

U.E. ARES

Architecture des Réseaux

Cours 4/6 : Couche réseau

Olivier Fourmaux
(olivier.fourmaux@upmc.fr)

Version 5.4



Couche R  seau

La **Couche R  seau** achemine les paquets de la source vers les destinataires en effectuant des sauts entre les diff  rents **n  uds interm  diaires**

- de bout-en-bout (*end-to-end*)
- connaissance de la topologie
- calcul du chemin (**routage**)
- adressage virtuel
- abstraction des technologies sous-jacentes
 - ✓ encapsulation sur chaque technologie
 - ✓ fragmentation
 - ✓ conversion d'adresses



Plan

Rappels sur la couche r  seau

La couche r  seau dans TCP/IP

Structure du paquet IPv4

Adressage classique IPv4

Adressage CIDR

Translation d'adresses

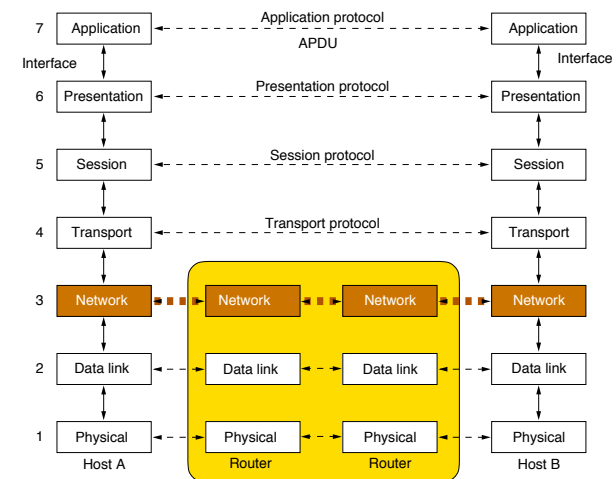
Messages de contr  le

Autoconfiguration

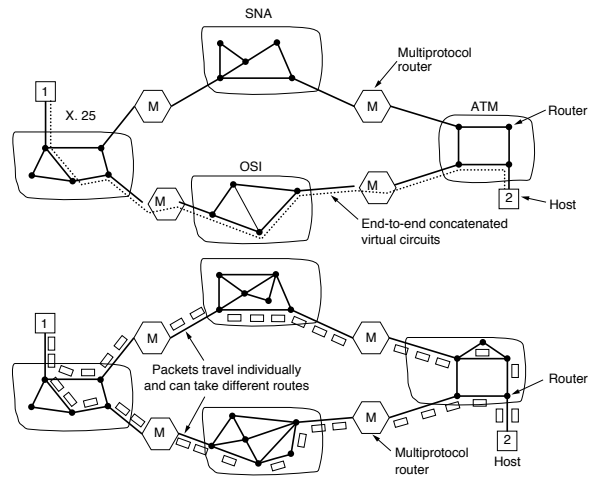
Tunnel et pare-feu



Couche R  seaux : OSI

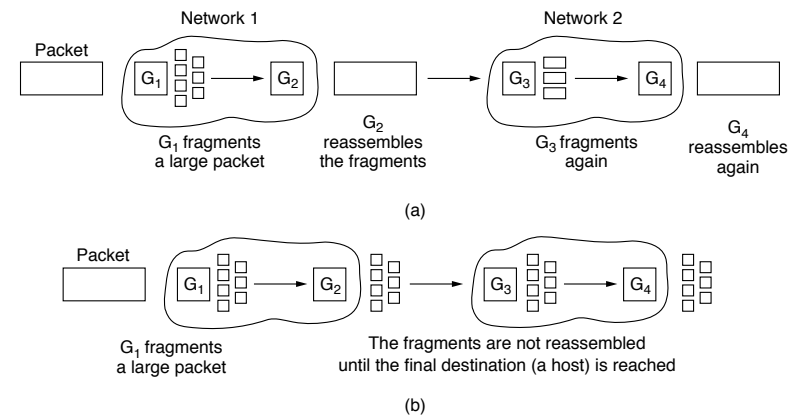


Couche Réseau : approche circuit virtuel ou datagramme



pictures from TANENBAUM A. S. *Computer Networks 3rd edition*

Couche Réseau : Fragmentation

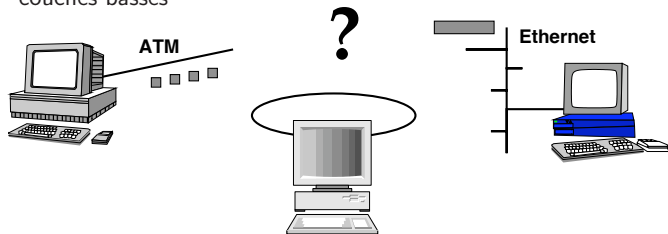


pictures from TANENBAUM A. S. *Computer Networks 3rd edition*

Couche Réseau : Encapsulation

La couche réseau fait abstraction des technologies sous-jacentes

- les données doivent pouvoir circuler de réseaux en réseaux
- les couches supérieures ne doivent faire aucune hypothèse sur les couches basses

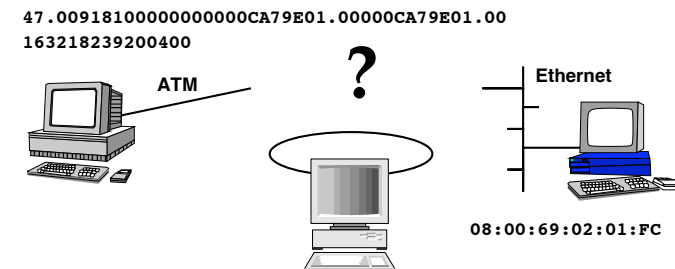


➡ sera approfondie dans les cours sur les **Architectures supports**

Couche Réseau : Adressage

La couche réseau définit un **adressage virtuel** valide sur tous les réseaux

- identification unique d'un équipement
- masquage des mécanismes d'adressages spécifiques à une technologie
✓ nécessite la mise en correspondance des adresses



➡ sera aussi approfondi dans les cours sur les **Architectures supports**

Couche Réseau : Routage

Calcul du chemin

- initial (circuits virtuels)
- à chaque paquet (sans mémoire)

Décisions de routage basée :

- table de routage
 - ✓ statique
 - ✓ dynamique
 - ☞ algorithmes de routage
 - ☞ protocoles de routage...

➡ sera approfondi dans le cours sur le **Routage**

Plan

Rappels sur la couche réseau

La couche réseau dans TCP/IP

Structure du paquet IPv4

Adressage classique IPv4

Adressage CIDR

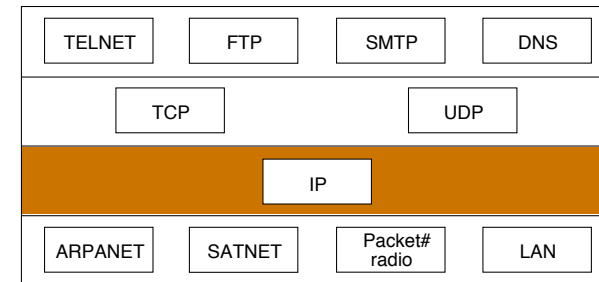
Translation d'adresses

Messages de contrôle

Autoconfiguration

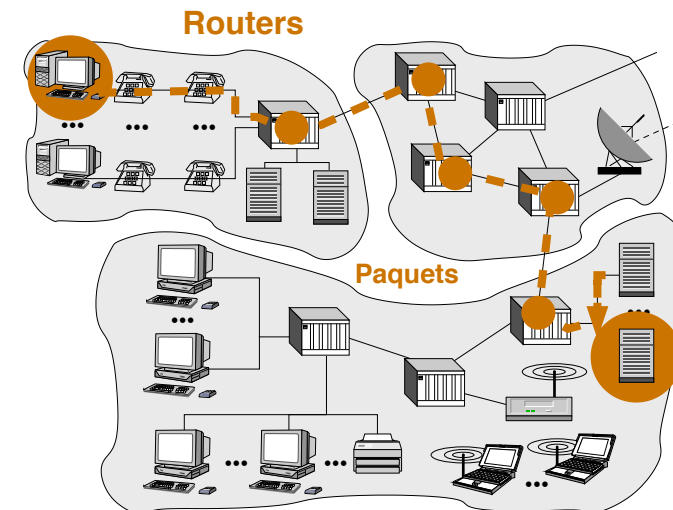
Tunnel et pare-feu

Couche Réseaux : TCP/IP



➡ IP est l'interface universelle

IPv4

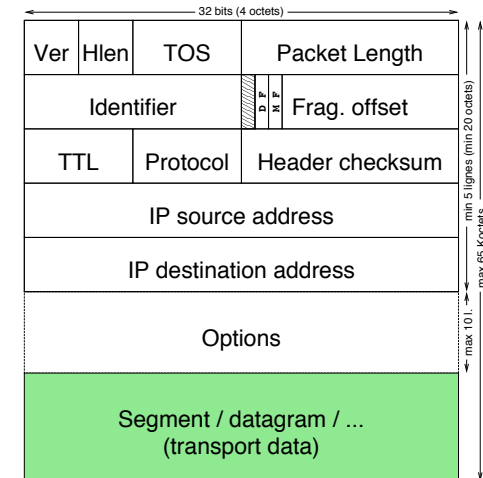


IPv4 : Service

Paquets en mode non connecté

Service à remise non garantie (*best effort*)

IPv4 : Structure



Plan

Rappels sur la couche réseau

La couche réseau dans TCP/IP

Structure du paquet IPv4

Adressage classique IPv4

Adressage CIDR

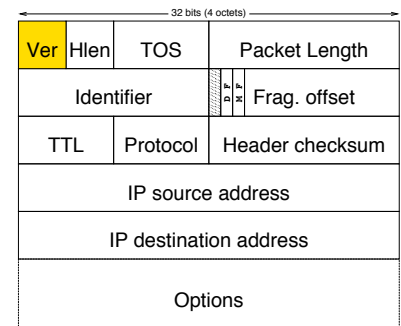
Translation d'adresses

Messages de contrôle

Autoconfiguration

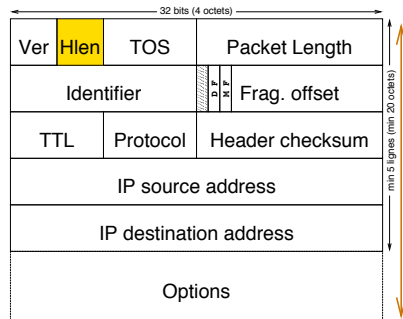
Tunnel et pare-feu

IPv4 : Version



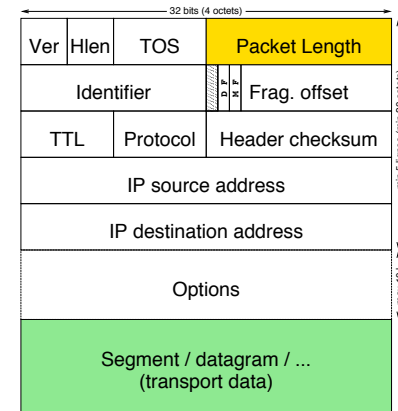
- 4 bits
- IP actuel : version 4
- IP *next génération* : version 6
➡ voir l'U.E. ING

IPv4 : Longueur de l'entête



- 4 bits (valeur 15 max)
 - ✓ indique le nombre de lignes de 32 bits dans l'entête IP
 - ↳ nécessaire car le champ option est de longueur variable (20 à 60 octets)
 - ↳ valeur de 5 (pas d'options) à 15 (10 lignes d'options, soit 40 octets)

IPv4 : Taille du paquet



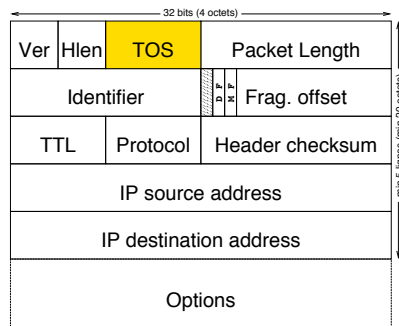
- 16 bits (64 Koctets maximum)
 - ✓ taille totale du paquet **avec entête**
 - ✓ exprimé en octets
 - ↳ le réseau support doit accepter un MTU¹ ≥ **576 octets**²

