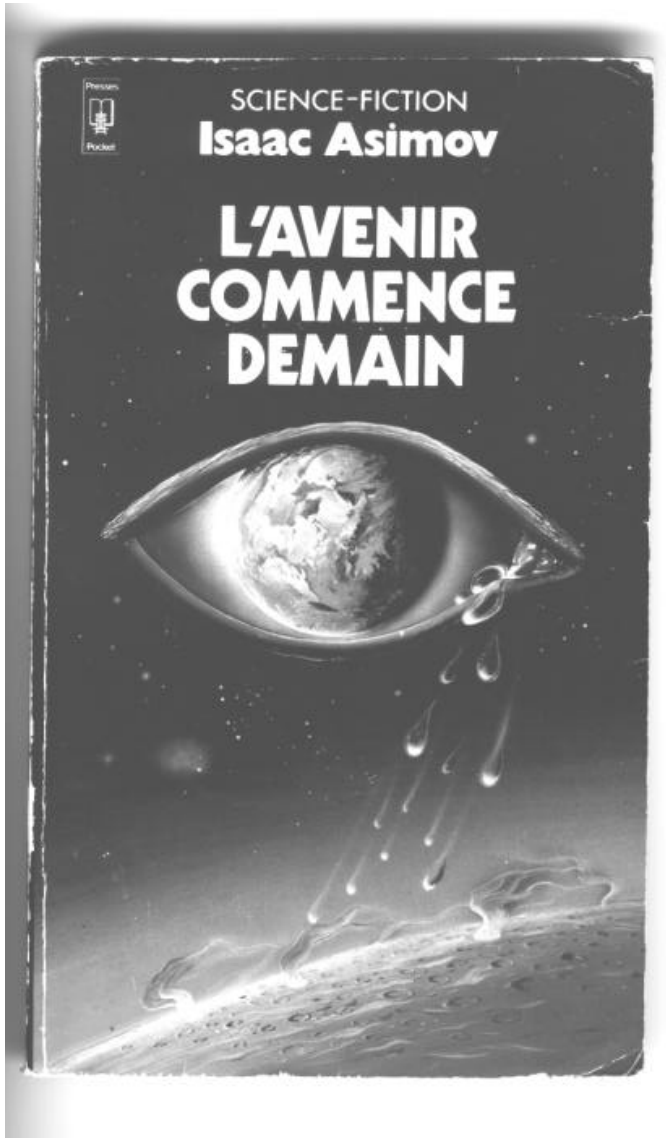
Les réseaux TCP/IP

Les réseaux TCP/IP

Laurent Toutain

EVOLUTION DES RÉSEAUX INFORMATIQUES



2

TOUS LES ENNUIS DU MONDE

L'INDUSTRIE la plus vaste de la Terre était centrée sur Multivac... Multivac, l'ordinateur géant qui en cinquante ans avait grandi jusqu'à remplir de ses diverses branches Washington et ses faubourgs, puis avait étendu ses tentacules à toutes les villes et bourgs du globe.

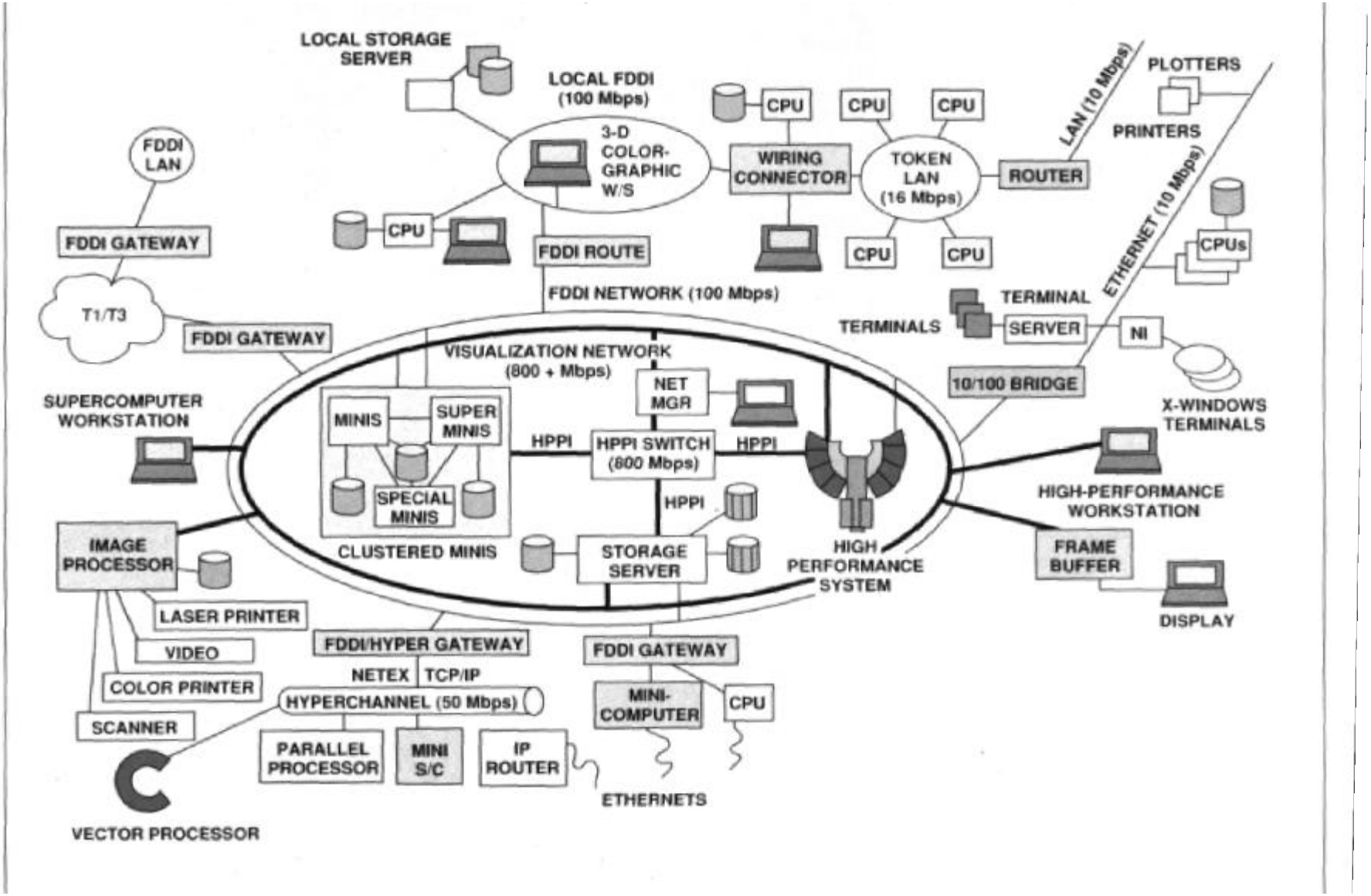
Une armée de fonctionnaires l'approvisionnaient sans cesse de renseignements alors qu'une autre armée interprétait ses réponses et établissait les corrélations entre elles. Un véritable corps du génie patrouillait à l'intérieur, tandis que mines et usines s'acharnaient à maintenir continuellement au complet, avec précision, pour donner pleine et entière satisfaction, les stocks de réserve de pièces de rechange.

Multivac dirigeait l'économie de la Terre et venait en aide à sa science. Mais il était avant tout le centre de recueil et de classement de tous les faits connus relatifs à chaque individu de la Terre.

Et tous les jours, une partie des devoirs de Multivac consistait à examiner les quatre milliards d'ensembles de faits portant sur chacun des êtres humains qui garnissaient ses entrailles pour extrapoler une nouvelle journée à partir de cette base. Tous les Services de Rectification de la Terre recevaient

11

ORGANISATION ACTUELLE - CAS D'ÉCOLE



TYPES DE RÉSEAUX

EN FONCTION DE LA DISTANCE

Caractéristiques de distance	LAN Local Area Network Réseau Local	MAN Metropolitan Area Network Réseau de campus	WAN Wide Area Network Réseau public
Taille géographique	1 mètre à 2 kilomètres	1 mètre à 200 kilomètres	quelques milliers de kilomètres
Nombre d'abonnés	2 à 200	2 à 1000	plusieurs milliers
Opérateur	l'utilisateur ou son service	regroupement d'utilisateurs	différent des utilisateurs
Facturation	gratuit	forfait	volume et durée
Débits	1 à 100 Mbits/s	1 à 100 Mbits/s	de 50 b/s à 2 Mbits/s
Taux d'erreur	inférieur à 10-9	inférieur à 10-9	10-3 à 10-6
Délai	1 à 100 ms	10 à 100 ms	inférieur à 0.5 s

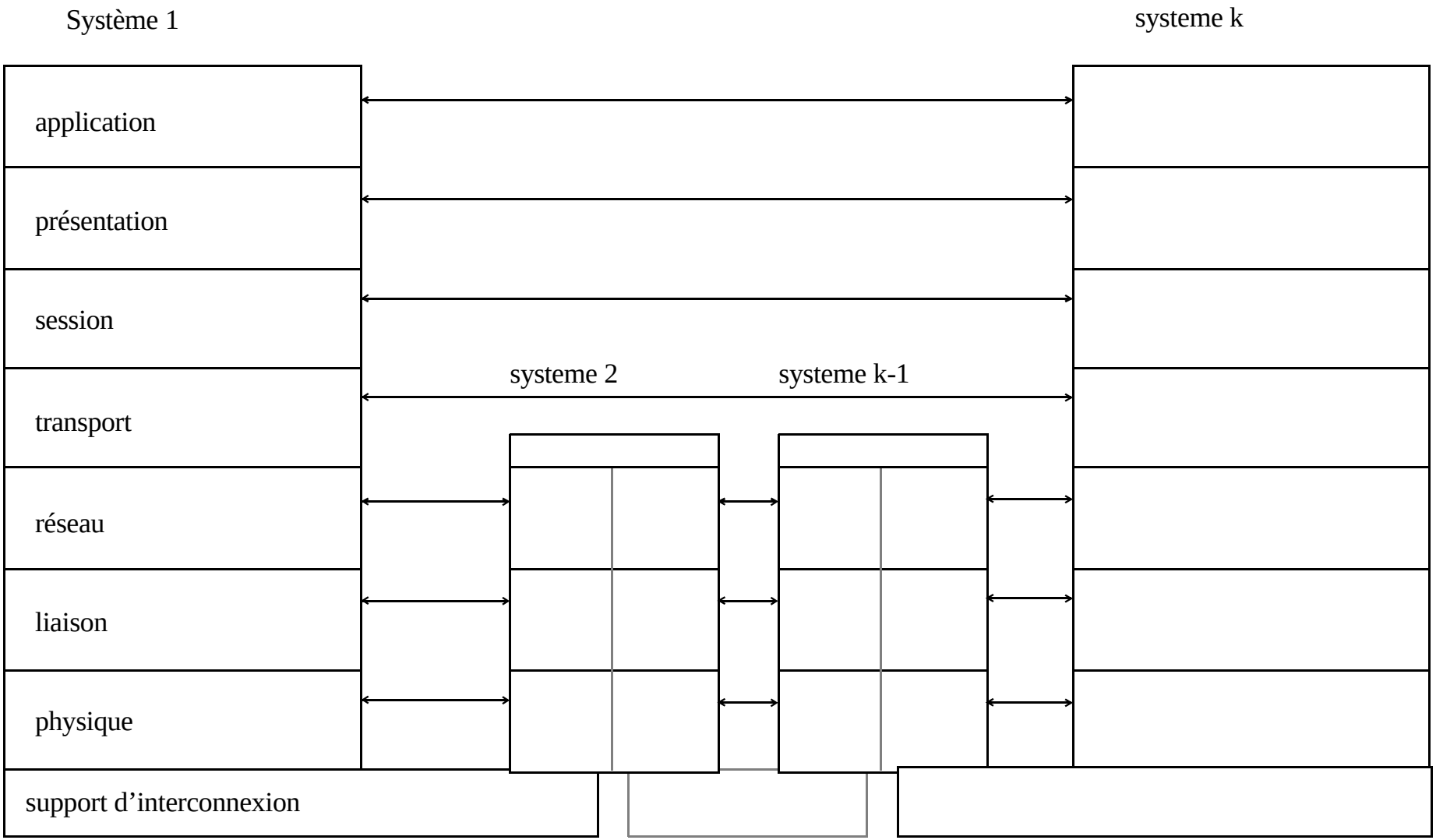
EN FONCTION DE LA MÉTHODE D'ACCÈS :

- Diffusion : Ethernet, Token Ring, FDDI.
- **N**on **B**roadcast **M**ultiple **A**ccess : réseau téléphonique, X.25, ATM
- Point-à-point : liaison spécialisée, téléphone rouge, Circuit Virtuel ou Permanent
- Un équipement maître peut diffuser, les autres ne peuvent dialoguer qu'avec le maître : GSM, RNIS.

