0 255.255.224.0 10.75.96.1
ip route 10.75.64.0 255.255.224.0 10.75.96.6

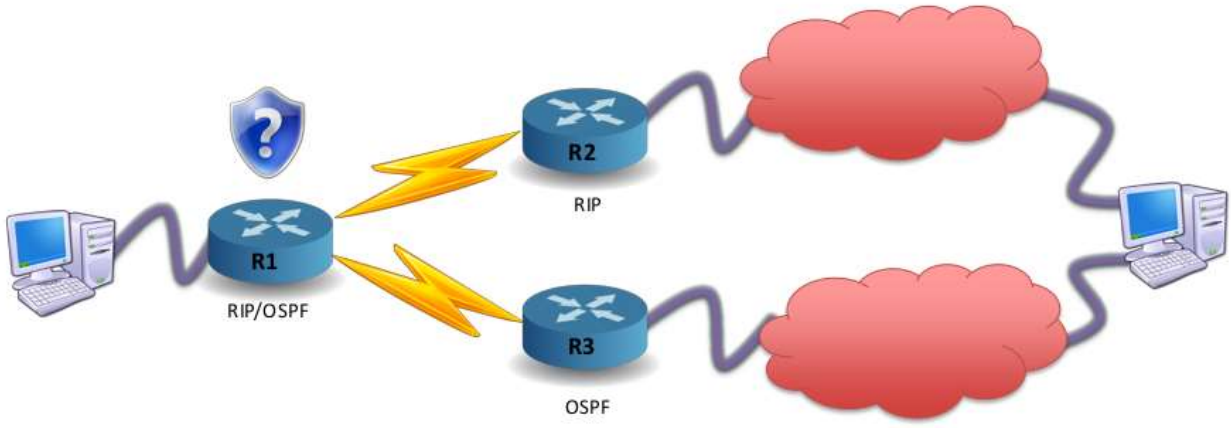
Routeur 3 :
ip route 10.75.0.0 255.255.224.0 10.75.96.5
ip route 10.75.32.0 255.255.224.0 10.75.96.5

Si nous ajoutons un quatrième routeur, il faudra ajouter de nouvelles routes.

III.4.2 Routage dynamique

Les routeurs sont susceptibles d'apprendre des routes pour une même destination provenant de sources différentes. Ces dernières peuvent avoir été apprises manuellement (routage statique) ou par un protocole de routage (routage dynamique).

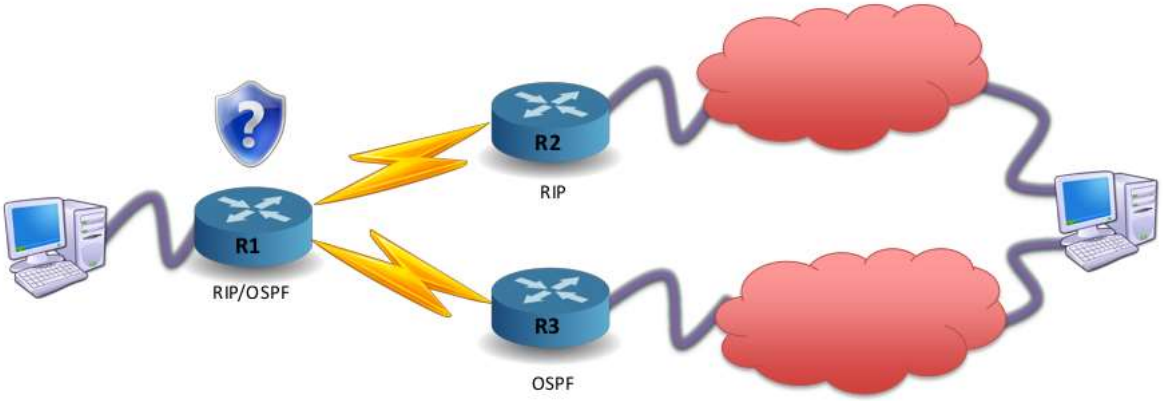
Pour résoudre ce dilemme, il est attribué à chaque source de routes un indice de confiance appelé « distance administrative ».



Plus cette distance est faible, plus la source est de confiance.

Source	Distance administrative
Interface connectée	0
Route statique	1
EIGRP	90
IGRP	100
OSPF	110
RIP (v1ou v2)	120
« External » EIGRP	170
Inconnue	255

S'il reste plusieurs routes candidates, le routeur va alors sélectionner la ou les routes présentant la plus faible métrique.



La métrique est un coût associé à chaque route. Chaque protocole de routage a son propre algorithme pour calculer la métrique : pour certains il s'agit du nombre de routeurs à traverser pour atteindre la destination, pour d'autres d'une fonction de la bande passante...

La distance administrative et la métrique permettent donc aux routeurs de sélectionner les meilleures routes pour chaque destination.

III.4.3 RIP

RIP (Routing Information Protocol) est un protocole de routage IP à vecteur de distances (Vector Distance) s'appuyant sur l'algorithme de détermination des routes décentralisé Bellman-Ford. Il permet à chaque routeur de communiquer aux routeurs voisins la métrique c'est-à-dire la distance qui les sépare d'un réseau IP déterminé quant au nombre de sauts (hops).

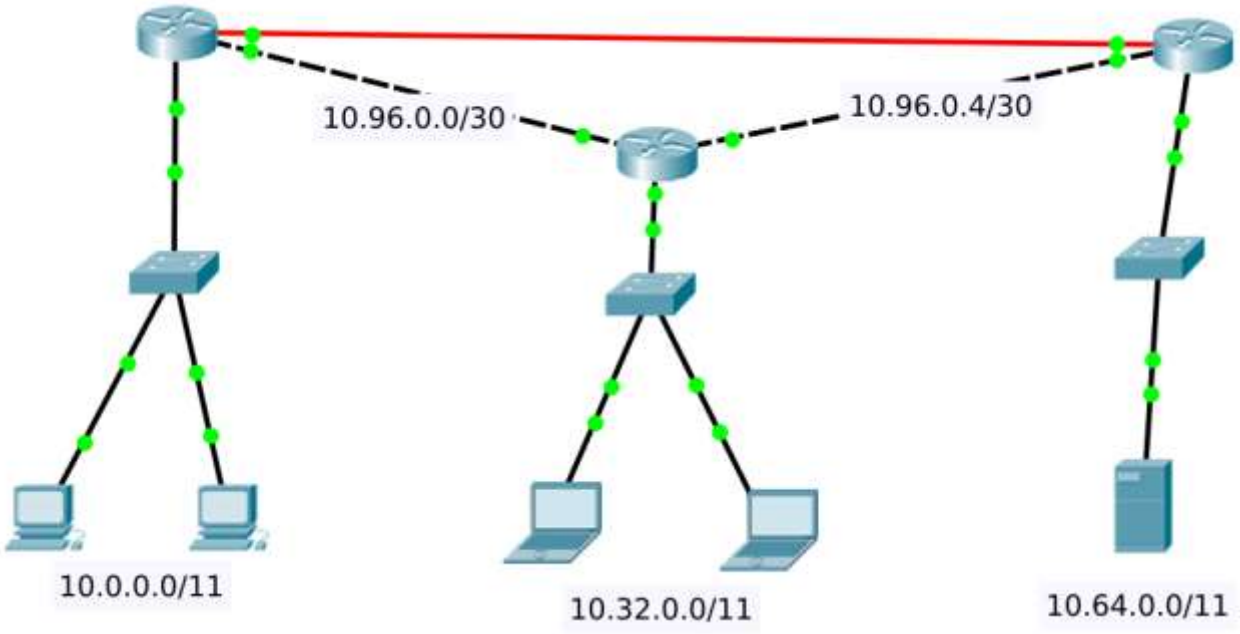
Pour chaque réseau IP connu, chaque routeur conserve l'adresse du routeur voisin dont la métrique est la plus petite. Ces meilleures routes sont diffusées toutes les 30 secondes.

Pour éviter les boucles de routage, le nombre de sauts est limité à 15. Au-delà, les paquets sont supprimés.

Il existe 2 versions :

- RIPv1, défini dans la RFC 1058, est un protocole classful. Cette version ne prend pas en charge les masques de sous-réseau de longueur variable ni l'authentification des routeurs. Les routes sont envoyées en broadcast.
- RIPv2, défini dans la RFC 2453, est un protocole classless. Cette version a été conçue pour répondre aux contraintes des réseaux actuels telles que les découpages des réseaux IP en sous-réseaux, l'authentification par mot de passe, etc. Les routes sont envoyées à l'adresse multicast 224.0.0.9.

Exemple :

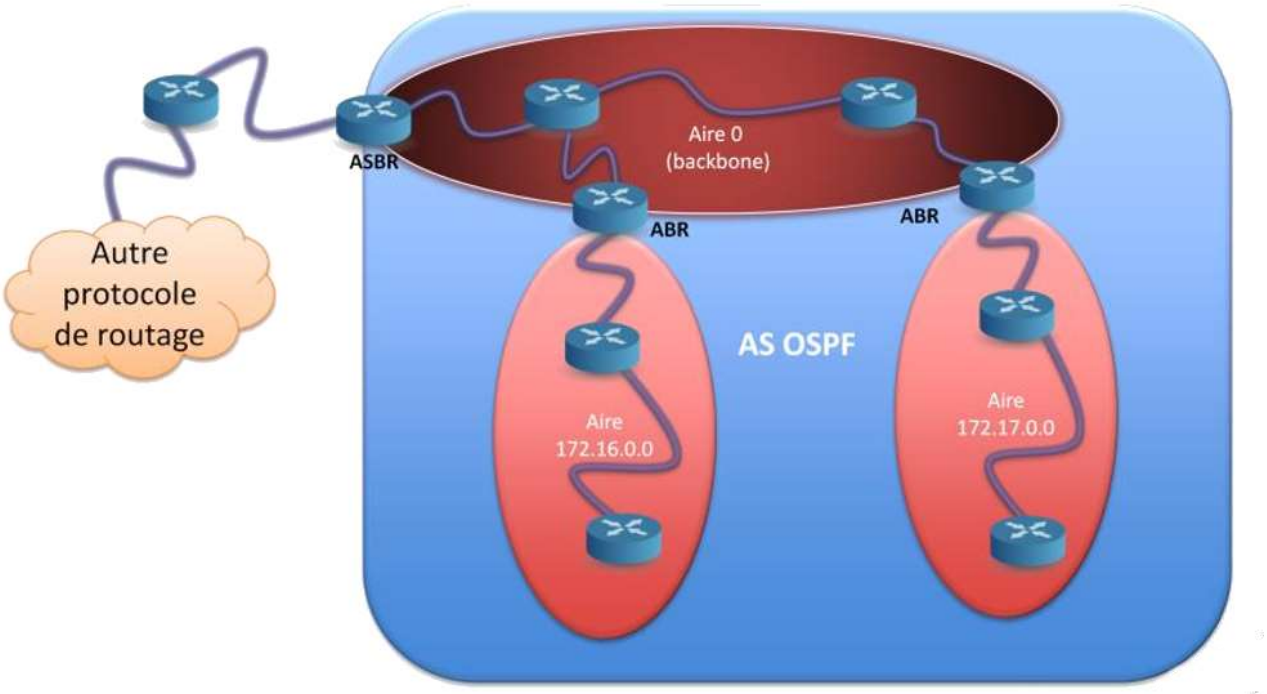


Sur chaque routeur, nous devons taper les commandes suivantes :

router rip	Active le routage dynamique RIP
Version 2	Précise la version du protocole
network 10.0.0.0	Alimente la table de routage avec les routes qui concerne le réseau 10.0.0.0

III.4.4 OSPF

OSPF (Open Shortest Path First) est un protocole de routage IGP de type « à état de liens » qui a été développé au sein de l'IETF (Internet Engineering Task Force) à partir de 1987. La version actuelle d'OSPF version 2 est décrite dans la RFC 2328 en 1997. Une version 3 est définie depuis 2008 dans la RFC 5340 qui permet, en autre, l'utilisation d'OSPF dans un réseau IPv6.

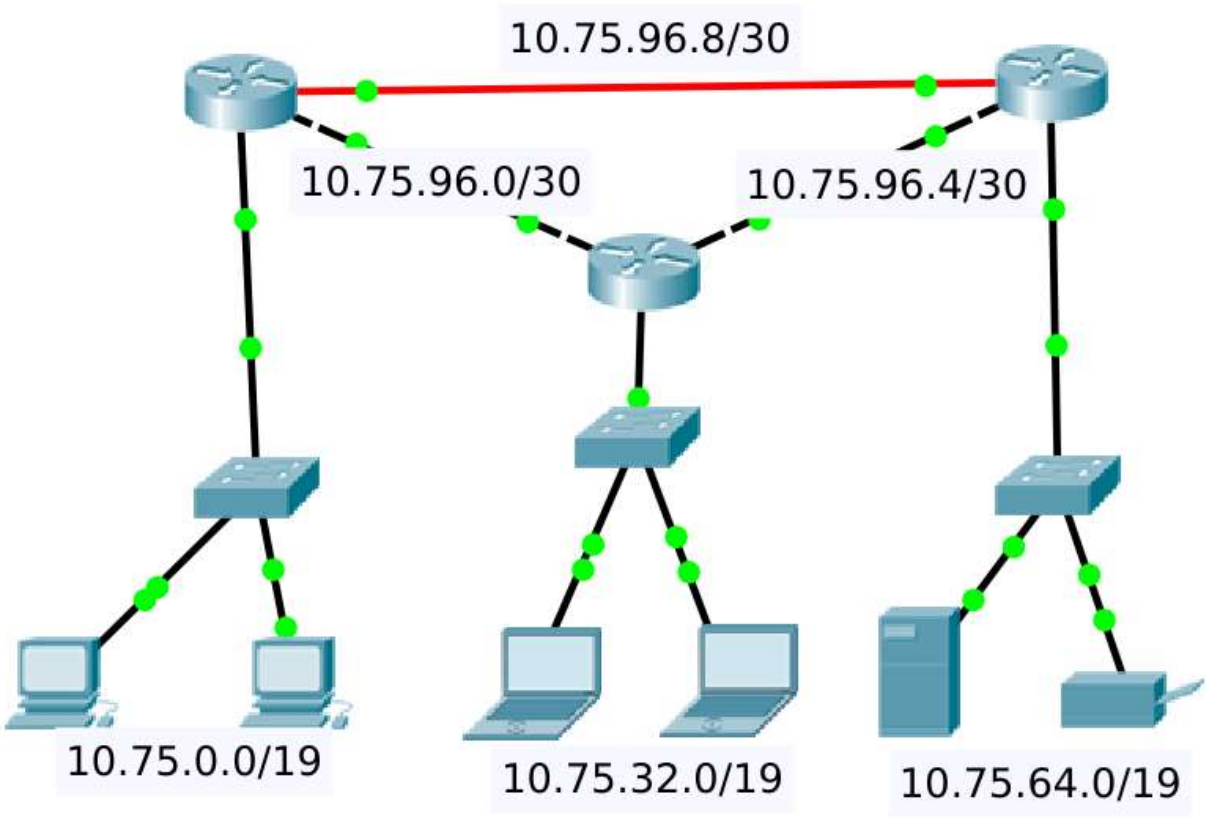


Afin d'éviter qu'un problème survenant sur un lien ait un impact sur l'ensemble des routeurs, le domaine de routage OSPF peut être subdivisé en Aires.

Dans une aire, les routeurs partagent la même base de liens. Par contre, les échanges effectués entre les aires ne sont pas constitués des états des liens, mais de résumés. C'est-à-dire d'adresses destinations, des masques associés et des métriques. C'est le rôle des routeurs se trouvant en bordure d'aire (ABR) d'émettre des résumés.

Chaque routeur établit des relations d'adjacence avec ses voisins immédiats en envoyant des messages « hello » à intervalle régulier. Chaque routeur communique ensuite la liste des réseaux auxquels il est connecté par des messages LSA (Link-state advertisements) propagés de proche en proche à tous les routeurs du réseau. L'ensemble des LSA forme une base de données de l'état des liens LSDB (Link-State Database) pour chaque aire, qui est identique pour tous les routeurs participants dans cette aire. L'algorithme de Dijkstra, SPF (Shortest Path First), est ensuite utilisé par les routeurs pour déterminer la route la plus rapide vers chacun des réseaux connus dans la LSDB.

Exemple :



Sur chaque routeur, nous devons taper les commandes suivantes :

router ospf 1	Active le routage dynamique OSPF La valeur 1 identifie le processus
log-adjacency-changes	Configure le routeur pour envoyer un message « syslog » quand l'état d'un voisin change
network 10.75.0.0 0.0.255.255 area 0	Alimente la table de routage avec les routes qui concerne le sous-réseau 10.75.0.0 de l'aire 0. Le wildcard 0.0.255.255 permet de masquer les 2 derniers octets de l'adresse IP, à savoir la partie hosts. Cela permet d'identifier la partie réseau.

III.5. NETWORK ADDRESS TRANSLATION

III.5.1 Principe

Le NAT (Network Address Translation) est un mécanisme permettant de cacher la véritable adresse IP d'une ou plusieurs machines. Ainsi, les adresses visibles à l'extérieur du réseau ne sont pas les véritables adresses.

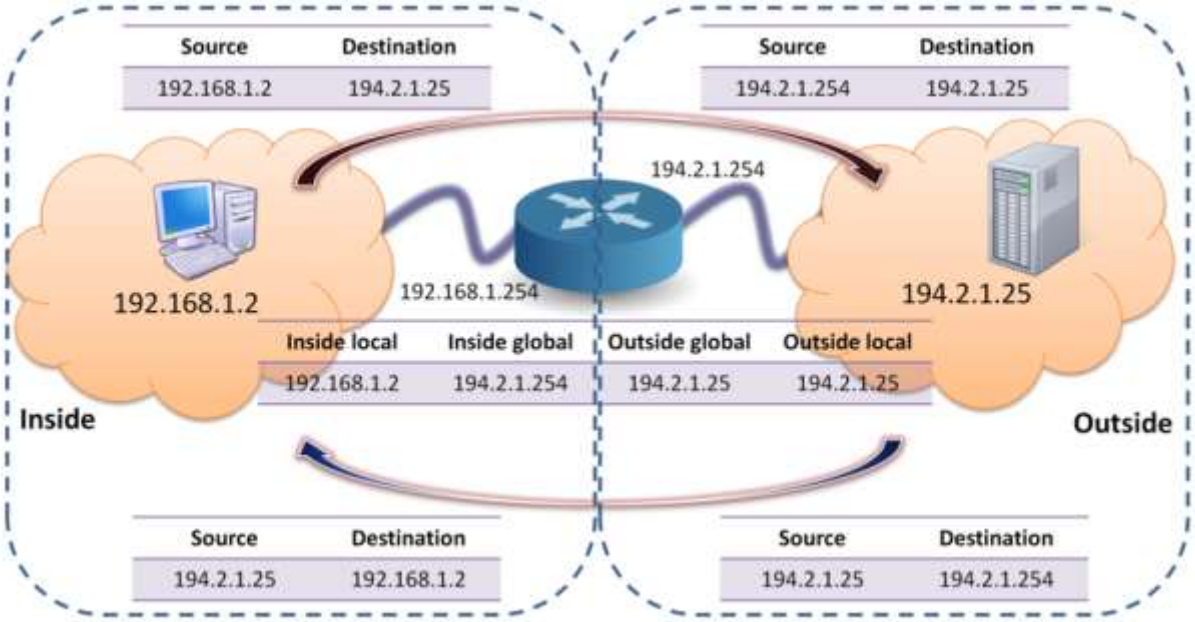
Il existe plusieurs motivations à l'utilisation du NAT :

- La structure interne du réseau est complètement cachée : des machines se trouvant sur des sous-réseaux différents sont vues à l'extérieur comme appartenant au même sous réseau. Il est possible de changer la véritable adresse de la machine sans que cela ait un impact pour les clients externes.
- La plupart des entreprises utilisent pour leurs réseaux internes des adresses issues de la RFC 1918 qui définit les adresses IP privées. Ces adresses ne sont pas routables sur le réseau public Internet. Il est donc nécessaire de traduire (translater) les adresses IP privées en adresses publiques.

Nous pouvons distinguer trois catégories de NAT :

- NAT statique
- NAT dynamique
- PAT (Port Address Translation)

III.5.2 Les termes Cisco

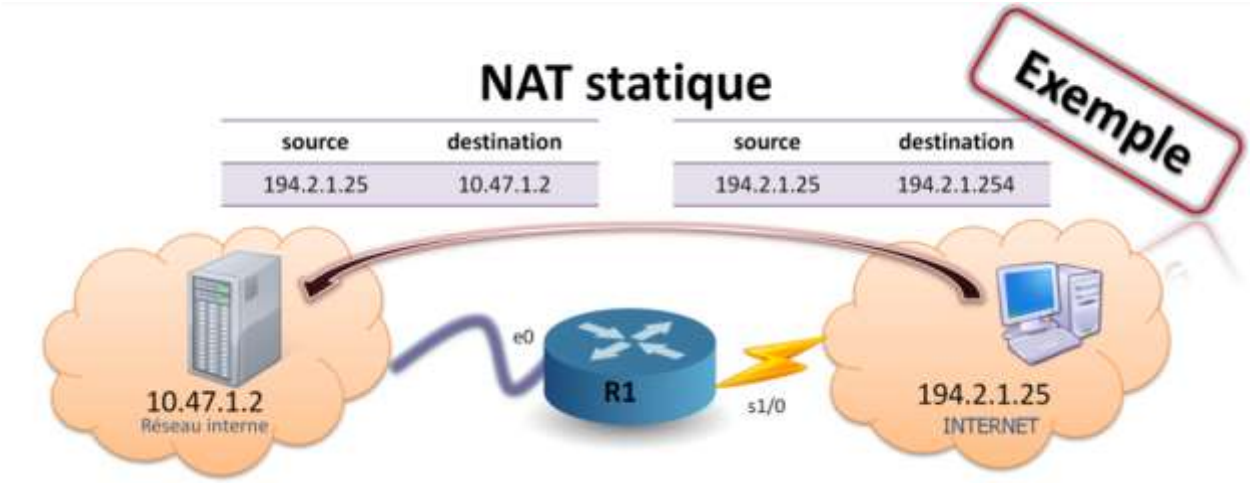


La translation d'adresse s'effectue généralement entre un réseau local appartenant à une entreprise et un réseau externe. Les termes utilisés par Cisco pour définir les réseaux et les adresses sont les suivants :

- **inside** : réseaux appartenant à l'entreprise dont les adresses doivent être translatées. Les machines de ces réseaux apparaîtront en interne avec des adresses IP appartenant à un espace d'adressage appelé **LOCAL**. Leurs adresses du côté **outside** appartenant à l'espace d'adressage **GLOBAL**.
- **outside** : Réseaux avec lesquels les réseaux inside doivent communiquer.

La translation aura toujours lieu entre une interface dite « inside » et une interface « outside ».

III.5.3 NAT statique

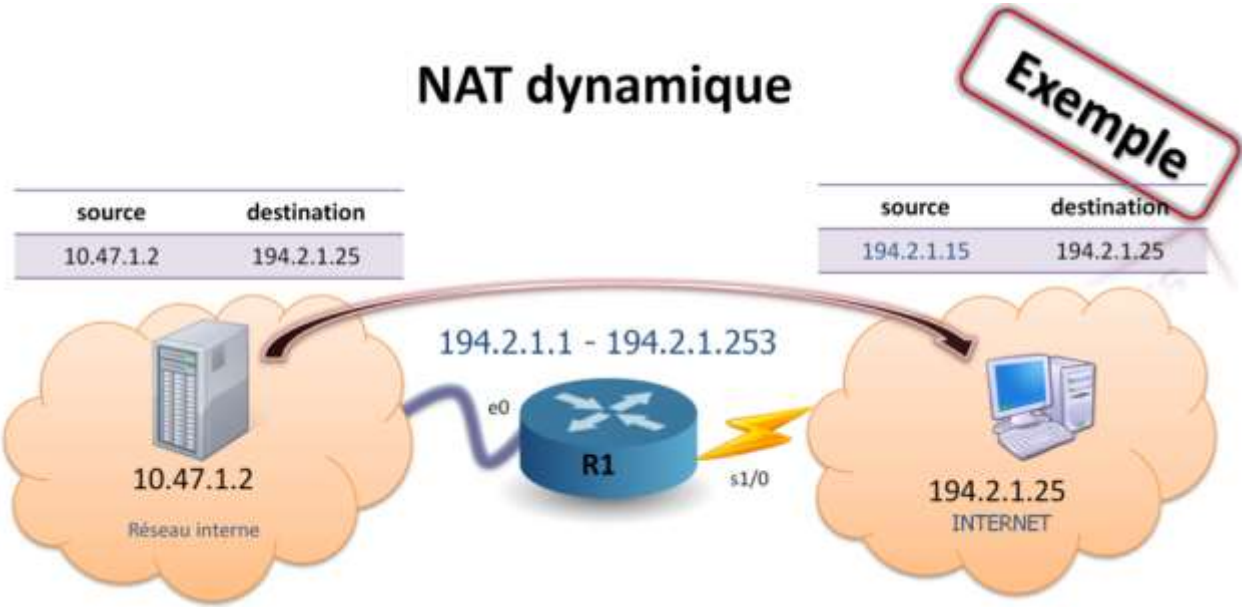


A une adresse IP interne correspond une et une seule adresse IP externe. Cette correspondance est définie par l'administrateur dans la configuration du routeur. Cela permet à des machines externes d'ouvrir des connexions vers les machines internes. Le NAT statique est généralement utilisé pour les serveurs.