¹MTU : Maximum Transmission Unit

²576 octets = 512 de données applicative + 64 de surcoût protocolaires (entêtes IP et transport)

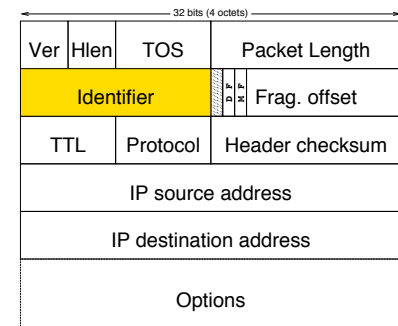
IPv4 : TOS

Type Of Service



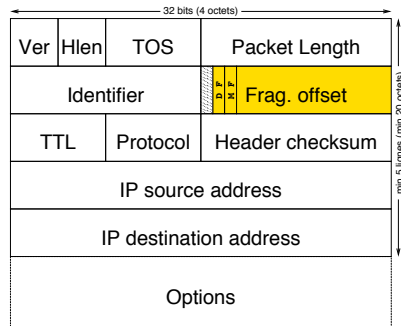
- 8 bits :
 - ✓ 3 bits de **priorité** (*precedence*)
 - ↳ 000 : Routine
 - ↳ 001 : Priority
 - ↳ 010 : Immediate
 - ↳ 011 : Flash
 - ↳ 100 : Flash override
 - ↳ 110 : Internetwork control
 - ↳ 111 : Network control
 - ✓ 3 bits de **service**
 - ↳ Delay
 - ↳ Throughput
 - ↳ Reliability
 - ↳ (Cost)
- non utilisé...
 - ... pour le moment
 - ➡ voir l'U.E. **ING** (*DiffServ Byte*)

IPv4 : Identificateur



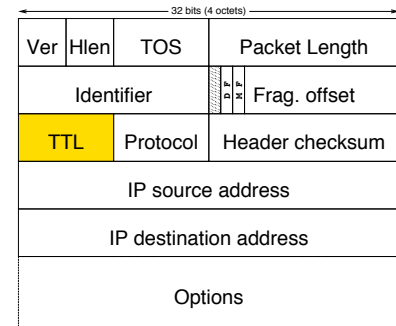
- 16 bits (boucle tous les 64 Kpaquets)
- défini de manière **unique** pour chaque paquet
- pour réassembler les fragments d'un **même** paquet
- habituellement, **incrément** d'un compteur pour chaque paquet successif

IPv4 : Fragmentation



- 1 bit réservé
- 1 bit DF : *Don't fragment* (=1 interdit la fragmentation)
- 1 bit MF : *More fragment* (=0 pour le dernier fragment)
- 13 bits *fragment offset* en octets/8 (shift 3)
 - ✓ exemples :
 - ☞ 0x0000 paquet entier (*offset*=0)
 - ☞ 0x2000 premier fragment (*offset*=0)
 - ☞ 0x20A0 fragment central (*offset*=1280)
 - ☞ 0x00B0 dernier fragment (*offset*=1408)

IPv4 : Temps de vie

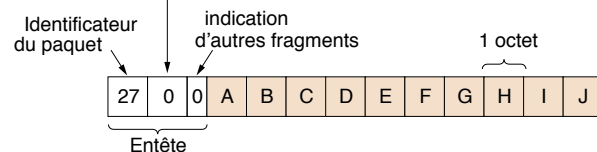


Time To Live

- 8 bits
 - ✓ unité initiale : **seconde**
 - ✓ valeur maximum fixé par l'émetteur (255, 128, 64...)
 - ✓ décrémentation dans chaque routeur
 - ☞ minimum 1 par routeur
 - ☞ nombre de **sauts**
 - ✓ max 255 secondes ou sauts
 - ☞ **évite les boucles**

IPv4 : Fragmentation

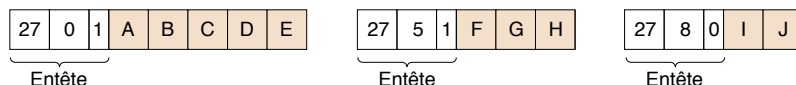
Numéro du premier élément du segment contenu dans ce paquet



(a)

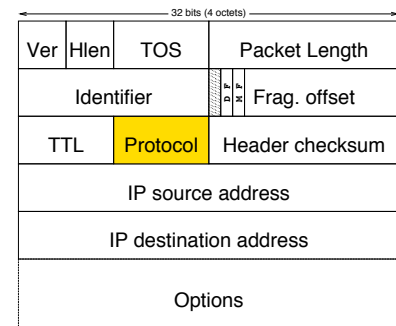


(b)



(c)

IPv4 : Protocole transporté

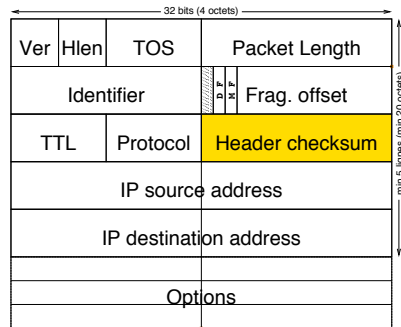


- démultiplexage vers les protocoles de la couche supérieure :


```

Unix> cat /etc/protocols
icmp 1 # internet control message protocol
ggp 3 # gateway-gateway protocol
ipencap 4 # IP encapsulated in IP
st 5 # ST datagram mode
tcp 6 # transmission control protocol
egp 8 # exterior gateway protocol
udp 17 # user datagram protocol
rdp 27 # "reliable datagram" protocol
iso-tp4 29 # ISO Transport Protocol class 4
xtp 36 # Xpress Transfer Protocol
idrp 45 # Inter-Domain Routing Protocol
rsvp 46 # Reservation Protocol
gre 47 # General Routing Encapsulation
ospf 89 # Open Shortest Path First IGP
...
            
```
- 8 bits

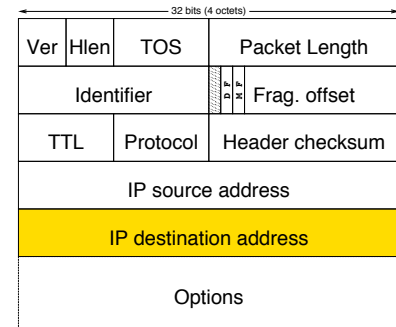
IPv4 : Contrôle d'erreur sur l'entête



- 16 bits
- contrôle d'erreur
 - ✓ idem UDP/TCP mais que sur l'entête
- émetteur :
 - ✓ entête IP = suite mot_{16bits}
 - ✓ $checksum^3 = \sum mot_{16bits}$
- récepteur :
 - ✓ recalcul de $\sum mot_{16bits}$
 - ✓ $checksum = 0$: pas d'erreur détectée toujours possible...
 - ✓ $checksum \neq 0$: erreur (destruction silencieuse)

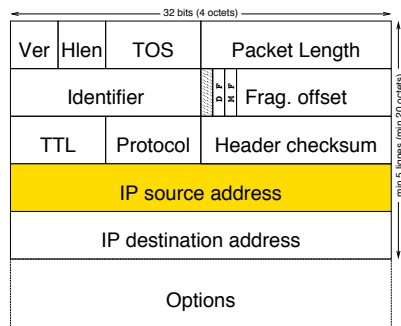
³Somme binaire sur 16 bits avec report de la retenue débordante ajoutée au bit de poids faible

IPv4 : Adresse destination



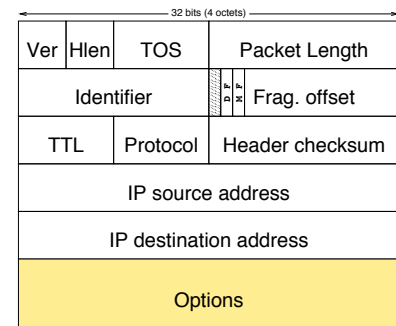
- adresse IP 32 bits
- utilisée pour le routage
 - ✓ indique le réseau (ou l'agrégation de réseau) du destinataire
 - ✓ identifie l'**interface** du destinataire dans son réseau

IPv4 : Adresse source



- adresse IP 32 bits
- identifie l'émetteur du paquet
- permet de retourner un message à l'émetteur (ICMP, UDP...)

IPv4 : Options



- système TLV identique à TCP
- analysées dans **chaque routeur**
- exemple :
 - ✓ enregistrement de la route
 - ✓ routage à la source strict
 - ✓ routage à la source relâché
 - ✓ estampilles temporelles
 - ✓ sécurité
 - ✓ ...
- 0 à 40 octets (alignés sur 32 bits)

⇒ A éviter !

Plan

Rappels sur la couche réseau

La couche réseau dans TCP/IP

Structure du paquet IPv4

Adressage classique IPv4

Adressage CIDR

Translation d'adresses

Messages de contrôle

Autoconfiguration

Tunnel et pare-feu



Adressage : Masques

Application de masques binaires

classe	masque binaire	netmask	prefixe
A	11111111000000000000000000000000	255.0.0.0	/8
B	11111111111111110000000000000000	255.255.0.0	/16
C	11111111111111111111111100000000	255.255.255.0	/24

Extraction du netId

132.227. 60.135	netId.hostId
&& 255.255. 0. 0	&& netmask
132.227. 0. 0	netId. 0. 0

Extraction du hostId

132.227. 60.135	netId.hostId
&& 0. 0.255.255	&& !netmask
60.135	hostId



Adressage : Classes

		32 Bits		
		<div></div>	Range of host addresses	
Class				
A	0	Network	Host	1.0.0.0 to 127.255.255.255
B	10	Network	Host	128.0.0.0 to 191.255.255.255
C	110	Network	Host	192.0.0.0 to 223.255.255.255
D	1110	Multicast address		224.0.0.0 to 239.255.255.255
E	11110	Reserved for future use		240.0.0.0 to 247.255.255.255

pictures from TANENBAUM A. S. Computer Networks 3rd edition



Adressage : Adresses particulières

Adresses particulières :

0 0	This host
0 0 . . . 0 0	Host
1 1	Broadcast on the local network
Network 1 1 1 1 . . . 1 1 1 1	Broadcast on a distant network
127 (Anything)	Loopback

pictures from TANENBAUM A. S. Computer Networks 3rd edition



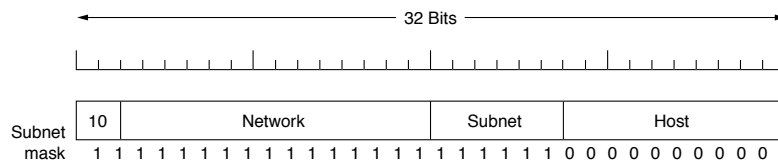
Adressage : Subneting (1)

Taille de l'identifiant de réseau (netId) par défaut :

- classe A : /8 (255.0.0.0)
- classe B : /16 (255.255.0.0)
- classe C : /24 (255.255.255.0)

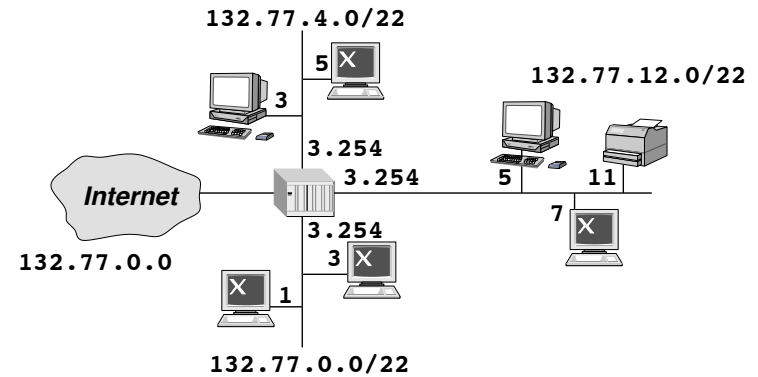
Subdivision possible :