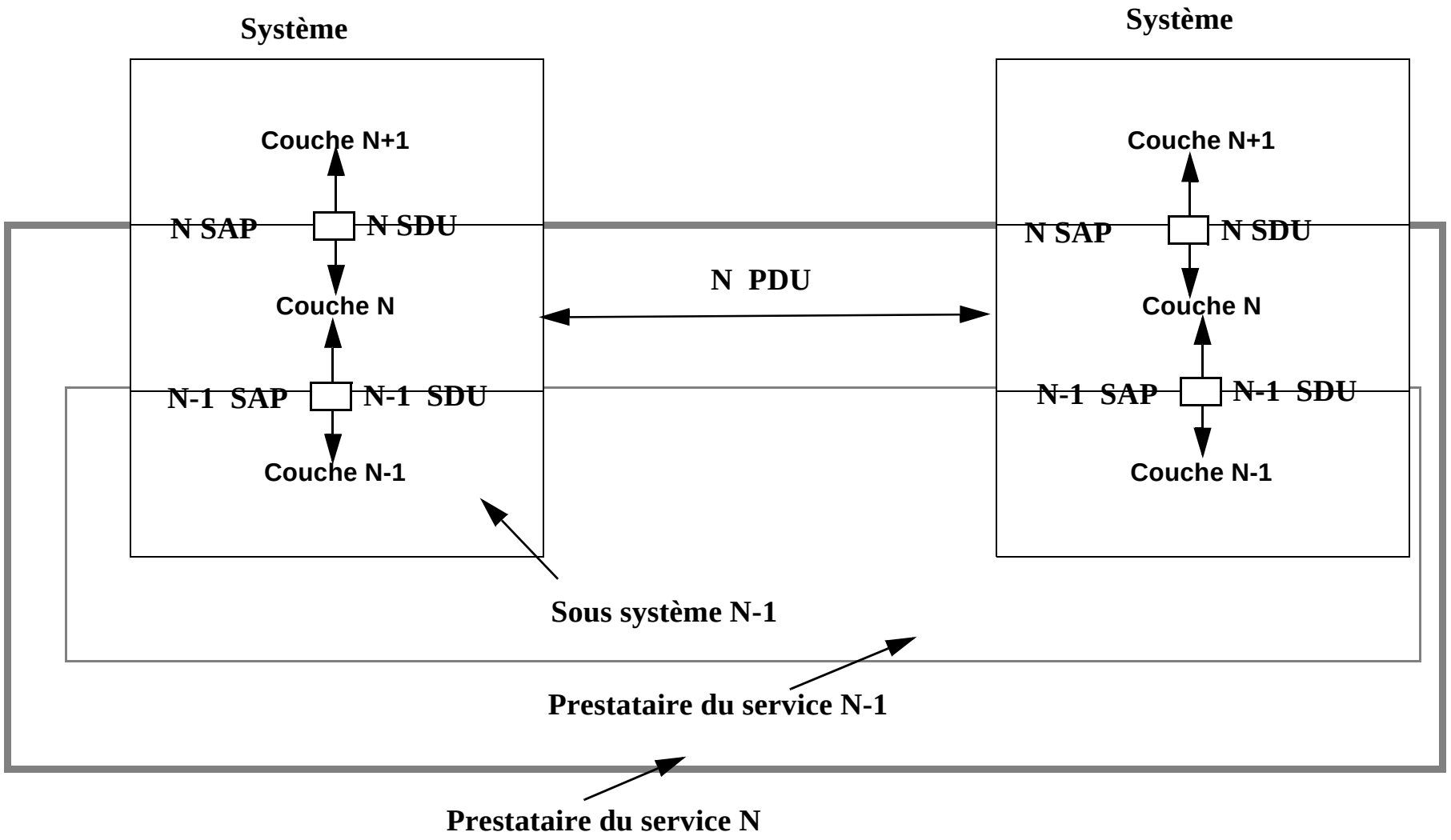
+ Pour quel type de réseau, un adressage est-il nécessaire ?

+ Quels sont les types de réseaux qui résistent au facteur d'échelle ?

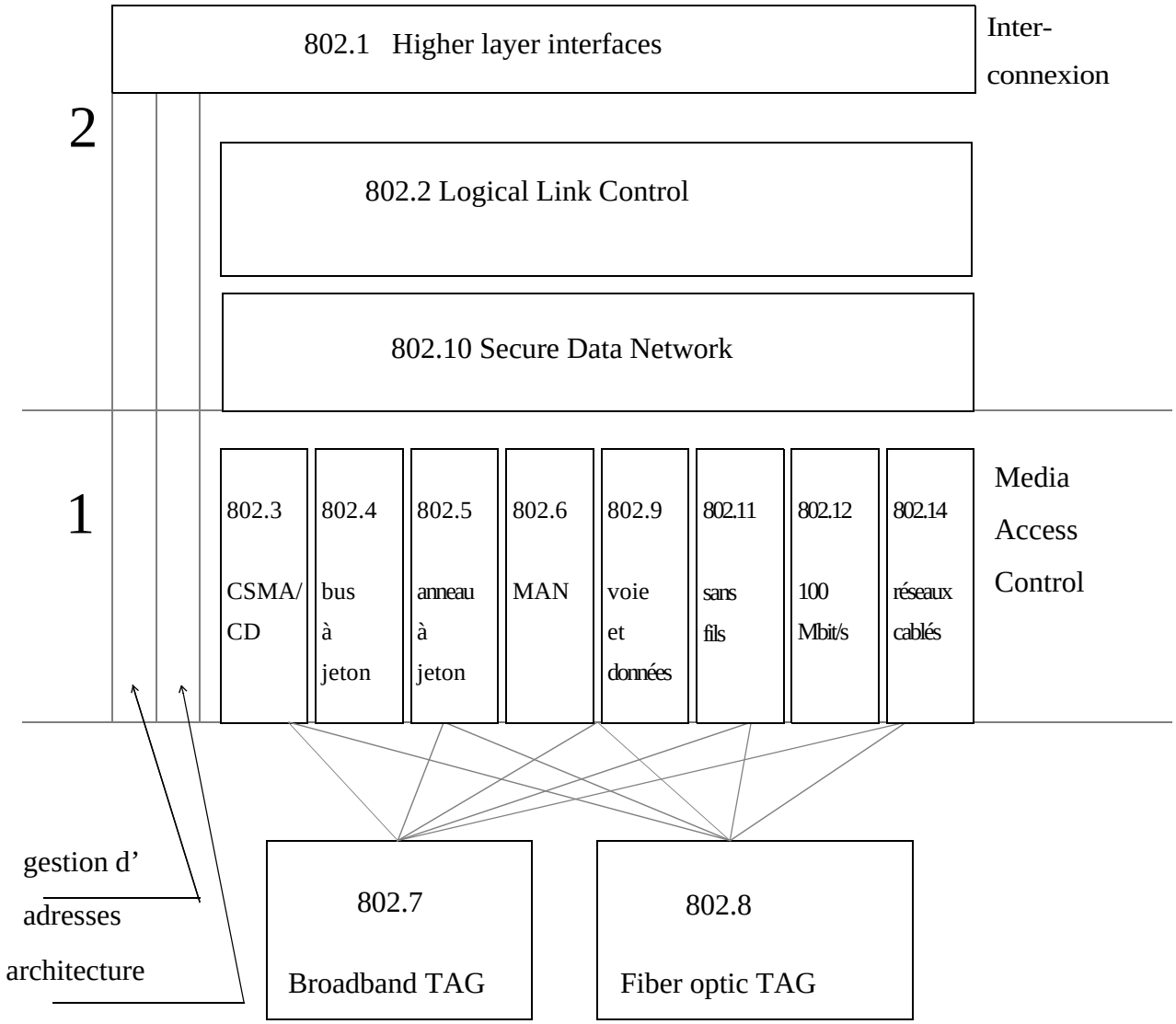
MODÈLE DE RÉFÉRENCE DE L'ISO



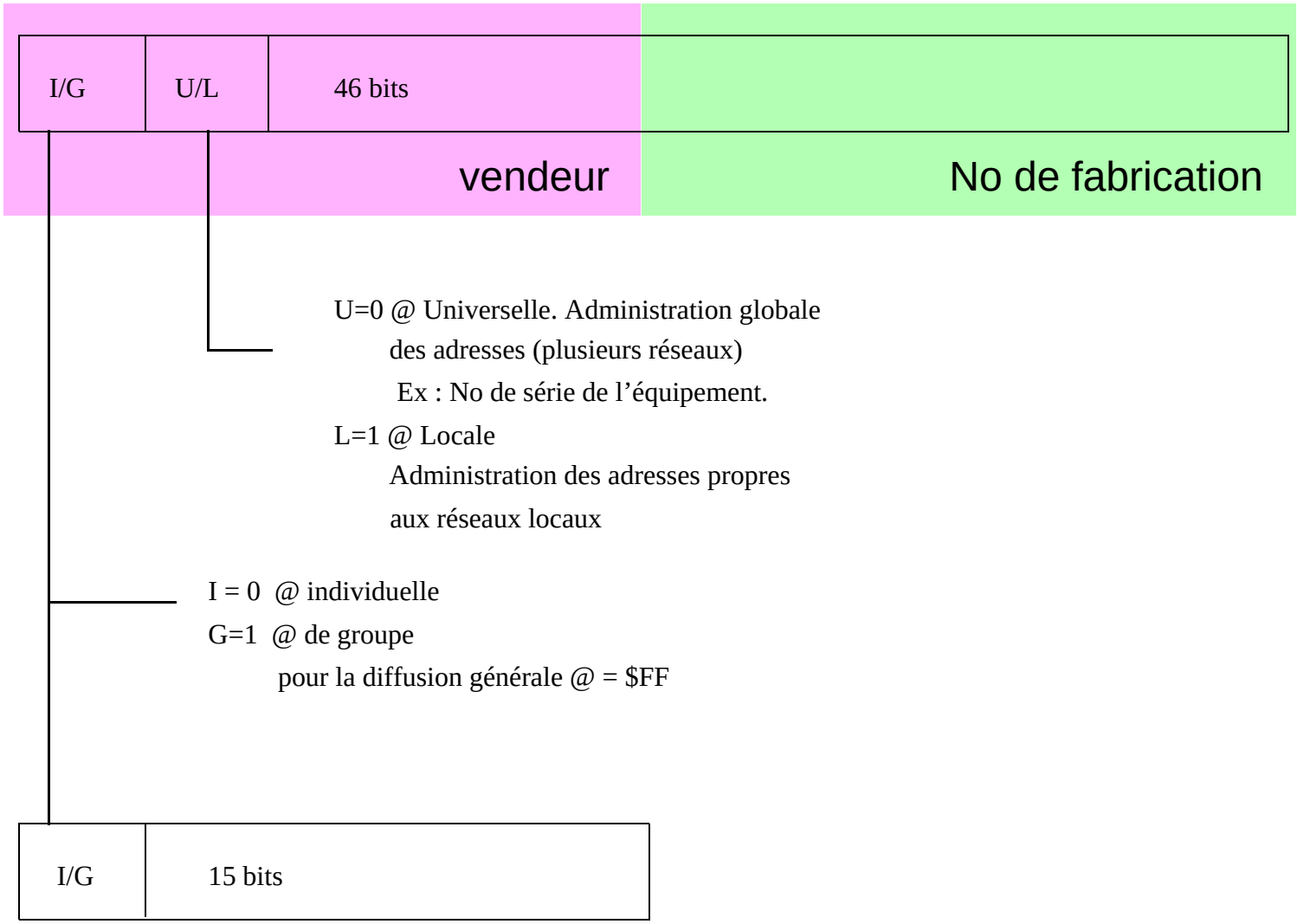
RÈGLES DE L'ARCHITECTURE DU MODÈLE



MODULES CONSTITUTIFS DU STANDARD IEEE 802



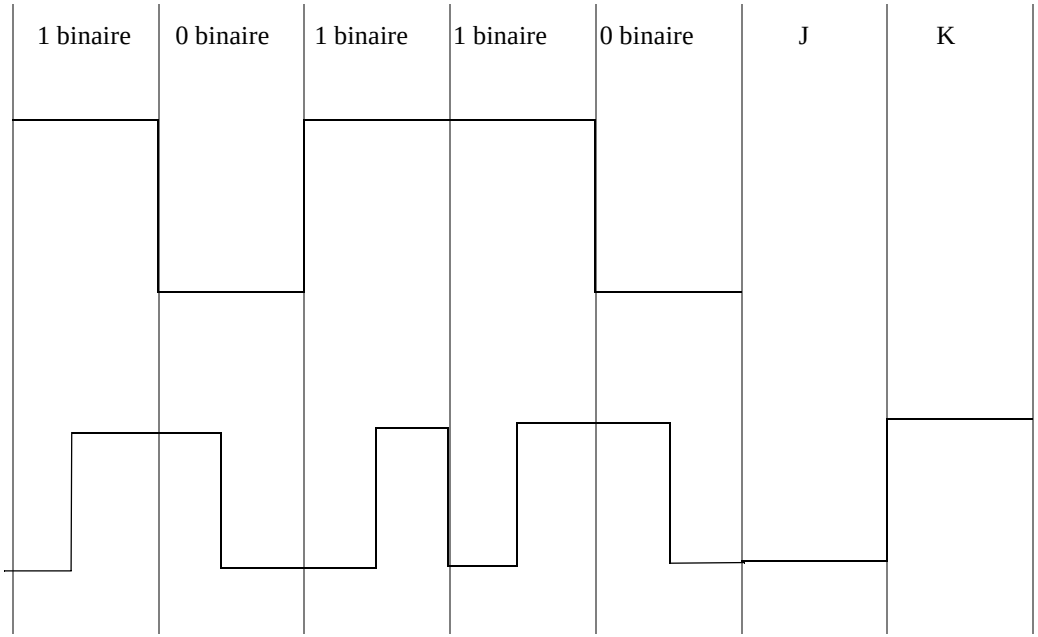
ADRESSES MAC (IEEE 802.1)