Dans le routeur, nous devons écrire :

interface e0 ip address 10.47.1.254 255.255.255.0 ip nat inside	Déclaration de l'interface du réseau local en tant que « inside »
interface s1/0 ip address 194.2.1.254 255.255.255.0 ip nat outside	Déclaration de l'interface publique en tant que «outside»
ip nat inside source static 10.47.1.2 194.2.1.254	Translation de l'adresse ip 10.47.1.2 en 194.2.1.254

III.5.4 NAT dynamique



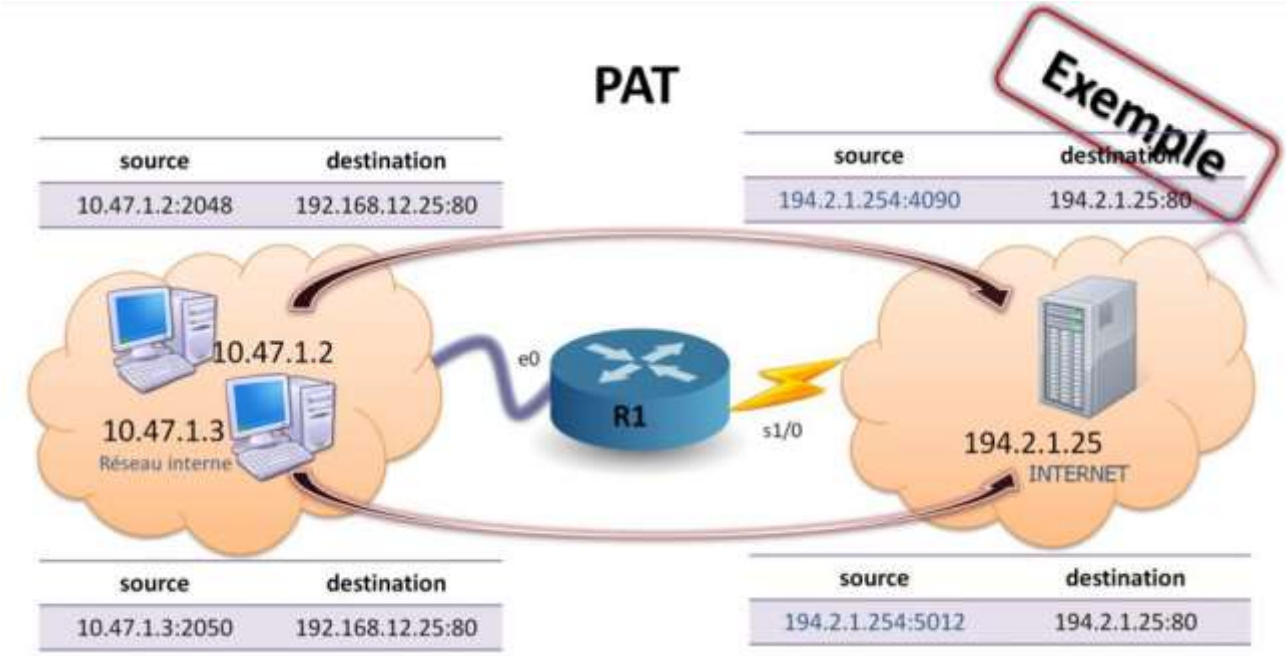
Lorsqu'une machine interne désire ouvrir une connexion vers l'extérieur, une adresse IP officielle libre est prise dans un pool d'adresses. Cette correspondance restera tant que la machine interne dialoguera avec l'extérieur.

Par exemple, si le pool contient 10 adresses IP publiques, seules 10 hôtes du réseau local pourront dialoguer simultanément avec l'extérieur.

Dans le routeur, nous devons écrire :

interface e0 ip address 10.47.1.254 255.255.255.0 ip nat inside	Déclaration de l'interface du réseau local en tant que « inside »
interface s1/0 ip address 194.2.1.254 255.255.255.0 ip nat outside	Déclaration de l'interface publique en tant que «outside»
ip nat pool adrpub 194.2.1.10 194.2.1.15 netmask 255.255.255.0	Définition du pool d'adresses IP publiques nommé adrpub
access-list 10 permit 10.47.1.0 0.0.0.255	Définition d'une ACL (Access Control List) pour autoriser la translation des adresses IP du réseau 10.47.1.0/24
ip nat inside source list 10 pool adrpub	Définition du NAT dynamique basé sur l'ACL 10 et utilisant le pool adrpub

III.5.5 PAT



Il permet d'économiser les adresses IP publiques car une seule adresse IP publique est nécessaire. Plusieurs machines du réseau local peuvent utiliser une seule adresse IP publique à un instant T.

Dans le routeur, nous devons écrire :

interface e0 ip address 10.47.1.254 255.255.255.0 ip nat inside	Déclaration de l'interface du réseau local en tant que « inside »
interface s1/0 ip address 194.2.1.254 255.255.255.0 ip nat outside	Déclaration de l'interface publique en tant que «outside»
access-list 10 permit 10.47.1.0 0.0.0.255	Définition d'une ACL (Access Control List) pour autoriser la translation des adresses IP du réseau 10.47.1.0/24
ip nat inside source list 10 interface s1/0 overload	Définition du PAT basé sur l'ACL 10. Nous avons mentionné l'interface s1/0 plutôt que son adresse IP. Cette dernière peut changer si elle est fournie par un serveur DHCP.

IV. DEPANNAGE

Objectifs

- Étudier les pannes classiques
- Utiliser la commande adaptée au problème
- Capture et analyse de trame avec Wireshark

IV.1. PROBLEME AU NIVEAU DE LA COUCHE PHYSIQUE

IV.1.1 Symptômes des problèmes

Voici les principaux symptômes :

- Performances inférieures à la ligne de base

Il est recommandé pour une bonne gestion de réseau d'effectuer la surveillance et de constituer une documentation sur le réseau. L'un des moyens les plus efficaces pour surveiller les performances d'un réseau et le dépanner consiste à établir une ligne de base du réseau.

Une ligne de base est un processus permettant d'étudier le réseau à intervalles réguliers pour s'assurer qu'il fonctionne comme prévu. Elle implique le contrôle des performances et du comportement du réseau sur une période donnée permettant l'établissement d'un point de référence pour tout contrôle ultérieur des performances.

Des logiciels comme Nagios, Shinken, Icinga, Munin, etc. permettent de recueillir les informations.

- Perte de connectivité

Les problèmes de connectivité réseau ont différentes causes mais elles se produisent généralement en raison des cartes réseaux incorrectes, des paramètres de commutateurs inappropriés, du matériel défectueux ou des problèmes de pilote. Certains problèmes de connectivité sont intermittents et ne pointent pas clairement à l'une de ces causes.

Les causes courantes des problèmes de connectivité sont les suivantes :

- Les cartes réseau et les ports de commutateur n'est pas de niveau de duplex ou les paramètres de vitesse de transfert.
- Les cartes réseau ou les commutateurs avec transmission des taux de 10/100 mégabits par seconde (Mbits/s) ne commutent pas correctement. Certains paramètres de détection automatique ne peuvent pas détecter correctement la vitesse de certaines cartes réseau.
- La carte réseau est incompatible avec la carte mère ou avec d'autres composants ou pilotes matériel ou logiciels.
- Nombre de collisions élevé

Cela concerne uniquement un réseau Ethernet half duplex puisque dans un réseau Ethernet en full-duplex, la détection de collision est désactivée.

Une collision est le mécanisme utilisé par Ethernet pour contrôler l'accès et pour allouer la bande passante partagée entre les stations qui veulent transmettre en même temps sur un support partagé. Puisque le support est partagé, un mécanisme doit exister selon lequel deux stations peuvent détecter qu'elles veulent transmettre en même temps. Ce mécanisme est la détection de collision. L'Ethernet utilise CSMA/CD (Carrier Sense Multiple Access/Collision Detect) comme méthode de détection de collision.

Les collisions sont une manière de distribuer la charge de trafic dans le temps par l'accès arbitraire au support partagé. Les collisions ne sont pas mauvaises ; elles sont essentielles pour corriger le fonctionnement d'Ethernet.

La quantité maximale de tranches de temps est limitée à 1024 et la quantité maximale de retransmissions pour la même trame dans le mécanisme de collision est 16. Si elle échoue 16 fois consécutives, elle est comptée comme collision excessive.

- **Goulots d'étranglement (bottleneck) sur le réseau**

Un de goulot d'étranglement de la bande passante nuit aux performances du réseau. Il survient lorsqu'il n'y a pas assez de bande passante disponible pour que tous les paquets de données soient acheminés sur le réseau en temps opportun. Les goulots d'étranglement ainsi formés peuvent provoquer une lenteur anormale et un manque de réactivité chez certaines applications.

Un bon outil de gestion de la bande passante vous indique combien le système en consomme, détecte les embouteillages réseau et vous aide à identifier les ressources à mettre en œuvre pour que la capacité de la bande passante corresponde à vos besoins effectifs.

PRTG Network Monitor, par exemple, est un logiciel de surveillance du réseau et de la bande passante qui s'appuie sur SNMP (Simple Network Management Protocol). Cet outil permet à l'utilisateur d'observer la consommation de bande passante Ethernet, de contrôler le trafic réseau et de surveiller la disponibilité, le temps de fonctionnement et les performances des composants réseau.

MRTG (Multi Router Traffic Grapher) est un logiciel développé sous licence GNU/GPL Ce logiciel permet de surveiller et mesurer le trafic réseau. Il utilise le protocole SNMP pour interroger des équipements réseaux tels que des routeurs, commutateurs, ou bien encore serveurs, disposant d'une MIB.

- **Forte utilisation du CPU**
  - Vérifiez qu'il n'y a pas de problème de sécurité. Généralement, l'utilisation élevée du CPU est provoquée par un problème de sécurité, tel qu'un ver ou un virus présent dans votre réseau. C'est d'autant plus probable s'il n'y a pas eu de changements récents apportés à votre réseau.
  - Assurez-vous que toutes les commandes de débogage de votre routeur sont arrêtées en lançant les commandes **undebug all** ou **no debug all**.
  - Collectez périodiquement la sortie de la commande **show process cpu** qui montre si l'utilisation élevée du CPU est provoquée par des interruptions ou par un certain processus.
  - Messages d'erreur sur la console
  - Relevez le message affiché et le type d'équipement. Ensuite, il faut consulter sur le site Web du constructeur la signification de ce message dans la base de connaissance.

IV.1.2 Causes des problèmes

Les causes les plus courantes sont les problèmes d'alimentation, des défauts matériels, des défauts de câblage, une atténuation ou du bruit. Cela peut être également des erreurs de configuration de l'interface, un dépassement des limites de conception ou une surcharge du CPU.

IV.1.3 Dépannage de la couche physique (couche 1)

Vérifiez que :

- Les câbles ou les connexions sont corrects,
- La norme de câblage en vigueur est respectée sur l'ensemble du réseau,
- Le câblage des périphériques est correct,
- Les configurations d'interfaces sont correctes,
- Consultez les statistiques d'utilisation et les taux d'erreurs des données.

IV.2. PROBLEME AU NIVEAU DE LA COUCHE LIAISON

IV.2.1 Symptômes des problèmes

Les symptômes les plus courants sont :

- Pas de fonctionnalité ni de connectivité au niveau de la couche réseau ou à un niveau supérieur,
- Les performances réseau sont inférieures à la ligne de base,
- Nombre excessif de diffusions (broadcasts),
- Messages sur la console.

IV.2.2 Causes des problèmes

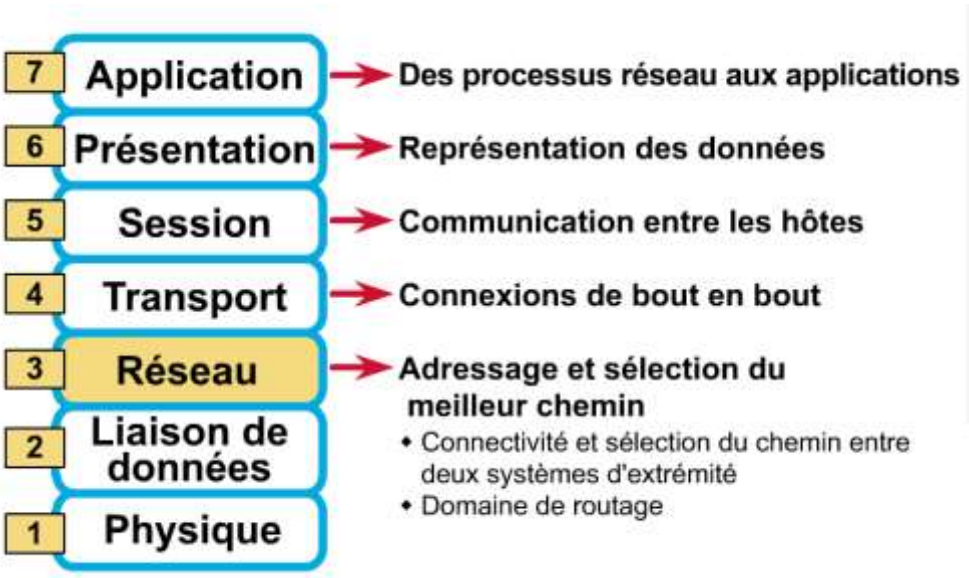
Vous pouvez rencontrer des erreurs d'encapsulation, des erreurs de mappage d'adresses, des erreurs de trames ou des boucles STP (Spanning Tree Protocol).

IV.2.3 Dépannage de la couche liaison (couche 2)

Dans le modèle OSI ou TCP/IP, la couche de liaison de données est juste au-dessus de la couche physique. À cette couche, vous devriez vérifier les éléments suivants :

- Lien statut : la plupart des systèmes d'exploitation offrent une sorte d'outils pour vérifier l'état de la liaison. Assurez-vous que la carte réseau est activée dans le système d'exploitation.
- Vérifiez les voyants d'état : la plupart des cartes réseau ont un ou deux voyants. Ces lumières vont identifier l'état de la liaison et de la vitesse ainsi que l'activité sur le lien.
- Essayer d'utiliser Address Resolution Protocol (ARP): testez pour voir si d'autres ordinateurs du réseau sont visibles à tous. ARP est un outil de ligne de commande qui permet d'afficher les adresses MAC de tous les appareils avec lesquels vous avez récemment communiquées. Cela signifie que vous pouvez tenter de communiquer avec un autre appareil, puis vérifiez votre cache ARP.
- Si vous pouvez voir d'autres ordinateurs, alors votre connexion à la couche de liaison de données est susceptible fonctionne correctement.

IV.3. PROBLEME AU NIVEAU DE LA COUCHE RESEAU



IV.3.1 Configuration IP dynamique

Vous pouvez vérifier :

- Si l'adresse IP commence par 169.254 avec un masque de sous-réseau /16 alors vous avez obtenu une adresse APIPA.
- Si la configuration n'est pas obtenue par un serveur DHCP, vérifier les paramètres de ce dernier.
- Vérifier aussi si le serveur DHCP est sur le même segment réseau ou si un relai DHCP est nécessaire.

IV.3.2 Configuration IP statique

Si vous avez saisi la bonne adresse IP et le bon masque de sous-réseau.

Si vous n'avez pas un conflit d'adresse IP avec un autre ordinateur.

Si vous avez la bonne adresse de passerelle par défaut (gateway).

IV.3.3 Dépannage de la couche réseau (couche 3)

Avec la commande **ARP** (Address Resolution Protocol), vous pouvez vérifier la traduction d'une adresse IP en une adresse MAC.

Les commandes **ifconfig** pour Unix, **ip address** pour les distributions Linux et **ipconfig** pour Microsoft Windows permettent de consulter la configuration.

La commande **ping** permet de vérifier la connectivité.

Les commandes **tracert** pour Microsoft Windows, **traceroute** pour Unix/Linux trace l'itinéraire des paquets sur le réseau.

IV.4. CAPTURER ET ANALYSER UN TRAFIC RESEAU AVEC WIRESHARK

IV.4.1 C'est quoi une capture réseau ?

Une capture réseau est une photo à un instant T (ou sur une période de temps) de ce qui transite sur un réseau informatique. Si l'on choisit bien le point de capture, on pourra sauvegarder l'ensemble d'un trafic intéressant pour ensuite l'analyser à tête reposée.

IV.4.2 Une capture réseau pour quoi faire ?

Les principales applications d'une capture puis d'une analyse réseau sont :

- Identifier les flux consommateurs sur une liaison Internet, WAN ou locale
- Caractériser le trafic généré par une nouvelle application
- Identifier les causes d'un problème de performance sur un réseau (tempête de broadcast, problème physique sur un équipement...)

Cette liste est bien sûr loin d'être exhaustive…

IV.4.3 Où effectuer la capture réseau ?

C'est un des points critiques. Il faut effectivement se positionner à un endroit où l'on va capturer l'ensemble du trafic « intéressant » par rapport à votre besoin sans pour autant sauver des gigaoctets d'informations.

Pour analyser les flux d'une liaison WAN, le plus simple est de faire la capture au plus proche de l'interface réseau de sortie vers cette liaison. Si cette liaison n'est pas une liaison Ethernet, alors il faudra se positionner en amont du routeur gérant cette liaison.

Pour analyser une nouvelle application, le plus simple est de capturer le trafic sur un des PC clients ou sur le serveur dans le cas d'une architecture client/serveur.

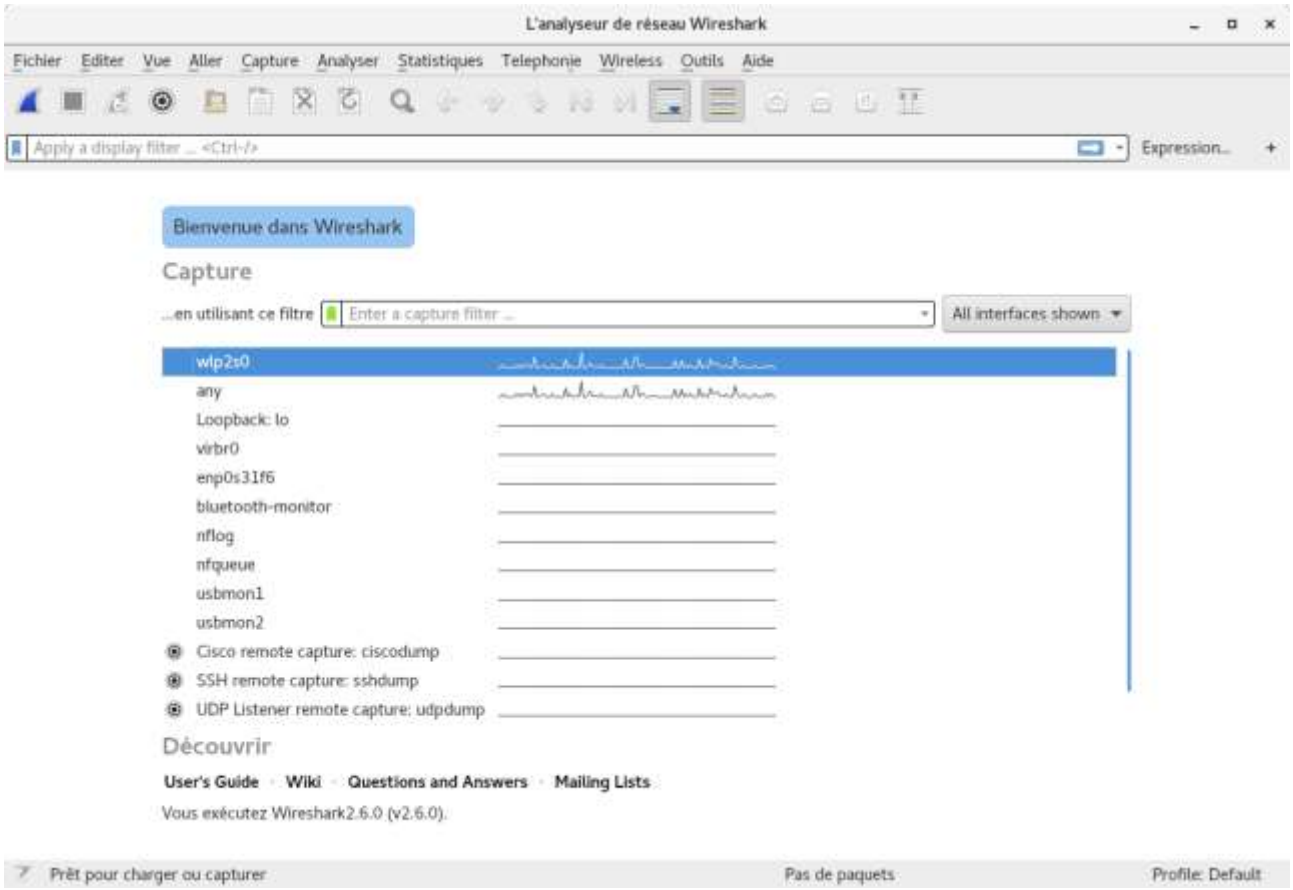
Sur un réseau LAN commuté, vous pouvez soit utiliser les fonctions de redondance de port au niveau d'un commutateur central, soit positionner un bon vieux HUB à un endroit stratégique de votre réseau.

IV.4.4 Comment faire la capture réseau ?

Wireshark (GUI) et tshark (CLI) sont deux logiciels qui permettent la capture de trafic réseau. Le 1er fonctionne sur les principales plateformes informatiques (Windows, Mac Osx, Linux…) tandis que le second est sur Linux.

Exemples :

Lancer une capture avec Wireshark :



Choisissez l'interface sur laquelle vous désirez capturer.

Lancer une capture avec tshark :

```
$ sudo tshark -i eth0 -w ./capture.pcap
```

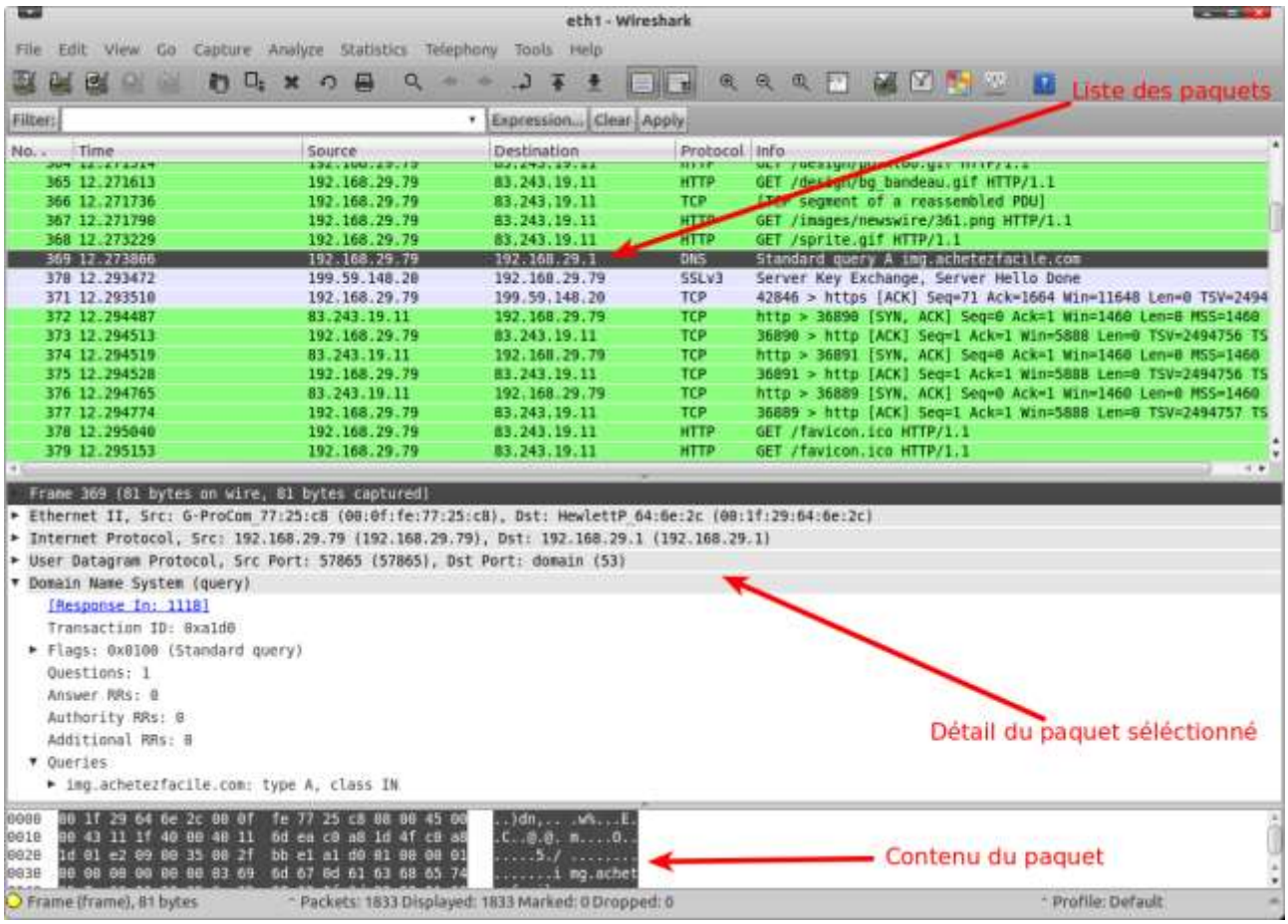
IV.4.5 Comment analyser la capture réseau ?