- 132.77.12.0/22 (notation par **préfixe**)
- 132.77.12.0 netmask 255.255.252.0 (notation par **masque**)

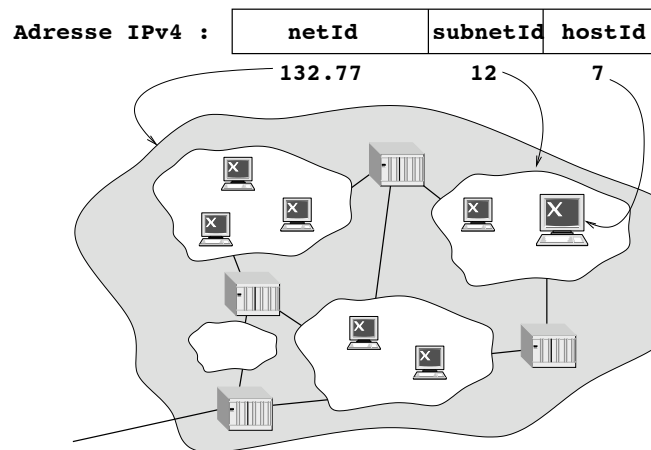


pictures from TANENBAUM A. S. *Computer Networks 3rd edition*

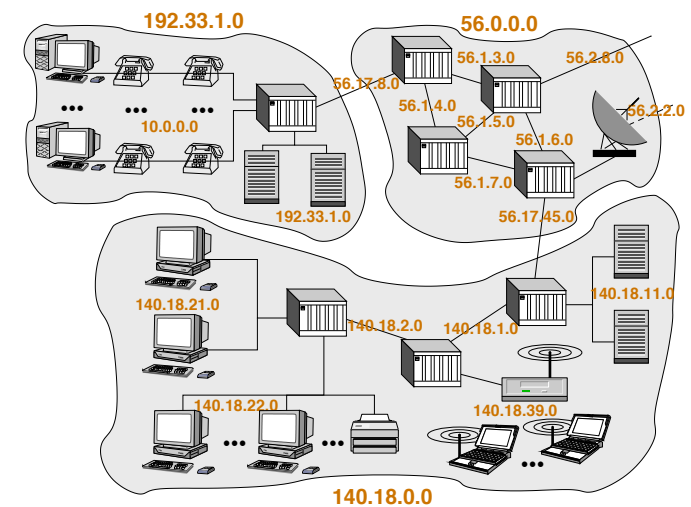
Adressage : Subneting (3)



Adressage : Subneting (2)

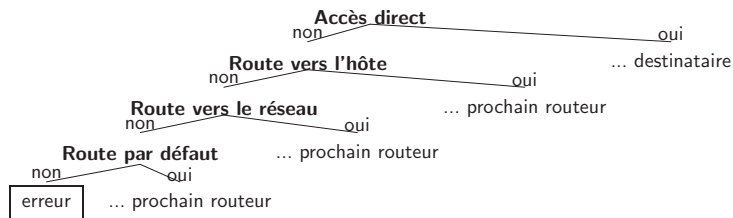


Adressage : affectation



IPv4 : Logique de routage

Selon l'adresse destination, envoi au ...



Unix> route -n

Kernel IP routing table

Destination	Gateway	Genmask	Flags	Metric	Ref	Use	Iface
192.33.182.0	0.0.0.0	255.255.255.0	U	0	0	0	eth0
10.0.0.0	0.0.0.0	255.0.0.0	U	0	0	0	atm0
154.18.2.0	0.0.0.0	255.255.255.0	U	0	0	0	eth1
132.77.0.0	154.18.2.254	255.255.0.0	UG	0	0	0	eth1
default	192.33.182.254	0.0.0.0	UG	0	0	0	eth0

Plan

Rappels sur la couche réseau

La couche réseau dans TCP/IP

Structure du paquet IPv4

Adressage classique IPv4

Adressage CIDR

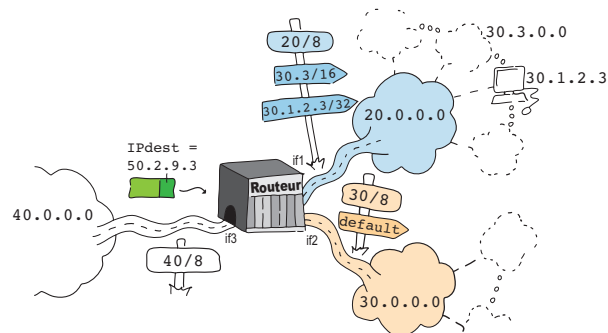
Translation d'adresses

Messages de contrôle

Autoconfiguration

Tunnel et pare-feu

Routeur : Longest Prefix Match



Destination	Gateway	Genmask	Flags	Metric	Ref	Use	Iface
20.0.0.0	0.0.0.0	255.0.0.0	U	0	0	0	if1
30.0.0.0	0.0.0.0	255.0.0.0	U	0	0	0	if2
40.0.0.0	0.0.0.0	255.0.0.0	U	0	0	0	if3
30.3.0.0	20.1.2.3	255.255.0.0	UG	0	0	0	if1
30.1.2.3	20.1.0.1	255.255.255.255	UGH	0	0	0	if1
60.126.6.0	30.0.0.1	255.255.255.0	UG	0	0	0	if2
default	30.0.0.1	0.0.0.0	UG	0	0	0	if2

Adressage sans classe

L'attribution des adresses IP avec classe est **inefficace**

- adresses allouées par blocs de 256, 65K ou 16M
 - ✓ les sous-réseaux permettent une meilleure gestion
- un adressage **sans classe** augmente la souplesse dans l'attribution des adresses :
 - ✓ les adresses :
 - ☞ 192.77.16.0/24
 - ☞ 192.77.17.0/24
 - ☞ 192.77.18.0/24
 - ☞ 192.77.19.0/24
 - ✓ peuvent être regroupées en :
 - ☞ notation par **préfixe** : 192.77.16.0/22
 - ☞ notation par **masque** : 192.77.16.0 netmask 255.255.252.0

Adressage : Supernetting

CIDR (*Classless InterDomain Routing*)

- utilisé pour agréger des blocs d'adresses contigües
 - permet aux routeurs de maintenir une seule entrée de table de routage
 - utilisé initialement par les ISP pour grouper des adresses de classe C
 - initialement décrit en réduisant la taille du préfixe réseau
 - ✓ le préfixe réseau par défaut pour la classe C est /24
 - ✓ les valeurs de préfixes réseau /23, /22, /21, etc. décrivent des agrégations d'adresses de classe C
- ☞ exemples :

197.88.0.0/16 agrège 256 adresses de classe C

81.152.12.0/22 agrège ??

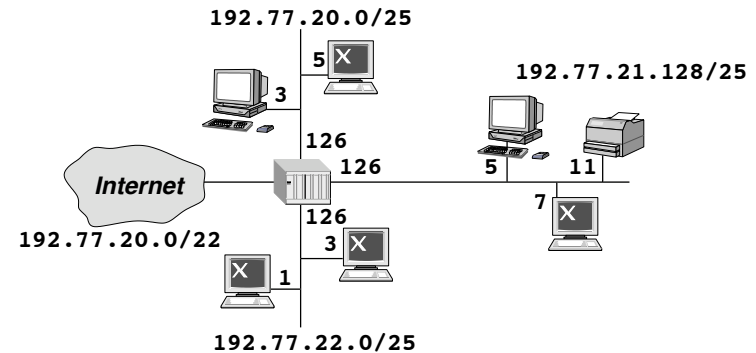
Adressage : Calcul CIDR

Un bloc CIDR est donc l'agrégation d'un ensemble d'adresses

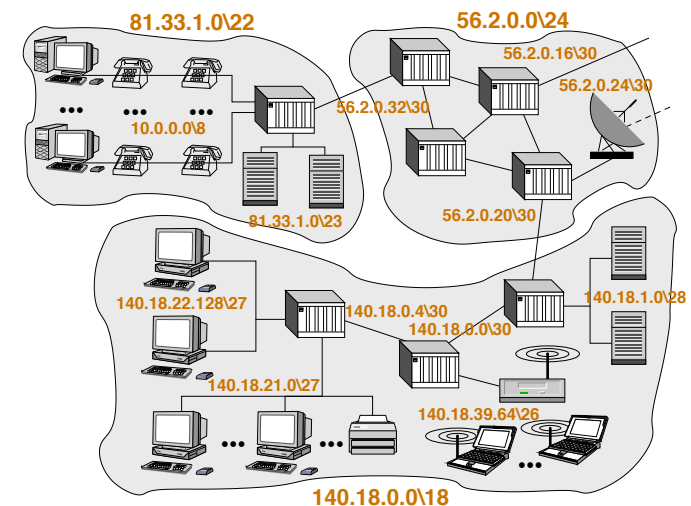
- **bits réseau** (netId) d'un bloc CIDR correspondent aux N bits les plus à gauche (/N définit le masque réseau du bloc CIDR)
 - **bits hôte** (hostId) du bloc CIDR correspondent aux $32 - N$ bits restants
 - ensemble des adresses attribuables dans un bloc CIDR :
 - ✓ premier hôte : hostId = 000...0001
 - ✓ dernier hôte : hostId = 111...1110
 - ✓ adresse de diffusion : hostId = 111...1111
- ☞ exemple :

Bloc CIDR -> 192.77.20.0/22
 @ premier hôte : 192.77.20.1
 ...
 @ dernier hôte : 192.77.23.254
 @ de diffusion : 192.77.23.255

Adressage : Subnetting des agrégats



Adressage : Affectation



Adressage : Synthèse

Observation sur le découpage des plages d'adresses en sous-réseaux (*subnetting*) ou en agrégats (*supernetting*) :

- **Attention** aux analyses simplistes...
 - ✓ N = nombre de bits réseau
 - ✓ H = nombre de bits hôte
 - ✓ D = préfixe réseau par défaut (8 pour la classe A, 16 pour la classe B, 24 pour la classe C)
 - ☞ si $N = D$, pas de *subnetting* ni de *supernetting*
 - ☞ si $N > D$, *subnetting* (sous-réseau)
 - ☞ si $N < D$, *supernetting* (CIDR)
- ... **FAUX**, car on peut combiner *subnetting* et *supernetting* !

Plan

Rappels sur la couche réseau

La couche réseau dans TCP/IP

Structure du paquet IPv4

Adressage classique IPv4

Adressage CIDR

Translation d'adresses

Messages de contrôle

Autoconfiguration

Tunnel et pare-feu

IPv4 : Adresses privées

Deux types d'adressage :

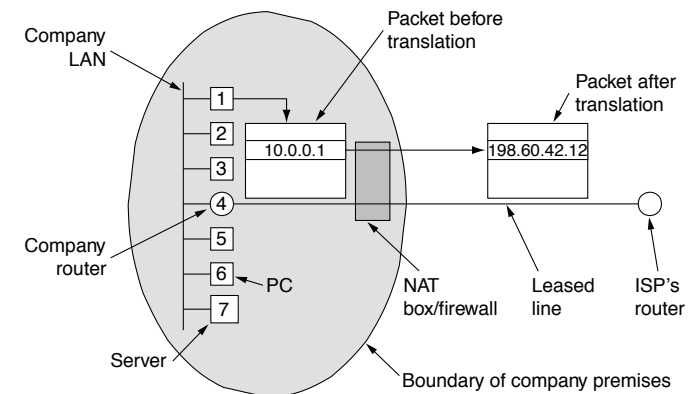
Public : tout hôte connecté à l'Internet doit avoir une adresse unique valide

Privé : pour un usage de TCP/IP non connecté à l'Internet

- gestion autonome d'un plan d'adressage (avec adresses uniques)
- utilisation de plages d'adresses spécifiques **recommandée** :
 - ✓ **adresses non routées** (adresses privées) :
 - 10.0.0.0/8 (1 ex-classe A)
 - 172.16.0.0/12 (16 ex-classe B)
 - 192.168.0.0/16 (256 ex-classes C)
 - 169.254.0.0/16 (*link local block* pour l'auto-configuration)
 - ☞ utilisable dans chaque *internet* privé
 - ☞ même en cas de connexion à l'Internet, ce trafic n'est pas relayé
 - ☞ possibilité de "sortir" du réseau privé à l'aide de :
 - ☛ serveurs proxys
 - ☛ conversion d'adresses **NAT**

IPv4 : Translation des adresses

Network Address Translation (NAT)



pictures from TANENBAUM A. S. *Computer Networks 4rd edition*

IPv4 : NAT, DNAT et NAPT

Plusieurs approches de la conversion d'adresses :

NAT statique : correspondance fixe d'adresses

NAT dynamique : correspondance dynamique d'adresses

☞ table d'adresses dynamique :

Adresse entrante	adresse sortante
10.0.0.3	192.33.182.117
10.0.0.4	192.33.182.118
...	...