EXEMPLE D'ADRESSE DE VENDEUR (RFC 1700)

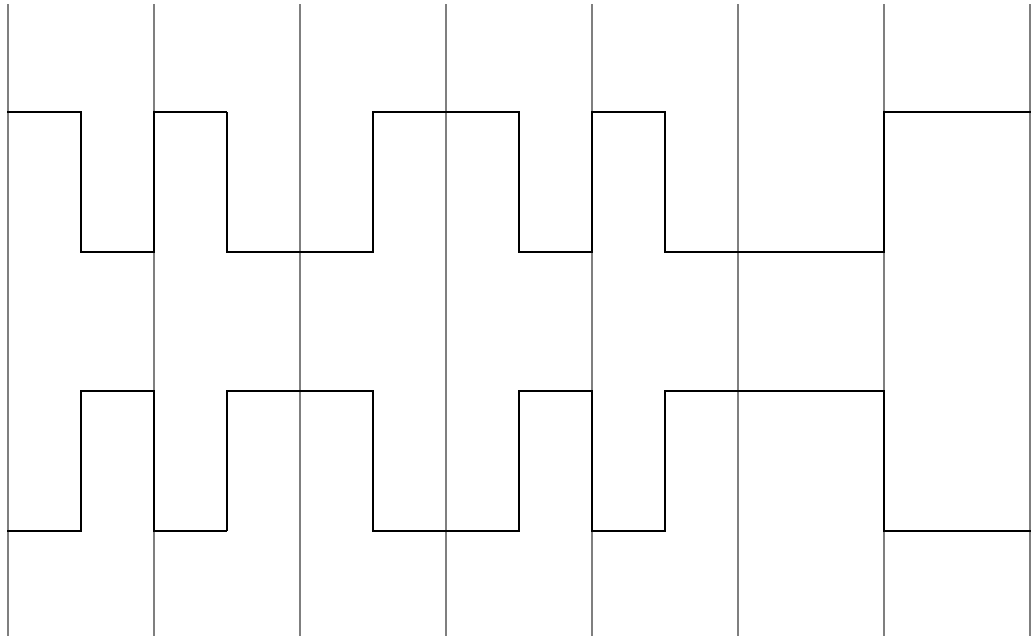
Début d'adresse MAC (en hexadécimal)	Vendeur
00:00:0C	Cisco
00:00:1D	Cabletron
08:00:20	Sun
08:00:2B	DEC
08:00:5A	IBM

CODAGE DE L'INFORMATION : CODAGE MANCHESTER



- Front montant : 1
- Front Descendant : 0

CODAGE DE L'INFORMATION : CODAGE MANCHESTER DIFFÉRENTIEL



- même état = 1
- changement d'état = 0

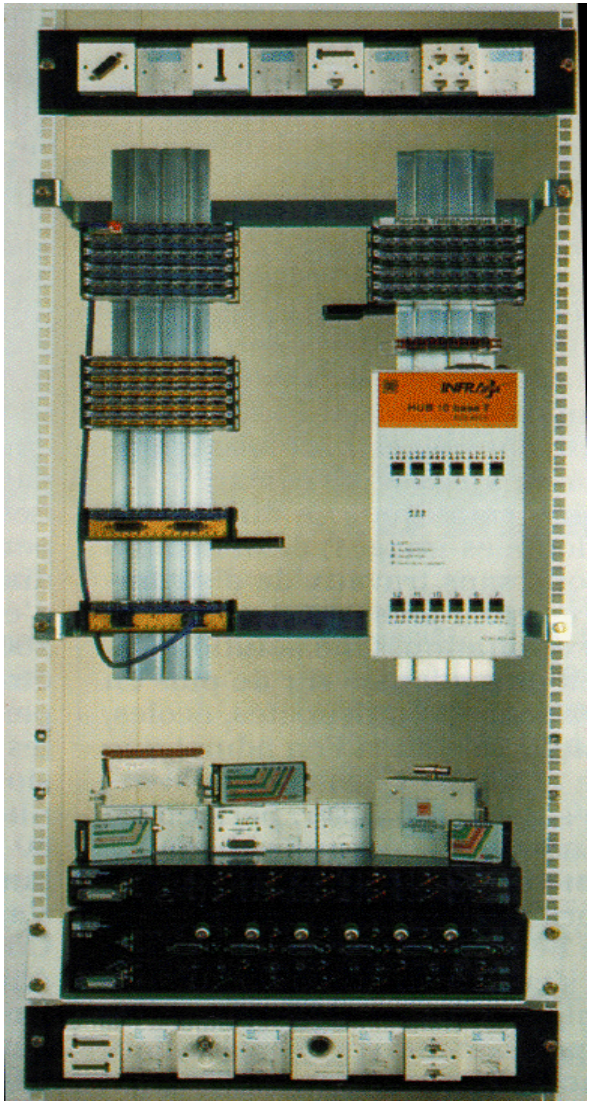
AVANTAGE DU CODAGE MANCHESTER

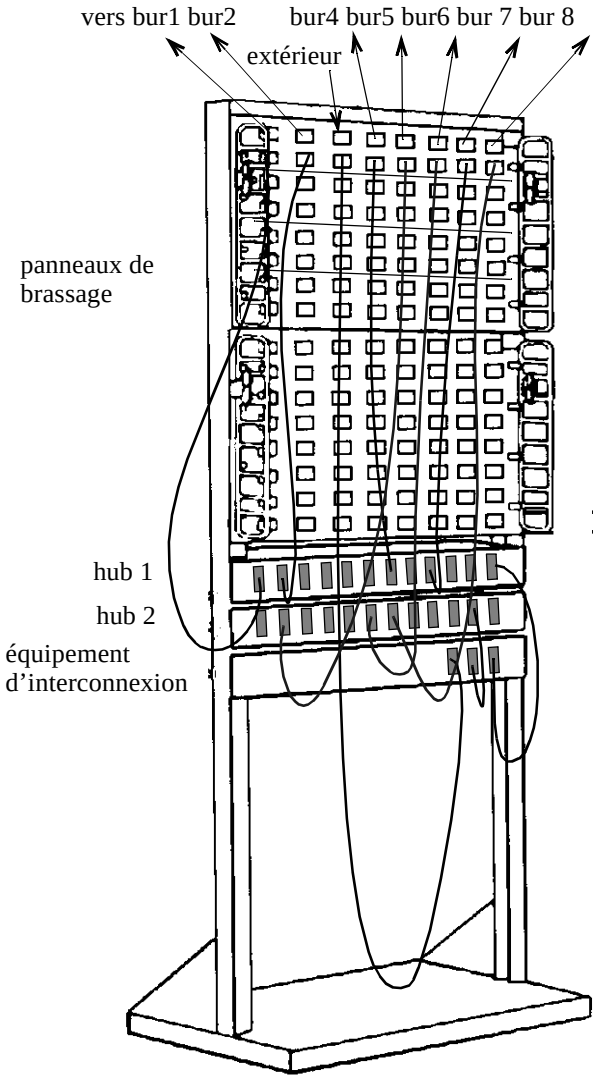
- introduction de transitions
- transport de courant d'alimentation d'équipement réseau
- pas de composante continue
- détection facile des erreurs de transmission quand aucun changement de polarité n'est détecté pendant une longue période.
- présence de deux symboles J et K autorisant la signalisation des trames,
- le codage différentiel permet d'ignorer la polarité du signal

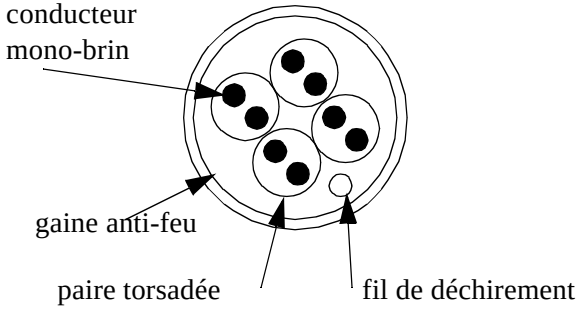
CÂBLAGE

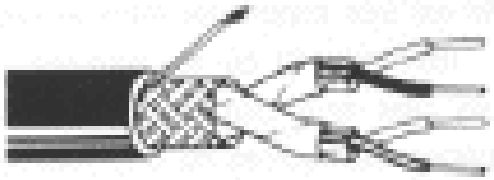
- le 100 Ohms (ATT EIA/TIA-568A (US) IEC/ISO 11801 (Europe)). En plus de la voix :
  - la catégorie 3 : bande passante de 16 Mhz,
  - la catégorie 4 : bande passante de 20 Mhz,
  - la catégorie 5 : bande passante de 100 Mhz,
- le 150 Ohms (IBM) :
  - le type 1 ou STP (*Shielded Twisted Pair*) seul type pour les réseaux à 16 Mbits/s.
  - le type 2 deux paires torsadées blindées (données) et quatre paires torsadées non blindées (téléphonie).
  - le type 3 ou UTP (*Unshielded Twisted Pair*) appartient à la catégorie 3 de EIA/TIA.
  - le type 5 ou fibre optique.
  - le type 6 ou câble de jarretière HF.
  - le type 8 ou câble extra-plat.
- le 120 Ohms (France Télécom, Deutsche Telekom) avec le câblage COREL L120 est aussi une alternative dans la norme ISO 11801. Son impédance peut poser des problèmes.