Dès que la capture est terminée ou bien après avoir chargé la capture par le menu **File**, **Open** puis sélectionner le fichier contenant la capture.

Par exemple : /data/capture.pcap.

La fenêtre de Wireshark est divisée en 3 sections :

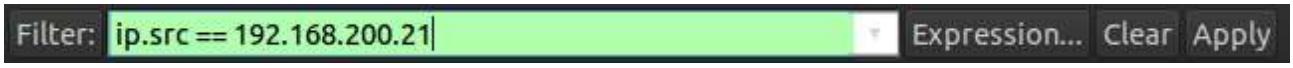
- La première affiche une liste des paquets IP capturés
- La seconde donne le détail du paquet IP sélectionné dans la première section
- La troisième affiche le contenu, en hexadécimal, du paquet IP sélectionné dans la première section



IV.4.6 Appliquer un filtre

Selon votre capture, il peut être utile d'appliquer un filtre qui ne va afficher que certains paquets. Il est également possible d'effectuer ce filtrage lors de la capture mais c'est peu recommandé car nous pourrions louper des informations complémentaires…

Pour mettre en place un filtrage comme « afficher uniquement les paquets dont l'adresse source est 192.168.200.21 » il faut saisir le filtre dans la section **Filter** puis cliquer sur le bouton **Apply** :



La liste exhaustive des filtres d'affichage disponible est disponible sur le site Internet de Wireshark : https://www.wireshark.org/docs/dfref/

Note :

Si le fond de la zone de saisie du filtre n'est pas vert alors vous avez une erreur de syntaxe.

V. TCP/IP : COUCHE TRANSPORT

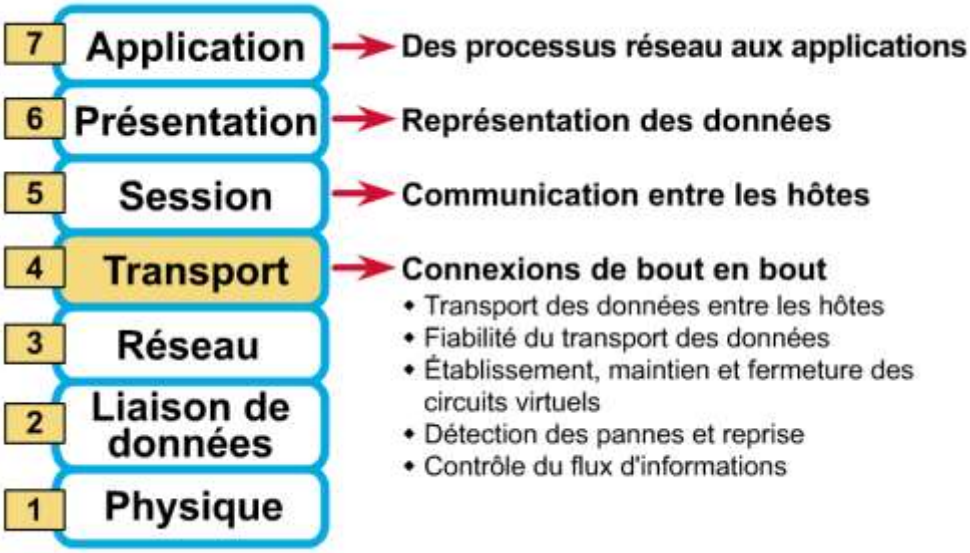
Objectifs

- Présenter les 2 principaux protocoles de la couche 4 : TCP et UDP

Références

- https://tools.ietf.org/html/rfc793
- https://www.iana.org/assignments/service-names-port-numbers/service-names-port-numbers.xhtml
- https://tools.ietf.org/html/rfc768
- http://www.frameip.com/liste-des-ports-tcp-udp/

V.1. TCP ET UDP



V.1.1 Introduction

Comme nous l'avons dit dans le module introduction, il existe plusieurs protocoles dans la couche 4 transport. Nous allons étudier maintenant les 2 principaux protocoles :

- **TCP (Transmission Control Protocol)**

Ce protocole est dit fiable et orienté connexion. Il permet l'acheminement sans erreur de paquets issus d'une machine d'un Internet à une autre machine du même Internet. Son rôle est de fragmenter les messages en paquets à transmettre de manière à pouvoir le faire passer sur la couche 3 (couche réseaux) donc au protocole IP. À l'inverse, sur la machine destination, TCP replace dans l'ordre les paquets transmis sur la couche Internet pour reconstruire le message initial. Il s'occupe également du contrôle de flux de la connexion.

- **UDP (User Datagram Protocol)**

Ce protocole en revanche plus « simple » que TCP. Il est non fiable et sans connexion. Son utilisation présuppose que l'on n'a pas besoin ni du contrôle de flux, ni de la conservation de l'ordre de remise des paquets. Par exemple, il est utilisé lorsque la couche application se charge de la remise en ordre des messages. Une autre utilisation d'UDP est la transmission de la voix, de la vidéo ou de données particulières dont la latence et la taille est faible c'est-à-dire lorsqu'il est nécessaire d'être rapide dans l'envoi de paquets.

V.2. TCP (TRANSMISSION CONTROL PROTOCOL)

V.2.1 Introduction

Défini dans la RFC 793, le protocole TCP fonctionne en 3 phases :

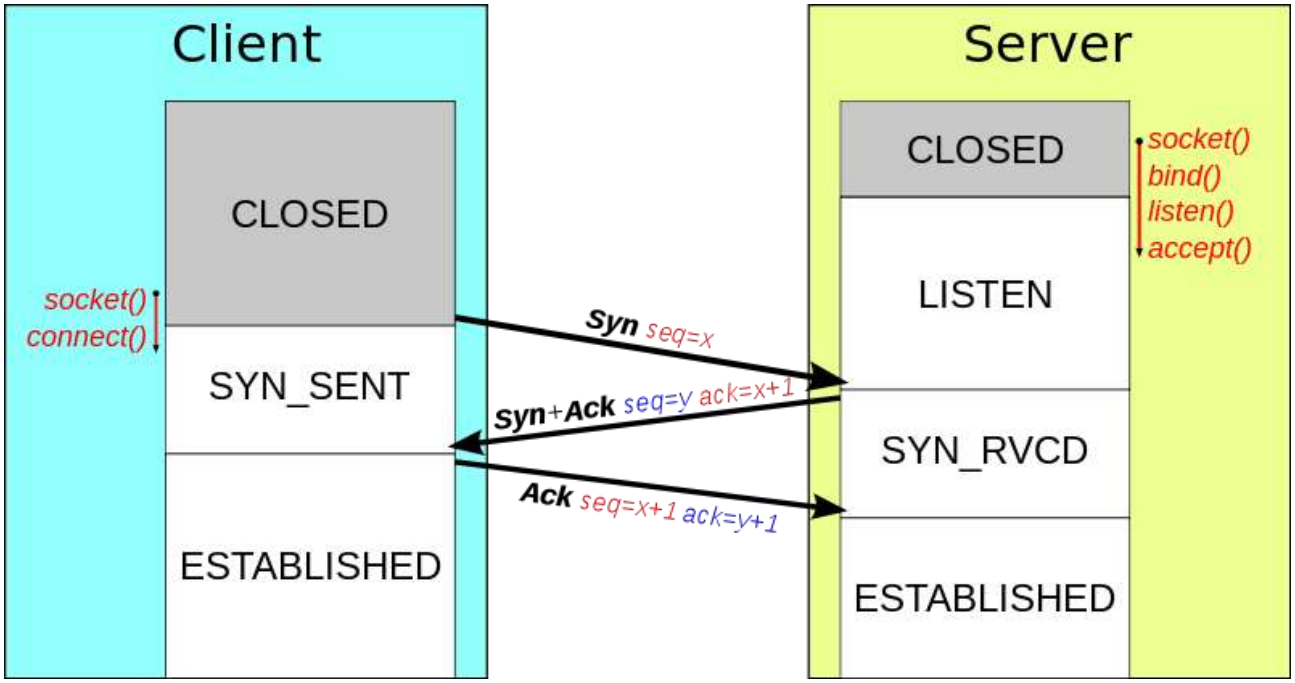
- L'établissement de la connexion ;
- Les transferts de données ;
- La fin de la connexion.

V.2.2 Établissement d'une connexion

L'établissement de la connexion se fait par un handshaking (établissement d'une liaison) en 3 temps.

Même s'il est possible pour 2 systèmes d'établir une connexion entre eux simultanément, dans le cas général, un système ouvre une « socket » (point d'accès à une connexion TCP) et se met en attente passive de demandes de connexion d'un autre système. Ce fonctionnement est communément appelé ouverture passive, et est utilisé par le côté serveur de la connexion. Le côté client de la connexion effectue une ouverture active en 3 temps :

- Le client envoie un segment SYN au serveur,
- Le serveur lui répond par un segment SYN/ACK,
- Le client confirme par un segment ACK.



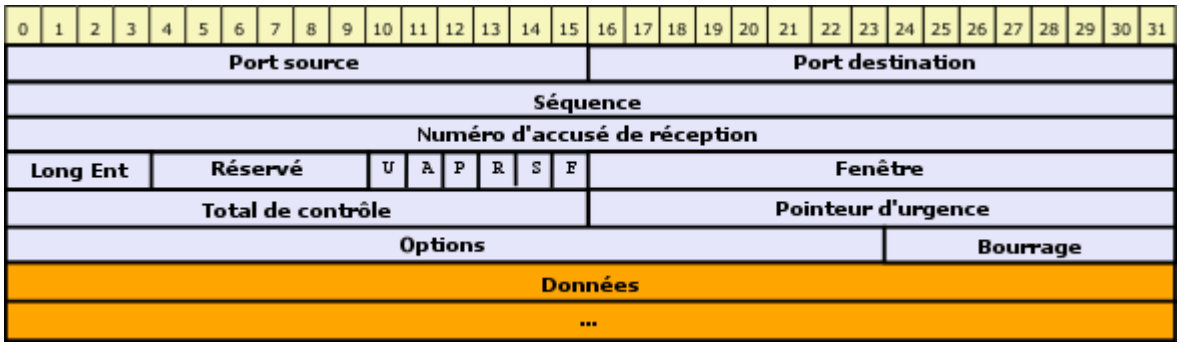
Durant cet échange initial, les numéros de séquence des deux parties sont synchronisés :

- Le client utilise son numéro de séquence initial dans le champ "Numéro de séquence" du segment SYN (x par exemple),
- Le serveur utilise son numéro de séquence initial dans le champ "Numéro de séquence" du segment SYN/ACK (y par exemple) et ajoute le numéro de séquence du client plus un (x+1) dans le champ "Numéro d'acquittement" du segment,

Le client confirme en envoyant un ACK avec un numéro de séquence augmenté de « un » (x+1) et un numéro d'acquittement correspondant au numéro de séquence du serveur plus un (y+1).

V.2.3 Structure d'un segment TCP

Un segment TCP est constitué comme suit :



- Signification des champs

Port Source (16 bits)

Port relatif à l'application en cours sur la machine source

Port Destination (16 bits)

Port relatif à l'application en cours sur la machine de destination

Numéro d'ordre (32 bits)

Lorsque le drapeau SYN est à 0, le numéro d'ordre est celui du premier mot du segment en cours. Lorsque SYN est à 1, le numéro d'ordre est égal au numéro d'ordre initial utilisé pour synchroniser les numéros de séquence (ISN)

Numéro d'accusé de réception (32 bits)

Le numéro d'accusé de réception également appelé numéro d'acquittement correspond au numéro (d'ordre) du prochain segment attendu, et non le numéro du dernier segment reçu.

Décalage des données (4 bits)

Il permet de repérer le début des données dans le paquet. Le décalage est ici essentiel car le champ d'options est de taille variable

Réservé (6 bits)

Champ inutilisé actuellement mais prévu pour l'avenir

Drapeaux (flags) (6x1 bit)

Les drapeaux représentent des informations supplémentaires :

URG

Si ce drapeau est à 1 le paquet doit être traité de façon urgente.

ACK

Si ce drapeau est à 1 le paquet est un accusé de réception.

PSH (PUSH)

Si ce drapeau est à 1, le paquet fonctionne suivant la méthode PUSH.

RST

Si ce drapeau est à 1, la connexion est réinitialisée.

SYN

Le Flag TCP SYN indique une demande d'établissement de connexion.

FIN

Si ce drapeau est à 1 la connexion s'interrompt.

Fenêtre (16 bits)

Champ permettant de connaître le nombre d'octets que le récepteur souhaite recevoir sans accusé de réception.

Somme de contrôle (Checksum ou CRC)

La somme de contrôle est réalisée en faisant la somme des champs de données de l'en-tête, afin de pouvoir vérifier l'intégrité de l'en-tête.

Pointeur d'urgence (16 bits)

Indique le numéro d'ordre à partir duquel l'information devient urgente.

Options (Taille variable)

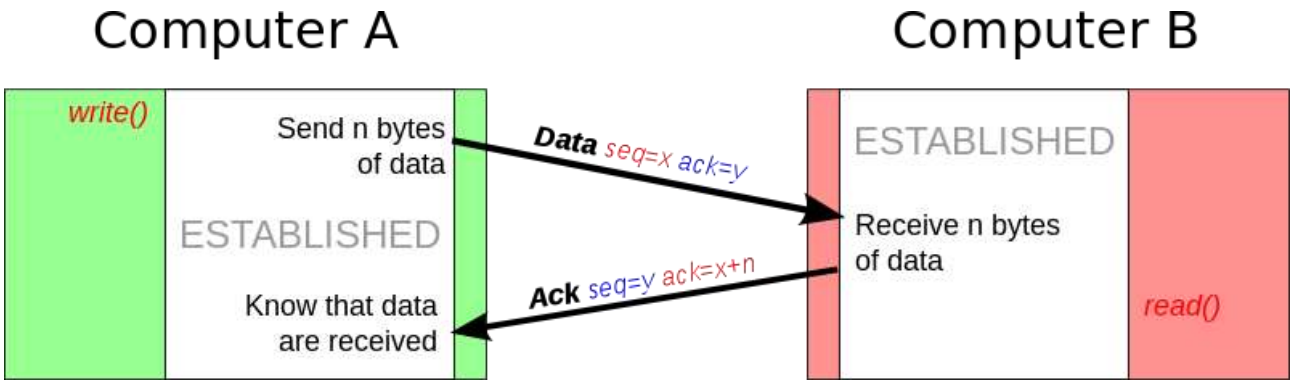
Des options diverses

Remplissage

On remplit l'espace restant après les options avec des zéros pour avoir une longueur multiple de 32 bits.

V.2.4 Transferts de données

Pendant la phase de transferts de données, certains mécanismes clefs permettent d'assurer la robustesse et la fiabilité de TCP. En particulier, les numéros de séquence sont utilisés afin d'ordonner les segments TCP reçus et de détecter les données perdues, les sommes de contrôle permettent la détection d'erreurs, et les acquittements ainsi que les temporisations permettent la détection des segments perdus ou retardés.



V.2.5 Numéros de séquence

Grâce aux numéros de séquence et d'acquittement, les systèmes terminaux peuvent remettre les données reçues dans l'ordre à l'application destinataire.

Les numéros de séquence sont utilisés pour décompter les données dans le flux d'octets. On trouve toujours deux de ces nombres dans chaque segment TCP, qui sont le numéro de séquence et le numéro d'acquittement. Le numéro de séquence représente le propre numéro de séquence de l'émetteur TCP, tandis que le numéro d'acquittement représente le numéro de séquence du destinataire. Afin d'assurer la fiabilité de TCP, le destinataire doit acquitter les segments reçus en indiquant qu'il a reçu toutes les données du flux d'octets jusqu'à un certain numéro de séquence.

Le numéro de séquence indique le premier octet des données.

Prenons un exemple :

Lors d'un échange de segments par Telnet, l'hôte A envoie un segment à l'hôte B contenant un octet de données, un numéro de séquence égal à 43 (Seq = 43) et un numéro d'acquittement égal à 79 (Ack = 79),

L'hôte B envoie un segment ACK à l'hôte A. Le numéro de séquence de ce segment correspond au numéro d'acquittement de l'hôte A (Seq = 79), et le numéro d'acquittement au numéro de séquence de « A » tel que reçu par B, augmenté de la quantité de données en bytes reçue (Ack = 43 + 1 = 44).

L'hôte « A » confirme la réception du segment en envoyant un ACK à l'hôte B, avec comme numéro de séquence son nouveau numéro de séquence, à savoir 44 (Seq = 44) et comme numéro d'acquittement le numéro de séquence du segment précédemment reçu, augmenté de la quantité de données reçue (Ack = 79 + 1 = 80).

Les numéros de séquence sont des nombres entiers non signés sur 32 bits, qui reviennent à zéro après avoir atteint (2^32)-1. Le choix du numéro de séquence initial est une des clefs de la robustesse et de la sécurité des connexions TCP.

Une amélioration de TCP, nommée acquittement sélectif (selective acknowledgement ou SACK), autorise le destinataire TCP à acquitter des blocs de données reçus dans le désordre.

V.2.6 Somme de contrôle

Une somme de contrôle sur 16 bits, constituée par le complément à un de la somme complémentée à un de tous les éléments d'un segment TCP (en-tête et données), est calculée par l'émetteur, et incluse dans le segment émis. Le destinataire recalcule la somme de contrôle du segment reçu, et si elle correspond à la somme de contrôle reçue, on considère que le segment a été reçu intact et sans erreur.

V.2.7 Temporisation

La perte d'un segment est gérée par TCP en utilisant un mécanisme de temporisation et de retransmission. Après l'envoi d'un segment, TCP va attendre un certain temps la réception du ACK correspondant. Un temps trop court entraîne un grand nombre de retransmissions inutiles et un temps trop long ralentit la réaction en cas de perte d'un segment.

Le délai avant retransmission doit être supérieur au RTT (ROUND-TRIP TIME) moyen d'un segment c'est-à-dire autant que prend un segment pour effectuer l'aller-retour entre le client et le serveur. Comme cette valeur peut varier dans le temps, des échantillons sont prélevés à intervalle régulier et une moyenne pondérée est calculée :

$$\text{RTT moyen} = (1 - \alpha) * \text{RTT moyen} + \alpha * \text{RTT échantillon}$$

Une valeur typique pour α est 0.125. L'influence des échantillons diminue de manière exponentielle dans le temps.

Le délai à utiliser est obtenu à partir de cette estimation du RTT moyen et en y ajoutant une marge de sécurité. Plus la différence entre un échantillon et la moyenne est grande, plus la marge de sécurité à prévoir est importante. Le calcul se fait à partir de la variance pondérée entre l'échantillon et la moyenne :

$$\text{Variance RTT} = (1 - \beta) * \text{Variance RTT} + \beta * |\text{RTT échantillon} - \text{RTT moyen}|$$

Une valeur typique pour β est 0.25. Le délai à utiliser est finalement donné par la formule suivante :

$$\text{Délai} = \text{RTT moyen} + 4 * \text{Variance RTT}$$

L'Algorithme de Karn permet de mieux évaluer le délai en présence d'acquittements ambigus. Si un segment envoyé a été perdu, les segments ultérieurs provoqueront des acquittements où figurera le numéro du premier octet manquant et, de ce fait, on ne sait plus à quel segment envoyé correspondent ces acquittements.

Parfois, quand le délai est trop long, il est avantageux de ne pas attendre avant de retransmettre un segment. Si un hôte reçoit 3 ACKs pour le même segment, alors il considère que tous les segments transmis après le segment acquitté ont été perdus et les retransmet donc immédiatement (Fast retransmit).

NOTE :
Le round-trip time (RTT) ou round-trip delay time (RTD) est le temps que met un signal pour parcourir l'ensemble d'un circuit fermé.

V.2.8 Contrôle de flux

Chaque partenaire dans une connexion TCP dispose d'un tampon de réception dont la taille n'est pas illimitée. Afin d'éviter qu'un hôte ne surcharge l'autre, TCP prévoit plusieurs mécanismes de contrôle de flux. Ainsi, chaque segment TCP contient la taille disponible dans le tampon de réception de l'hôte qui l'a envoyé. En réponse, l'hôte distant va limiter la taille de la fenêtre d'envoi afin de ne pas le surcharger.

Il existe d'autres algorithmes tels que Nagle ou Clarck qui facilitent aussi le contrôle du flux.

V.2.9 Contrôle de congestion

La congestion intervient lorsque trop de sources tentent d'envoyer trop de données trop vite pour que le réseau soit capable de les transmettre. Ceci entraîne la perte de nombreux paquets et de longs délais.

Les acquittements des données émises, ou l'absence d'acquittements, sont utilisés par les émetteurs pour interpréter de façon implicite l'état du réseau entre les systèmes finaux. À l'aide de temporisations, les émetteurs et destinataires TCP peuvent modifier le comportement du flux de données. C'est ce qu'on appelle généralement le contrôle de congestion.

Il existe une multitude d'algorithmes d'évitement de congestion pour TCP comme par exemple :

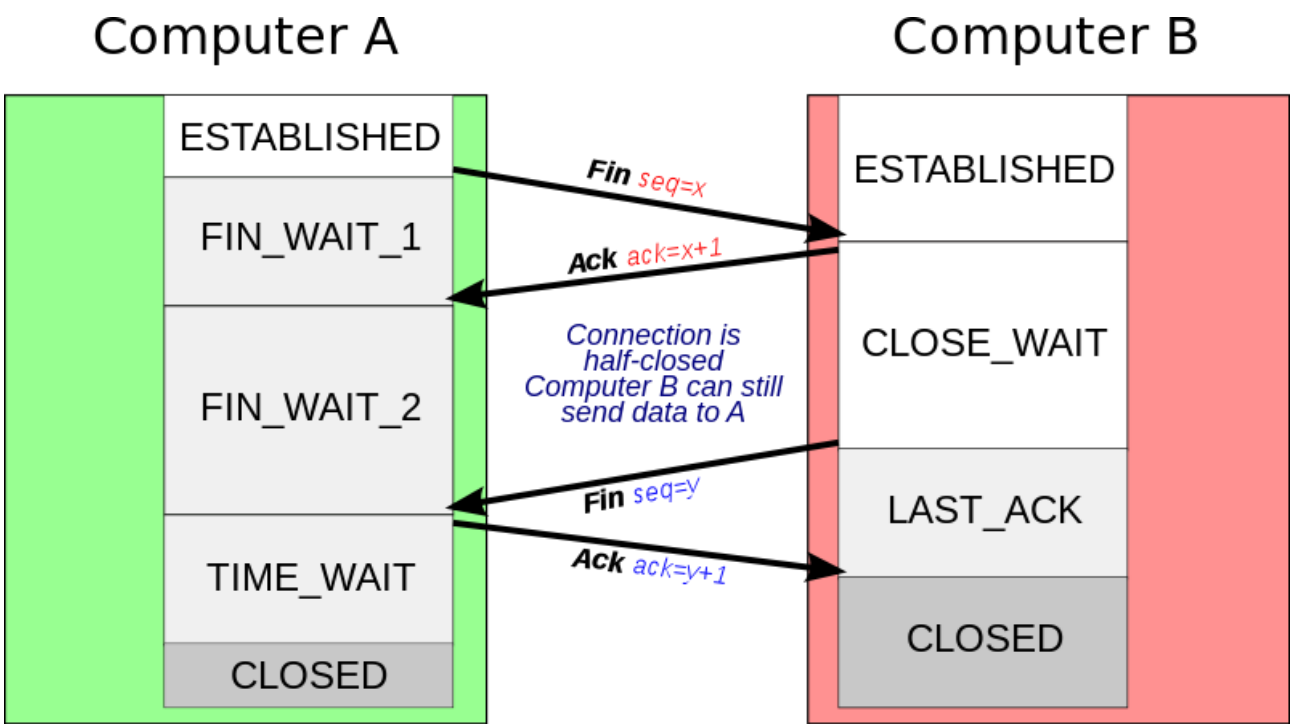
- TCP tahoe,
- TCP Reno,
- TCP new Reno,
- TCP Vegas...

V.2.10 Pour conclure

Le protocole TCP utilise un certain nombre de mécanismes afin d'obtenir une bonne robustesse et des performances élevées. Ces mécanismes comprennent l'utilisation d'une fenêtre glissante, l'algorithme de démarrage lent (slow start), l'algorithme d'évitement de congestion (congestion avoidance), les algorithmes de retransmission rapide (fast retransmit) et de récupération rapide (fast recovery), etc. Des recherches sont menées actuellement afin d'améliorer TCP pour traiter efficacement les pertes, minimiser les erreurs, gérer la congestion et être rapide dans des environnements très haut débit.

V.2.11 Terminaison d'une connexion

La phase de terminaison d'une connexion utilise un handshaking en 4 temps. Chaque extrémité de la connexion effectuant sa terminaison de manière indépendante. La fin d'une connexion nécessite une paire de segments FIN et ACK pour chaque extrémité.



V.2.12 Ports TCP

TCP, utilise le numéro de port pour identifier les applications. À chaque extrémité (client/serveur) de la connexion TCP est associé un numéro de port sur 16 bits (de 1 à 65535) assigné à l'application émettrice ou réceptrice.