NAPT (*NAT overload*) : correspondance dynamique vers une adresse (ou plusieurs adresses) avec surcharge

☞ utilisation des ports

☞ table dynamique (pour chaque protocole) :

Proto	Adresse entrante	Port entrant	adresse sortante	Port sortant
TCP	10.0.0.3	1027	192.33.182.117	1027
TCP	10.0.0.4	1027	192.33.182.117	1028
UDP	10.0.0.4	31765	192.33.182.117	31765
...

IPv4 : NAT et IETF

Un standard publié : RFC 1631

- NAPT **fortement utilisé** actuellement
 - ✓ entreprises (flexibilité)
 - ✓ fournisseurs de services (manque d'adresses)
 - ✓ particuliers (n'ont qu'une adresse)
- pose qqs **problèmes**
 - ✓ architecturaux :
 - ☞ les ports doivent identifier des processus et non des machines
 - ☞ les routeurs modifient les paramètres de la couche transport
 - ☞ **principe de bout-en-bout** : deux hôtes doivent communiquer directement
 - ✓ sécuritaires : incompatible avec les mécanismes d'**authentification**
 - ✓ techniques : comment "entrer" dans le réseau translaté
- **solutions**
 - ✓ court terme ➡ conversions statiques, serveurs intermédiaires (UDP)
 - ✓ long terme ➡ IPv6

IPv4 : Mécanismes NAPT

Où sont modifiées les adresses ?

➡ au niveau de la carte d'interface :

NAT en entrée ➡ processus de routage ➡ NAT en sortie

Modifications annexes :

- le *checksum* des entêtes doit être recalculé
 - ✓ NAT IP, TCP et UDP (adresse + *pseudo-header*)
 - ✓ NAPT IP, TCP et UDP (adresse + *pseudo-header* + port)
- les adresses et ports paramètres de protocoles applicatifs doivent être aussi modifiées (commande PORT de FTP)
- les messages ICMP sont analysés

Plan

Rappels sur la couche réseau

La couche réseau dans TCP/IP

Structure du paquet IPv4

Adressage classique IPv4

Adressage CIDR

Translation d'adresses

Messages de contrôle

Autoconfiguration

Tunnel et pare-feu

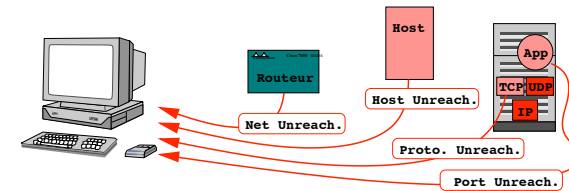
IPv4 : ICMP

Internet Control Message Protocol (RFC792)

- encapsulé dans un paquet IP (mais appartient à la couche 3)
- test et diagnostic du réseau :

ICMP Type	Code	Description
0	0	↔ echo reply
3	0	destination network unreachable
3	1	destination host unreachable
3	2	destination protocol unreachable
3	3	destination port unreachable
3	6	destination network unknown
3	7	destination host unknown
4	0	source quench
8	0	→ echo request
9	0	router advertisement
10	0	router discovery
11	0	TTL expired
11	1	reassembly time exceeded
12	0	IP header bad

ICMP : Destination inaccessible

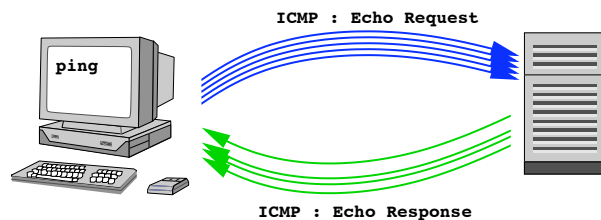


Type	Code	Checksum	Unused	Data
3	0 (Net Unreachable) 1 (Host Unreachable) 2 (Protocol Unreachable) 3 (Port Unreachable)			IP Header + 64 bits
1 octet	1	4	2	(IHL * 4) + 8

Messages émis lorsque la destination n'est pas accessible.

- l'entête IP et une partie de la couche transport sont retournés
 - ✓ @ source = créateur du message ICMP
 - ✓ @ destination = @ source de l'émetteur du paquet en cause

ICMP : ECHO

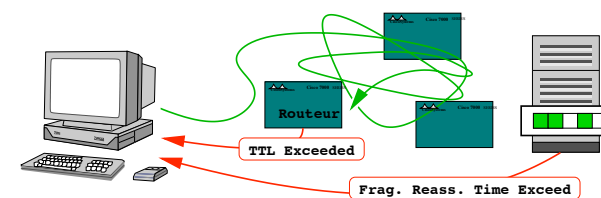


Type	Code	Checksum	Identifier	Seq. Num.	Data
8 (Echo Request)	0				
0 (Echo Response)	0				
1 octet	1	2	2	2	...

Teste l'accessibilité d'un équipement

- utilisé par la commande ping :
 - ✓ indique la connectivité et la disponibilité d'IP chez le destinataire
 - ✓ plusieurs messages permettent d'estimer le RTT et le taux de perte

ICMP : Timeout



Type	Code	Checksum	Unused	Data
11	0 (Time To Live Exceeded) 1 (Frag. Reass. Time Exceeded)			IP Header + 64 bits
1 octet	1	4	2	(IHL * 4) + 8

Messages émis lorsque le temps de vie ou de réassemblage est dépassé.

- l'entête IP et une partie de la couche transport sont retournés
 - ✓ @ source = créateur du message ICMP
 - ✓ @ destination = @ source de l'émetteur du paquet en cause
- utilisé par la commande traceroute

ICMP : Autres messages

- **Source Quench** (Type 4)
 - ✓ indique une congestion à la source
 - ☞ pas de signalisation de fin de congestion
- **Redirection** (Type 5)
 - ✓ indique si une meilleure route est disponible
 - ☞ configuration minimale des hôtes
- autres messages principalement pour l'**autoconfiguration**

Plan

Rappels sur la couche réseau

La couche réseau dans TCP/IP

Structure du paquet IPv4

Adressage classique IPv4

Adressage CIDR

Translation d'adresses

Messages de contrôle

Autoconfiguration

Tunnel et pare-feu

IPv4 : RARP

Reverse Adresse Résolution Protocol (RFC 903)

inverse du protocole ARP (réseaux à diffusion)

- obtention d'une @ IP à partir de @ MAC au démarrage
 - ✓ hôtes sans disques (terminaux X, imprimantes...)
 - ✓ hôtes mobiles (portable changé de réseau...)
- utilisation d'un **serveur** (rarpd)
 - ✓ mise en correspondance de /etc/ethers et de /etc/hosts
- format des trames identique à ARP
 - ✓ type Ethernet : 0x8035
 - ☞ code 3 pour une requête RARP
 - ☞ code 4 pour une réponse RARP
- exemple d'autoconfiguration :
 - ✓ la nouvelle station déclenche un échange **RARP**
 - ✓ la station demande le *netmask* par un échange **ICMP**
 - ✓ la station demande au serveur RARP son programme de démarrage par TFTP

IPv4 : BOOTP

BOOT Protocol (RFC 951 et RTF 1542)

- Protocole **portable**, sur UDP
 - ✓ requête sur le port **68**, réponse sur le port **67**
 - ✓ *quelles adresses IP utiliser lorsqu'on n'en connaît aucunes ?*
 - ☞ @ IP de diffusion (255.255.255.255)
 - ☞ @ IP par défaut (0.0.0.0)
 - ✓ permet d'atteindre un serveur sur un autre réseau
 - ☞ à travers des agents BOOTP relais
 - ✓ nombreuses extensions (RFC 1533)
 - ☞ *netmask*
 - ☞ liste des **routeurs** du sous-réseau
 - ☞ liste de **serveurs NTP**
 - ☞ liste des **serveurs de noms** (DNS)
 - ☞ liste des serveurs d'impression (LPD et autres)
 - ☞ *hostname* et *domainname*
 - ☞ TTL par défaut ...

IPv4 : DHCP(1)

Dynamic Host Configuration Protocol (RFC 2131)

Extension compatible de BOOTP avec **gestion dynamique des @ IP**

- attribution dynamique par **bail** (*lease*) limité dans le temps
 - ✓ bail renouvelé périodiquement si nécessaire
- nouvelles **options DHCP** (extensions BOOTP)

DHCPDISCOVER	C → S	localisation du serveur
DHCPOFFER	S → C	proposition au client
DHCPREQUEST	C → S	confirmation d'une proposition
DHCPACK	S → C	validation d'une configuration
DHCPNACK	S → C	invalidation d'une configuration
DHCPDECLINE	C → S	refus d'une configuration invalide
DHCPRELEASE	C → S	libération d'une configuration
DHCPINFORM	C → S	demande d'information autre que @ IP
DHCPFORCERENEW	S → C	demande de reconfiguration

Plan

Rappels sur la couche réseau

La couche réseau dans TCP/IP

Structure du paquet IPv4

Adressage classique IPv4

Adressage CIDR

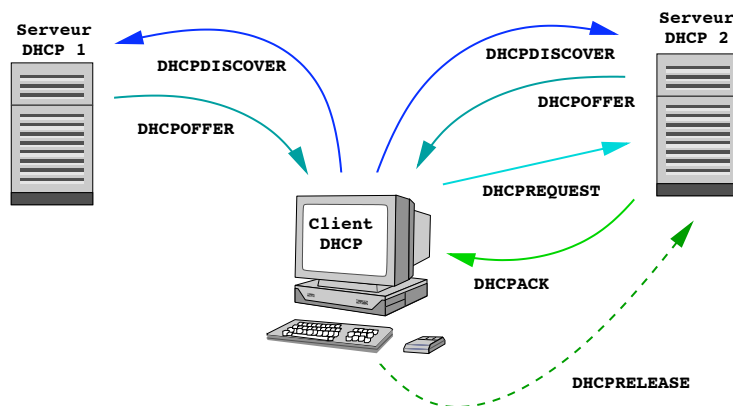
Translation d'adresses

Messages de contrôle

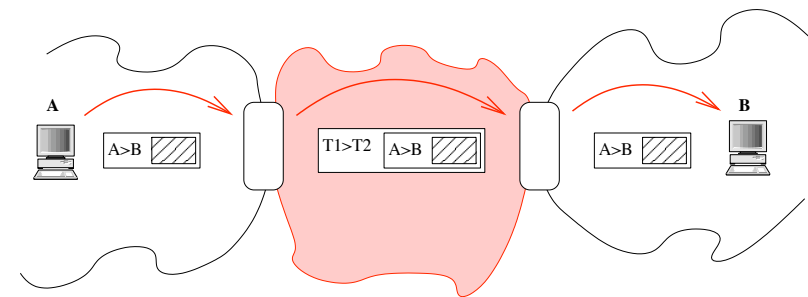
Autoconfiguration

Tunnel et pare-feu

IPv4 : DHCP(2)



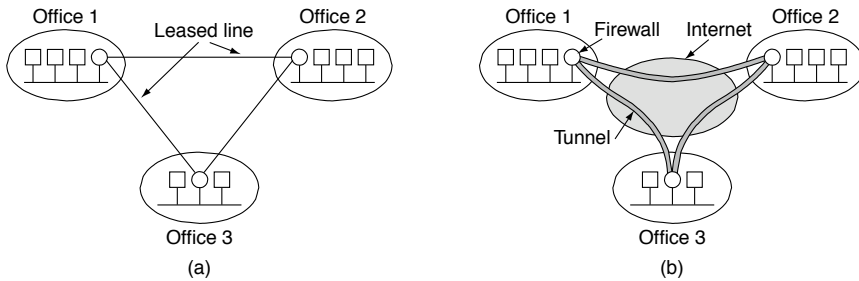
Tunneling



- **encapsulation** alternative à la traduction (*translation*)
- traversées de zones avec des protocoles différents
 - ✓ ex : relier des îlots avec des protocoles non généralisés (IPmulticast, IPv6...)
- contrôle du flux de T1 à T2 (IPv4 dans IPv4, VPN...)
 - ✓ VPN...