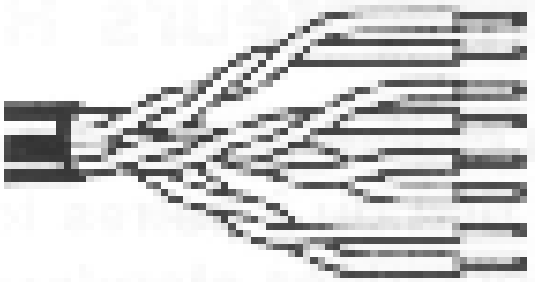
EXEMPLE DE CÂBLES











INTERCONNEXION DE RÉSEAUX

- Définition du vocabulaire :

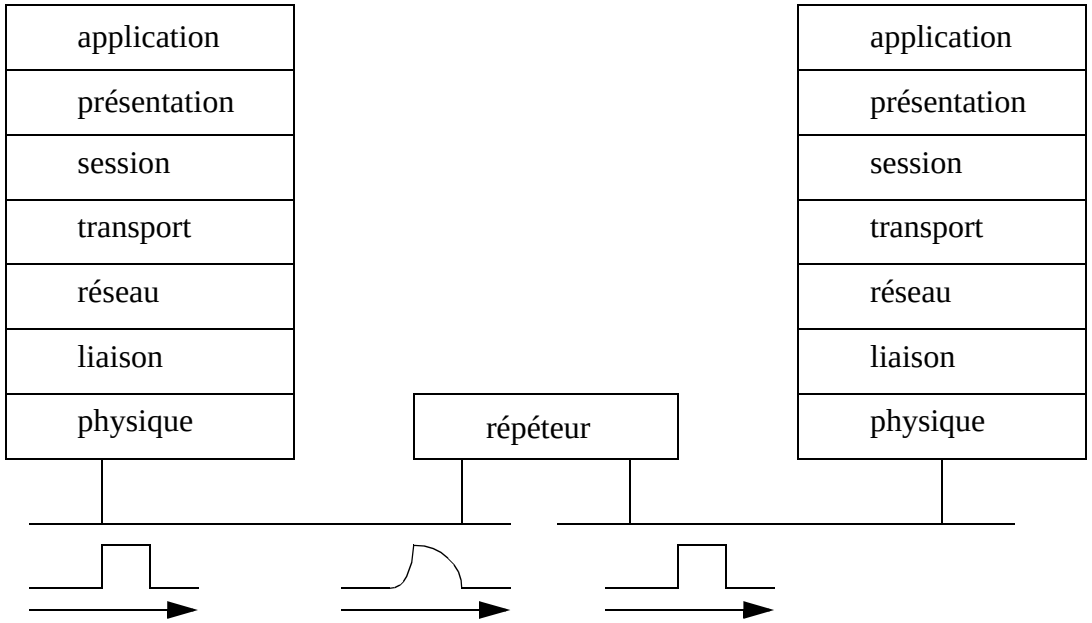
  - Au niveau physique : Le répéteur

  - Au niveau liaison : Le pont

  - Au niveau réseau : Le routeur

Aux niveaux liaison ou réseau : Le B-Router

LES RÉPÉTEURS

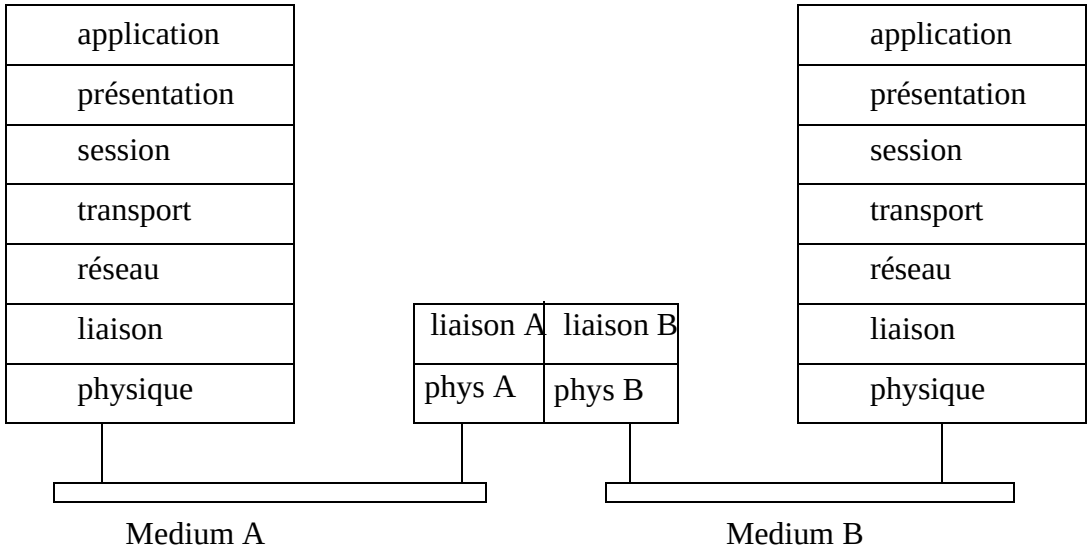


- Travaille au niveau du bit,
- Régéneration du signal,
- Changement de type de support.

Problème :

- nombre limité sur Ethernet

LES PONTS

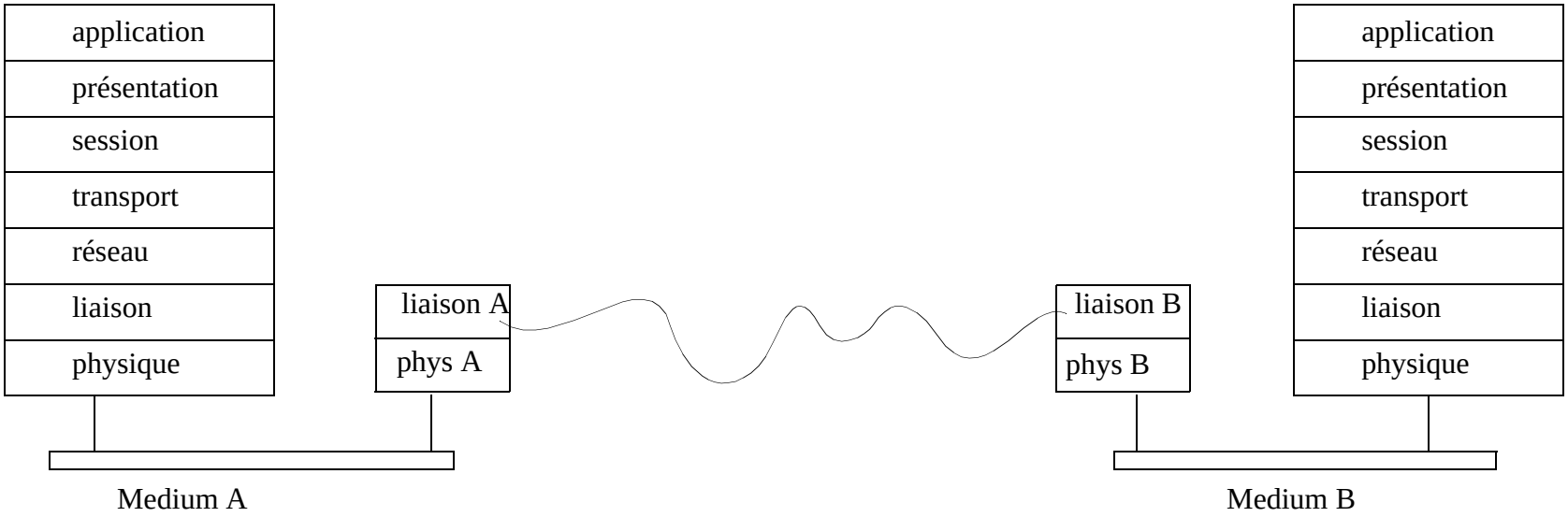


- Travaille sur les trames,
- Recopie et mémorise.

Problèmes :

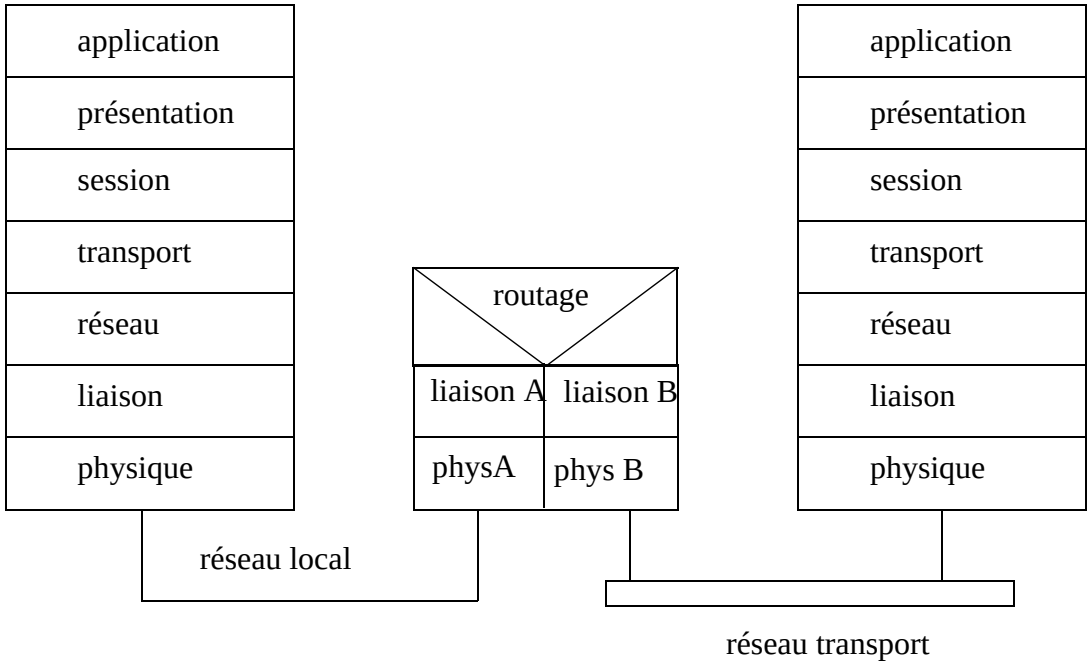
- le format des trames est différent,
- le bit le moins significatif est transmis en premier ou en dernier,
- la taille maximale est différente,
- le débit est différent.

LES DEMI-PONTS



- Même principe qu'un pont, mais il existe une liaison entre les deux équipements,
- La liaison est point-à-point,
- Protocole entre les deux demi-ponts est souvent propriétaire.

LES ROUTEURS

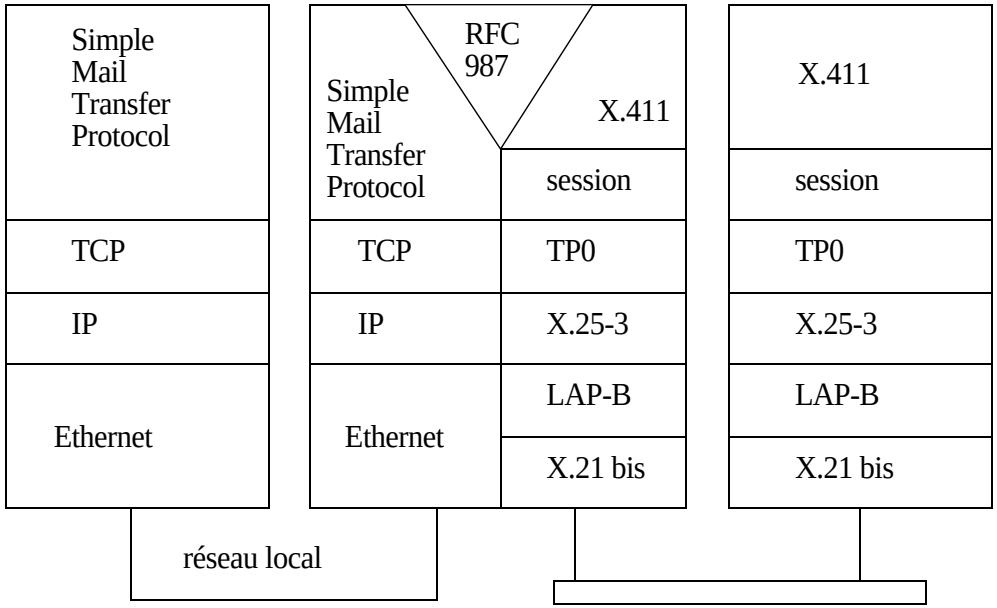


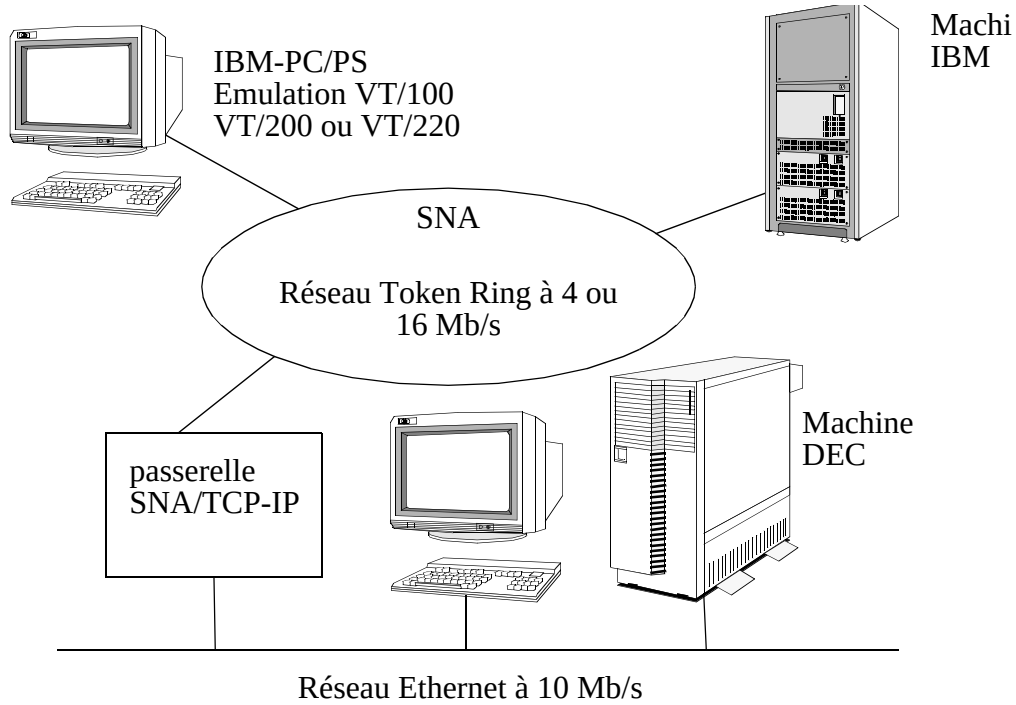
- Travaille au niveau du paquet
- Permet de changer d'espace d'adressage

Problème :

- demande de l'administration.