Les ports bien connus (well-known ports) sont assignés par l'IANA dans la plage 0-1023, et sont souvent utilisés par des processus système ou ayant des droits privilégiés. Les applications qui fonctionnent en tant que serveur et sont en attente de connexions utilisent généralement ces types de ports comme :

- FTP (21)
- SSH (22)
- Telnet (23)
- SMTP (25)
- HTTP (80)
- POP3 (110)
- LDAP (389)

V.3. UDP (USER DATAGRAM PROTOCOL)

V.3.1 Introduction

Le protocole UDP, défini dans la RFC 768, est l'un des principaux protocoles de la suite Internet que la couche transport (couche 4).

Le rôle de ce protocole est de permettre la transmission de données sous forme de datagrammes de manière très simple entre deux entités, chacune étant définie par une adresse IP et un numéro de port. Aucune communication préalable n'est requise pour établir la connexion contrairement à TCP qui utilise le procédé de « handshaking ». UDP utilise un mode de transmission sans connexion. Ainsi, UDP est considéré comme un protocole de transmission non fiable.

L'intégrité des données est assurée par une somme de contrôle sur l'en-tête. L'utilisation de cette somme est cependant facultative en IPv4 (mais obligatoire avec IPv6). Si un hôte n'a pas calculé la somme de contrôle d'un datagramme émis, la valeur de celle-ci est fixée à zéro. La somme de contrôle inclut également les adresses IP de la source et de la destination.

À cause de l'absence de mécanisme de « handshaking », ce protocole expose le programme qui l'utilise aux problèmes éventuels de fiabilité du réseau. Il n'existe pas de garantie de protection quant à la livraison, l'ordre d'arrivée ou la duplication éventuelle des datagrammes. Si des fonctionnalités de correction d'erreur sont requises, une application peut donc se tourner vers les protocoles TCP ou SCTP, qui sont conçus à cet effet. UDP est donc adapté à un usage pour lequel la détection et la correction d'erreurs ne sont pas nécessaires ou sont effectuées directement par l'application.

La nature du protocole UDP le rend utile pour transmettre rapidement de petites quantités de données, depuis un serveur vers de nombreux clients ou bien dans des cas où la perte éventuelle d'un datagramme est préférée à l'attente de sa retransmission. Le DNS, la voix sur IP, le streaming, TFTP, NTP, SNMP, DHCP ou les jeux en ligne utilisent typiquement ce protocole.

V.4. STRUCTURE D'UN DATAGRAMME UDP

Le paquet UDP est encapsulé dans un paquet IP. Il comporte un en-tête suivi des données proprement dites à transporter.

En-tête IP	En-tête UDP	Données

L'en-tête d'un datagramme UDP est plus simple que celui de TCP :

Port Source (16 bits)	Port Destination (16 bits)
Longueur (16 bits)	Somme de contrôle (16 bits)
Données (longueur variable)	

La signification des 4 champs :

- ***Port Source***

Indique depuis quel port le paquet a été envoyé.

- ***Port de Destination***

Indique à quel port le paquet doit être envoyé.

- ***Longueur***

Indique la longueur totale de l'en-tête et de données exprimée en octets du segment UDP. La longueur minimale est donc de 8 octets qui est la taille de l'en-tête.

- ***Somme de contrôle***

Elle permet de s'assurer de l'intégrité du paquet reçu quand elle est différente de zéro. Elle est calculée sur l'ensemble de l'en-tête UDP et des données mais aussi sur un pseudo en-tête (extrait de l'en-tête IP).

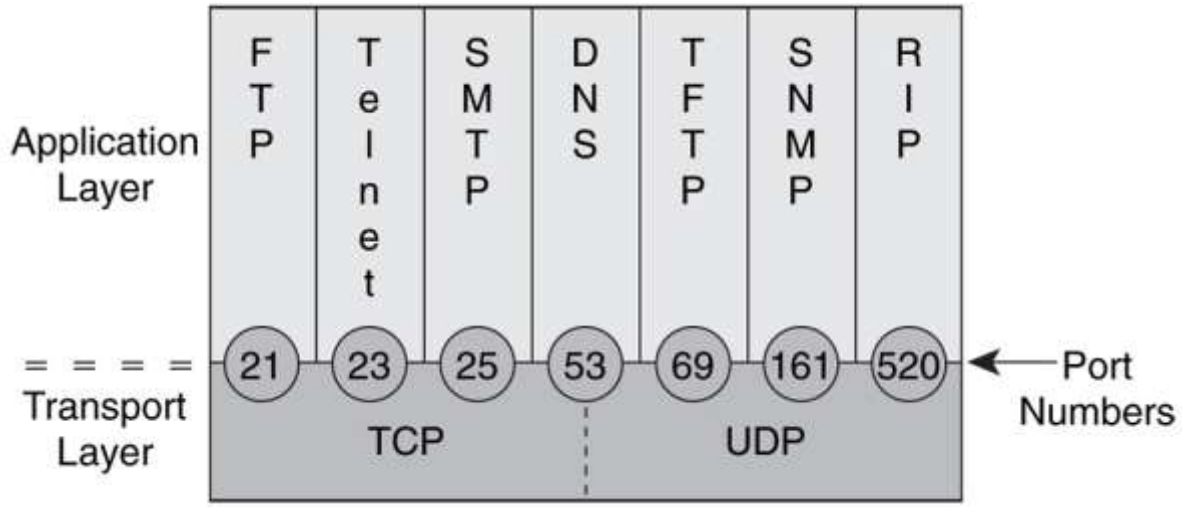
Note :

La présence de ce pseudo en-tête, interaction entre les deux couches IP et UDP, est une des raisons qui font que le modèle TCP/IP ne s'applique pas parfaitement au modèle OSI.

V.4.1 Ports UDP

UDP utilise, tout comme TCP, un numéro de port pour identifier les applications sur 16 bits (de 1 à 65535). Il est donc important de ne pas confondre le port 110/UDP et le port 110/TCP lorsque vous écrivez une ACL ou une règle dans un firewall (pare-feu).

V.5. PORTS TCP ET UDP



V.5.1 Liste des ports

Numéro du port	Type	Description
1	TCP	TCPMUX - TCP Port Service Multiplexer
2	TCP	compressnet - Management Utility
3	TCP	compressnet - Compression Process
5	TCP	rje - Remote job entry
7	TCP	echo
9	TCP	Discard Protocol
11	TCP	SYSTAT - Active Users
13	TCP	Daytime Protocol RFC 867
17	TCP	qotd - Quote of the Day
18	TCP	msp - Message send protocol
19	TCP	CHARGEN - Character Generator Protocol
20	TCP	ftp-data - File Transfer Protocol [flux de données]
21	TCP	ftp - File Transfer Protocol (le flux de contrôle pour le transfert de fichiers), voir Diagramme des flux FTP (port 20 et 21)
22	TCP	SSH - Secure Shell - SFTP
23	TCP	telnet
24	TCP	any private mail system
25	TCP	smtp - Simple Mail Transfer Protocol RFC 5321
27	TCP	nsw-fe - NSW User System FE
29	TCP	msg-icp
31	TCP	msg-auth - MSG Authentication

Numéro du port	Type	Description
33	TCP	dsp - Display Support Protocol
35	TCP	any private printer server
37	TCP	Time Protocol RFC 868
38	TCP	rap - Route Access Protocol
39	TCP	rlp - Resource Location Protocol
41	TCP	graphics
42	TCP	nameserver - Host Name Server
43	TCP	nicname - Who Is
44	TCP	mpm-flags - MPM FLAGS Protocol
45	TCP	mpm - Message Processing Module [recv]
46	TCP	mpm - Message Processing Module [default send]
47	TCP	ni-ftp
48	TCP	auditd - Digital Audit Daemon
49	TCP	login - Login Host Protocol (Terminal Access Controller Access-Control System)
50	UDP	re-mail-ck - Remote Mail Checking Protocol
51	TCP	la-maint - IMP Logical Address Maintenance
52	TCP	xns-time - XNS Time Protocol
53	UDP	domain - Domain Name System (DNS)
54	UDP	xns-ch - XNS Clearinghouse
55	TCP	isi-gl - ISI Graphics Language
56	TCP	xns-auth - XNS Authentication
57	TCP	any private terminal access
58	TCP	xns-mail - XNS Mail
59	TCP	any private file service
61	TCP	ni-mail
62	TCP	acas - ACA Services
64	TCP	covia - Communications Integrator (CI)
65	TCP	tacacs-ds - TACACS-Database Service
67	UDP	bootps - Bootstrap Protocol Server, ce port est aussi utilisé par une extension de bootp : DHCP, pour la recherche d'un serveur DHCP
68	UDP	bootpc - Bootstrap Protocol Client, ce port est aussi utilisé par une extension de bootp : DHCP, pour le dialogue entre le serveur DHCP et le client (attribution d'un bail pour une adresse IP)
69	UDP	tftp - Trivial File Transfer Protocol
70	TCP	gopher
71	TCP	netrjs-1 Remote Job Service
72	TCP	netrjs-2 Remote Job Service
73	TCP	netrjs-3 Remote Job Service

Numéro du port	Type	Description
74	TCP	netrjs-4 Remote Job Service
75	TCP	any private dial out service
76	TCP	deos - Distributed External Object Store
77	TCP	any private RJE service
78	TCP	vetTCP
79	TCP	finger
80	TCP	www-http - World Wide Web HTTP
81	TCP	host2-ns - HOSTS2 Name Server
82	TCP	xfer - XFER Utility
83	TCP	mit-ml-dev
84	TCP	ctf - Common Trace Facility
85	TCP	mit-ml-dev
86	TCP	mfcobol - Micro Focus Cobol
87	TCP	any private terminal link
88	TCP	kerberos
89	TCP	su-mit-tg - SU/MIT Telnet Gateway
90	TCP	dnsix - DNSIX Security Attribute Token Map
91	TCP	mit-dov - MIT Dover Spooler
92	TCP	npp - Network Printing Protocol
93	TCP	dcp - Device Control Protocol
94	TCP	objcall - Tivoli Object Dispatcher
95	TCP	supdup
96	TCP	dixie - DIXIE Protocol Specification
97	TCP	swift-rvf - Swift Remote Virtual File Protocol
98	TCP	tacnews
99	TCP	metagram - Metagram Relay
100	TCP	newacct - [unauthorized use]
101	TCP	hostname - NIC Host Name Server
102	TCP	iso-tsap
103	TCP	gppitnp - Genesis Point-To-Point Trans Net
104	TCP	acr-nema - ACR-NEMA Digital Imag. & Comm. 300
105	TCP	csnet-ns - Mailbox Name Nameserver
106	TCP	3com-tsmux
107	TCP	rtelnet - Remote Telnet Service
108	TCP	snagas - SNA Gateway Access Server
109	TCP	pop2 - Post Office Protocol - Version 2 RFC 937
110	TCP	pop3 - Post Office Protocol - Version 3 RFC 1939
111	TCP	sunrpc - SUN Remote Procedure Call
112	TCP	mcidas - McIDAS Data Transmission Protocol
113	TCP	auth - Authentication Service

Numéro du port	Type	Description
114	TCP	audionews - Audio News Multicast
115	TCP	sftp - Simple File Transfer Protocol (Ubuntu)
116	TCP	ansanotify - ANSA REX Notify
117	TCP	uucp-path - UUCP Path Service
118	TCP	sqlserv - SQL Services
119	TCP	nntp - Network News Transfer Protocol RFC 3977
120	TCP	cfdptkt
121	TCP	erpc - Encore Expedited Remote Pro.Call
122	TCP	smakynet
123	UDP	ntp - Network Time Protocol RFC 5905
124	TCP	ansatrader - ANSA REX Trader
125	TCP	locus-map - Locus PC-Interface Net Map Server
126	TCP	unitary - Unisys Unitary Login
127	TCP	locus-con - Locus PC-Interface Conn Server
128	TCP	gss-xlicen - GSS X License Verification
129	TCP	pwdgen - Password Generator Protocol
130	TCP	cisco-fna - cisco FNATIVE
131	TCP	cisco-tna - cisco TNATIVE
132	TCP	cisco-sys - cisco SYSMAINT
133	TCP	statsrv - Statistics Service
135	TCP	loc-srv - Location Service
136	TCP	profile - PROFILE Naming System
137	TCP	netbios-ns - NETBIOS Name Service
138	TCP	netbios-dgm - NETBIOS Datagram Service
139	TCP	netbios-ssn - NETBIOS Session Service
140	TCP	emfis-data - EMFIS Data Service
141	TCP	emfis-cntl - EMFIS Control Service
142	TCP	bl-idm - Britton-Lee IDM
143	TCP	imap2, imap4 - Internet Message Access Protocol v4 RFC 3501
144	TCP	news
145	TCP	uaac
146	TCP	iso-tp0
147	TCP	iso-ip
148	TCP	cronus - CRONUS-SUPPORT
149	TCP	aed-512 - AED 512 Emulation Service
150	TCP	sql-net
151	TCP	hems
152	TCP	bftp - Background File Transfer Program
153	TCP	sgmp
154	TCP	netsc-prod

Numéro du port	Type	Description
155	TCP	netsc-dev
156	TCP	sqlsrv - SQL Service
157	TCP	knet-cmp - KNET/VM Command/Message Protocol
158	TCP	pcmail-srv - PCMail Server
159	TCP	nss-routing
160	TCP	sgmp-traps
161	UDP	SNMP - Simple Network Management Protocol
162	UDP	snmptrap - Simple Network Management Protocol Trap
163	TCP	cmip-man - CMIP/TCP Manager
164	TCP	cmip-agent - CMIP/TCP Agent
165	TCP	xns-courier - Xerox
166	TCP	s-net - Sirius Systems
167	TCP	namp
168	TCP	rsvd
169	TCP	send
170	TCP	print-srv - Network PostScript
171	TCP	multiplex - Network Innovations Multiplex
172	TCP	cl/1 - Network Innocations CL/1
173	TCP	xyplex-mux - Xyplex
174	TCP	mailq
175	TCP	vmnet
176	TCP	genrad-mux
177	TCP	xdmcp - X Display Manager Control Protocol
178	TCP	nextstep - NeXTSTEP Window Server
179	TCP	bgp - Border Gateway Protocol
180	TCP	ris - Intergraph
181	TCP	unify
182	TCP	audit - Unisys Audit SITP
183	TCP	ocbinder
184	TCP	ocserver
185	TCP	remote-kis
186	TCP	kis - KIS Protocol
187	TCP	aci - Application Communication Interface
188	TCP	mumps - Plus Five's MUMPS
189	TCP	qft - Queued File Transport
190	TCP	gacp - Gateway Access Protocol
191	TCP	prospero - Prospero Directory Service
192	TCP	osu-nms - OSU Network Monitoring System
193	TCP	srmp - Spider Remote Monitoring Protocol
194	TCP	Internet relay chat (*IRC*)

Numéro du port	Type	Description
195	TCP	dn6-nlm-aud - DNSIX Network Level Module Audit
196	TCP	dn6-nlm-red - DNSIX Session Mgt Module Audit Redir
197	TCP	dls - Directory Location Service
198	TCP	dls-mon - Directory Location Service Monitor
199	TCP	smux
200	TCP	src - IBM System Resource Controller
201	TCP	at-rtmp - AppleTalk Routing Maintenance
202	TCP	at-nbp - AppleTalk Name Binding
203	TCP	at-3 - AppleTalk Unused
204	TCP	at-echo - AppleTalk Echo
205	TCP	at-5 - AppleTalk Unused
206	TCP	at-zis - AppleTalk Zone Information
207	TCP	at-7 - AppleTalk Unused
208	TCP	at-8 - AppleTalk Unused
209	TCP	tam - Trivial Mail Authentication Protocol
210	TCP	z39.50
211	TCP	914c/g - Texas Instruments 914C/G Terminal
212	TCP	anet - ATEXSSTR
213	TCP	ipx
214	TCP	vmpwscs - VM PWSCS
215	TCP	softpc - Insignia Solutions
216	TCP	atls - Access Technology License Server
217	TCP	dbase - dBASE Unix
218	TCP	mpp - Netix Message Posting Protocol
219	TCP	uarps - Unisys ARPs
220	TCP	imap3 - Interactive Mail Access Protocol v3 RFC 1203
221	TCP	fln-spx - Berkeley rlogind with SPX auth
222	TCP	rsh-spx - Berkeley rshd with SPX auth
223	TCP	cdc - Certificate Distribution Center
243	TCP	sur-meas - Surveet Measurement
245	TCP	link
246	TCP	dsp3270 - Display Systems Protocol
264	UDP	BGMP - Border Gateway Multicast Protocol
344	TCP	pdap - Prospero Data Access Protocol
345	TCP	pawserv - Perf Analysis Workbench
346	TCP	zserv - Zebra server
347	TCP	fatserv - Fatmen Server
348	TCP	csi-sgwp - Cabletron Management Protocol
371	TCP	clearcase
372	TCP	ulistserv - Unix Listserv

Numéro du port	Type	Description
373	TCP	legent-1 - Legent Corporation
374	TCP	legent-2 - Legent Corporation
375	TCP	hassle
376	TCP	nip - Amiga Envoy Network Inquiry Proto
377	TCP	tnETOS - NEC Corporation
378	TCP	dsETOS - NEC Corporation
379	TCP	is99c - TIA/EIA/IS-99 modem client
380	TCP	is99s - TIA/EIA/IS-99 modem server
381	TCP	hp-collector - hp performance data collector
382	TCP	hp-managed-node - hp performance data managed node
383	TCP	hp-alarm-mgr - hp performance data alarm manager
384	TCP	arns - A Remote Network Server System
385	TCP	ibm-app - IBM Application
386	TCP	asa - ASA Message Router Object Def.
387	TCP	aurp - AppleTalk Update-Based Routing Pro.
388	TCP	unidata-ldm - Unidata LDM Version 4
389	TCP	Lightweight Directory Access Protocol (LDAP)
390	TCP	uis
391	TCP	synotics-relay - SynOptics SNMP Relay Port
392	TCP	synotics-broker - SynOptics Port Broker Port
393	TCP	dis - Data Interpretation System
394	TCP	embl-ndt - EMBL Nucleic Data Transfer
395	TCP	NETscout Control Protocol
396	TCP	netware-ip - Novell NetWare encapsulé dans IP
397	TCP	mptn - Multi Protocol Trans. Net.
398	TCP	kryptolan
400	TCP	work-sol - Worksation Solutions
401	TCP	ups - Uninteruptible Power Supply
402	TCP	genie - Genie Protocol
403	TCP	decap
404	TCP	nced
407	TCP	Timbuktu (software)
408	TCP	prm-sm - Prospero Resource Manager Sys. Man.
409	TCP	prm-nm - Prospero Resource Manager Node Man.
410	TCP	decladebug - DECLadebug Remote Debug Protcol
411	TCP	rmt - Remote MT Protocol
412	TCP	synoptics-trap - Trap Convetion Port
413	TCP	smsp
414	TCP	infoseek
415	TCP	bnet

Numéro du port	Type	Description
416	TCP	silverplatter
417	TCP	onmux
418	TCP	hyper-g
419	TCP	ariel1
420	TCP	smpte
421	TCP	ariel2
422	TCP	ariel3
423	TCP	opc-job-start - IBM Operations Planning and Control Start
424	TCP	opc-job-track - IBM Operations Planning and Control Track
425	TCP	icad-el - ICAD
426	TCP	smartsdp
427	TCP	svrloc - Server Location
428	TCP	ocs_cmu
429	TCP	ocs_amu
430	TCP	utmpsd
431	TCP	utmpcd
432	TCP	iasd
433	TCP	nnsp
434	TCP	mobileip-agent
435	TCP	mobileip-mn
436	TCP	dna-cml
437	TCP	comscm
438	TCP	dsfgw
439	TCP	dasp
440	TCP	sgcp
441	TCP	decvms-sysmgt
442	TCP	cvc_hostd
443	TCP	https (SSL/TLS)
444	TCP	snpp - Simple Network Paging Protocol
445	TCP/UDP	microsoft-ds SMB, anciennement appelé CIFS (Common Internet File System)
446	TCP	ddm-rdb
447	TCP	ddm-dfm
448	TCP	ddm-byte
449	TCP	as-servermap - AS Server Mapper
450	TCP	tserver
465	TCP	smtp sécurisé (ssl) (non officiel)
497	TCP	retrospect - Retrospect Backup software
500	TCP	ISAKMP (Internet Security Association and Key Management Protocol), un des composants d'IPsec
502	TCP	Modbus sur TCP.

Numéro du port	Type	Description
514	UDP	Syslog RFC 3164 NB : ce service n'est pas listé habituellement dans le fichier *etc\services*
515	TCP	printer - spooler
517	TCP	talk
518	TCP	ntalk
520	UDP	Routing
525	TCP	timed - timeserver
526	TCP	tempo - newdate
546	UDP	DHCP - Dynamic Host Configuration Protocol v6 (for IPv6)
548	TCP	AppleShare IP Server
554	TCP	RTSP (Real Time Streaming Protocol) RFC 2326
563	TCP	nntp sécurisé (ssl) RFC 4642
587	TCP	Message Submission for Mail RFC 5068 RFC 6409
631	TCP	Internet Printing Protocol
636	TCP	LDAP encapsulé dans SSL/TLS
706	TCP	Secure Internet live conferencing
706	UDP	Secure Internet live conferencing
873	TCP	rsync
902	TCP/UDP	VMware Vsphere client et Vcenter heartbeat
989	TCP	Port de données File Transfer Protocol Secure
990	TCP	Port de connexion serveur FTPS
993	TCP	imap sécurisé (ssl)
995	TCP	pop3 sécurisé (ssl)
1080	TCP	SOCKS
1135	UDP/TCP	OmniVision Communication Service
1194	UDP/TCP	OpenVPN
1337	TCP/UDP	Port utilisé par les développeurs pour tester des applications (Leet speak)
1352	TCP	Lotus Notes
1414	TCP	IBM MQSeries
1433	TCP	Microsoft SQL Server
1434	TCP	Microsoft SQL Monitor
1521	TCP	Serveur Oracle NB : ce numéro de port était (ou est) aussi utilisé par ncube-lm (nCUBE)
1524	TCP	Ingreslock, voir Ingres (base de données)
1720	TCP	H.323
1723	TCP	PPTP
1863	TCP	MSN (tchat)
1883	TCP	MQTT
2049	TCP	NFSv4

Numéro du port	Type	Description
2106	TCP/UDP	L2J LoginServer
2164	TCP/UDP	Dynamic DNS
2382	TCP	ms-olap3
2427	UDP	MGCP
3000	TCP	First Class Server
3051	TCP	AMS (Agency Management System)
3074	TCP/UDP	nintendo server
3306	TCP	Mysql Server
3389	TCP	Microsoft Terminal Server (RDP)
3632	TCP	distcc (compilation partagée)
3644	TCP	Evidian, ssowatch
3725	TCP/UDP	Netia NA-ER Port
3979	TCP/UDP	Serveur Transport Tycoon deluxe en open source (openttd)
4569	UDP	IAX2 Inter-Asterisk-Exchange (rfc5456)
5003	TCP/UDP	FileMaker network sharing
5009	TCP/UDP	Apple AirPort Admin Utility, AirPort Express Assistant, Xwis
5060	TCP/UDP	serveur SIP (Session Initiation Protocol)
5062	TCP/UDP	Localisation access
5222	TCP	serveur Jabber
5223	TCP	serveur Jabber sécurisé (ssl)
5269	TCP	serveur à serveur (server to server) Jabber
5280	TCP	serveur BOSH
5353	TCP/UDP	Multicast DNS
5432	TCP	serveur PostgreSQL
5498	TCP	Hotline Tracker
5500	TCP	Hotline Server
5501	TCP	Hotline Server
5900	TCP	VNC Server
5984	TCP	CouchDB Server
6000	TCP/UDP	X11
6112	TCP	Warcraft III
6522	TCP	Gobby Server (Sobby)
6557	TCP	Monitoring
6667	TCP	Serveur IRC
6697	TCP	Serveur IRC sécurisé (ssl)
7000	TCP	Serveur IRC sécurisé (ssl) alternatif
7648	TCP	CU-SeeMe
7725	TCP/UDP	Port par défaut de Faronics Deep Freeze (console/serveur)
7777	TCP/UDP	Serveur Terraria / L2J Gameserver
8000	TCP	Hotline

Numéro du port	Type	Description
8006	TCP	Dell AppAssure Replication/Management Interface
8008	TCP	Serveur CalDAV
8009	TCP	Port AJP utilisé par Tomcat
8080	TCP	http alternatif (Web cache)
8098	TCP	Administration Microsoft Windows Server 2003
8140	TCP	Serveur Puppet (port par défaut pour le master)
8443	TCP	Serveur CalDAV sécurisé (ssl)
8787	TCP	Serveur RStudio
8883	TCP	MQTT sécurisé (SSL)
8888	TCP	VolumeOnLan client/serveur
9000	TCP/UDP	cslistener
9009	TCP	Pichat - Peer to peer Web chat software
9443	TCP	VMware HTTPS (TCP) Accès et administration de vCenter Server via l'interface Web
10001	TCP/UDP	SCP-Config
10011	TCP	Accès distant Livebox sécurisé HTTPS
11371	TCP/UDP	OpenPGP - OpenPGP HTTP Keyserver
20003	TCP/UDP	WriteEasy Office: Coordination affichage utilisateur et serveur local | Serveur API pour logiciels tiers
22000	TCP	Syncthing Block Exchange Protocol
22067	TCP	Syncthing Relay
22070	TCP	Syncthing Relay Status
22968	TCP	Plex media server
25444	TCP	Serveur Unturned
25565	TCP/UDP	Serveur Minecraft
27000 à 27050	TCP/UDP	Serveur Steam
27017	TCP	MongoDB
47808 à 47823	UDP	BACnet (BAC0 à BACF)

Les informations qui suivent sont en général reprises par chaque système dans un fichier :

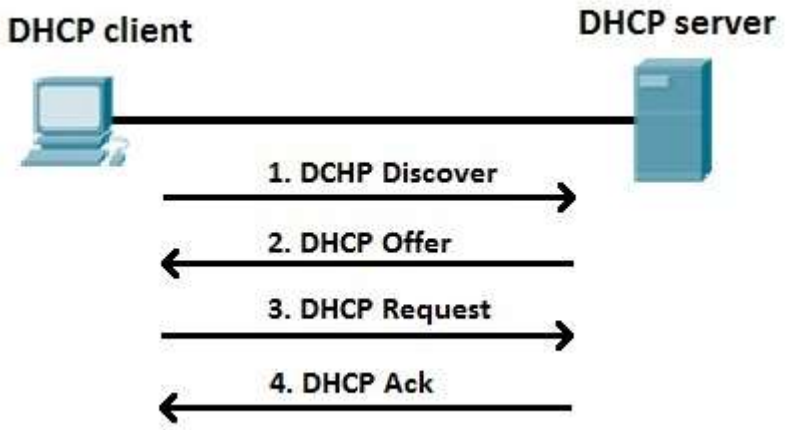
- Sur Linux et MacOs : /etc/services
- Sur Windows : %SystemRoot%\system32\drivers\etc\services.