Couche IPv4 : VPN

- intégration avec des mécanismes de sécurité, automatisation
 - ✓ IPSEC : confidentialité et intégrité (RFC 4301 à 4309)
 - ✓ AAA (*Authentication, Autorisation, Accounting*)
- autres approches VPN au niveau de la couche 2 (PPP)...



pictures from TANENBAUM A. S. *Computer Networks 4rd edition*

Fin

Document réalisé avec L^AT_EX.

Classe de document foils.

Dessins réalisés avec xfig.

Olivier Fourmaux, olivier.fourmaux@upmc.fr

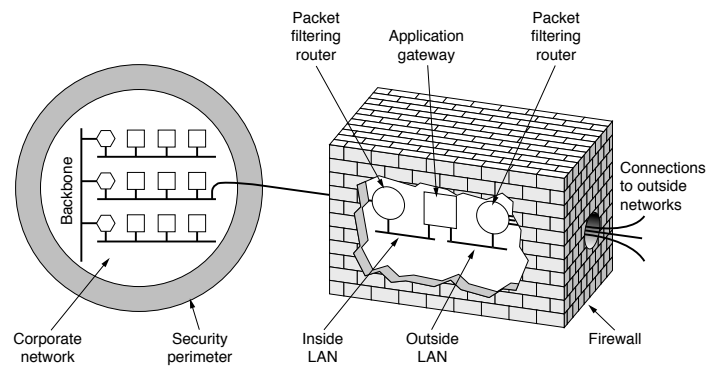
<http://www-rp.lip6.fr/~fourmaux>

Ce document est disponible en format PDF sur le site :

<http://www-master.ufr-info-p6.jussieu.fr/>

IPv4 : Filtrage d'adresses

Firewall...



pictures from TANENBAUM A. S. *Computer Networks 3rd edition*

U.E. ARES

Architecture des Réseaux

Cours 5/6 : Routage interne - externe

Olivier Fourmaux
(olivier.fourmaux@upmc.fr)

Version 5.4



Couche R  seau

La **Couche R  seau** achemine les paquets de la source vers les destinataires en effectuant des sauts entre les diff  rents **n  uds interm  diaires**

- acheminement de bout-en-bout (*end-to-end*)
 - ✓ adressage virtuel
- connaissance locale de la topologie
 - ✓ besoin d'informations pour orienter les PDU
 -    statique : configuration manuelle
 -    dynamique : algorithmes et protocoles de routage
- adaptation    la taille du r  seau
 - ✓ structure hi  rarchique (AS)
 -    routage interne : RIP, EIGRP, OSPF, IS-IS
 -    routage externe : BGP-4

Plan

Rappel sur le routage

Algorithmes de base

Hi  rarchie de routage

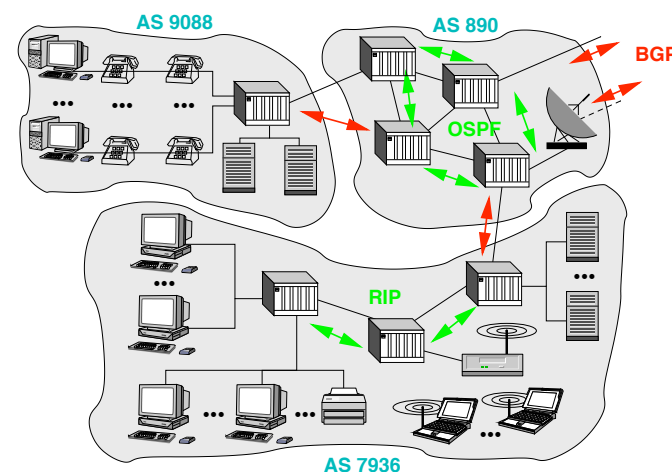
Routage interne : OSPF

Routage externe : BGP-4



U.E. ARES Cours 5/6 v5.4 olivier.fourmaux@upmc.fr 2

Routage



U.E. ARES Cours 5/6 v5.4 olivier.fourmaux@upmc.fr 1



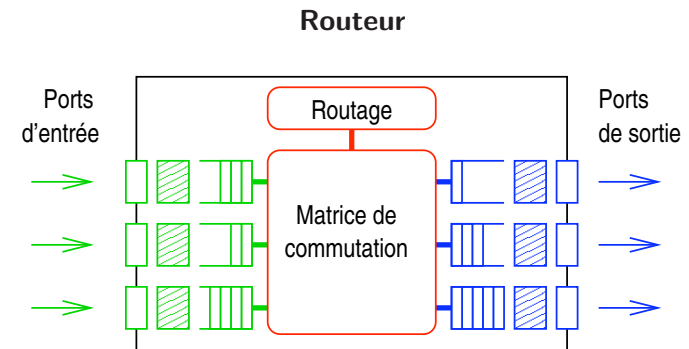
U.E. ARES Cours 5/6 v5.4 olivier.fourmaux@upmc.fr 3

Routeur

Routeur

```
Unix> /sbin/ifconfig eth0
eth0      Link encap:Ethernet  HWaddr 00:20:ED:87:FD:E6
          inet addr:132.227.61.122  Bcast:132.227.61.255  Mask:255.255.255.0
          UP BROADCAST NOTRAILERS RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:1115393 errors:0 dropped:0 overruns:0 frame:0
          TX packets:966470 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:100
          RX bytes:445681702 (425.0 Mb)  TX bytes:370060277 (352.9 Mb)
          Interrupt:9 Base address:0x6f00
```

```
Unix> /sbin/route
Kernel IP routing table
Destination Gateway Genmask Flags Metric Ref Use Iface
132.227.61.0 * 255.255.255.0 U 0 0 0 eth0
127.0.0.0 * 255.0.0.0 U 0 0 0 lo
default 132.227.61.200 0.0.0.0 UG 0 0 0 eth0
Unix>
```



Routeur et "relayage" (*forwarding*)

- interfaces (terminaisons physiques, encapsulation...)
- files d'attente
- système de **relayage** (mémoire partagée, bus ou *crossbar*)
- système de **roulage**
 - ✓ table, algorithmes et protocoles de roulage

Routeur

Routeur

```
C:\Program Files\Support Tools>ipconfig
Ethernet carte Connexion au réseau local :
  Suffixe DNS spéc. à la connexion. :
  Adresse IP. . . . . : 132.227.61.136
  Masque de sous-réseau . . . . . : 255.255.255.0
  Passerelle par défaut . . . . . : 132.227.61.200
```

```
C:\Program Files\Support Tools>route print
```

```
=====
Liste d'Interfaces
```

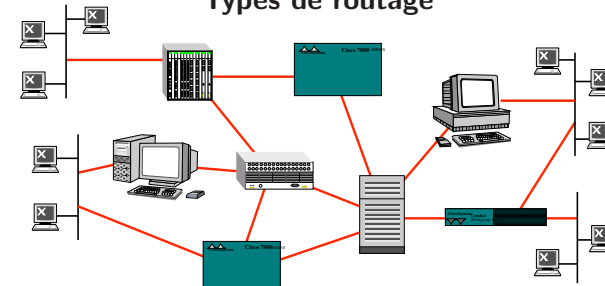
```
0x1 ..... MS TCP Loopback interface
0x1000003 ...00 03 47 7c b9 d5 ..... Intel(R) PRO Adapter
=====
```

```
Itinéraires actifs :
```

Destination réseau	Masque réseau	Adr. passerelle	Adr. interface	Métrique
0.0.0.0	0.0.0.0	132.227.61.200	132.227.61.136	1
127.0.0.0	255.0.0.0	127.0.0.1	127.0.0.1	1
132.227.61.0	255.255.255.0	132.227.61.136	132.227.61.136	1
132.227.61.136	255.255.255.255	127.0.0.1	127.0.0.1	1
132.227.61.255	255.255.255.255	132.227.61.136	132.227.61.136	1
224.0.0.0	224.0.0.0	132.227.61.136	132.227.61.136	1
255.255.255.255	255.255.255.255	132.227.61.136	132.227.61.136	1

```
Passerelle par défaut : 132.227.61.200
=====
```

Types de routage



Configuration du routeur

- statique
- dynamique (en particulier lorsqu'il y a des liens redondants)
 - ✓ protocoles et algorithmes de routage
 - ☞ ordinateurs : Unix avec logiciels *routed*, *gated*, GNU *Zebra*, *Quagga*...
 - ☞ matériels dédiés : Cisco, Juniper, Alcatel, Hp...

Plan

Rappel sur le routage

Algorithmes de base

- vecteurs de distance
- état des liaisons

Hiérarchie de routage

Routage interne : OSPF

Routage externe : BGP-4

Algorithmes de routage

Optimisation d'un critère

- plus court chemin
 - ✓ vecteurs de distance
 - ✓ état des liaisons
 - routage politique
 - ✓ vecteurs de chemin
 - routage multipoint
 - ✓ plus court chemin
 - ✓ coût minimum (arbre de steiner)
 - ✓ arbres centrés
- ☞ voir le module ING

Plan

Rappel sur le routage

Algorithmes de base

- **vecteurs de distance**
- état des liaisons

Hiérarchie de routage

Routage interne : OSPF

Routage externe : BGP-4

Routage par vecteurs de distance

Distance Vector Routing

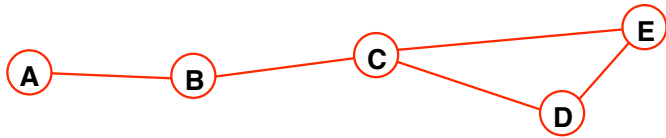
Algorithme simple basé sur :

- l'échange d'informations entre routeurs adjacents (liaison directe)
 - ✓ vecteur de distance (\neq table de routage)
- propagation de proche en proche de l'accessibilité du réseau

... mais limité à des réseaux de taille réduite

- utilisé sur des sites avec quelques routeurs pour éviter les configurations manuelles
- problème avec les informations de seconde main

Principe du routage à vecteur de distance



Les routeurs ne connaissent initialement que leurs propres liaisons. Ils diffusent leurs vecteurs de distance (table de routage sans les interface) à leur voisins

➡ Algorithme de Bellman-Ford distribué (ou Ford-Fulkerson 1962)

A la réception d'un vecteur, un routeur intègre l'information dans sa table :

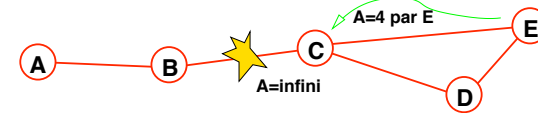
- rajout des entrées nouvelles en indiquant l'interface d'arrivée
- modifier le coût des entrées
 - ✓ si un plus court chemin est proposé
 - ✓ si un plus long chemin est proposé par la même interface que celle de la table

Les échanges successifs doivent amener à la **convergence**

Problème des algorithmes de routage à vecteur de distance

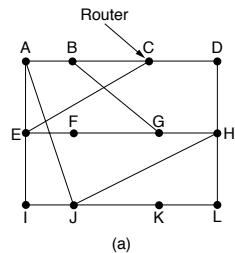
Plusieurs problèmes sont apparus avec ces algorithmes :

- convergence lente
- risques de boucle
 - ✓ horizon coupé (split horizon)



- envoi de vecteurs avec tous les réseaux de la table de routage
 - ✓ taille de réseau limitée

Exemple de table issue des vecteurs de distance



To	A	I	H	K	New estimated delay from J
A	0	24	20	21	8 A
B	12	36	31	28	20 A
C	25	18	19	36	28 I
D	40	27	8	24	20 H
E	14	7	30	22	17 I
F	23	20	19	40	30 I
G	18	31	6	31	18 H
H	17	20	0	19	12 H
I	21	0	14	22	10 I
J	9	11	7	10	0 -
K	24	22	22	0	6 K
L	29	33	9	9	15 K

JA delay is 8
JI delay is 10
JH delay is 12
JK delay is 6

Vectors received from J's four neighbors

New routing table for J

Plan

Rappel sur le routage

Algorithmes de base

- vecteurs de distance
- **état des liaisons**

Hierarchie de routage

Routage interne : OSPF

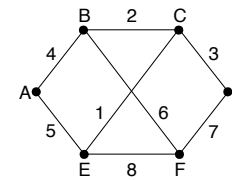
Routage externe : BGP-4

Routage par état des liaisons

Link State Routing

Comment s'adapter à des réseaux importants tout en évitant la propagation des informations de proche en proche ?