LES PASSERELLES





+ Les proxy sont aussi des passerelles :

- sécurité : vérification de la syntaxe des commandes
- audit
- cache (mémorisation des informations déjà demandées).

QUESTIONS

- Que mémorisent : un répéteur, un pont, un demi-pont, un routeur, une passerelle ?
- Quels équipements permettent d'interconnecter des stations qui sont dans des espaces d'adressage différents ?
- FDDI, Ethernet, l'anneau à jeton sont-t-ils dans le même espace d'adressage ?
- Combien d'attachement aux réseaux ont un pont ou un routeur ?

QUESTIONS

+ Quelle est la différence entre un demi-pont et un routeur ?

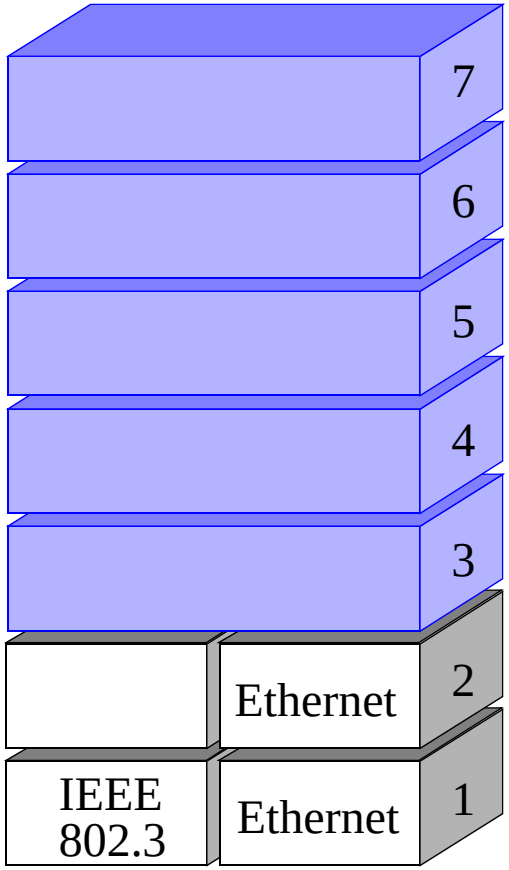
+ Que fait un pont quand il ne connaît pas le destinataire ?

+ Que fait un routeur quand il ne connaît pas le destinataire ?

+ Comment sont traités les trames de diffusion par un pont ?

+ Comment sont traités les trames de diffusion par un routeur ?

ETHERNET / IEEE 802.3



CARTE D'IDENTITÉ DU PROTOCOLE

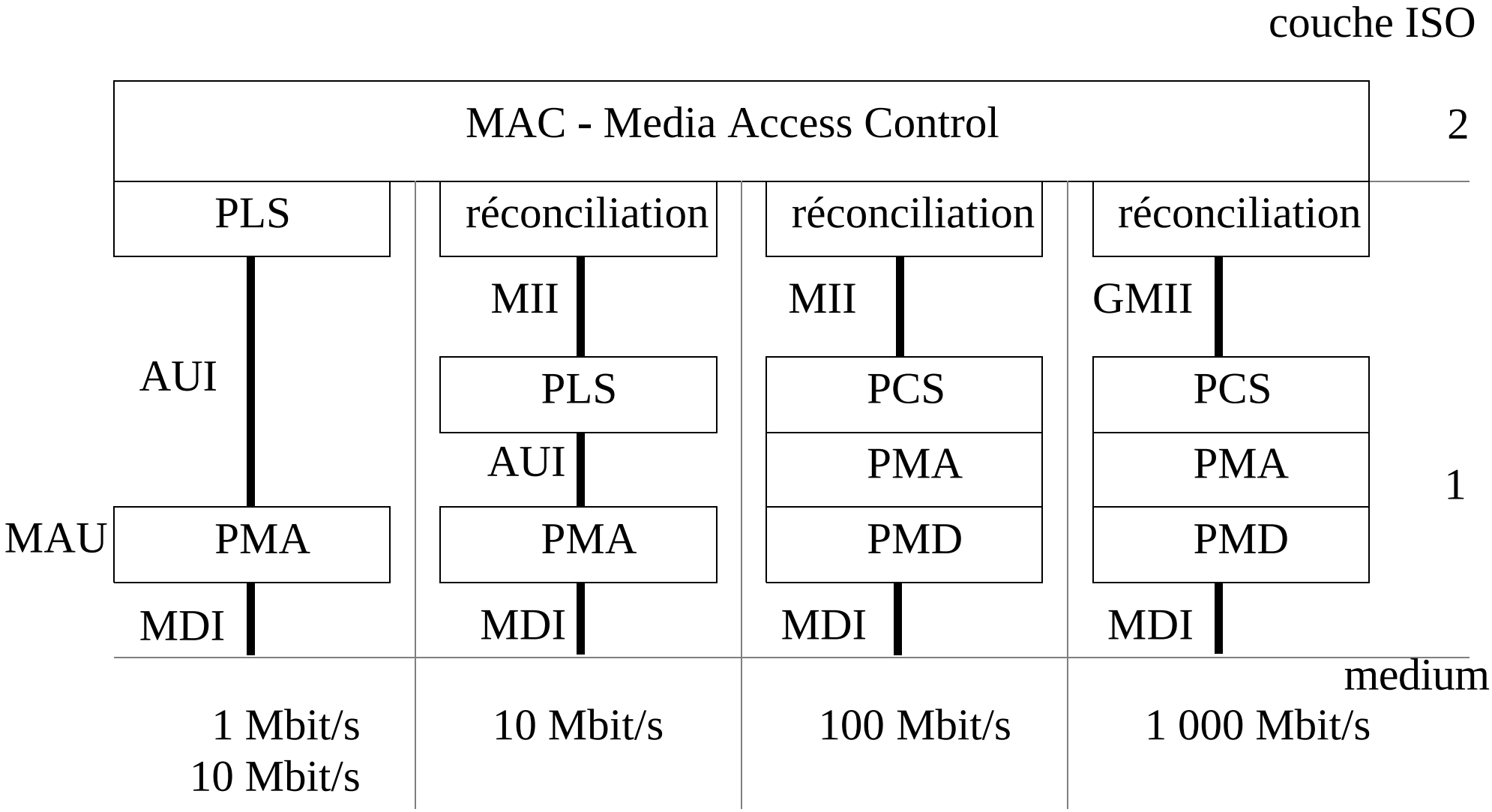
- Ethernet proposé par Digital, Intel et Xerox (DIX).
- Normalisé par IEEE : IEEE 802.3.
- Ethernet vs IEEE 802.3 :
  - même format de trame
  - affectation des champs différente => Deux protocoles INCOMPATIBLES
- Protocole évolutif :
  - vitesse de transmission (10, 100, 1 000, projets à 10 000 Mbit/s)
  - support (coaxial, paire torsadée, fibre optique)
  - methode d'accès (CDMA/CD, half duplex, full duplex)

ETAT DE LA NORMALISATION

1 Mbit/s	Paire Torsadée	1BASE5		HD
10 Mbit/s	coaxial	10BASE5 *Thick Ethernet*		HD
		10BASE2 *Thin Ethernet*		HD
	TV	10BROAD36		HD
	Fibre Optique	10BASE-F	10BASE-FB	HD / FD
			10BASE-FL	HD / FD
			10BASE-FP	HD / FD
	Paire Torsadée	10BASE-T		HD / FD
100 MBit/s	PT	100BASE-T	100BASE-T2	HD / FD
			100BASE-T4	HD
			100BASE-X 100BASE-TX	HD / FD
	Fibre Optique		100BASE-X 100BASE-FX	HD / FD
1000 Mbit/s	Fibre Optique	1000BASE-X	1000BASE-LX	HD / FD
			1000BASE-SX	HD / FD
	Paire Torsadée		1000BASE-CX	HD / FD
		1000BASE-TX		HD / FD