VI. TCP/IP : APPLICATIONS

Objectifs

- Adressage IP dynamique avec le service serveur DHCP et le relai DHCP
- Résoudre les noms d'hôtes avec le fichier hosts et le service DNS
- Administration à distance avec le service SSH
- Transférer des fichiers avec FTP, SFTP ou TFTP
- Transférer des pages hypertexte avec HTTP
- Étudier les protocoles associés à la messagerie : POP3, IMAP4, SMTP
- Administration des réseaux IP avec SNMP et MIB

VI.1. ADRESSAGE IP DYNAMIQUE



VI.1.1 Introduction

DHCP (Dynamic Host Configuration Protocol) est un protocole réseau dont le rôle est d'assurer la configuration automatique des paramètres IP d'une station ou d'une machine, notamment en lui attribuant automatiquement une adresse IP et un masque de sous-réseau. DHCP peut aussi configurer l'adresse de la passerelle par défaut, des serveurs de noms DNS.

La conception initiale d'IP supposait la pré-configuration de chaque ordinateur connecté au réseau avec les paramètres TCP/IP adéquats : c'est l'adressage statique (nommée également IP fixe). Sur des réseaux de grandes dimensions ou étendues, où des modifications interviennent souvent, l'adressage statique engendre une lourde charge de maintenance et des risques d'erreurs. En outre, les adresses assignées ne peuvent être utilisées si l'ordinateur qui détient cette charge n'est pas en service : un cas typique où ceci pose problème est celui des fournisseurs d'accès à Internet, qui ont en général plus de clients que d'adresses IP à leur disposition, mais dont les clients ne sont jamais tous connectés en même temps.

DHCP apporte une solution à ces trois inconvénients :

- Seuls les ordinateurs en service utilisent une adresse de l'espace d'adressage ;
- Toute modification des paramètres (adresse de la passerelle, des serveurs de noms) est répercutée sur les stations lors du redémarrage ;
- La modification de ces paramètres est centralisée sur les serveurs DHCP.

Le protocole a été présenté pour la première fois en octobre 1993 et est défini par la RFC 1531, modifiée et complétée par les RFC 1534, RFC 2131 et RFC 2132.

Ce protocole peut fonctionner avec IPv4 ; il fonctionne aussi avec IPv6 et il est alors appelé DHCPv6. Toutefois, en IPv6, les adresses peuvent être auto-configurées sans DHCP.

VI.1.2 Fonctionnement

L'ordinateur équipé de carte réseau, mais dépourvu d'adresse IP, envoie en diffusion Broadcast un datagramme (DHCP DISCOVER) qui s'adresse au port 67 de n'importe quel serveur à l'écoute sur ce port. Ce datagramme comporte entre autres l'adresse physique (MAC) du client.

Tout serveur DHCP ayant reçu ce datagramme, s'il est en mesure de proposer une adresse sur le réseau auquel appartient le client, envoie une offre DHCP (DHCP OFFER) à l'attention du client (sur son port 68), identifié par son adresse physique. Cette offre comporte l'adresse IP du serveur, ainsi que l'adresse IP et le masque de sous-réseau qu'il propose au client. Il se peut que plusieurs offres soient adressées au client.

Le client retient une des offres reçues (la première qui lui parvient), et diffuse sur le réseau un datagramme de requête DHCP (DHCP REQUEST). Ce datagramme comporte l'adresse IP du serveur et celle qui vient d'être proposée au client. Elle a pour effet de demander au serveur choisi, l'assignation de cette adresse, l'envoi éventuel des valeurs des paramètres, est d'informer les autres serveurs qui ont fait une offre qui n'a pas été retenue.

Le serveur DHCP élabore un datagramme d'accusé de réception (DHCP ACK pour acknowledgement) qui assigne au client l'adresse IP et son masque de sous-réseau, la durée du bail de cette adresse (dont découlent deux valeurs T1 et T2 qui déterminent le comportement du client en fin de bail), et éventuellement d'autres paramètres :

- Adresse IP de la passerelle par défaut,
- Adresse IP des serveurs DNS,
- Adresse IP des serveurs NBNS (WINS),
- Adresses IP du serveur de temps,
- Adresses IP du serveur pour le boot PXE.

La liste des options que le serveur DHCP peut accepter est consultable dans la RFC 2132 : Options DHCP et Extensions fournisseur BOOTP, Chapitre RFC 1497 : Extensions fournisseur.

Les serveurs DHCP doivent être pourvus d'une adresse IP statique.

VI.1.3 Compatibilité

La plupart des systèmes d'exploitation ont des clients DHCP v4. C'est le cas pour les Windows Vista, 7, 8, 8.1 et 10.

Plusieurs clients et serveurs libres pour DHCP v4 et v6 sont disponibles pour les plateformes BSD (FreeBSD/NetBSD/OpenBSD/Apple MacOS) ainsi que les plateformes POSIX (Linux et UNIX-like). Là encore il convient de vérifier lesquelles gèrent IPv4 seulement ou IPv4 et IPv6.

VI.1.4 Renouvellement du bail

Les adresses IP dynamiques sont octroyées pour une durée limitée (durée du bail, ou lease time), qui est transmise au client dans l'accusé de réception qui clôture la transaction DHCP.

La valeur T1 (par défaut, 50 % de la durée du bail) qui l'accompagne détermine la durée après laquelle le client commence à demander périodiquement le renouvellement de son bail auprès du serveur qui lui a accordé son adresse. Cette fois, la transaction est effectuée par transmission IP classique, d'adresse à adresse.

Si, lorsque le délai fixé par la deuxième valeur, T2 (par défaut, 87,5 % de la durée du bail), est écoulé et que le bail n'a pas pu être renouvelé (par exemple, si le serveur DHCP d'origine est hors service), le client demande une nouvelle allocation d'adresse par diffusion.

Si, au terme du bail le client n'a pu ni en obtenir le renouvellement, ni obtenir une nouvelle allocation, l'adresse est désactivée et il perd la faculté d'utiliser le réseau TCP/IP de façon normale.

VI.1.5 Client et serveur sur des segments différents

Lorsque le serveur DHCP et le client ne figurent pas sur le même segment Ethernet, les diffusions émises par ce dernier ne parviennent pas au serveur parce que les routeurs ne transmettent pas les diffusions générales (broadcast). Dans ce cas, on utilise un agent de relais DHCP.

L'hôte particulier qui contient l'agent relais est configuré avec une adresse IP statique, et lors de la configuration de l'agent, on indique l'adresse d'un serveur DHCP auquel il faudra transmettre les découvertes DHCP qui lui parviennent sur le port 67 (écouté par le programme agent de relais). Il diffuse sur son segment (qui est aussi celui du client) les réponses qu'il reçoit du serveur DHCP.

L'agent relais est un programme que l'on active sur une ou plusieurs interfaces de l'hôte qui sera chargé de relayer la requête DHCP du client lorsque ce dernier n'est pas sur le même segment que le serveur DHCP. Pour remplir son rôle, l'agent relais place sa propre adresse IP dans le champ GIADDR de la trame DHCP qu'il a reçu du client. Il transmet ensuite cette trame en unicast, directement à l'adresse du serveur DHCP qu'on lui a paramétré. Le serveur DHCP utilise le champ GIADDR pour déterminer le sous-réseau et proposer une adresse dans la bonne étendue d'adresses IP. Quand le serveur répond à l'agent, il envoie la réponse à l'adresse GIADDR qu'avait indiqué l'agent relais, encore en unicast. L'agent relais transmet alors la réponse en broadcast sur le segment Ethernet du client demandeur.

L'agent relais peut être mis en œuvre sur un routeur ou un switch si, bien sûr, la fonctionnalité est intégrée. L'installation de l'agent peut s'effectuer également sur un serveur.

VI.1.6 Configuration du serveur DHCP

Pour qu'un serveur DHCP puisse servir des adresses IP, il est nécessaire de lui donner un « réservoir » d'adresses dans lequel il pourra puiser : c'est la plage d'adresses (address range). Il est possible de définir plusieurs plages, disjointes ou contiguës.

Les adresses du segment qui ne figurent dans aucune plage mise à la disposition du serveur DHCP ne seront en aucun cas distribuées, et peuvent faire l'objet d'affectations statiques (couramment : pour les serveurs nécessitant une adresse IP fixe, les routeurs, les imprimantes réseau…).

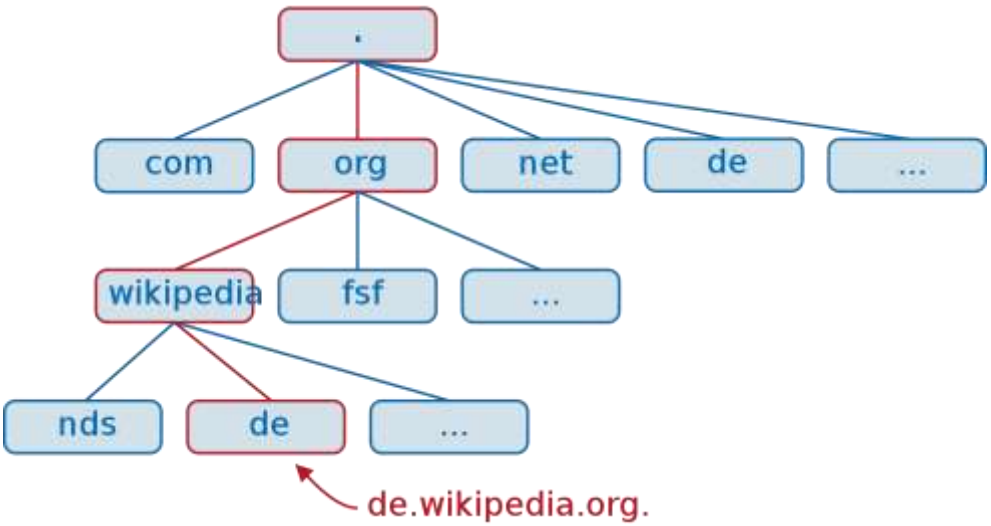
Il est également possible d'exclure pour un usage en adressage statique par exemple, des adresses ou blocs d'adresses compris dans une plage.

Enfin, on peut effectuer des réservations d'adresses en limitant la possibilité d'octroi de cette adresse au client possédant une adresse physique ou un « client identifier » donné. Ceci peut s'avérer utile pour des machines dont l'adresse doit rester fixe mais dont on veut contrôler de manière centrale et automatique les autres paramètres IP. Ce mécanisme est assuré par l'option 61 (voir RFC 2131).

Une autre option permet de donner toujours la même adresse IP à un équipement connecté à un port donné : option 82 (voir RFC 3046).

Lors de l'utilisation sur un même segment de plusieurs serveurs DHCP, l'intersection des plages d'adresses des différents serveurs doit être vide, sous peine d'ambiguïté dans les affectations et les renouvellements. En effet, les serveurs DHCP n'échangent aucune information relative aux baux qu'ils octroient.

VI.2. RESOUDRE LES NOMS D'HOTES



VI.2.1 Introduction

Le DNS (Domain Name System) est un service distribué utilisé pour traduire les noms de domaine Internet en adresse IP ou autres enregistrements.

À la demande de la DARPA (Defense Advanced Research Projects Agency), Jon Postel et Paul Mockapetris ont conçu le DNS en 1983 et en ont rédigé la première réalisation.

VI.2.2 Un système hiérarchique et distribué

VI.2.2.1 Hiérarchie du DNS

Le système des noms de domaine consiste en une hiérarchie dont le sommet est appelé la racine. On représente cette dernière par un point. Dans un domaine, on peut créer un ou plusieurs sous-domaines ainsi qu'une délégation pour ceux-ci, c'est-à-dire une indication que les informations relatives à ce sous-domaine sont enregistrées sur un autre serveur. Ces sous-domaines peuvent à leur tour déléguer des sous-domaines vers d'autres serveurs.

Tous les sous-domaines ne sont pas nécessairement délégués. Les délégations créent des zones, c'est-à-dire des ensembles de domaines et leurs sous-domaines non délégués qui sont configurés sur un serveur déterminé. Les zones sont souvent confondues avec les domaines.

Les domaines se trouvant immédiatement sous la racine sont appelés domaine de premier niveau (TLD : Top Level Domain). Les noms de domaines ne correspondant pas à une extension de pays sont appelés des domaines génériques (gTLD), par exemple .org ou .com. S'ils correspondent à des codes de pays (fr, be, ch…), ce sont des domaines de premier niveau national, aussi appelés ccTLD (country code TLD).

On représente un nom de domaine en indiquant les domaines successifs séparés par un point, les noms de domaines supérieurs se trouvant à droite. Par exemple, le domaine org. est un TLD, sous-domaine de la racine. Le domaine wikipedia.org. est un sous-domaine de .org. Cette délégation est accomplie en indiquant la liste des serveurs DNS associée au sous-domaine dans le domaine de niveau supérieur.

Les noms de domaines sont donc résolus en parcourant la hiérarchie depuis le sommet et en suivant les délégations successives, c'est-à-dire en parcourant le nom de domaine de droite à gauche.

Pour qu'il fonctionne normalement, un nom de domaine doit avoir fait l'objet d'une délégation correcte dans le domaine de niveau supérieur

VI.2.2.2 Résolution du nom par un hôte

Les hôtes n'ont qu'une connaissance limitée du système des noms de domaine. Quand ils doivent résoudre un nom, ils s'adressent à un ou plusieurs serveurs de noms dits récursifs, c'est-à-dire qu'ils vont parcourir la hiérarchie DNS et faire suivre la requête à un ou plusieurs autres serveurs de noms pour fournir une réponse. Les adresses IP de ces serveurs récursifs sont souvent obtenues via DHCP ou encore configurés en dur sur la machine hôte. Les fournisseurs d'accès à Internet mettent à disposition de leurs clients ces serveurs récursifs. Il existe également des serveurs récursifs ouverts comme ceux de Google Public DNS ou OpenDNS.

Quand un serveur DNS récursif doit trouver l'adresse IP de fr.wikipedia.org, un processus itératif démarre pour consulter la hiérarchie DNS. Ce serveur demande aux serveurs DNS appelés serveurs racine quels serveurs peuvent lui répondre pour la zone org. Parmi ceux-ci, le serveur va en choisir un pour savoir quels serveurs sont capables de lui répondre pour la zone wikipedia.org. C'est un de ces derniers qui pourra lui donner l'adresse IP de fr.wikipedia.org. S'il se trouve qu'un serveur ne répond pas, un autre serveur de la liste sera consulté.

Pour optimiser les requêtes ultérieures, les serveurs DNS récursifs font aussi office de DNS cache : ils gardent en mémoire (cache) la réponse d'une résolution de nom afin de ne pas effectuer ce processus à nouveau ultérieurement. Cette information est conservée pendant une période nommée Time to live et associée à chaque nom de domaine.

Un nom de domaine peut utiliser plusieurs serveurs DNS. Généralement, les noms de domaines en utilisent au moins deux : un primaire et un secondaire. Il peut y avoir plusieurs serveurs secondaires.

L'ensemble des serveurs primaires et secondaires font autorité pour un domaine, c'est-à-dire que la réponse ne fait pas appel à un autre serveur ou à un cache. Les serveurs récursifs fournissent des réponses qui ne sont pas nécessairement à jour, à cause du cache mis en place. On parle alors de réponse ne faisant pas autorité (non-authoritative answer).

Cette architecture garantit au réseau Internet une certaine continuité dans la résolution des noms. Quand un serveur DNS tombe en panne, le bon fonctionnement de la résolution de nom n'est pas remis en cause dans la mesure où des serveurs secondaires sont disponibles.

VI.2.2.3 Résolution inverse

Pour trouver le nom de domaine associé à une adresse IP, on utilise un principe semblable. Dans un nom de domaine, la partie la plus générale est à droite : org dans fr.wikipedia.org, le mécanisme de résolution parcourt donc le nom de domaine de droite à gauche. Dans une adresse IP V4, c'est le contraire : 213 est la partie la plus générale de 213.228.0.42. Pour conserver une logique cohérente, on inverse l'ordre des quatre termes de l'adresse et on la concatène au pseudo domaine in-addr.arpa. Ainsi, par exemple, pour trouver le nom de domaine de l'adresse IP 91.198.174.2, on résout 2.174.198.91.in-addr.arpa.

La déclaration inverse est importante sur les adresses IP publiques Internet puisque l'absence d'une résolution inverse est considérée comme une erreur opérationnelle (RFC 1912) qui peut entraîner le refus d'accès à un service. Par exemple, un serveur de messagerie électronique se présentant en envoi avec une adresse IP n'ayant pas de résolution inverse (PTR) a de grandes chances de se voir refuser, par l'hôte distant, la transmission du courrier (message de refus de type : IP lookup failed).

Cette résolution inversée est importante dans le cadre de la réalisation de diagnostics réseaux car c'est elle qui permet de rendre les résultats de la commande traceroute (ou tracert) humainement exploitables. Les dénominations des noms d'hôtes inverses sont souvent des composites de sous-domaines de localisation (ville, région, pays) et de domaines explicites indiquant le fournisseur d'accès Internet traversé comme francetelecom.net (XXXX.nctou202.Toulouse.francetelecom.net) et opentransit.net (XXXX.Aubervilliers.opentransit.net) pour France Télécom, ou encore proxad.net (XXXX.intf.routers.proxad.net) pour Free.

Une adresse IP peut être associée à différents noms de domaine via l'enregistrement de plusieurs entrées PTR dans le sous-domaine .arpa dédié à cette adresse (in-addr.arpa. pour IPv4 et ip6.arpa. pour IPv6). L'utilisation d'enregistrements PTR multiples pour une même adresse IP est éventuellement présente dans le cadre de l'hébergement virtuel de multiples domaines Web derrière la même adresse IP mais n'est pas recommandée dans la mesure où le nombre des champs PTR à renvoyer peut faire dépasser à la réponse la taille des paquets UDP de réponse et entraîner l'utilisation du protocole TCP (plus coûteux en ressources) pour envoyer la réponse à la requête DNS.

VI.2.2.4 Résolution inverse CIDR

Les délégations des zones inverses se font sur une frontière d'octet, ce qui fonctionne quand les blocs d'adresses sont distribués de façon classful mais pose des problèmes quand les blocs assignés sont de taille quelconque.

Par exemple, si deux clients A et B disposent chacun des blocs 192.168.0.0/25 et 192.168.0.128/25, il n'est pas possible de déléguer 0.168.192.in-addr.arpa. Au premier pour qu'il puisse définir les PTR correspondant à ses hôtes, car cela empêcherait le second de faire de même.

La RFC 2317 a défini une approche pour traiter ce problème, elle consiste à faire usage de domaines intermédiaires et de CNAME.

```
$ORIGIN 0.168.192.in-addr.arpa.
0/25 NS ns.clientA.fr.
128/25 NS ns.clientB.fr.

0 CNAME 0.0/25.0.168.192.in-addr.arpa.
1 CNAME 1.0/25.0.168.192.in-addr.arpa.
...
127 CNAME 127.0/25.0.168.192.in-addr.arpa.
128 CNAME 128.128/25.0.168.192.in-addr.arpa.
...
255 CNAME 255.128/25.0.168.192.in-addr.arpa.
```

Le client A définit la zone 0/25.0.168.192.in-addr.arpa. :

```
$ORIGIN 0/25.0.168.192.in-addr.arpa.
1 PTR hote1.clientA.fr.
...
127 PTR hote127.clientA.fr.
```

Le client B fait de même pour 128/25.0.168.192.in-addr.arpa. et les adresses 128 à 255.

La résolution inverse de 192.168.0.1 aboutira aux requêtes suivantes :

```
1.0.168.192.in-addr.arpa. CNAME 1.0/25.0.168.192.in-addr.arpa.
1.0/25.0.168.192.in-addr.arpa. PTR hote1.clientA.fr.
```

Ce qui assure le fonctionnement de la résolution inverse, moyennant un niveau d'indirection supplémentaire.

VI.2.3 Serveurs DNS racine

Les serveurs racine sont gérés par douze organisations différentes : deux sont européennes, une japonaise et les neuf autres sont américaines. Sept de ces serveurs sont en réalité distribués dans le monde grâce à la technique anycast et neuf disposent d'une adresse IPv614. Grâce à anycast, plus de 200 serveurs répartis dans 50 pays du monde assurent ce service15. Il existe 13 autorités de nom appelées de a à m.root-servers.net16. Le serveur k reçoit par exemple de l'ordre de 40 000 à 60 000 requêtes par seconde en 2016.

Le DNS ne fournit pas de mécanisme pour découvrir la liste des serveurs racine, chacun des serveurs doit donc connaître cette liste au démarrage grâce à un encodage explicite. Cette liste est ensuite mise à jour en consultant l'un des serveurs indiqués. La mise à jour de cette liste est peu fréquente de façon que les serveurs anciens continuent à fonctionner.

VI.2.4 Fully Qualified Domain Name

On entend par FQDN (Fully qualified domain name) un nom de domaine écrit de façon absolue, y compris tous les domaines jusqu'au TLD (domaine de premier niveau), il est ponctué par un point final, « fr.wikipedia.org. » par exemple.

La norme prévoit qu'un élément d'un nom de domaine (appelé label) ne peut dépasser 63 caractères, un FQDN ne pouvant dépasser 253 caractères.

VI.2.5 Nom de domaine internationalisé

Dans leur définition initiale, les noms de domaines sont constitués des caractères de A à Z (sans casse : les lettres capitales ne sont pas différenciées), de chiffres et du trait d'union.

La RFC 3490 définit un format appelé Punycode qui permet l'encodage d'un jeu de caractère plus étendu.

VI.2.6 Les techniques du DNS Round-Robin pour la distribution de la charge

Lorsqu'un service génère un trafic important, celui-ci peut faire appel à la technique du DNS Round-Robin (en français tourniquet DNS), une des techniques de répartition de charge qui consiste à associer plusieurs adresses IP à un FQDN. Les différentes versions de Wikipedia, comme fr.wikipedia.org par exemple, sont associées à plusieurs adresses IP : 207.142.131.235, 207.142.131.236, 207.142.131.245, 207.142.131.246, 207.142.131.247 et 207.142.131.248. L'ordre dans lequel ces adresses sont renvoyées sera modifié d'une requête à la suivante. Une rotation circulaire entre ces différentes adresses permet ainsi de répartir la charge générée par ce trafic important entre les différentes machines ayant ces adresses IP. Il faut cependant nuancer cette répartition car elle n'a lieu qu'à la résolution du nom d'hôte et reste par la suite en cache sur les différents resolvers (client DNS).

VI.2.7 Principaux enregistrements DNS

Le type d'enregistrement de ressource (RR pour Resource Record) est codé sur 16 bits19, l'IANA conserve le registre des codes assignés. Les principaux enregistrements définis sont les suivants :

- A record ou address record (également appelé enregistrement d'hôte) qui fait correspondre un nom d'hôte ou un nom de domaine ou un sous-domaine à une adresse IPv4 de 32 bits distribués sur quatre octets ex: 123.234.1.2 ;
- AAAA record ou IPv6 address record qui fait correspondre un nom d'hôte à une adresse IPv6 de 128 bits distribués sur seize octets ;
- CNAME record ou canonical name record qui permet de faire d'un domaine un alias vers un autre. Cet alias hérite de tous les sous-domaines de l'original ;
- MX record ou mail exchange record qui définit les serveurs de courriel pour ce domaine ;
- PTR record ou pointer record qui associe une adresse IP à un enregistrement de nom de domaine, aussi dit « reverse » puisqu'il fait exactement le contraire du A record ;
- NS record ou name server record qui définit les serveurs DNS de ce domaine ;
- SOA record ou Start Of Authority record qui donne les informations générales de la zone : serveur principal, courriel de contact, différentes durées dont celle d'expiration, numéro de série de la zone ;
- SRV record qui généralise la notion de MX record, mais qui propose aussi des fonctionnalités avancées comme le taux de répartition de charge pour un service donné, standardisé dans la RFC 2782 ;
- NAPTR record ou Name Authority Pointer record qui donne accès à des règles de réécriture de l'information, permettant des correspondances assez lâches entre un nom de domaine et une ressource. Il est spécifié dans la RFC 3403 ;
- TXT record permet à un administrateur d'insérer un texte quelconque dans un enregistrement DNS (par exemple, cet enregistrement est utilisé pour implémenter la spécification Sender Policy Framework) ;
- D'autres types d'enregistrements sont utilisés occasionnellement, ils servent simplement à donner des informations (par exemple, un enregistrement de type LOC indique l'emplacement physique d'un hôte, c'est-à-dire sa latitude et sa longitude). Certaines personnes disent que cela aurait un intérêt majeur mais n'est que très rarement utilisé sur le monde Internet.