- **connaître son voisinage**
- construire une synthèse de l'info locale
- **diffuser l'info locale** à tous les routeurs
- construire un **graphe** représentant le réseau
- calculer le **plus court chemin** (SPF) vers tous les routeurs



(a)

Link		State		Packets	
A	B	C	D	E	F
Seq.	Seq.	Seq.	Seq.	Seq.	Seq.
Age	Age	Age	Age	Age	Age
B 4	A 4	B 2	C 3	A 5	B 6
E 5	C 2	D 3	F 7	C 1	D 7
	F 6	E 1		F 8	E 8

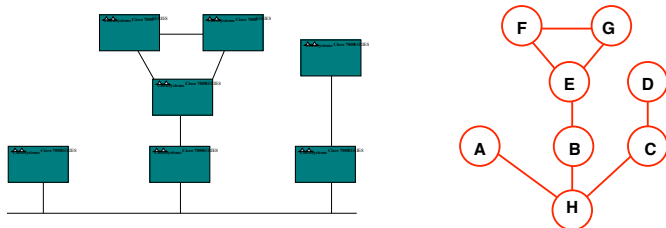
(b)

pictures from TANENBAUM A. S. *Computer Networks 3rd edition*

Etat des liaisons : Acquisition du voisinage

But : création d'un graphe équivalent

- envoi de paquets de détection sur les liaisons
- les supports partagés (LAN), sont remplacés par un seul nœud virtuel



Pour pondérer les liaisons, on peut faire des mesures sur ces liaisons

Etat des liaisons : Distribution des paquets

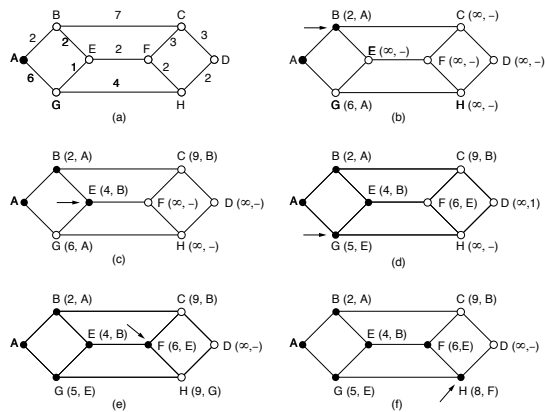
Les routeurs doivent recevoir les messages de **tous les routeurs** :

- besoin d'une distribution fiable
 - ✓ numéro de séquence
 - ✓ age de la connexion
- diffusion de routeur en routeur sans modification du contenu des messages

Problème de **consistance** pendant la diffusion de changements

Etat des liaisons : Calcul des routes

Algorithme du plus court chemin de Dijkstra



pictures from TANENBAUM A. S. *Computer Networks 3rd edition*

Plan

Rappel sur le routage

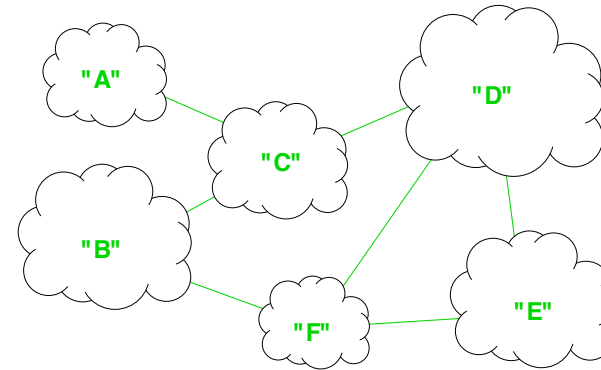
Algorithmes de base

Hiérarchie de routage

Routage interne : OSPF

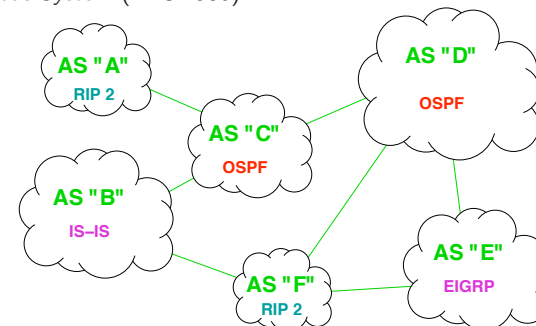
Routage externe : BGP-4

Organisation de gros réseaux : Internet



AS : Organisation interne

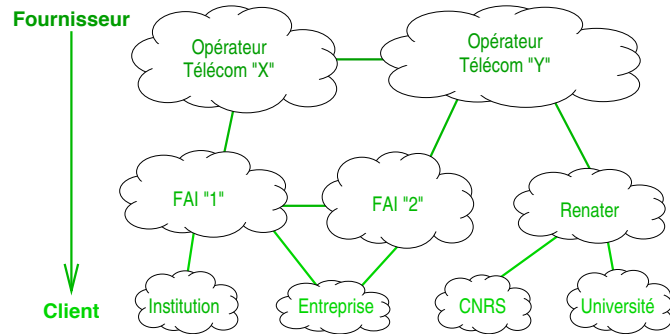
Autonomous System (RFC 1930)



Un AS est un ensemble d'un ou plusieurs préfixes IP interconnectés et gérés par un ou plusieurs opérateurs de réseaux qui fonctionnent avec une **unique** politique de routage **clairement définie**.

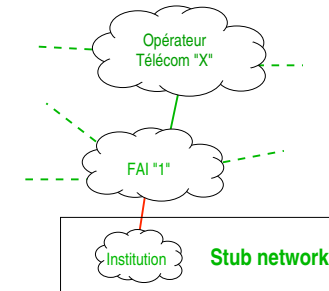
AS : Organisation externe (1)

Les relations entre AS sont basées sur la notion de **client/fournisseur**



AS : Routage simple

Pour un réseau d'extrémité (*stub network*) :

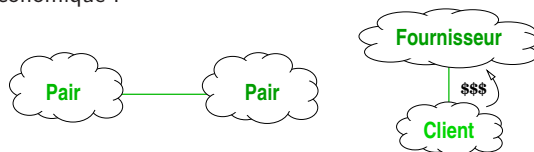


➡ Annonce directe :

- ses préfixes sont annoncés pour qu'il reçoive son trafic entrant
- le réseau d'extrémité envoie tout son trafic sortant vers le reste de l'Internet

AS : Organisation externe (2)

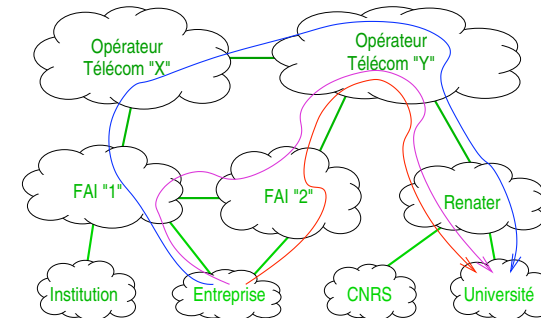
Relation économique :



- les fournisseurs font payer leurs clients
 - ✓ les pairs échangent gratuitement du trafic
 - ☞ les contrats sont secrets !
- Tier-1 : les plus gros fournisseurs
 - ✓ **ATT** (Worldnet), **MCI** (Worldcomm/UUnet), **Sprint**, **Level3** (Genuity/BBN), **Quest**, Global Crossing, CableWireless, BT, NTT (Verio), Telstra, Equant (SITA), Infonet...
 - ☞ infrastructure mondiale
 - ☞ possèdent leur réseau physique
 - ☞ (en général) ne payent personne !

AS : Routage entre multiples AS

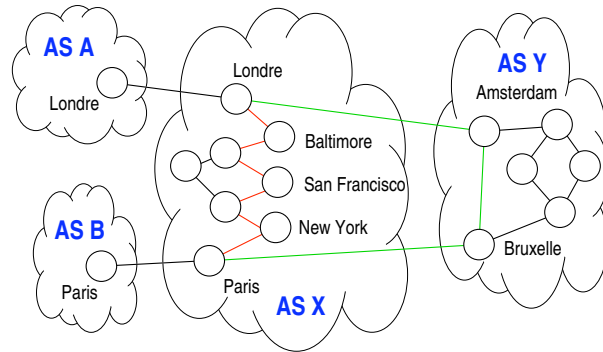
Pour les réseaux d'infrastructure (*transit network*) :



➡ Comment trouver son chemin à travers plusieurs possibilités ?

AS : Critère optimal du routage

Routage politique (critère commercial) :



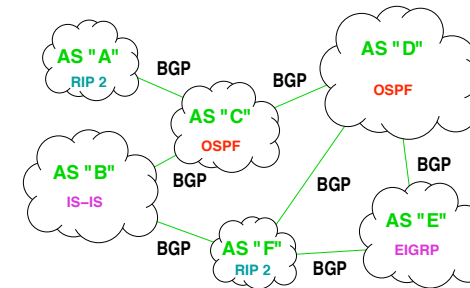
➡ Ce n'est pas forcément le plus court chemin !