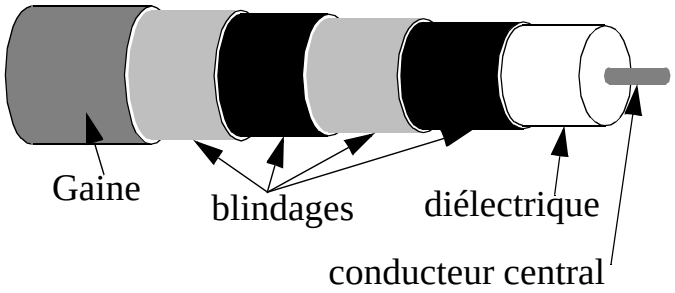
HD : Half Duplex - FD : Full Duplex

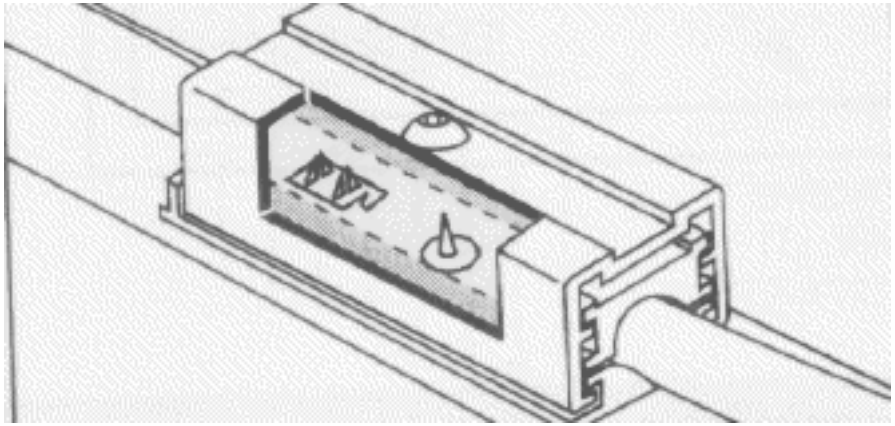
ARCHITECTURE

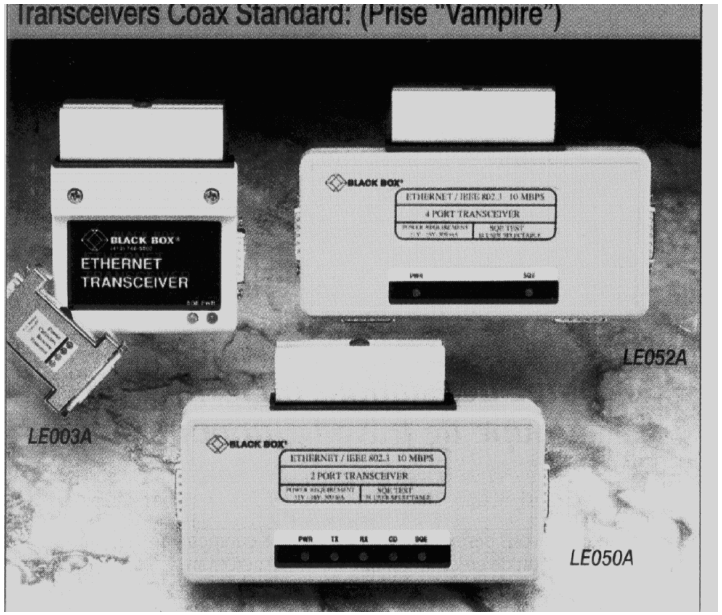


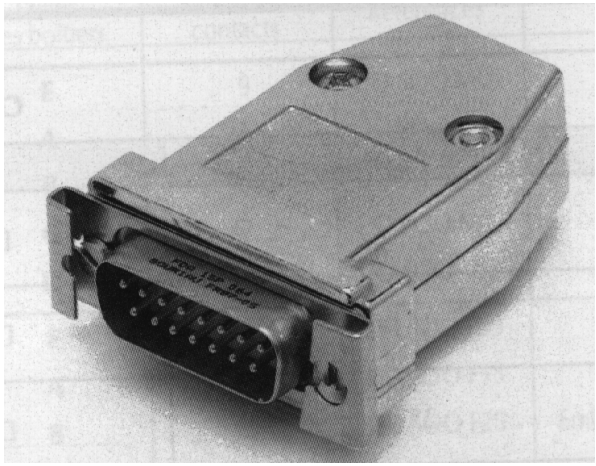
EQUIPEMENTS

10 BASE 5 (*THICK ETHERNET*)



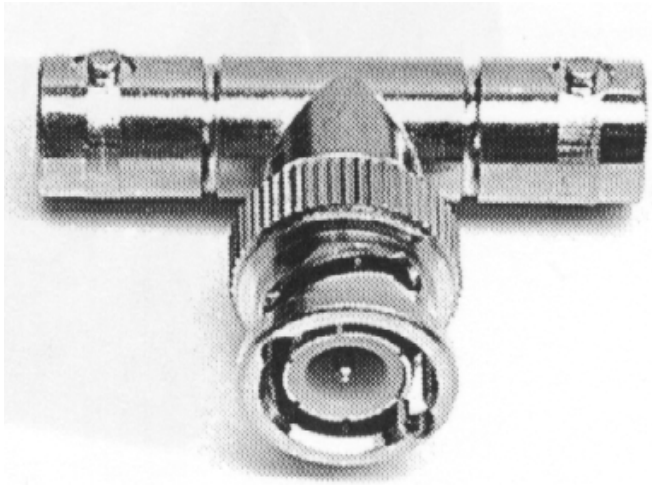




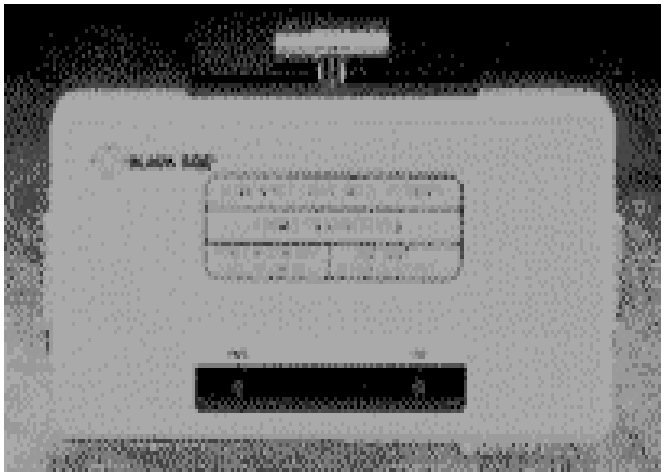


10 BASE 2 (*THIN ETHERNET*)

câble thin ethernet

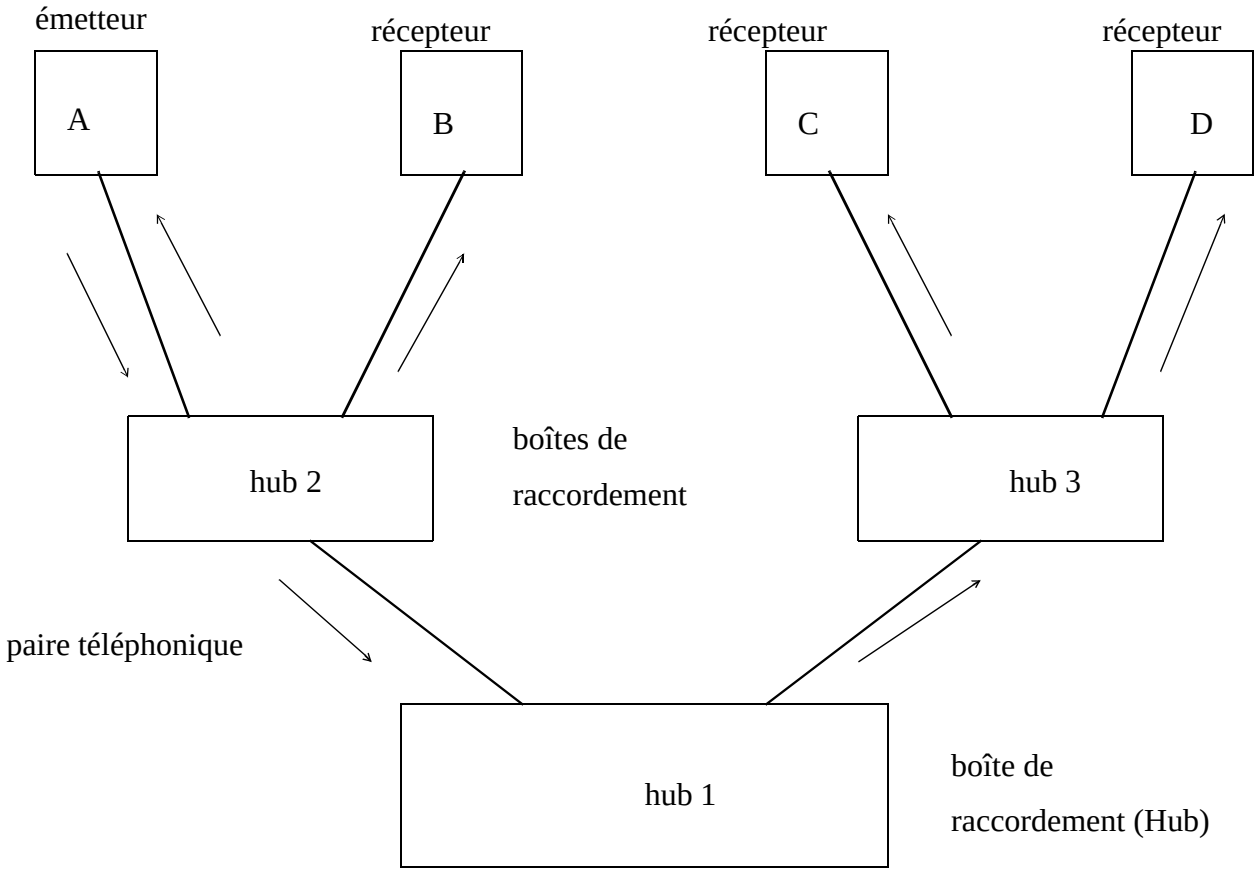


prise BNC en T

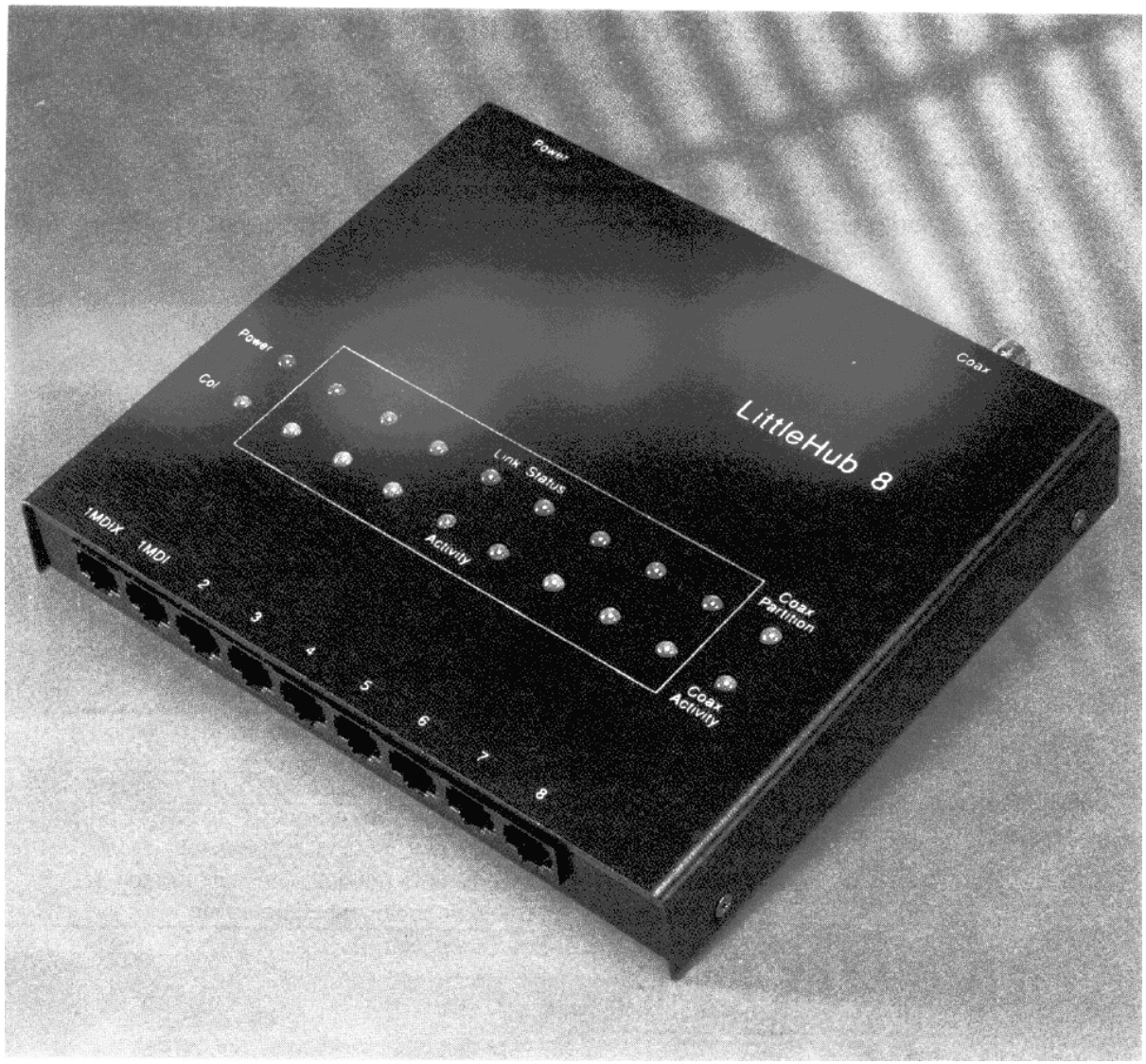


Transceiver

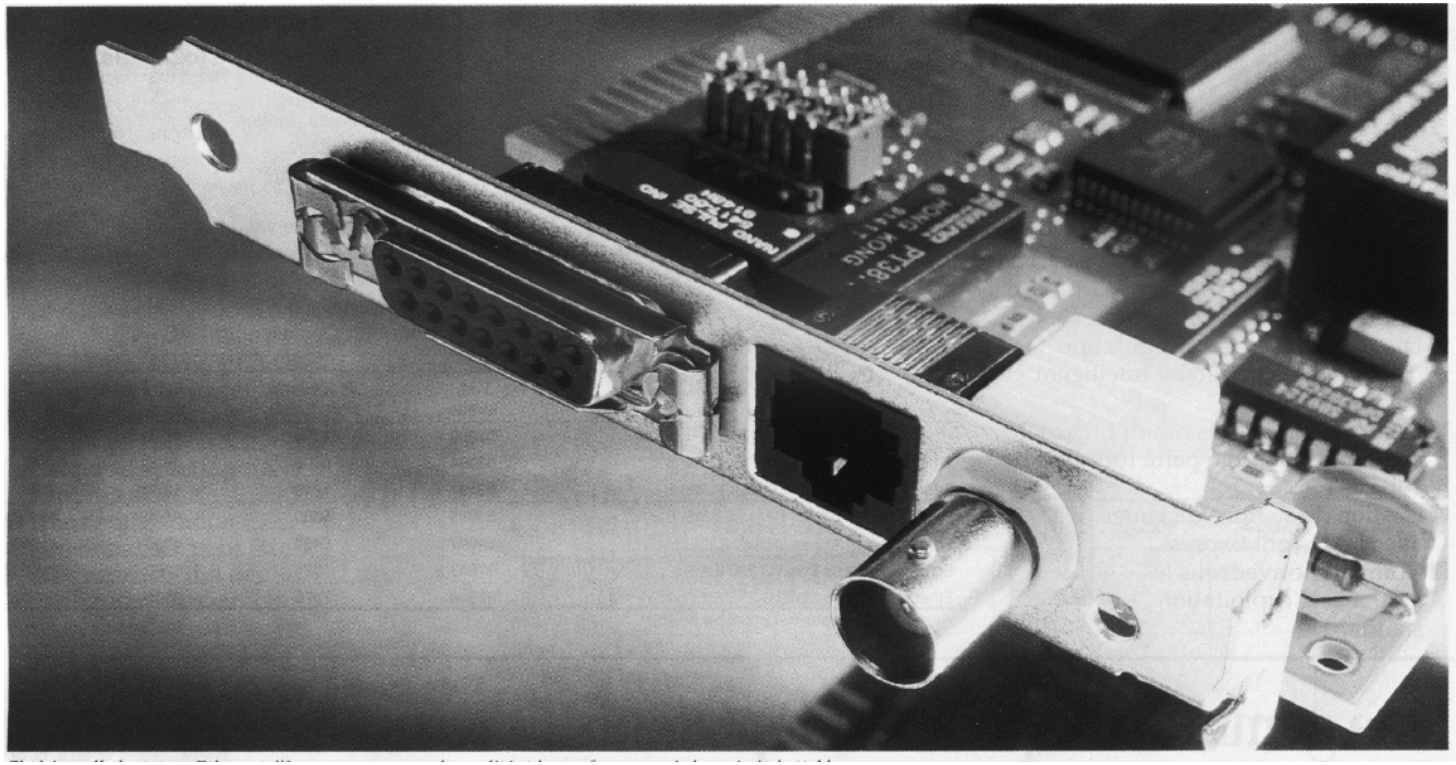
HUB : EMULATION D'UN BUS



MATÉRIEL

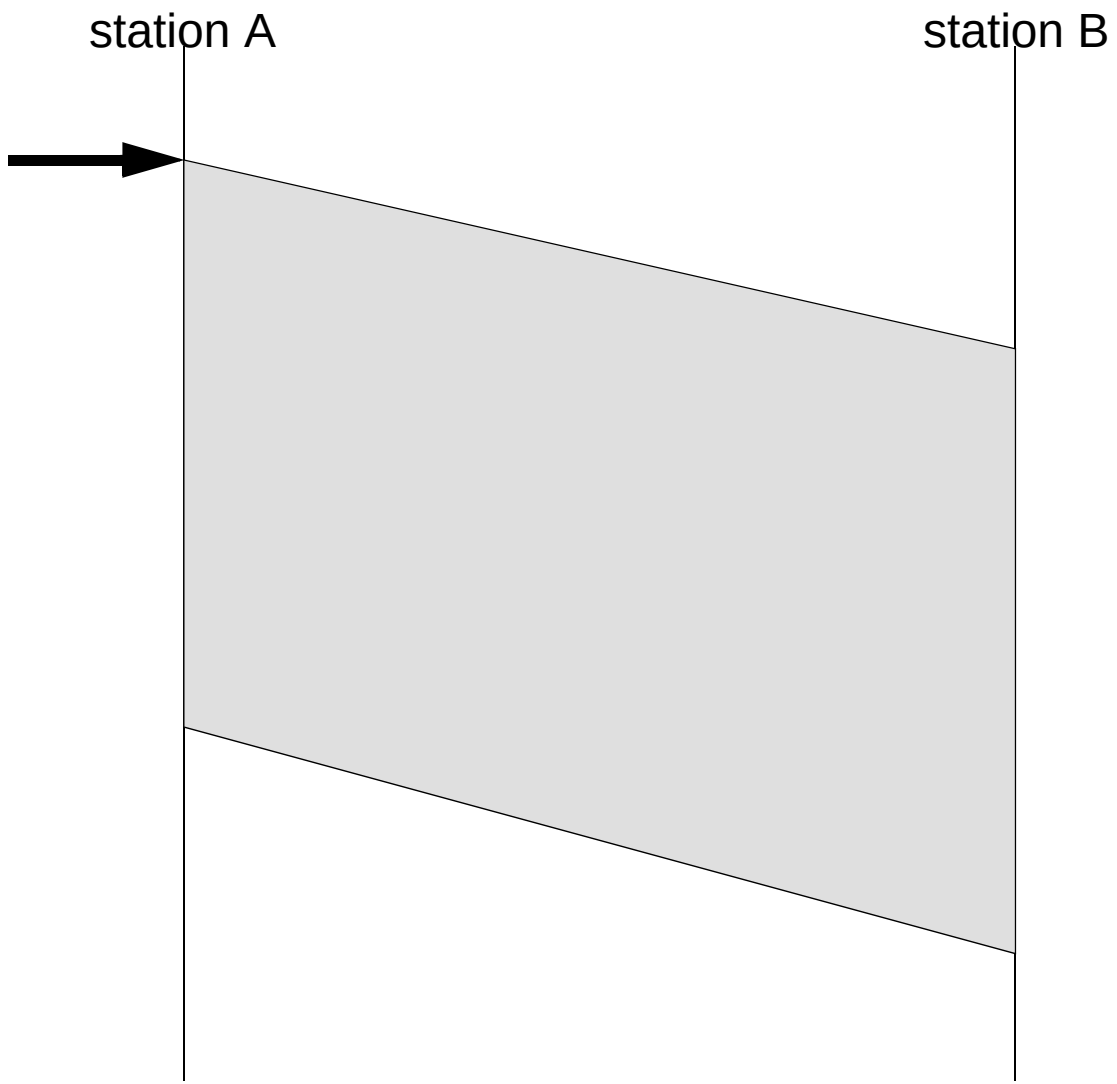