VI.2.7.1 NS record

L'enregistrement NS crée une délégation d'un sous-domaine vers une liste de serveurs.

Dans la zone org, les enregistrements NS suivants créent le sous-domaine Wikipédia et délèguent celui-ci vers les serveurs indiqués.

L'ordre des serveurs est quelconque. Tous les serveurs indiqués doivent faire autorité pour le domaine.

- Wikipédia NS ns1.wikimedia.org.
- Wikipédia NS ns2.wikimedia.org.
- Wikipédia NS ns0.wikimedia.org.

VI.2.7.2 PTR record

À l'inverse d'une entrée de type A ou AAAA, une entrée PTR indique à quel nom d'hôte correspond une adresse IPv4 ou IPv6. Si elle est spécifiée, elle doit contenir l'enregistrement inverse d'une entrée DNS A ou AAAA.

Par exemple (pour une adresse IPv4) cet enregistrement PTR est :

- 232.174.198.91.in-addr.arpa. IN PTR text.esams.wikimedia.org.

Correspond à cette entrée A :

- text.esams.wikimedia.org. IN A 91.198.174.232

Dans le cas d'une adresse IPv6, les entrées de type PTR sont enregistrées dans la zone ip6.arpa. (Pendant de la zone in-addr.arpa. Des adresses IPv4).

La règle permettant de retrouver l'entrée correspondant à une adresse IPv6 est similaire à celle pour les adresses IPv4 (renversement de l'adresse et recherche dans un sous-domaine dédié de la zone arpa.), mais diffère au niveau du nombre de bits de l'adresse utilisés pour rédiger le nom du domaine où rechercher le champ PTR : là où pour IPv4 le découpage de l'adresse se fait par octet, pour IPv6 c'est un découpage par quartet qui est utilisé.

Par exemple à l'adresse IPv6 :

- 2001:610:240:22::c100:68b
- Correspond le nom de domaine :
- b.8.6.0.0.0.1.c.0.0.0.0.0.0.0.0.2.2.0.0.0.4.2.0.0.1.6.0.1.0.0.2.ip6.arpa. PTRwww.ipv6.ripe.net.

VI.2.7.3 MX record

Une entrée DNS MX indique les serveurs SMTP à contacter pour envoyer un courriel à un utilisateur d'un domaine donné. Par exemple :

- wikimedia.org. IN MX 10 mchenry.wikimedia.org.
- wikimedia.org. IN MX 50 lists.wikimedia.org.

On voit que les courriels envoyés à une adresse en @wikimedia.org sont envoyés au serveur mchenry.wikimedia.org. ou lists.wikimedia.org. Le nombre précédant le serveur représente la priorité. Le serveur avec la priorité numérique la plus petite est employé en priorité. Ici, c'est donc mchenry.wikimedia.org. qui doit être utilisé en premier, avec une valeur de 10.

Les serveurs indiqués doivent avoir été configurés pour accepter de relayer les courriers pour le nom de domaine indiqué. Une erreur courante consiste à indiquer des serveurs quelconques comme serveurs secondaires, ce qui aboutit au rejet des courriers quand le serveur primaire devient inaccessible. Il n'est pas indispensable de disposer de serveurs secondaires, les serveurs émetteurs conservant les messages pendant un temps déterminé (typiquement, plusieurs jours) jusqu'à ce que le serveur primaire soit à nouveau disponible.

Les entrées MX sont généralisées par les entrées SRV qui permettent de faire la même chose mais pour tous les services, pas seulement SMTP (le courriel). L'avantage des entrées SRV, par rapport aux entrées MX, est aussi qu'elles permettent de choisir un port arbitraire pour chaque service ainsi que de faire de la répartition de charge plus efficacement. L'inconvénient, c'est qu'il existe encore peu de programmes clients qui gèrent les entrées SRV. Cependant, depuis 2009, avec l'augmentation de l'utilisation du protocole SIP sur les services de VoIP, les enregistrements SRV deviennent plus fréquents dans les zones DNS.

VI.2.7.4 CNAME record

L'enregistrement CNAME permet de créer un alias.

Par exemple :

- fr.wikipedia.org. IN CNAME text.wikimedia.org.
- text.wikimedia.org. IN CNAME text.esams.wikimedia.org.
- text.esams.wikimedia.org. IN A 91.198.174.232

Celui-ci exclut tout autre enregistrement (RFC 1034 section 3.6.2, RFC 1912 section 2.4), c'est-à-dire qu'on ne peut avoir à la fois un CNAME et un A record pour le même nom de domaine.

Par exemple, ceci est interdit :

- fr.wikipedia.org. IN CNAME text.wikimedia.org.
- fr.wikipedia.org. IN A 91.198.174.232

Par ailleurs, pour des raisons de performance, et pour éviter les boucles infinies du type

- fr.wikipedia.org. IN CNAME text.wikimedia.org.
- text.wikipedia.org. IN CNAME fr.wikimedia.org.

Les spécifications (RFC 1034 section 3.6.2, RFC 1912 section 2.4) recommandent de ne pas faire pointer un CNAME sur un autre CNAME ni sur un DNAME (alias pour un nom et tous ses sous-noms).

Ainsi, le premier exemple serait préférablement enregistré de la façon suivante :

- fr.wikipedia.org. IN CNAME text.esams.wikimedia.org.
- text.wikimedia.org. IN CNAME text.esams.wikimedia.org.
- text.esams.wikimedia.org. IN A 91.198.174.232

VI.2.7.5 NAPTR record

Peu répandus à l'heure actuelle (ils sont surtout utilisés par ENUM), ils décrivent une réécriture d'une clé (un nom de domaine) en URI. Par exemple, dans ENUM, des enregistrements NAPTR peuvent être utilisés pour trouver l'adresse de courrier électronique d'une personne, connaissant son numéro de téléphone (qui sert de clé à ENUM).

Ses paramètres sont dans l'ordre :

1. ***Order*** : indique dans quel ordre évaluer les enregistrements NAPTR ; tant qu'il reste des enregistrements d'une certaine valeur de « order » à examiner, les enregistrements des valeurs suivantes de « order » n'entrent pas en considération ;
2. ***Preference*** : donne une indication de priorité relative entre plusieurs enregistrements NAPTR qui ont la même valeur de « order » ;
3. ***Flags*** : indique par exemple si l'enregistrement décrit une réécriture transitoire (dont le résultat est un nom de domaine pointant sur un autre enregistrement NAPTR) ou une réécriture finale ; la sémantique précise du paramètre flags dépend de l'application DDDS ('Dynamic Delegation Discovery System', RFC 340123) employée (ENUM en est une parmi d'autres) ;
4. ***Services*** : décrit le service de réécriture ; par exemple dans ENUM, la valeur de services spécifie le type de l'URI résultante ; la sémantique précise de ce paramètre dépend également de l'application DDDS employée ;
5. ***Regexp*** : l'opération de réécriture elle-même, formalisée en une expression rationnelle ; cette expression rationnelle est à appliquer à la clé ; ne peut être fournie en même temps que replacement ;
6. ***Replacement*** : nom de domaine pointant sur un autre enregistrement NAPTR, permettant par exemple une réécriture transitoire par délégation ; ne peut être fourni en même temps que regexp.
7. L'enregistrement NAPTR est défini par la RFC 3403.

VI.2.7.6 SOA record

Cet enregistrement permet d'indiquer le serveur de nom maître (primaire), l'adresse e-mail d'un contact technique (avec @ remplacé par un point) et des paramètres d'expiration.

Il désigne l'autorité (start of authority) ou le responsable de la zone dans la hiérarchie DNS.

Ces paramètres sont dans l'ordre :

wikipedia.org. IN SOA ns0.wikimedia.org. hostmaster.wikimedia.org. 2010060311 43200 7200 1209600 3600

- ***Serial*** : indique un numéro de version pour la zone (32 bits non signé). Ce nombre doit être incrémenté à chaque modification du fichier zone ; on utilise par convention une date au format « yyyymmddnn » (« yyyy » pour l'année sur 4 chiffres, « mm » pour le mois sur 2 chiffres, « dd » pour le jour sur 2 chiffres, « nn » pour un compteur de révision si le numéro de série est modifié plusieurs fois dans un même jour. Cette convention évite tout débordement du 32 bits non signé jusqu'en l'an 4294) ;
- ***Refresh*** : l'écart en secondes entre les demandes successives de mise à jour réalisées depuis le serveur secondaire ou les serveurs esclaves ;
- ***Retry*** : le délai en seconde que doivent attendre le serveur secondaire ou les serveurs esclaves lorsque leur précédente requête a échoué ;
- ***Expire*** : le délai en seconde au terme duquel la zone est considérée comme invalide si le secondaire ou les esclaves ne peuvent joindre le serveur primaire ;
- ***Minimum ou negative TTL*** : utilisé pour spécifier, en seconde, la durée de vie pendant laquelle sont conservées en cache les réponses qui correspondent à des demandes d'enregistrements inexistants.

Les versions récentes de BIND (named) acceptent les suffixes M, H, D ou W pour indiquer un intervalle de temps en minutes, heures, jours ou semaines respectivement.

VI.2.7.7 Time to live

Chaque record est associé à un Time to live (TTL) qui détermine combien de temps il peut être conservé dans un serveur cache. Ce temps est typiquement d'un jour (86400 s) mais peut être plus élevé pour des informations qui changent rarement, comme des records NS. Il est également possible d'indiquer que des informations ne doivent pas être mises en cache en spécifiant un TTL de zéro.

Certaines applications, comme des navigateurs Web disposent également d'un cache DNS, mais qui ne respecte pas nécessairement le TTL du DNS. Ce cache applicatif est généralement de l'ordre de la minute, mais Internet Explorer, par exemple, conserve les informations jusqu'à 30 minutes24, indépendamment du TTL configuré.

VI.2.7.8 Glue records

Quand un domaine est délégué à un serveur de noms qui appartient à ce sous-domaine, il est nécessaire de fournir également l'adresse IP de ce serveur pour éviter les références circulaires. Ceci déroge au principe général selon lequel l'information d'un domaine n'est pas dupliquée ailleurs dans le DNS.

Par exemple, dans la réponse suivante au sujet des NS pour le domaine wikimedia.org :

- wikimedia.org. IN NS ns2.wikimedia.org.
- wikimedia.org. IN NS ns1.wikimedia.org.
- wikimedia.org. IN NS ns0.wikimedia.org.

Il est nécessaire de fournir également les adresses IP des serveurs indiqués dans la réponse (glue records25), car ils font partie du domaine en question :

- ns0.wikimedia.org. IN A 208.80.152.130
- ns1.wikimedia.org. IN A 208.80.152.142
- ns2.wikimedia.org. IN A 91.198.174.4

VI.2.8 Mise à jour dynamique

Une extension du DNS nommée DNS dynamique (DDNS) permet à un client de mettre à jour une zone avec des informations qui le concernent (RFC 2136). Ceci est utile quand des clients obtiennent une adresse IP par DHCP et qu'ils souhaitent que le DNS reflète le nom réel de la machine.

VI.2.8.1 Considérations opérationnelles

VI.2.8.2 Mise à jour du DNS

Les mises à jour se font sur le serveur primaire du domaine, les serveurs secondaires recopiant les informations du serveur primaire dans un mécanisme appelé transfert de zone. Pour déterminer si un transfert de zone doit avoir lieu, le serveur secondaire consulte le numéro de version de la zone et le compare à la version qu'il possède. Le serveur primaire détermine à quelle fréquence le numéro de version est consulté. Quand un changement est effectué, les serveurs envoient des messages de notification aux serveurs secondaires pour accélérer le processus.

Il se peut que des informations qui ne sont plus à jour soient cependant conservées dans des serveurs cache. Il faut alors attendre l'expiration de leur Time to live pour que ces informations cachées disparaissent et donc que la mise à jour soit pleinement effective. On peut minimiser le temps nécessaire en diminuant le TTL associé aux noms de domaines qui vont être modifiées préalablement à une opération de changement.

VI.2.8.3 Cohérence du DNS

Quand la liste des serveurs de noms change, ou quand une adresse IP qui fait l'objet d'un 'Glue_record' est modifiée, le gestionnaire du domaine de niveau supérieur doit effectuer la mise à jour correspondante.

VI.2.8.4 Robustesse du DNS

Pour éviter les points individuels de défaillance, on évite de partager l'infrastructure entre les serveurs qui font autorité. Un serveur secondaire sera de préférence délocalisé et routé différemment que le serveur primaire.

Bien que cela soit techniquement possible, on évite de mêler sur un même serveur le rôle de DNS récursif et celui de serveur qui fait autorité.

De même, un hôte sera configuré avec plusieurs serveurs récursifs, de sorte que si le premier ne répond pas à la requête, le suivant sera employé. En général, les serveurs récursifs fournis par les FAI refusent les requêtes émanant d'adresses IP appartenant à d'autres FAI.

Il existe des services de DNS récursifs ouverts, c'est-à-dire qu'ils acceptent les requêtes de tous les clients. Il est donc possible à un utilisateur de configurer ceux-ci en lieu et place de ceux fournis par le FAI. Ceci pose cependant les problèmes suivants :

- Il n'y a pas de garantie que les réponses fournies seront les mêmes qu'avec des serveurs récursifs habituels. Un tel service pourrait en effet faire référence à une autre hiérarchie depuis la racine, disposer de TLD additionnels non standard, restreindre l'accès à certains domaines, voire altérer certains records avant leur transmission au client.
- Il n'y a pas de garantie de confidentialité, c'est-à-dire que ce service pourrait déterminer à quels domaines un utilisateur a accès en conservant des traces des requêtes DNS.

VI.2.9 Sécurité du DNS

Le protocole DNS a été conçu avec un souci minimum de la sécurité. Plusieurs failles de sécurité du protocole DNS ont été identifiées depuis. Les principales failles du DNS ont été décrites dans le RFC 3833 publié en août 2004.

VI.2.9.1 Interception des paquets

Une des failles mises en avant est la possibilité d'intercepter les paquets transmis. Les serveurs DNS communiquent au moyen de paquets uniques et non signés. Ces deux spécificités rendent l'interception très aisée. L'interception peut se concrétiser de différentes manières, notamment via une attaque de type « man in the middle », de l'écoute des données transférées et de l'envoi de réponse falsifiée (cf. paragraphe ci-dessous).

VI.2.9.2 Fabrication d'une réponse

Les paquets des serveurs DNS étant faiblement sécurisés, authentifiés par un numéro de requête, il est possible de fabriquer de faux paquets. Par exemple, un utilisateur qui souhaite accéder au site http://mabanque.example.com fait une demande au site DNS. Il suffit, à ce moment, qu'un pirate informatique réponde à la requête de l'utilisateur avant le serveur DNS pour que l'utilisateur se retrouve sur un site d'hameçonnage.

VI.2.9.3 Corruption des données

La trahison par un serveur, ou corruption de données, est, techniquement, identique à une interception des paquets. La seule différence venant du fait que l'utilisateur envoie volontairement sa requête au serveur. Cette situation peut arriver lorsque, par exemple, l'opérateur du serveur DNS souhaite mettre en avant un partenaire commercial.

VI.2.9.4 Empoisonnement du cache DNS

L'empoisonnement du cache DNS ou pollution de cache DNS (en anglais, DNS cache poisoning) est une technique permettant de leurrer les serveurs DNS afin de leur faire croire qu'ils reçoivent une requête valide tandis qu'elle est frauduleuse.

VI.2.9.5 Déni de service

Une attaque par déni de service (ou attaque par saturation; en anglais, Denial of Service attack ou DoS attack) est une attaque sur un serveur informatique qui résulte en l'incapacité pour le serveur de répondre aux requêtes de ses clients.

VI.2.10 DNSSEC

Pour contrer ces vulnérabilités, le protocole DNSSEC a été développé.

VI.2.11 Détails du protocole

DNS utilise en général UDP et le port 53. La taille maximale des paquets utilisée est de 512 octets. Si une réponse dépasse cette taille, la norme prévoit que la requête doit être renvoyée sur le port TCP 53. Ce cas est cependant rare et évité, et les firewalls bloquent souvent le port TCP 53. Les transferts de zone s'effectuent par TCP sur le même numéro de port. Pour des raisons de sécurité, les serveurs restreignent généralement la possibilité de transférer des zones.

L'extension EDNS0 (RFC 2671) permet d'utiliser une taille de paquets plus élevée, sa prise en charge est recommandée pour IPv6 comme pour DNSSEC.

La norme prévoit qu'il existe une classe associée aux requêtes. Les classes IN (Internet), CH (Chaos) et HS (Hesiod) sont définies, seule la classe IN étant réellement utilisée en pratique. La classe chaos est utilisée par BIND pour révéler le numéro de version.

VI.2.12 Outils de diagnostics

Pour vérifier l'association entre un nom et une adresse IP, plusieurs commandes sont disponibles suivant les systèmes d'exploitation utilisés.

Par exemple sur Microsoft Windows la commande **nslookup** est disponible via l'invite de commande :

```
> nslookup www.google.fr
Serveur : geronimo
Address: 192.168.1.1

Réponse ne faisant pas autorité :
Nom : www.l.google.com
Addresses:
          209.85.229.104
          209.85.229.106
          209.85.229.103
          209.85.229.147
          209.85.229.105
          209.85.229.99
Aliases: www.google.fr
          www.google.com
```

Ou sur les systèmes Unix/Linux, la commande **dig** même si la commande **nslookup** est présente :

```
> dig www.google.com aaaa
; <<>> DiG 9.7.0-P1 <<>> www.google.com aaaa
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 47055
;; flags: qr rd ra; QUERY: 1, ANSWER: 7, AUTHORITY: 4, ADDITIONAL: 0

;; QUESTION SECTION:
;www.google.com.                IN      AAAA

;; ANSWER SECTION:
www.google.com.      422901  IN      CNAME   www.l.google.com.
www.l.google.com.    77      IN      AAAA    2a00:1450:8004::67
www.l.google.com.    77      IN      AAAA    2a00:1450:8004::68
www.l.google.com.    77      IN      AAAA    2a00:1450:8004::69
www.l.google.com.    77      IN      AAAA    2a00:1450:8004::6a
www.l.google.com.    77      IN      AAAA    2a00:1450:8004::93
www.l.google.com.    77      IN      AAAA    2a00:1450:8004::63

;; AUTHORITY SECTION:
google.com.          155633  IN      NS      ns2.google.com.
google.com.          155633  IN      NS      ns1.google.com.
google.com.          155633  IN      NS      ns3.google.com.
google.com.          155633  IN      NS      ns4.google.com.

;; Query time: 0 msec
;; SERVER: ::1#53(::1)
;; WHEN: Sun May 23 16:23:49 2010
;; MSG SIZE rcvd: 292
```

VI.3. ADMINISTRATION A DISTANCE AVEC LE SERVICE SSH

VI.3.1 Introduction

SSH (Secure Shell), créé en 1995, est à la fois un programme informatique et un protocole de communication sécurisé. Le protocole de connexion impose un échange de clés de chiffrement en début de connexion. Par la suite, tous les segments TCP sont authentifiés et chiffrés. Il devient donc impossible d'utiliser un sniffer pour voir ce que fait l'utilisateur.

Le protocole SSH a été conçu avec l'objectif de remplacer les différents protocoles non chiffrés comme rlogin, Telnet, rcp, ftp et rsh.

VI.3.2 Protocole SSH

Le protocole SSH existe en deux versions majeures : la version 1.0 et la version 2.0.

La première version permet de se connecter à distance à un ordinateur afin d'obtenir un Shell ou une ligne de commande. Cette version souffrait néanmoins de problèmes de sécurité dans la vérification de l'intégrité des données envoyées ou reçues, la rendant vulnérable à des attaques actives. En outre, cette version implémentait un système sommaire de transmission de fichiers, et du port tunneling.

La version 2 qui était à l'état de brouillon jusqu'en janvier 2006 est déjà largement utilisée à travers le monde.

Cette version est beaucoup plus sûre au niveau cryptographique, et possède en plus un protocole de transfert de fichiers complet, le SSH file transfer protocol.

Habituellement le protocole SSH utilise le port TCP 22. Il est particulièrement utilisé pour ouvrir un shell sur un ordinateur distant. Peu utilisé sur les stations Windows (quoiqu'on puisse l'utiliser avec PuTTY, mRemote, cygwin ou encore OpenSSH), SSH fait référence pour l'accès distant sur les stations Linux et Unix.

SSH peut également être utilisé pour transférer des ports TCP d'une machine vers une autre, créant ainsi un tunnel. Cette méthode est couramment utilisée afin de sécuriser une connexion qui ne l'est pas (par exemple le protocole de récupérations de courrier électronique POP3) en la faisant transférer par le biais du tunnel chiffré SSH.

Il est également possible de faire plusieurs sauts entre consoles SSH, c'est-à-dire ouvrir une console sur un serveur, puis, de là, en ouvrir une autre sur un autre serveur.

VI.3.3 SSH avec authentification par clés

Avec SSH, l'authentification peut se faire sans l'utilisation de mot de passe ou de phrase secrète en utilisant la cryptographie asymétrique. La clé publique est distribuée sur les systèmes auxquels on souhaite se connecter. La clé privée, qu'on prendra le soin de protéger par un mot de passe, reste uniquement sur le poste à partir duquel on se connecte. L'utilisation d'un « agent ssh » permet de stocker le mot de passe de la clé privée pendant la durée de la session utilisateur.

Note :

Cette configuration profite aussi à SCP (Secure Copy) et à SFTP (Secure File Transfert Protocol) qui se connectent au même serveur SSH.

VI.3.4 Implémentations logicielles

- OpenSSH, le projet libre d'outils SSH. OpenSSH est l'implémentation ssh la plus utilisée, y compris par les distributions GNU/Linux.
- Portable OpenSSH, une implémentation OpenSSH multiplateforme.
- lsh2, une implémentation distribuée par le projet GNU selon les termes de la licence GNU GPL.
- MacSSH3, une implémentation lsh pour MacOS classic 68k et PPC.
- FRESH4, une implémentation ssh en environnement JBoss.
- SSHWindows5, une implémentation pour Windows non maintenue.
- Dropbear6, une implémentation libre ayant pour but de remplacer OpenSSH sur les systèmes Unix ayant peu de ressources (processeur, mémoire, etc.) comme les systèmes embarqués.
- PuTTY, un client SSH multi-OS.
- TTyEmulator - Émulateur de terminal propriétaire, en français sous Windows incluant un grand nombre de fonctionnalités.

Il en existe bien évidement d'autres.

VI.4. TRANSFERT OU COPIE DE FICHIERS

VI.4.1 Introduction

Nous avons plusieurs solutions pour transférer ou de copier des fichiers à travers un réseau :

- FTP (File Transfert Protocol)
- TFTP (Trivail File Transfert Protocol)
- SFTP (Secure File Transfert Protocol) et SCP (Secure Copy)

VI.4.2 FTP

File Transfer Protocol (RFC 3659) est un protocole de communication destiné au partage de fichiers sur un réseau TCP/IP. Il permet, depuis un ordinateur, de copier des fichiers vers un autre ordinateur du réseau, ou encore de supprimer ou de modifier des fichiers sur cet ordinateur. Ce mécanisme de copie est souvent utilisé pour alimenter un site Web hébergé chez un tiers.

La variante de FTP protégée par les protocoles SSL ou TLS (SSL étant le prédécesseur de TLS) s'appelle FTPS.

FTP obéit à un modèle client-serveur, c'est-à-dire qu'une des deux parties, le client, envoie des requêtes auxquelles réagit l'autre, appelé serveur. En pratique, le serveur est un ordinateur sur lequel fonctionne un logiciel lui-même appelé serveur FTP, qui rend publique une arborescence de fichiers similaire à un système de fichiers UNIX. Pour accéder à un serveur FTP, on utilise un logiciel client FTP (possédant une interface graphique ou en ligne de commande).

FTP, qui appartient à la couche application du modèle OSI et du modèle TCP/IP, utilise une connexion TCP.

Par convention, deux ports sont attribués (well known ports) pour les connexions FTP : le port 21 pour les commandes et le port 20 pour les données. Pour le FTPS dit implicite, le port conventionnel est le 990.

Ce protocole peut fonctionner avec IPv4 et IPv6.

VI.4.2.1 Utilisation

Pour accéder à un serveur FTP, on utilise un logiciel (client FTP). Ces logiciels existent avec ligne de commande ou avec une interface graphique. Le standard FTP est si répandu que ces logiciels sont à présent inclus avec les dernières distributions Windows & Unix/Linux.