AS : Routage politique

Intégration des contraintes politiques :

- nouvelles règles ;
 - ✓ un AS accepte le trafic de ou vers ses clients
 - ✓ un AS n'accepte pas le trafic de transit entre deux clients de ses concurrents
 - ↳ besoin d'un nouveau type de routage !
- but simple :
 - ✓ un FAI route le trafic en provenance d'un des ses clients
 - ✓ le trafic est routé à un FAI pair ou à un FAI de niveau supérieur (*tier-1*)
 - ✓ le FAI du destinataire route le trafic vers son client destinataire
- mais plus complexe :
 - ✓ les AS peuvent être rattachés à plusieurs FAI (*multihoming*)
 - ✓ souvent plusieurs chemins possibles

AS : Routage hiérarchique



Deux catégories de protocole :

- **IGP** (*Interior Gateway Protocols*)
 - ✓ Routage à l'intérieur d'un AS (basé sur le plus court chemin)
 - ↳ RIP-2, EIGRP, IS-IS, **OSPF**
- **EGP** (*Exterior Gateway Protocols*)
 - ✓ Routage entre AS (basé sur les aspects politiques)
 - ↳ il n'y en a qu'un : **BGP-4**

Plan

Rappel sur le routage

Algorithmes de base

Hiérarchie de routage

Routage interne : OSPF

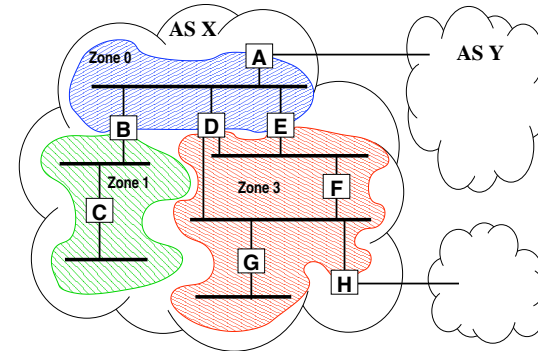
Routage externe : BGP-4

OSPF : Introduction

Open Shortest Path First

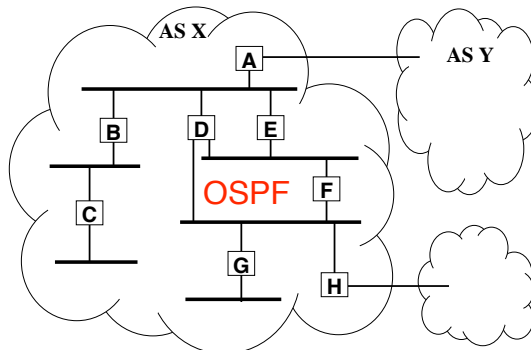
- conçu par l'IETF dès 1988 pour :
 - ✓ dépasser l'approche de RIP
 - ☞ converger rapidement
 - ☞ s'adapter aux réseaux de grande taille
 - ✓ s'adapter au cas général :
 - ☞ LAN (*broadcast*)
 - ☞ NBMA
 - ☞ point-à-point
- ✓ acquérir la topologie du réseau
- ✓ calculer le plus court chemin sur le graphe associé au réseau
- ✓ être non propriétaire

OSPF : Zones (2)



- 3 types de zone :
 - ✓ **terminale** (*stub area*) sans trafic de transit (Zone 1)
 - ✓ **pas si terminale** (*NSSA, Not So Stubby Area*)
 - ✓ **transit** (*transit area*) (Zones 0 et 3)

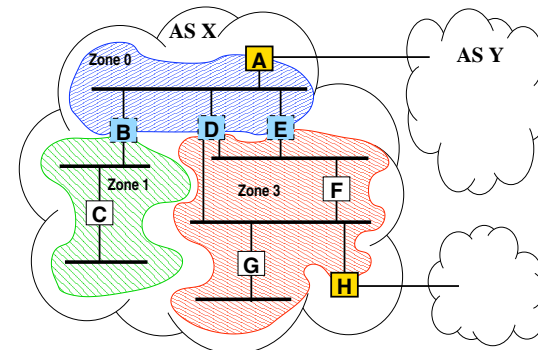
OSPF : Zones (1)



Pour limiter l'impact des changements (échanges, recalculs...)

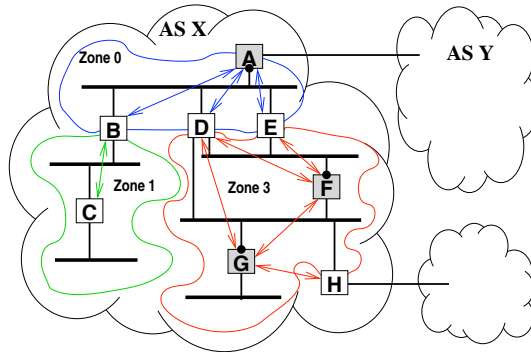
- **Zone** (*Areas*) : sous-parties de l'AS où fonctionne OSPF
 - ✓ identificateur sur 32 bits
 - ✓ contiguës à un *backbone* (Zone 0)

OSPF : Zones (3)



- 3 types de routeur :
 - ✓ **bordure d'AS** : échange de l'information avec l'extérieur (A et H)
 - ✓ **frontière de zone** : appartenant à deux zones (B, D et E)
 - ✓ **interne** : appartenant à 1 zone (C, F et G)

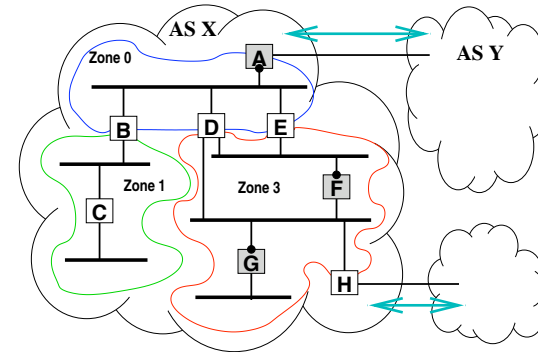
OSPF : Routage dans une zone



Diffusion de l'information dans sa zone

- LAN (*broadcast*) : routeur désigné
- **inondation** (ne pas propager une information déjà reçue)
- ✓ les annonces de G sont transmises à D par F inutilement

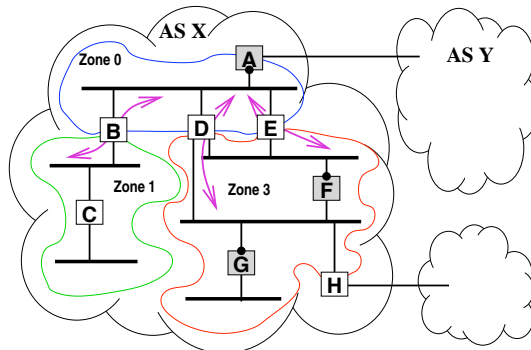
OSPF : Communication avec l'extérieur de l'AS



Echange d'annonces avec l'extérieur

- informe des accessibilités locales
- ✓ différencier les annonces externes pour ne pas transformer le réseau en réseau de transit

OSPF : Echange entre zone



Annonces entre zones

- la Zone 1 reçoit les annonces du *backbone* et de la Zone 3 par B
- ✓ B est le routeur par défaut
- la Zone 3 reçoit les annonces du *backbone* et de la Zone 1 par D et E
- ✓ permet de choisir D ou E

OSPF : Protocoles

Version 2 (RFC 2328) incompatible avec OSPF v1

- définition complexe avec plusieurs sous-protocoles
- ✓ **hello** : test des voisins et élection du routeur désigné (LAN)
- ✓ **transfert de base** : synchronisation
- ✓ **mise à jour** : envoi de l'état des liaisons
- ✓ **acquiescement** : confirmation des mises à jours
- ✓ **demande de l'état des liaisons** : connaissance des routeurs de la zone (NBMA)
- encapsulation directe dans un paquet IP (**protocole 89**)
- utilisation du multicast si disponible :
 - ✓ 224.0.0.5 : tous les routeurs du réseau
 - ✓ 224.0.0.6 : les routeurs désignés

OSPF : Entête générique

0	7	15	23	bit 31
Version		Type	Longueur du paquet	
Identité du routeur				
Indicateur de zone				
Checksum			Type d'authentification	
Authentification				
données				

- Version = 2
- Type = 1 (Hello), 2 (transfert de base), 3 (demande de l'état des liaisons), 4 (mise à jour), 5 (acquittement)
- Longueur du paquet = taille avec entête
- Identité du routeur = unique même si plusieurs interfaces
- Indicateur de zone = zone où se trouve le routeur
- Authentification = permet l'utilisation de MD5
- données... nombreuses structures : voir le RFC 2328

Plan

Rappel sur le routage

Algorithmes de base

Hiérarchie de routage

Routage interne : OSPF

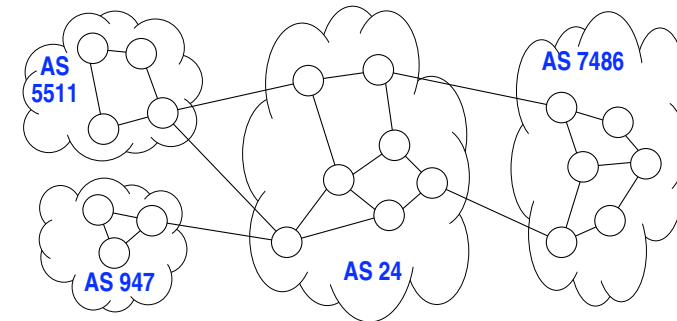
Routage externe : BGP-4

BGP : Introduction

Protocole de facto de routage externe

- chronologie des standards :
 - ✓ EGP (1984) : RFC 904
 - ✓ BGP-1 (1989) : RFC 1195
 - ✓ BGP-2 (1990) : RFC 1163
 - ✓ BGP-3 (1991) : RFC 1267
 - ✓ BGP-4 (1995) : RFC 1771, 1772 et 1773
 - ☞ support de CIDR
 - ☞ exploitation à grande échelle dès 95 avec la commercialisation d'Internet
- protocole à **vecteur de chemin** :
 - ✓ similaire aux protocoles à vecteur de distance
 - ✓ permet d'appliquer des contraintes politiques

BGP : Topologie



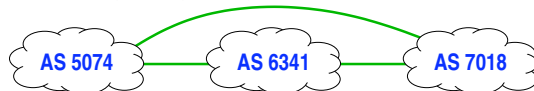
BGP se base sur un ensemble d'AS interconnectés.

- les AS sont représentés par des numéros sur 16 bits
 - ✓ attribués par les bureaux d'enregistrement (ARIN, RIPE-NCC...)
 - ☞ comme pour les préfixes de réseau
 - ✓ env. 25000 attribués (64512 à 65535 privés)

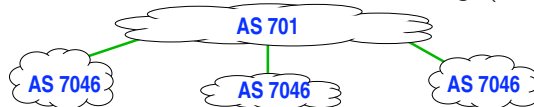
BGP : Correspondance AS/Réseau

Un AS ne correspond pas forcément à un réseau

- les *Tier-1* fractionnent souvent leur réseau :
 - ✓ ATT : 5074, 6341, 7018...
 - ✓ MCI (UUnet) : 284, 701, 702, 12199...
 - ✓ Sprint : 1239, 1240, 6211, 6242...

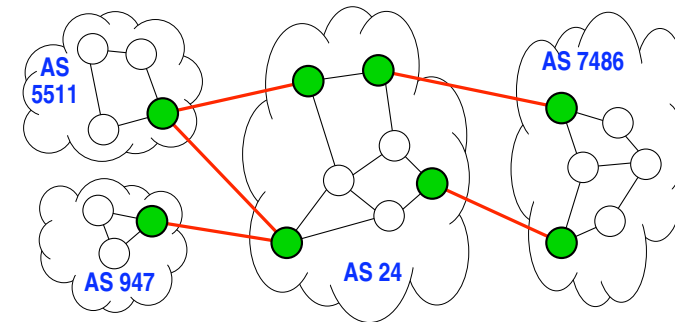


- un numéro d'AS peut être partagé :
 - ✓ AS 7046 : Crestar Bank + NJIT + Hood College (clients AS 701)



- et de nombreux réseaux d'extrémité n'ont pas besoin de BGP et de numéro d'AS (routage statique en bordure du réseau)

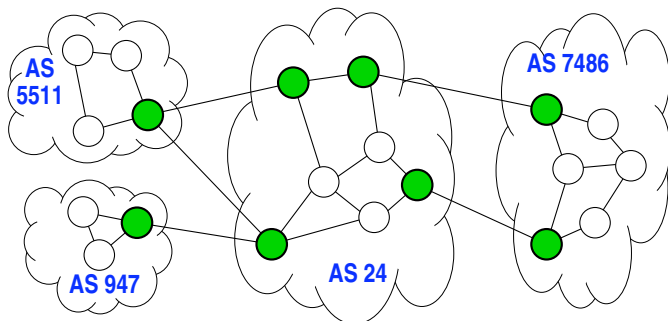
BGP : Connexion eBGP



exterior BGP

- interconnexion entre AS par les routeurs de frontière
- signalisation BGP sur connexion TCP (port 179) directe