Power
Power
Col
Link Status
Activity
LittleHub 8
Coax
1MDIX
1MDI
2
3
4
5
6
7
8
Coax Partition
Coax Activity

SUR UN PC

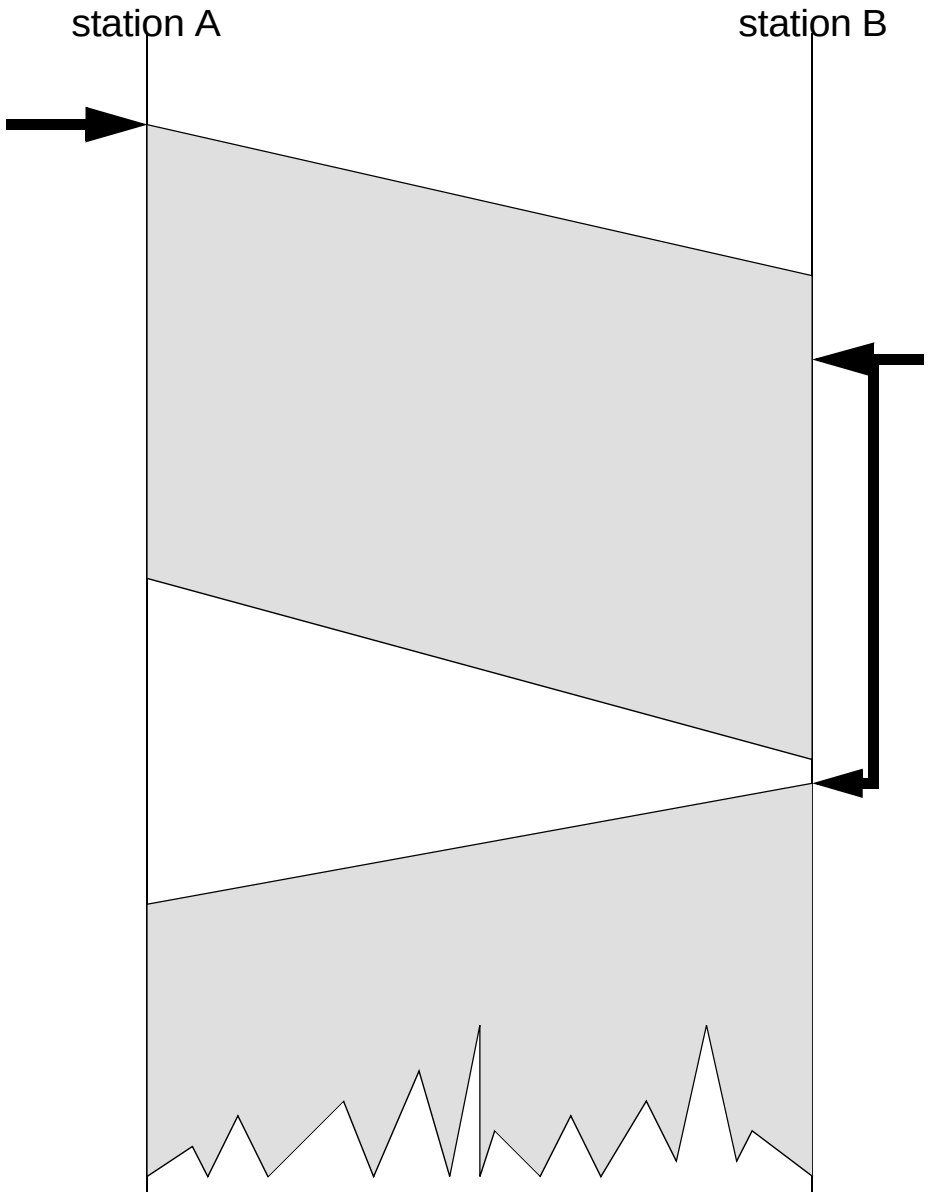


+ A quel type de réseau Ethernet correspondent les différentes prises de cette carte ?

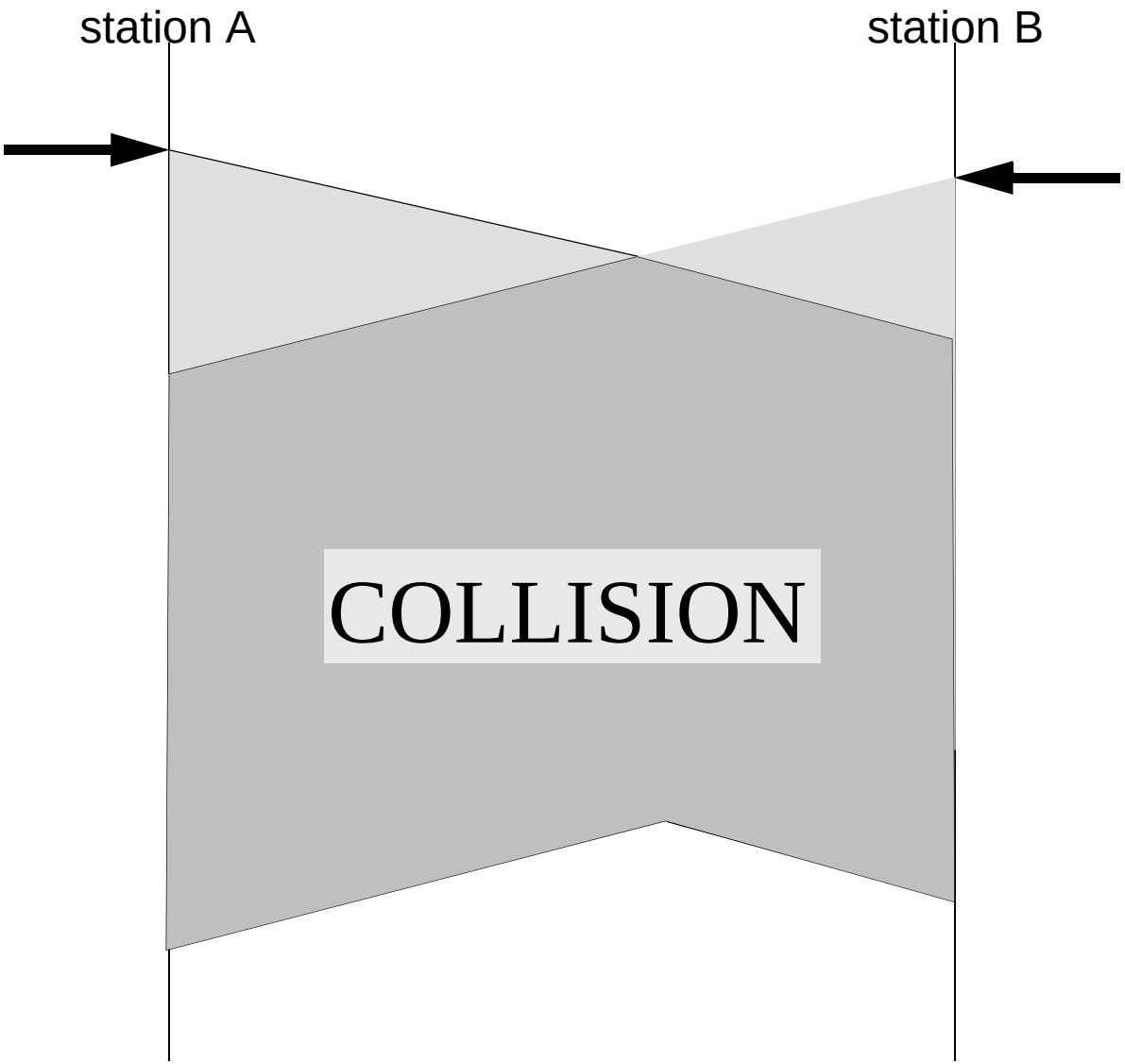
ALGORITHME DU **CSMA**/CD



ALGORITHME DU **CSMA**/CD



ALGORITHME DU **CSMA**/CD



LES COLLISIONS

- Il faut que TOUTES les stations soient dans le même état

⇒ La durée d'émission doit être d'au moins 2 fois la durée de propagation du signal

- Si la trame est trop courte, il faut ajouter des bits de bourrage.