L'utilisation en ligne de commande, sous Windows comme sous Linux, se fait généralement au moyen de la commande ftp adresse_du_serveur saisie dans une console (ici, ftp est le nom du logiciel, et le paramètre le nom du serveur).

Dans les interfaces graphiques, comme les navigateurs Web, la forme usuelle est utilisée, à savoir : ftp://adresse_du_serveur. Ici, ftp est le nom du protocole, suivi du nom du serveur.

Sous Microsoft Windows Vista et les versions suivantes, un logiciel client FTP est installé sur la machine. On peut y accéder à travers le dossier Ordinateur puis la commande Ajouter un emplacement réseau.

VI.4.2.2 Mise en œuvre

Logiciels clients de FTP :

- GNU inetutils : paquet logiciel GNU contenant un client FTP en ligne de commande
- ftp (en ligne de commande sous Unix/Linux/Windows)
- cURL (en ligne de commande sous Linux/OS X/Windows)
- Cyberduck (pour Mac OS X et Windows)
- FileZilla (pour Linux, Mac OS X et Windows)
- FireFTP (extension pour Firefox)
- gFTP (pour GNOME)
- NcFTP (Windows et systèmes de type UNIX)
- Yafc
- WinSCP

Logiciels serveurs de FTP :

- FileZilla Server (Windows)
- VsFTPd (Unix)
- ProFTPd (Unix)
- Pure-FTPd (Unix)

VI.4.2.3 Interopérabilité

Le protocole FTP ne permet pas toujours d'assurer l'interopérabilité entre plateformes différentes et régions différentes par une gestion adéquate de l'encodage des noms de fichiers. Seuls les logiciels serveur et client respectant le standard RFC 26407 en donnent la garantie 8, grâce à l'utilisation de l'encodage UTF-8 et accessoirement d'une nouvelle commande LANG permettant de choisir la langue des messages retournés par le serveur lors de la session FTP. L'encodage UTF-8 permet d'encoder les noms des fichiers provenant de n'importe quel pays, bien qu'un encodage plus spécifique puisse toujours être utilisé localement par le serveur, la conversion vers l'UTF-8 restant à sa discrétion.

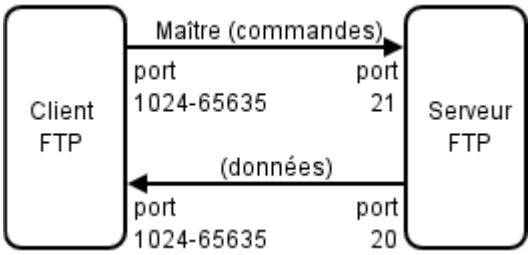
VI.4.2.4 Le protocole

Le protocole utilise deux types de connexions TCP :

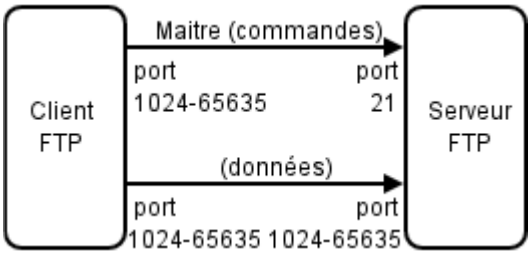
- Une connexion de contrôle initialisée par le client, vers le serveur (port 21 en général), pour transmettre les commandes concernant les fichiers (suppression de fichiers, renommage, liste des fichiers…).
- Une connexion de données initialisée par le client ou le serveur pour transférer les données requises (contenu des fichiers, liste de fichiers).

- **Établissement des connexions**

FTP peut s'utiliser de deux façons différentes.



En mode actif, c'est le client FTP qui détermine le port de connexion à utiliser pour permettre le transfert des données. Ainsi, pour que l'échange des données puisse se faire, le serveur FTP initialisera la connexion de son port de données (port 20) vers le port spécifié par le client. Le client devra alors configurer son pare-feu pour autoriser les nouvelles connexions entrantes afin que l'échange des données se fasse. De plus, il peut s'avérer problématique pour les utilisateurs essayant d'accéder à des serveurs FTP lorsque ces utilisateurs sont derrière une passerelle NAT. Étant donnée la façon dont fonctionne le NAT, le serveur FTP lance la connexion de données en se connectant à l'adresse externe de la passerelle NAT sur le port choisi. Certaines passerelles NAT n'ayant pas de correspondance pour le paquet reçu dans la table d'état, le paquet sera ignoré et ne sera pas délivré au client.



En mode passif, le serveur FTP détermine lui-même le port de connexion à utiliser pour permettre le transfert des données (data connexion) et le communique au client. En cas de présence d'un pare-feu devant le serveur, celui-ci devra être configuré pour autoriser la connexion de données. L'avantage de ce mode est que le serveur FTP n'initialise aucune connexion. Ce mode fonctionne sans problème avec des clients derrière une passerelle NAT. Dans les nouvelles implémentations, le client initialise et communique directement par le port 21 du serveur ; cela permet de simplifier les configurations des pare-feu serveur.

- **Mode de transfert**

Lors du transfert de fichier sur la connexion de données, deux modes peuvent être utilisés :

- Le mode binaire : le fichier est transmis tel quel.
- Le mode ASCII : uniquement destiné aux fichiers texte. Le fichier est examiné et des transformations apportées pour conserver un format correct. Par exemple, la fin de ligne est représentée par le caractère <LF> sur un système UNIX, et par la paire <CR><LF> sous Windows. Une machine Windows recevant un fichier texte par FTP récupère donc finalement un fichier avec des <CR><LF> en mode ASCII et des <LF> en mode binaire. Ce mode a donc ses avantages, mais peut être source de corruption de fichiers (non texte) pendant le transfert si on utilise un client ancien / en ligne de commande, incapable de s'adapter au type de fichier. Il faut alors basculer en mode binaire (en utilisant généralement la commande BIN) avant le transfert, afin de le conserver intact.

VI.4.3 TFTP

Trivial File Transfer Protocol est un protocole simplifié de transfert de fichiers.

Il fonctionne en UDP sur le port 69, au contraire du FTP qui utilise lui TCP. L'utilisation d'UDP, protocole « non fiable », implique que le client et le serveur doivent gérer eux-mêmes une éventuelle perte de paquets. En termes de rapidité, l'absence de fenêtrage nuit à l'efficacité du protocole sur les liens à forte latence. On réserve généralement l'usage du TFTP à un réseau local.

Les principales simplifications visibles du TFTP par rapport au FTP sont qu'il ne gère pas le listage de fichiers, et ne dispose pas de mécanismes d'authentification, ni de chiffrement. Il faut connaître à l'avance le nom du fichier que l'on veut récupérer. De même, aucune notion de droits de lecture/écriture n'est disponible en standard.

À cause de ces fonctionnalités absentes, FTP lui est généralement préféré. TFTP reste très utilisé pour la mise à jour des logiciels embarqués sur les équipements réseaux (routeurs, pare-feu, etc.) ou pour démarrer un PC à partir d'une carte réseau.

La dernière version de ce protocole est la version 2, définie dans RFC 1350. Elle est la plus utilisée.

VI.4.4 SFTP

Dans le contexte de SSH (Secure Shell, SFTP décrit ces deux choses-ci :

Un protocole de communication fonctionnant au-dessus de SSH pour transférer et gérer des fichiers à distance ;

Un programme en ligne de commande qui implémente la partie cliente de ce protocole de communication, comme celui fourni par OpenSSH.

Comparé à SCP, le protocole SFTP supporte beaucoup plus d'opérations sur des fichiers à distance. Il se comporte plus comme un protocole de système de fichiers. Il est censé être plus indépendant de la plate-forme d'utilisation ; par exemple, avec scp, l'extension des wildcards (*) spécifiés par le client sont à la charge du serveur, qui en fait ce qu'il veut, alors que l'architecture de SFTP évite ce genre de problèmes.

Le programme sftp apporte une interface similaire au programme ftp. Le protocole SFTP n'est pas FTP au-dessus de SSL (visitez File Transfer Protocol over SSL), c'est un nouveau protocole conçu intégralement par le groupe de travail IETF SECSH. Il n'existe pas de RFC décrivant le protocole SFTP, mais seulement un brouillon (draft).

Certaines implémentations du programme scp utilisent en fait le protocole SFTP à la place du protocole scp.

sftp est souvent associé au protocole (et au programme) SSH-2, parce qu'ils ont été conçus en même temps par le même groupe. Cependant, il est possible de le faire fonctionner sur SSH-1, et certaines implémentations le font.

VI.4.5 SCP

Le protocole SCP (Secure Copy) est basé sur le protocole BSD RCP. À la différence de ce dernier, les données sont chiffrées lors du transfert pour éviter les extractions d'informations utilisables provenant de paquets de données sniffés. Le protocole ne fournit pas lui-même l'authentification et la sécurité nécessaires au transfert, il s'appuie pour cela sur le protocole sous-jacent SSH.

SCP peut demander à l'utilisateur tout mot de passe nécessaire à l'établissement d'une connexion distante, alors que RCP échoue dans cette situation.

Le protocole SCP implémente uniquement le transfert de fichiers. Pour cela, une connexion est établie en utilisant SSH, puis un serveur SCP est lancé. Le programme serveur est le même que le programme client.

Lors d'un téléversement, le client envoie au serveur les fichiers à téléverser en comprenant éventuellement leurs attributs de base (permissions, horodatage). C'est un avantage par rapport au protocole FTP, qui ne prévoit pas d'inclure les attributs lors d'un téléchargement.

Lors d'un téléchargement descendant, le client envoie une requête pour obtenir les fichiers et dossiers à télécharger. Le serveur alimente ensuite le client en incluant les sous-répertoires. Le téléchargement étant piloté par le serveur, il existe un risque de sécurité une fois connecté à un serveur malveillant.

Dans la plupart des applications, le protocole SCP est supplanté par le protocole SFTP lui aussi fondé sur SSH.

VI.5. TRANSFERER DES PAGES HYPERTEXTES

VI.5.1 Protocole HTTP

L'Hypertext Transfer Protocol, plus connu sous l'abréviation HTTP, est un protocole de communication client-serveur développé pour le World Wide Web. HTTPS est la variante du HTTP sécurisée par l'usage des protocoles SSL ou TLS.

HTTP est un protocole de la couche application. Il peut fonctionner sur n'importe quelle connexion fiable, dans les faits on utilise le protocole TCP comme couche de transport. Un serveur HTTP utilise alors par défaut le port 80 et HTTPS le port 443.

Les clients HTTP les plus connus sont les navigateurs Web permettant à un utilisateur d'accéder à un serveur contenant les données. Il existe aussi des systèmes pour récupérer automatiquement le contenu d'un site tel que les aspirateurs de site Web ou les robots d'indexation

Une version du protocole HTTP existait au début des années 90 qui n'avait pas de numéro de version. On la nomma HTTP/0.9.

En mai 1996, HTTP/1.0 voit le jour et est décrit dans la RFC 1945. Cette version supporte les serveurs HTTP virtuels, la gestion de cache et l'identification.

En janvier 1997, HTTP/1.1 devient finalement standard de l'IETF. Il est décrit dans la RFC 2068 de l'IETF, puis dans la RFC 2616 en juin 1999. Cette version ajoute le support du transfert en pipeline et la négociation de type de contenu tel que le format de données, la langue.

En mars 2012, les travaux à propos de HTTP/2.0 démarrent à l'IETF adoptant SPDY comme matériel de départ.

En février 2014, la spécification de HTTP/1.1 a été republiée. Elle a été éclatée en plusieurs RFC et corrigée pour toutes ses imprécisions, RFC 7230 à RFC 7237.

VI.5.2 Serveur HTTP

Un serveur HTTP ou bien un serveur Web est un logiciel servant des requêtes respectant le protocole de communication client-serveur HyperText Transfer Protocol (HTTP) qui a été développé pour le World Wide Web.

CERN httpd est le premier serveur HTTP, inventé en même temps que le World Wide Web en 1990 au CERN de Genève. Il est rapidement devenu obsolète en raison de l'évolution exponentielle des fonctionnalités du protocole.

D'autres serveurs HTTP ont vu le jour :

- Apache HTTP Server de l'Apache Software Foundation, successeur du NCSA HTTPd ;
- Google Web Server de Google ;
- Internet Information Services (IIS) de Microsoft ;
- lighttpd de Jan Kneschke ;
- nginx d'Igor Sysoev…

VI.5.3 Navigateur Web

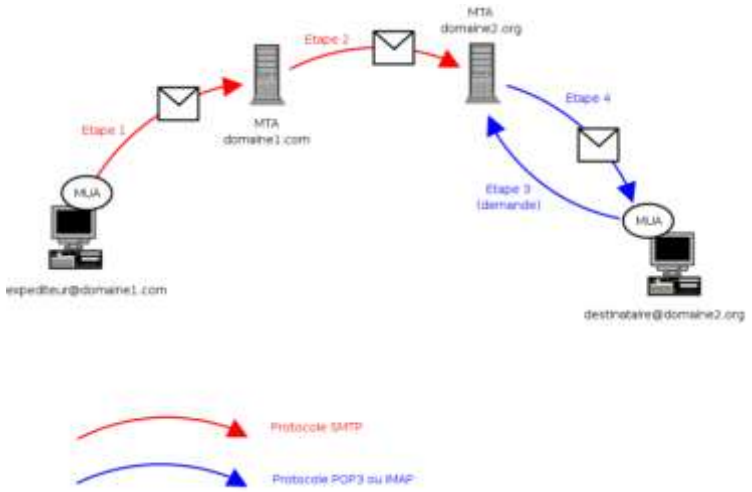
Un navigateur Web (browser) est un logiciel conçu pour consulter et afficher le World Wide Web. Techniquement, c'est au minimum un client HTTP.

Il existe de nombreux navigateurs Web, pour toutes sortes de matériels (ordinateur personnel, tablette tactile, téléphones mobiles...) et pour différents systèmes d'exploitation (GNU/Linux, Windows, MacOS, iOS et Android).

Dans les années 2010, les plus utilisés sont :

- Google Chrome ;
- Mozilla Firefox ;
- Internet Explorer/Edge ;
- Safari ;
- Opera.

VI.6. ÉTUDIER LES PROTOCOLES ASSOCIES A LA MESSAGERIE



VI.6.1 Introduction

Un serveur de messagerie électronique est un logiciel serveur de courrier électronique. Il a pour vocation de transférer les messages électroniques d'un serveur à un autre. Un utilisateur n'est jamais en contact direct avec ce serveur mais utilise soit un client de messagerie, soit une messagerie Web, qui se charge de contacter le serveur pour envoyer ou recevoir les messages.

La plupart des serveurs de messagerie actuels disposent des fonctions d'envoi et de réception, mais elles sont indépendantes, et peuvent être dissociées physiquement.

VI.6.2 Envoi

Le transfert de messages entre serveurs de messagerie électronique se fait généralement sur le port TCP 25 qui est le port standard enregistré auprès de l'IANA. Les serveurs utilisent les enregistrements MX des serveurs DNS pour acheminer le courrier.

Les clients de messagerie utilisaient aussi le port TCP 25 pour soumettre des messages en utilisant le protocole SMTP. Mais la nécessité de mieux contrôler les envois des clients, en particulier par l'authentification, a conduit à l'attribution du port 587 (submission).

Les administrateurs de serveur peuvent choisir si les clients utilisent le port TCP 25 ou le port 587 (submission), tel que formalisé dans la RFC 6409 (RFC 2476 précédemment), pour relayer le courrier sortant vers un serveur de messagerie. Les spécifications et de nombreux serveurs supportent les deux. Bien que certains serveurs ont longtemps pris en charge le port historique 465 (smtps, aussi appelé submissions) pour le SMTP sécurisé, en violation des spécifications jusqu'à fin 2017, il est préférable d'utiliser les ports standard et les commandes ESMTP (Extended SMTP) standard selon la RFC 3207, si une session sécurisée doit être utilisée entre le client et le serveur. Cependant, début 2018, la RFC 8314 a finalement affecté officiellement le port 465 au protocole SMTP avec TLS implicite.

VI.6.3 Livraison

La livraison d'un courrier électronique se déroule elle aussi en deux temps. Le serveur reçoit le message du serveur de l'expéditeur, il doit donc gérer des problèmes comme un disque plein ou une corruption de la boîte aux lettres et signaler au serveur expéditeur toute erreur dans la livraison. Il communique avec ce dernier par l'intermédiaire des canaux d'entrée/sortie standard ou à l'aide d'un protocole spécialisé comme LMTP. Cette fonction de livraison est appelée Mail Delivery Agent (MDA).

Finalement, lorsque le destinataire désire accéder à ses messages, il envoie une requête au serveur, qui en retour lui transmet ses messages, généralement via le protocole POP ou IMAP. La plupart des clients de messagerie peuvent être configurés de manière à interroger régulièrement le serveur de messagerie (par exemple toutes les 20 minutes), ce qui rend l'étape 3 de l'acheminement du courrier complètement transparente pour le destinataire.

VI.6.3.1 Protocole POP

POP (Post Office Protocol) est un protocole qui permet de récupérer les courriers électroniques situés sur un serveur de messagerie électronique. En dehors d'un paramétrage spécifique, POP se connecte au serveur de messagerie, s'authentifie, récupère le courrier, efface éventuellement le courrier sur le serveur, et se déconnecte.

Ce protocole a été réalisé en plusieurs versions; respectivement POP1, POP2 et actuellement POP3.

Cette opération transite sur un réseau TCP/IP et utilise le protocole de transfert TCP via le port 110. Ce protocole est défini par la RFC 1939.

POP3S (POP3 over SSL ou POPS) permet de chiffrer la communication avec le serveur au moyen de TLS. Ce protocole est défini par la RFC 2595. Selon cette dernière, l'usage d'un port spécifique pour ces communications chiffrées (initialement TCP 995 avec le chiffrement SSL) est maintenant déconseillé.

VI.6.3.2 Protocole IMAP

IMAP (Internet Message Access Protocol) est un protocole qui permet d'accéder à ses courriers électroniques directement sur les serveurs de messagerie. Son fonctionnement est donc à l'opposé de POP qui, lui, récupère les messages localement (depuis le poste de travail) via un logiciel spécialisé. L'évolution des différentes versions d'IMAP (IMAP 4) en fait aujourd'hui un protocole permettant également de récupérer les messages localement.

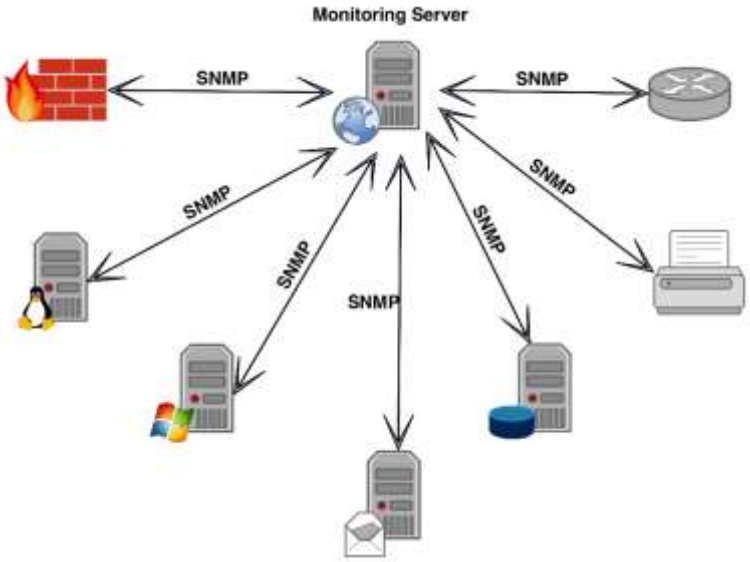
Ce protocole permet de laisser les courriels sur le serveur dans le but de pouvoir les consulter de différents clients de messagerie ou Web mail. Il comporte des fonctionnalités avancées comme la possibilité de créer des dossiers ou de manipuler les messages directement sur le serveur. Il offre aussi la possibilité de trier ses courriels sur le serveur6. Le langage Sieve a été conçu pour permettre de filtrer des messages sur des serveurs sur lesquels l'utilisateur n'a pas le droit d'exécuter des tâches.

Le fait que les messages soient archivés sur le serveur fait que l'utilisateur peut y accéder depuis n'importe où sur le réseau et que l'administrateur peut facilement faire des copies de sauvegarde.

L'inconvénient, est qu'IMAP requiert une connexion permanente. Cependant, depuis IMAP4, de nombreux clients de messagerie proposent un mode « hors-ligne » pour pallier ce problème. D'autre part, il limite l'utilisation de la capacité du réseau car il permet de ne récupérer qu'une partie des messages (par exemple les en-têtes, sans le corps du message). Les messages peuvent être déplacés ou effacés sans être entièrement récupérés par le client.

IMAP utilise le port TCP 143. L'utilisation de TLS permet l'accès sécurisé au serveur. La RFC 2595, qui décrit le fonctionnement de TLS avec IMAP, déconseille l'utilisation du port 993 qui avait été préalablement enregistré pour IMAPS (IMAP over SSL).

VI.7. ADMINISTRATION DES RESEAUX IP



VI.7.1 Introduction

SNMP (Simple Network Management Protocol) est un protocole de communication qui permet aux administrateurs réseau de gérer les équipements du réseau, de superviser et de diagnostiquer des problèmes réseaux et matériels à distance.

VI.7.2 SNMP

Les systèmes de gestion de réseau sont basés sur trois éléments principaux : un superviseur (manager), des nœuds (nodes) et des agents. Dans la terminologie SNMP, le synonyme manager est plus souvent employé que superviseur. Le superviseur est la console qui permet à l'administrateur réseau d'exécuter des requêtes de gestion (management). Les agents sont des entités qui se trouvent au niveau de chaque interface, connectant au réseau l'équipement géré (nœud) et permettant de récupérer des informations sur différents objets.

Commutateurs (switches), concentrateurs (hubs), routeurs, postes de travail et serveurs (physiques ou virtuels) sont des exemples d'équipements contenant des objets gérables. Ces objets gérables peuvent être des informations matérielles, des paramètres de configuration, des statistiques de performance et autres objets qui sont directement liés au comportement en cours de l'équipement en question. Ces objets sont classés dans une sorte de base de données arborescente définie par l'ISO appelée MIB (Management Information Base). SNMP permet le dialogue entre le superviseur et les agents afin de recueillir les objets souhaités dans la MIB.

L'architecture de gestion du réseau proposée par le protocole SNMP est donc fondée sur trois principaux éléments :

Les équipements gérés (managed devices) sont des éléments du réseau (ponts, commutateurs, concentrateurs, routeurs ou serveurs), contenant des « objets de gestion » (managed objects) pouvant être des informations sur le matériel, des éléments de configuration ou des informations statistiques ;

Les agents, c'est-à-dire les applications de gestion de réseau résidant dans un périphérique, sont chargés de transmettre les données locales de gestion du périphérique au format SNMP ;

Les systèmes de gestion de réseau (network management systems notés NMS), c'est-à-dire les consoles à travers lesquelles les administrateurs peuvent réaliser des tâches d'administration.

VI.7.3 MIB

Une MIB (management information base) est un ensemble d'informations structuré sur une entité réseau, par exemple un routeur, un commutateur ou un serveur. Ces informations peuvent être récupérées ou parfois modifiées par un protocole comme SNMP.

La structure de la MIB est hiérarchique : les informations sont regroupées en arbre. Chaque information a un ***object identifier,*** une suite de chiffres séparés par des points, qui l'identifie de façon unique et un nom, indiqué dans le document qui décrit la MIB.

Par exemple, 1.3.6.1.2.1.2.2.1.2 est l'***object identifier ifDescr*** qui est la chaîne de caractères décrivant une interface réseau (comme eth0 sur Linux ou Ethernet0 sur un routeur Cisco).

Une des MIB les plus connues est MIB-II, décrite dans le RFC 1213, et qui est mise en œuvre dans quasiment tous les équipements TCP/IP. Elle compte dix groupes, "system", "interfaces" (dont fait partie ifDescr, citée plus haut), "Address Translation", "IP", "ICMP", "TCP", "UDP", "EGP", "transmission" et "SNMP".

VI.7.4 Logiciels

Un grand nombre de logiciels libres et propriétaires utilisent SNMP pour interroger régulièrement les équipements et produire des graphes rendant compte de l'évolution des réseaux ou des systèmes informatiques :

- Centreon,
- NetCrunch,
- MRTG,
- Cacti,
- Shinken,
- Nagios,
- Zabbix)...

VII. VERS IPV6

Objectifs

- Découvrir les points essentiels du protocole TCP/IPv6

Références

https://tools.ietf.org/html/rfc2460

VII.1. IPv6

VII.1.1 Introduction

IPv6 (Internet Protocol version 6) est un protocole réseau sans connexion de la couche 3 du modèle OSI (Open Systems Interconnection).

IPv6 est l'aboutissement des travaux menés au sein de l'IETF au cours des années 1990 pour succéder à IPv4 et ses spécifications ont été finalisées dans la RFC 24601 en décembre 1998. IPv6 a été standardisé dans la RFC 82002 en juillet 2017.

Grâce à des adresses de 128 bits au lieu de 32 bits, IPv6 dispose d'un espace d'adressage bien plus important qu'IPv4. Cette quantité d'adresses considérable permet une plus grande flexibilité dans l'attribution des adresses et une meilleure agrégation des routes dans la table de routage d'Internet. La traduction d'adresse, qui a été rendue populaire par le manque d'adresses IPv4, n'est plus nécessaire.