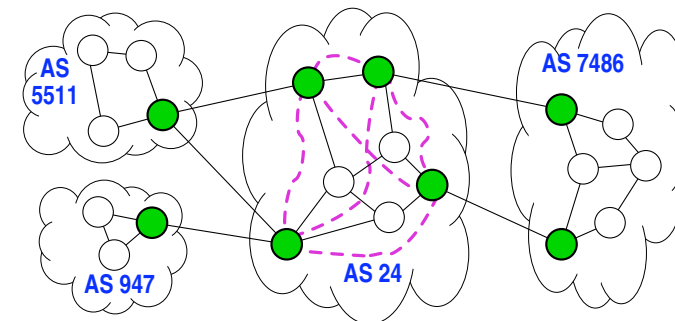
BGP : Routeur de frontière



Border Gateway Routers

- passages vers les autres AS
- associés à deux types de connexion :
 - ✓ externe (eBGP)
 - ✓ interne (iBGP)

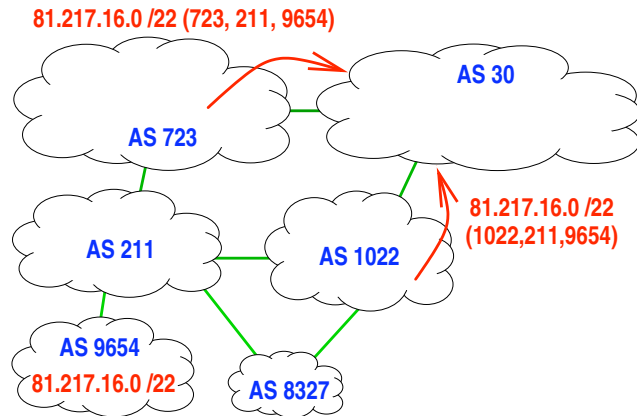
BGP : Connexion iBGP



interior BGP

- interconnexion entre les routeurs de frontière dans un AS
- connexion TCP (port 179) routée avec l'IGP de l'AS
- maillage complet (*full mesh*)

BGP : Informations échangées



Quelles sont les informations échangées entre AS ?

- principalement les **préfixes** IP et les **chemins** des AS vers ceux-ci

BGP : Attributs (1)

Value	Code	Reference
1	ORIGIN	[RFC1771]
2	AS_PATH	[RFC1771]
3	NEXT_HOP	[RFC1771]
4	MULTI_EXIT_DISC	[RFC1771]
5	LOCAL_PREF	[RFC1771]
6	ATOMIC_AGGREGATE	[RFC1771]
7	AGGREGATOR	[RFC1771]
8	COMMUNITY	[RFC1997]
9	ORIGINATOR_ID	[RFC1998]
10	CLUSTER_LIST	[RFC1998]
11	DPA	[Chen]
12	ADVERTISER	[RFC1863]
13	RCID_PATH / CLUSTER_ID	[RFC1863]
14	MP_REACH_NLRI	[RFC2283]
15	MP_UNREACH_NLRI	[RFC2283]
16	EXTENDED COMMUNITIES	[Rosen]
17	NEW_AS_PATH	[E.Chen]
18	NEW_AGGREGATOR	[E.Chen]
19-254	Unassigned	
255	reserved for development	

Annonce = préfixe + quelques attributs (pas tous)

BGP : Messages

Seulement 4 messages BGP :

- OPEN** : ouverture de la connexion
- KEEPALIVE** : maintien de la connexion
 - ✓ envois périodiques
- NOTIFICATION** : terminaison de la connexion
- UPDATE** : échange de **préfixes** avec **attributs**
 - ✓ toute l'information initialement
 - ✓ mise à jours ensuite
 - ☞ **annonce** (*announcing*) de nouvelles routes
 - ☞ **abandon** (*withdrawing*) de route déjà annoncées

BGP : Attributs (2)

ORIGIN : d'où provient la connaissance du préfixe

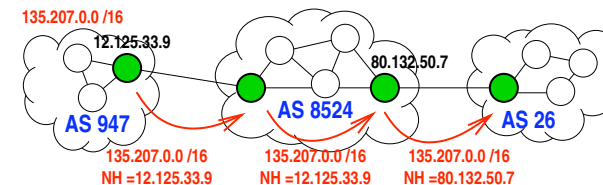
- IGP = vient de l'intérieur de l'AS
- EGP = vient de l'extérieur de l'AS
- INCOMPLETE = configuré manuellement

AS_PATH : suite de numéro d'AS parcouru par l'annonce

- permet de détecter les **boucles** (*Interdomain loop prevention*)

NEXT_HOP : vers qui orienter le trafic du préfixe annoncé

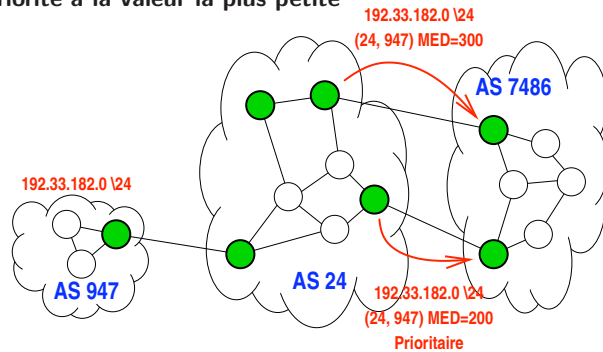
- dernier routeur de l'AS précédent



BGP : Attributs (3)

MULTI_EXIT_DISC : lorsqu'il y a plusieurs points de sortie d'un AS

- **priorité à la valeur la plus petite**



BGP : Annonces

Emission d'un message **UPDATE**

- quels préfixes annoncer ?
✓ **choix de l'émetteur**
- quelles valeurs d'attribut associer ?
✓ dépend de l'attribut
 - ☞ AS_PATH = AS_PATH précédent + numéro de l'AS actuel
 - ☞ MULTI_EXIT_DISC = dépend du choix de l'émetteur
 - ☞ ...

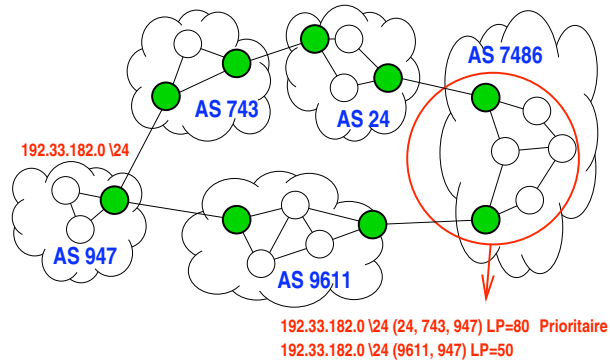
Réception d'un message **UPDATE**

- quels informations prendre en compte ?
✓ **choix de préfixes** (filtrage)
✓ possibilité de modifier les attributs
- que faire des informations acceptées ?
✓ **choisir les routes**
 - ☞ utilisation d'un algorithme de décision...

BGP : Attributs (4)

LOCAL_PREF : préférence administrative

- **priorité à la valeur la plus élevée**



BGP : Algorithme de choix des routes

Critères de choix du plus fort au plus faible :

1. LOCAL_PREF le plus élevé
2. AS_PATH le plus court
 - mais pas forcément le plus court chemin
3. MULTI_EXIT_DISC le plus petit
4. priorité aux chemins appris par iBGP que par eBGP
5. chemin le plus court pour atteindre le NEXT_HOP (métrique IGP)
6. identifiant de routeur le plus petit

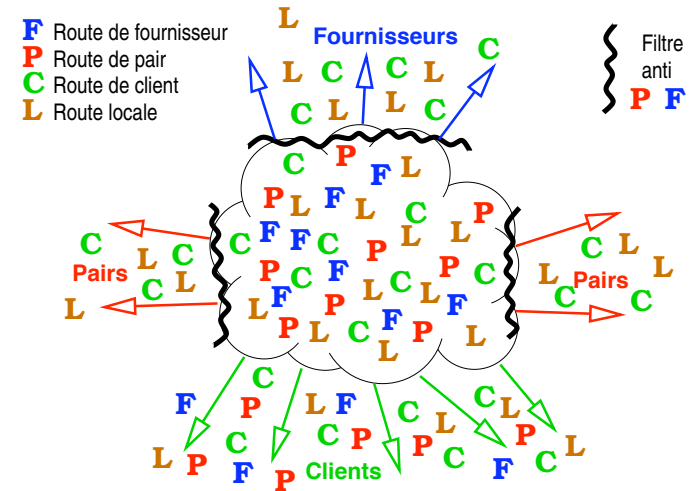
BGP : Et le choix politique ?

Encore un attribut...

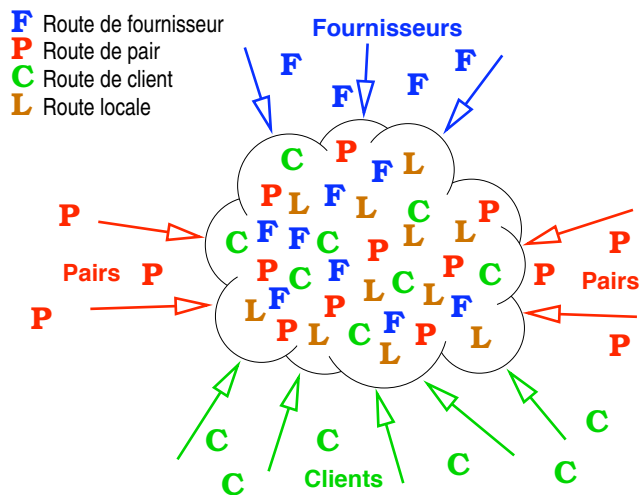
COMMUNITY : permet de "colorier" les routes

- liste de valeurs indiquant à quelles communautés appartient un préfixe
 - ✓ 32bits (16bits AS colorieur + 16bits au choix)
 - ✓ les annonces sont généralement coloriées à l'entrée de l'AS
 - ☞ communauté client
 - ☞ communauté pair
 - ☞ communauté fournisseur
 - ✓ permet de **filtrer** à la sortie de l'AS
 - ☞ exemple : ne pas injecter les préfixes d'un pair à un autre pair (et ainsi se transformer en AS de transit)

BGP : Export de routes



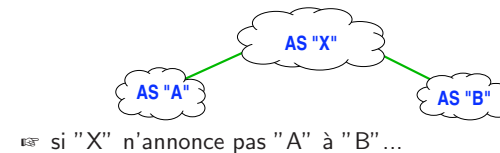
BGP : Import de routes



BGP : Connectivité

BGP garantit-il la connectivité ?

- **non**, certains réseaux peuvent être injoignables
 - ✓ dépend des politiques rencontrées sur le chemin des annonces :



BGP : Convergence

BGP garantit-il la convergence pour un routage stable ?

- sans changement, il peut y avoir des oscillations (*route flapping*)
 - ✓ un routeur annonce un préfixe puis l'abandonne
↳ lié à des liens défaillants
- avec changement, le nombre d'annonces est élevé
 - ✓ certains AS peuvent observer plus 10^6 UPDATE par jours

Fin

Document réalisé avec L^AT_EX.

Classe de document foils.

Dessins réalisés avec xfig.

Olivier Fourmaux, olivier.fourmaux@upmc.fr

<http://www-rp.lip6.fr/~fourmaux>

Ce document est disponible en format PDF sur le site :

<http://www-master.ufr-info-p6.jussieu.fr/>

BGP : Problèmes

- les erreurs ont une portée globale (sur tout l'Internet)
 - ✓ un AS avec une mauvaise configuration peut indiquer qu'il a la meilleur route pour tout les destinataires...
- croissance exponentielle du nombre des annonces
 - ✓ de plus en plus d'AS
 - ✓ préfixes de plus en plus petits
 - ✓ pas d'agrégation à cause du *multihoming*
- supervision complexe
 - ✓ le graphe des AS dépend du point de vue
- tentative d'amortissement du *route flapping*
 - ✓ utilisation du *route dampening*