IPv6 dispose également de mécanismes d'attribution automatique des adresses et facilite la renumérotation. La taille du sous-réseau, variable en IPv4, a été fixée à 64 bits en IPv6. Les mécanismes de sécurité comme IPSec font partie des spécifications de base du protocole. L'en-tête du paquet IPv6 a été simplifié et des types d'adresses locales facilitent l'interconnexion de réseaux privés.

Le déploiement d'IPv6 sur Internet est compliqué en raison de l'incompatibilité des adresses IPv4 et IPv6. Les traducteurs d'adresses automatiques se heurtent à des problèmes pratiques importants (RFC 4966). Pendant une phase de transition où coexistent IPv6 et IPv4, les hôtes disposent d'une double pile, c'est-à-dire qu'ils disposent à la fois d'adresses IPv6 et IPv4, et des tunnels permettent de traverser les groupes de routeurs qui ne prennent pas encore en charge IPv6.

En 2011, seules quelques sociétés ont entrepris de déployer la technologie IPv6 sur leur réseau interne, notamment Google.

Au début de l'année 2016, le déploiement d'IPv6 est encore limité, la proportion d'utilisateurs Internet en IPv6 étant estimée à 10 %, et ce en dépit d'appels pressants à accélérer la migration adressée aux fournisseurs d'accès à Internet et aux fournisseurs de contenu de la part des registres Internet régionaux et de l'ICANN, l'épuisement des adresses IPv4 publiques disponibles étant imminent.

VII.1.2 Fonctionnement d'IPv6

Le fonctionnement d'IPv6 est très similaire à celui d'IPv4. Les protocoles TCP et UDP sont pratiquement inchangés.

VII.1.2.1 Adresse IPv6

Une adresse IPv6 est longue de 128 bits (16 octets) contre 32 bits (4 octets) pour IPv4. La notation décimale pointée employée pour les adresses IPv4 (par exemple 172.16.32.1) est abandonnée au profit d'une écriture hexadécimale, où les 8 groupes de 2 octets (16 bits par groupe) sont séparés par un signe deux points :

- **2001:0db8:0000:85a3:0000:0000:ac1f:8001**

Il est permis d'omettre d'un à trois chiffres zéros non significatifs dans chaque groupe de quatre chiffres hexadécimaux. Ainsi, l'adresse IPv6 ci-dessus est équivalente à la suivante :

- **2001:db8:0:85a3:0:0:ac1f:8001**

De plus, une unique suite d'un ou plusieurs groupes consécutifs de 16 bits tous nuls peut être omise, en conservant toutefois les signes deux-points de chaque côté de la suite de chiffres omise, c'est-à-dire une paire de deux points « :: » (RFC 2373). Ainsi, l'adresse IPv6 ci-dessus peut être abrégée en la suivante :

- **2001:db8:0:85a3::ac1f:8001**

Une même adresse IPv6 peut être représentée de plusieurs façons différentes, comme 2001:db8::1:0:0:1 et 2001:db8:0:0:1::1. La RFC 5952 recommande une représentation canonique.

Les réseaux sont identifiés en utilisant la notation CIDR : la première adresse du réseau est suivie par une barre oblique « / » puis par un entier compris entre 0 et 128, lequel indique la longueur en bits du préfixe du réseau, à savoir de la partie commune des adresses déterminées par ledit réseau.

Voici des exemples d'adresses réseau IPv6 avec leurs ensembles d'adresses déterminées :

Préfixe	Ensemble des adresses
2001:db8:1f89::/48	De 2001:db8:1f89:0:0:0:0:0 à 2001:db8:1f89:ffff:ffff:ffff:ffff:ffff
2000::/3	De 2000:0:0:0:0:0:0:0 à 3fff:ffff:ffff:ffff:ffff:ffff:ffff:ffff
fc00::/7	de fc00:0:0:0:0:0:0:0 à fdff:ffff:ffff:ffff:ffff:ffff:ffff:ffff
fe80::/10	de fe80:0:0:0:0:0:0:0 à febf:ffff:ffff:ffff:ffff:ffff:ffff:ffff

Certains préfixes d'adresses IPv6 jouent des rôles particuliers :

Préfixe	Description
::/8	Adresses réservées
2000::/3	Adresses unicast routables sur Internet
fc00::/7	Adresses locales uniques
fe80::/10	Adresses locales lien
ff00::/8	Adresses multicast

Deux des adresses réservées de ::/8 peuvent être remarquées :

- ::/128 est l'adresse non spécifiée. On peut la trouver comme adresse source initiale, à l'instar de 0.0.0.0 en IPv4, dans une phase d'acquisition de l'adresse réseau ;
- ::1/128 est l'adresse de boucle locale (dite aussi localhost). Elle est semblable à 127.0.0.1 en IPv424.

Les adresses de 2000::/3 peuvent être distinguées comme suit :

- Les adresses permanentes (2001::/16) sont ouvertes à la réservation depuis 1999 :
  - La plage 2001::/32 est utilisée pour Teredo ;
  - La plage 2001:db8::/32 est dédiée à un adressage de réseau IPv6 au sein de la documentation technique impliquant de tels réseaux. Cet usage réservé est spécifié dans la RFC 3849 ;
  - Les adresses 6to4 (2002::/16) permettent d'acheminer le trafic IPv6 via un ou plusieurs réseaux IPv4 ;
  - Toutes les autres adresses routables (plus des trois quarts de la plage 2000::/3) sont actuellement réservées à un usage ultérieur.

VII.1.2.2 Structure de l'adresse IPv6 unicast globale

champ	*Préfixe de routage global*	*Identificateur de sous-réseau*	*identificateur d'interface*
bits	n	64-n	64

Le préfixe de routage global, de taille variable, représente la topologie publique de l'adresse, autrement dit celle qui est vue à l'extérieur d'un site. La partie sous-réseau constitue la topologie privée. La RFC 4291 indique que toutes les adresses unicast globales doivent avoir une taille d'identificateur d'interface (IID) égale à 64 bits, à l'exception de celles qui débutent par 000 en binaire. Pour les liens point-à-point, il est cependant possible d'utiliser un /127 (RFC 6164). La RFC 7421 explique le choix architectural de cette taille uniforme d'identificateur d'interface qui semble dépasser largement les besoins d'adressage dans un sous-réseau.

VII.1.2.3 Scope

Le scope d'une adresse IPv6 consiste en son domaine de validité et d'unicité.

On distingue :

Les adresses unicast :

- L'adresse loopback ::1/128 a une validité limitée à l'hôte ;
- Les adresses link-local, uniques sur un lien donné ;
- Les autres adresses, y compris les adresses locales uniques, ont un scope global, c'est-à-dire qu'elles sont uniques dans le monde et peuvent être utilisées pour communiquer avec d'autres adresses globalement uniques, ou des adresses link-local sur des liens directement connectés,
- Les adresses anycast, dont le scope est identique aux adresses unicast ;
- Les adresses multicast ff00::/8, pour lesquels les bits 13 à 16 déterminent le scope : local, lien, organisation ou global.
- Toutes les interfaces où IPv6 est actif ont au moins une adresse de scope link-local (fe80::/10).

VII.1.2.4 Indice de zone

Il peut exister plusieurs adresses link-local sur des liaisons différentes d'une même machine, on lève les ambiguïtés en fournissant un indice de zone (RFC 4007) qu'on ajoute à l'adresse après un signe pourcent : fe80::3%eth0 correspondra à l'adresse link-local fe80::3 sur l'interface eth0 par exemple.

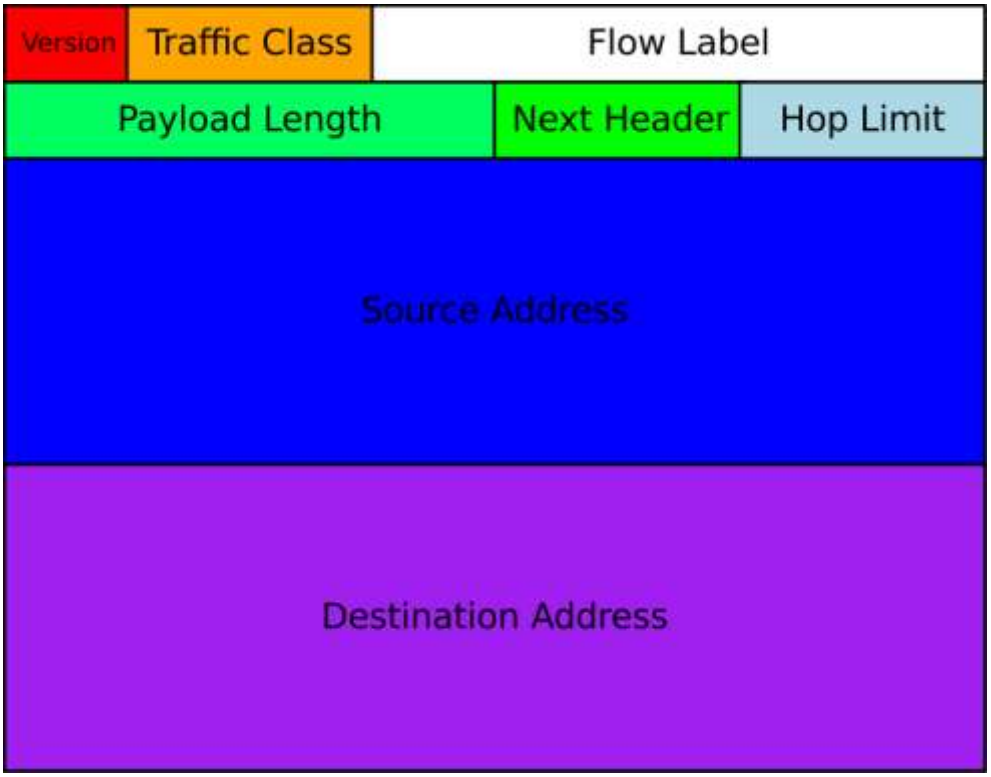
VII.1.3 Attribution des blocs d'adresses IPv6

Dans l'espace d'adresse unicast global (2000::/3), l'IANA attribue des blocs dont la taille varie de /12 a /23 aux registres Internet régionaux30, comme le RIPE NCC en Europe. Ces derniers distribuent des préfixes /32 aux registres Internet locaux qui les attribuent ensuite sous forme de bloc /48 à /64 aux utilisateurs finaux (RFC 6177).

Chaque utilisateur final se voit attribuer un bloc dont la taille varie de /64 (un seul sous-réseau) à /48 (65 536 sous-réseaux), chacun des sous-réseaux pouvant accueillir un nombre d'hôtes virtuellement illimité (264). Dans le bloc 2000::/3 qui représente ⅛e de l'espace d'adressage disponible en IPv6, on peut donc créer 229, soit 500 millions de blocs /32 pour des fournisseurs d'accès à Internet, et 245, soit 35 000 milliards de réseaux d'entreprise typiques (/48).

VII.1.4 En-tête IPv6

L'en-tête du paquet IPv6 est de taille fixe à 40 octets, tandis qu'en IPv4 la taille minimale est de 20 octets, des options pouvant la porter jusqu'à 60 octets, ces options demeurant rares en pratique.



La signification des champs est la suivante :

Version (4 bits)

Fixé à la valeur du numéro de protocole Internet, 6

Traffic Class (8 bits)

Utilisé dans la qualité de service.

Flow Label (20 bits)

Permet le marquage d'un flux pour un traitement différencié dans le réseau.

Payload length (16 bits)

Taille de la charge utile en octets.

Next Header (8 bits)

Identifie le type de header qui suit immédiatement selon la même convention qu'IPv4.

Hop Limit (8 bits)

Décrémenté de 1 par chaque routeur, le paquet est détruit si ce champ atteint 0 en transit.

Source Address (128 bits)

Adresse source

VII.1.4.1 Fragmentation et option jumbo

En IPv4, les routeurs qui doivent transmettre un paquet dont la taille dépasse le MTU du lien de destination ont la tâche de le fragmenter, c'est-à-dire de le segmenter en plusieurs paquets IP plus petits. Cette opération complexe est coûteuse en termes de CPU pour le routeur ainsi que pour le système de destination et nuit à la performance des transferts, d'autre part les paquets fragmentés sont plus sensibles aux pertes : si un seul des fragments est perdu, l'ensemble du paquet initial doit être retransmis.

En IPv6, les routeurs intermédiaires ne fragmentent plus les paquets et renvoient un paquet ICMPv6 Packet Too Big en lieu et place, c'est alors la machine émettrice qui est responsable de fragmenter le paquet. L'utilisation du Path MTU discovery est cependant recommandée pour éviter toute fragmentation.

Ce changement permet de simplifier la tâche des routeurs, leur demandant moins de puissance de traitement.

La MTU minimale autorisée pour les liens a également été portée à 1 280 octets (contre 68 pour l'IPv4). Si des liens ne peuvent pas soutenir ce MTU minimal, il doit exister une couche de convergence chargée de fragmenter et de réassembler les paquets.

Comme pour IPv4, la taille maximale d'un paquet IPv6 hors en-tête est de 65 535 octets. IPv6 dispose cependant d'une option jumbogram (RFC 2675) permettant de porter la taille maximale d'un paquet à 4 Go et profiter ainsi des réseaux avec un MTU plus élevé.

VII.1.4.2 En-têtes d'extension

L'en-tête IPv6 peut être suivi d'un certain nombre d'en-tête d'extensions. Ceux-ci se succèdent, chaque en-tête indiquant la nature du suivant. Quand ils sont présents, leur ordre est le suivant :

Nom	Type	Taille	Description	RFC
Options Hop-By-Hop	0	variable	Contient les options qui doivent être honorées par tous les routeurs de transit, par exemple l'option jumbogram.	2460 2675
Routage	43	variable	Permet de modifier le routage à partir de la source, qui est utilisé notamment par Mobile IPv6	2460 3775 5095
Fragment	44	64 bits	Contient les informations relatives à la fragmentation.	2460
Authentication Header (AH)	51	variable	Contient les informations nécessaires à l'authentification de l'en-tête, voir IPsec.	4302
Encapsulating Security Payload (ESP)	50	variable	Contient les informations relatives au chiffrement du contenu, voir IPsec.	4303
Options de destination	60	variable	Options qui doivent être traitées par la destination finale.	2460
No Next Header	59	vide	Indique qu'il n'y a aucune charge utile qui suit.	2460

VII.1.5 Neighbor Discovery Protocol

Le Neighbor Discovery Protocol (ND, RFC 4861) associe les adresses IPv6 à des adresses MAC sur un segment, comme ARP pour IPv4. Il permet également de découvrir les routeurs et les préfixes routés, le MTU, de détecter les adresses dupliquées, les hôtes devenus inaccessibles et l'auto configuration des adresses et éventuellement les adresses des serveurs DNS récursifs (RDNSS, RFC 5006). Il s'appuie sur ICMPv6.

VII.1.6 Attribution des adresses IPv6

Construction d'une adresse d'interface EUI-64 modifiée à partir d'une adresse MAC.

Dans un sous-réseau, il existe plusieurs méthodes d'attribution des adresses :

VII.1.6.1 Configuration manuelle

L'administrateur fixe l'adresse. Les adresses constituées entièrement de 0 ou de 1 ne jouent pas de rôle particulier en IPv6.

VII.1.6.2 Configuration automatique

Auto configuration sans état (Stateless Address Autoconfiguration, SLAAC) basée sur l'adresse MAC qui utilise le Neighbor Discovery Protocol (NDP) (RFC 4862).

- **Auto configuration avec tirage pseudo aléatoire (RFC 4941)**

Utilisation d'adresses générées cryptographiquement (CGA, RFC 3972), qui lient l'adresse à la clé publique du client et qui peuvent être utilisées par SEND,

- **Attribution par un serveur DHCPv6 (RFC 3315)**

L'utilisation de l'adresse MAC d'une carte réseau pour construire une adresse IPv6 a suscité des inquiétudes quant à la protection des données personnelles54 dans la mesure où l'adresse MAC permet d'identifier de façon unique le matériel. Pour pallier cet inconvénient, il est possible d'utiliser des adresses temporaires générées de façon pseudo-aléatoire et modifiées régulièrement (RFC 4941) ou bien d'utiliser un service d'attribution automatique des adresses IPv6 par un serveur, de façon similaire à ce qui existe pour IPv4, avec DHCPv6.

VII.1.7 Multicast

Le multicast, qui permet de diffuser un paquet à un groupe, fait partie des spécifications initiales d'IPv6. Cette fonctionnalité existe également en IPv4 où il a été ajouté par la RFC 988 en 1986.

Il n'y a plus d'adresse broadcast en IPv6, celle-ci étant remplacée par une adresse multicast spécifique à l'application désirée. Par exemple, l'adresse ff02::101 permet de contacter les serveurs NTP sur un lien. Les hôtes peuvent ainsi filtrer les paquets destinés à des protocoles ou des applications qu'ils n'utilisent pas, et ce sans devoir examiner le contenu du paquet.

Au niveau Ethernet, une série de préfixes OUI est réservée aux adresses IPv6 multicast (33:33:xx). L'adresse MAC du groupe multicast consistera en ces 16 bits que l'on fait suivre par les 32 derniers bits de l'adresse IPv6 multicast. Par exemple, l'adresse ff02::3:2 correspondra à l'adresse MAC 33:33:00:03:00:02. Bien que de nombreux groupes multicast partagent la même adresse MAC, ceci permet déjà un filtrage efficace au niveau de la carte réseau.

Bien que la prise en charge de multicast au niveau des liens soit obligatoire pour IPv6, le routage des paquets multicast au-delà du segment requiert l'utilisation de protocoles de routage comme PIM, à la discrétion du fournisseur d'accès à Internet.

Le protocole Multicast Listener Discovery permet d'identifier les groupes actifs sur un segment, à l'instar d'IGMP pour IPv4.

Les commutateurs Ethernet les plus simples traitent les trames multicast en les diffusant comme des trames broadcast. Ce comportement est amélioré avec MLD Snooping qui limite la diffusion aux seuls hôtes manifestant un intérêt pour le groupe, à l'instar d'IGMP Snooping pour IPv4.

Alors qu'en IPv4 il est difficile de réserver des adresses multicast globales, la RFC 3306 associe un bloc d'adresses multicast /96 pour chaque préfixe routable sur Internet, c'est-à-dire que chaque organisation dispose automatiquement de 4 milliards d'adresses multicast publiques. La RFC 3956 simplifie également la réalisation de points de rendez-vous pour les interconnexions multicast.

VII.1.8 DNS

Dans le Domain Name System, les noms d'hôtes sont associés à des adresses IPv6 grâce à l'enregistrement AAAA.

www.ipv6.ripe.net. IN AAAA 2001:610:240:22::c100:68b

L'enregistrement inverse est réalisé sous ip6.arpa en inversant l'adresse écrite sous forme canonique (RFC 3596) :

b.8.6.0.0.0.1.c.0.0.0.0.0.0.0.0.2.2.0.0.0.4.2.0.0.1.6.0.1.0.0.2.ip6.arpa. IN PTR www.ipv6.ripe.net.

La première mouture de la norme prévoyait d'utiliser le suffixe ip6.int.

Le mécanisme utilisé pour construire le nom de domaine inverse est similaire à celui employé en IPv4, à la différence que les points sont utilisés entre chaque nibble (groupe de 4 bits), ce qui allonge le domaine.

Les plus complexes bitlabels (RFC 2673), DNAME et A6 (RFC 2874), qui permettent de s'affranchir de la contrainte de la délégation sur une frontière de nibble, sont considérés comme expérimentaux et leur support est rare (RFC 3363, l'enregistrement A6, inusité, est relégué à l'état « historique » par la RFC 6563 en 2012).

La résolution inverse peut être utilisée par des systèmes de contrôle d'accès ainsi que par des outils de diagnostic comme traceroute.

VII.1.9 Traduction d'adresse

Le recours à la traduction d'adresse est découragé en IPv6 pour préserver la transparence du réseau, son utilisation n'est plus nécessaire pour économiser des adresses.

VII.1.10 IPv6 et mobilité

IPv6 prévoit des mécanismes pour conserver une même adresse IPv6 pour une machine pouvant être connectée à des réseaux différents, tout en évitant autant que possible le routage triangulaire.

VII.1.11 Technologies de transition pour l'accès à l'Internet IPv6

Les adresses IPv4 et IPv6 ne sont pas compatibles, la communication entre un hôte ne disposant que d'adresses IPv6 et un hôte ne disposant que d'adresses IPv4 constitue donc un problème. La transition consiste à doter les hôtes IPv4 d'une double pile, c'est-à-dire à la fois d'adresses IPv6 et IPv4.

La manière la plus simple d'accéder à IPv6 est lors de l'abonnement de choisir un FAI qui offre de l'IPv6 nativement, c'est-à-dire sans recours à des tunnels.

À défaut, et pendant une phase de transition, il est possible d'obtenir une connectivité IPv6 via un tunnel. Les paquets IPv6 sont alors encapsulés dans des paquets IPv4, qui peuvent traverser le réseau du FAI jusqu'à un serveur qui prend en charge IPv6 et IPv4, et où ils sont décapsulés. Le recours à des tunnels, et donc à un réseau overlay, est de nature à nuire aux performances.

VII.1.12 Tunnels statiques

Plusieurs services du type « tunnel broker » sont disponibles, nécessitant en général une inscription. On peut citer SixXS64, ou Hurricane Electric65.

Les protocoles utilisés peuvent être :

- 6in4 (RFC 4213) fait usage du numéro de protocole 41 d'IP et est donc parfois bloqué par des pare-feux et les NAT.
- AYIYA67 permet le transport sur UDP ou TCP et gère le changement d'adresse IP.
- GRE utilise le numéro de protocole 47.

Le Tunnel Setup Protocol (RFC 5572) facilite la création des tunnels et permet la mobilité et l'authentification. Le Tunnel Information and Control Protocol, utilisé par AICCU (en), automatise la création des tunnels.

VII.1.13 Tunnels automatiques

6to4 (RFC 3056) si une adresse IPv4 publique (de préférence fixe) est disponible, 6to4 est simple à mettre en place. Pour l'accès aux adresses IPv6 hors du préfixe 2002::/16 (réservé pour 6to4), une adresse relais anycast est réservée, 192.88.99.1.

6rd (RFC 5569) est similaire au précédent. Il ne fait pas usage du préfixe 2002::/16 mais d'un préfixe spécifique au fournisseur d'accès.

6over4 (RFC 2529) permet la connexion à travers un réseau IPv4 qui prend en charge multicast

ISATAP (RFC 5214), une amélioration du précédent qui ne requiert pas le support multicast.

Teredo (RFC 4380) utilisable dans un réseau d'adresses IPv4 privées, relié à Internet via un routeur assurant une traduction d'adresses. Une implémentation de Teredo fait partie de la pile IPv6 des systèmes Windows, et une implémentation pour Linux et les systèmes BSD est miredo74.

VII.1.14 Passerelles applicatives

Il est possible de faire usage de serveurs qui disposent d'une double pile et qui font office de passerelle applicative (Application-Level gateway, ALG), un serveur mandataire Web, par exemple.

NAT-PT combine la traduction d'adresse réseau et un serveur DNS pour permettre la communication entre des systèmes IPv4 et des systèmes IPv6. Il est défini dans la RFC 2766 mais a été rendu obsolète par la RFC 4966 en raison de problèmes causés.

VII.1.15 Multihoming

Le multihoming consiste, pour un réseau, à disposer de plusieurs fournisseurs de transit dans le but d'augmenter la fiabilité de l'accès Internet. En IPv4, ceci est généralement accompli en disposant d'un numéro d'AS propre, d'une plage d'adresse IP de type Provider Independent (PI) et en utilisant BGP pour échanger des routes de façon dynamique avec chacun des fournisseurs d'accès.

Cette façon de réaliser le multihoming consomme des numéros d'AS et augmente la taille de la table de routage Internet en raison de préfixes PI qu'il n'est pas possible d'agréger.

La standardisation du multihoming en IPv6 a tardé, une des ambitions initiales de l'architecture IPv6 étant de n'utiliser que des adresses de type Provider Aggregatable (PA) pour réduire la taille de la table de routage Internet. Dans cette optique, le multihoming était réalisé en attribuant autant d'adresses PA qu'il y a de fournisseurs, les mécanismes d'IPv6 comme l'attribution automatique et la durée de vie limitée des adresses facilitant les changements d'adresses liées aux changements de fournisseurs. Par conséquent, les registres Internet régionaux ne distribuaient pas de bloc PI pour IPv6 jusqu'à récemment.

En 2009, les RIR, comme le RIPE NCC, ont modifié leur politique en acceptant d'attribuer des blocs PI aux entreprises qui veulent se connecter à plusieurs fournisseurs77, la taille minimale du bloc PI est de /48, alors que la taille des blocs PA est /32. Ceci permet de réaliser le multihoming de la même façon qu'en IPv4.

D'autre approches possibles sont basées sur la séparation de l'identificateur et du localisateur (Identifier / Locator Separation) :

- SHIM6 (RFC 5533)
- Host Identity Protocol (RFC 442379, RFC 5102)
- Stream Control Transmission Protocol
- GSE/8+881,82.
- Locator/Identifier Separation Protocol (LISP)83
- NPTv6, soit la traduction de préfixe (RFC 6296)

Ceci est un sujet de recherche confié au groupe de travail Routing Research de l'Internet Research Task Force.

Découvrez également l'ensemble des stages à votre disposition sur notre site

http://www.m2iformation.fr