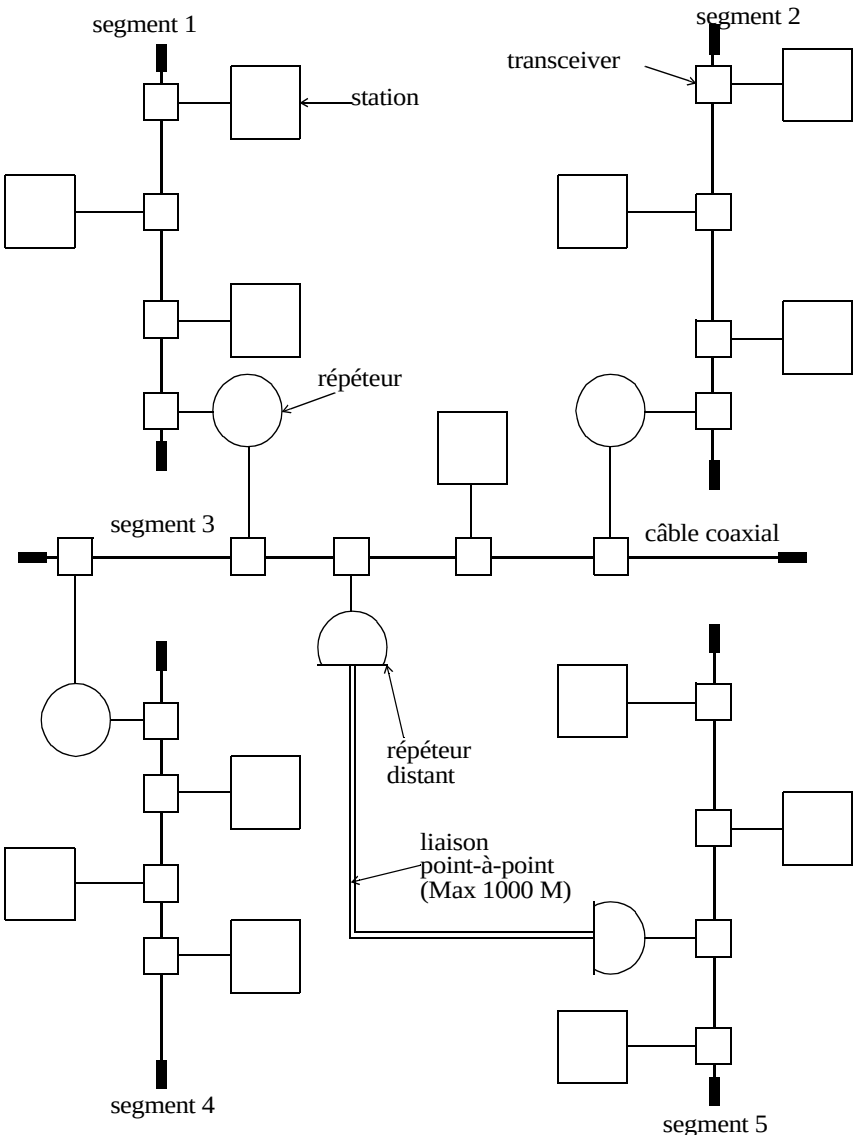
⇒ La topologie doit être limitée pour éviter des durées de propagation qui forcerait à allonger la longueur des trames

- La durée minimale d'émission est de 51.2 µs (soit 64 octets pour Ethernet à 10 Mbits/s).

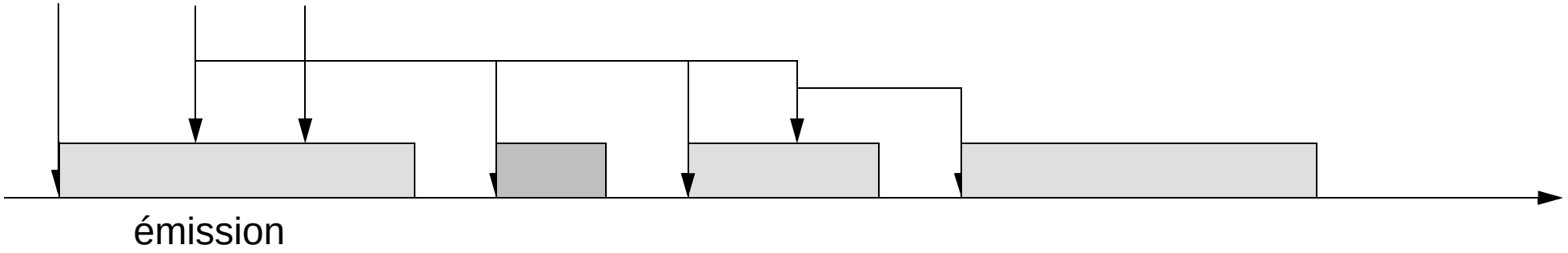
⇒ Les règles de câblage sont (pour le 10 base 5):

  - des segments de 500 mètres maximum,
  - traversée de 4 répéteurs.

TOPOLOGIE D'UN RÉSEAU ETHERNET



RÉSOLUTION DES COLLISIONS



- si collision attente de 0 ou 1 Tranche Canal
- si de nouveau collision : attente de 0, 1 2 ou 3 Tranche Canal
- si > 10 collisions : attente de 0 à 2^{10} Tranche Canal
- si > 16 collisions : arrêt de la tentative

\+ Pas de garantie d'émission (ni de borne)

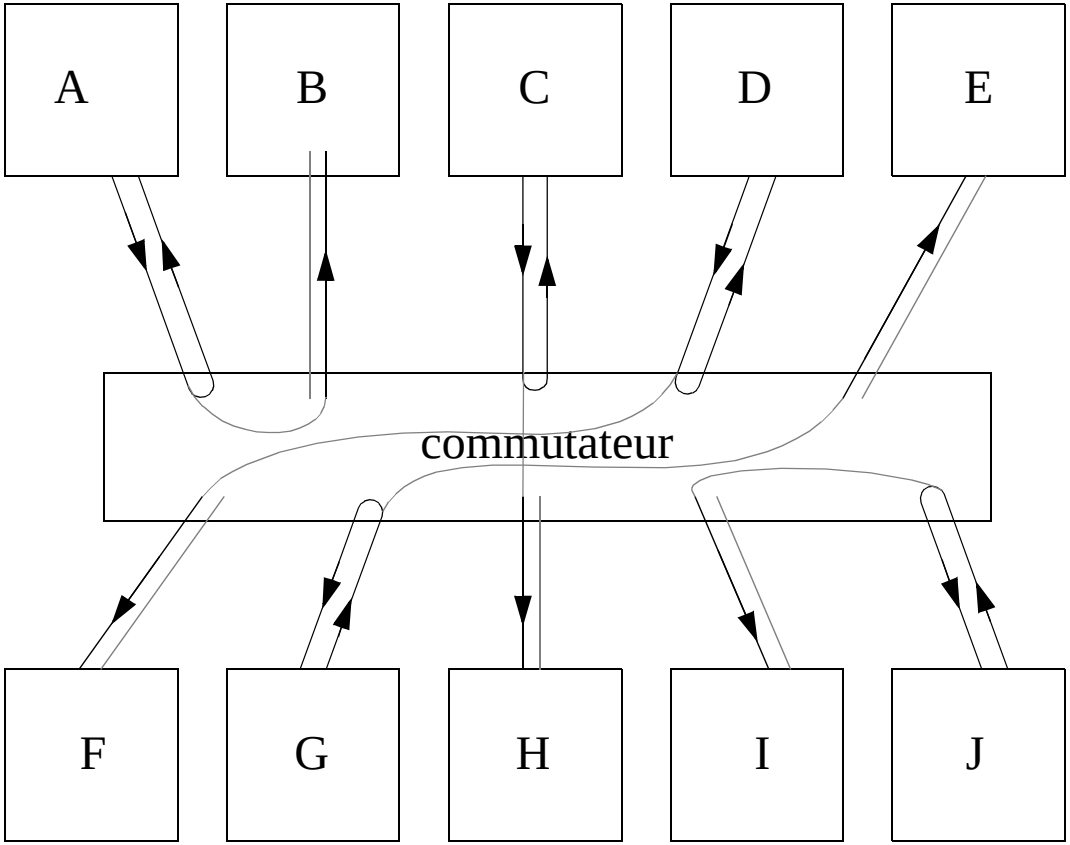
QUAND S'APPLIQUE LE CSMA/CD ?

- topologie en bus
- Emulation d'un bus avec un Hub
- Avec un commutateur et un équipement Half-duplex (utilisé pour faire du contrôle de flux)

\+ Le CSMA/CD n'est pas mis en œuvre dans le cas d'un commutateur et d'équipements terminaux full-duplex :

- suppression des contraintes de câblage
- augmentation des débits
- problème de contrôle de flux

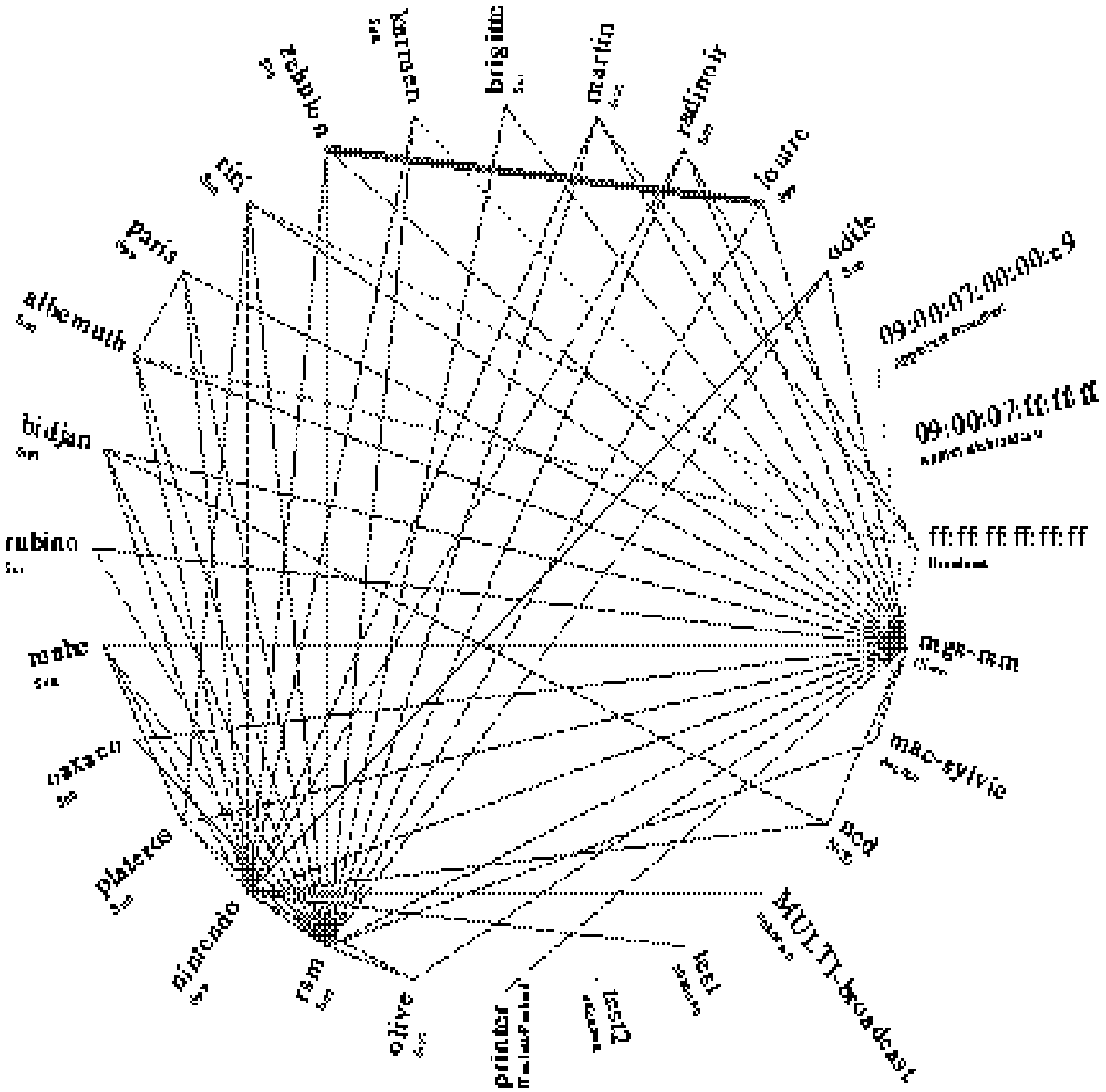
LES COMMUTATEURS (SWITCHS)



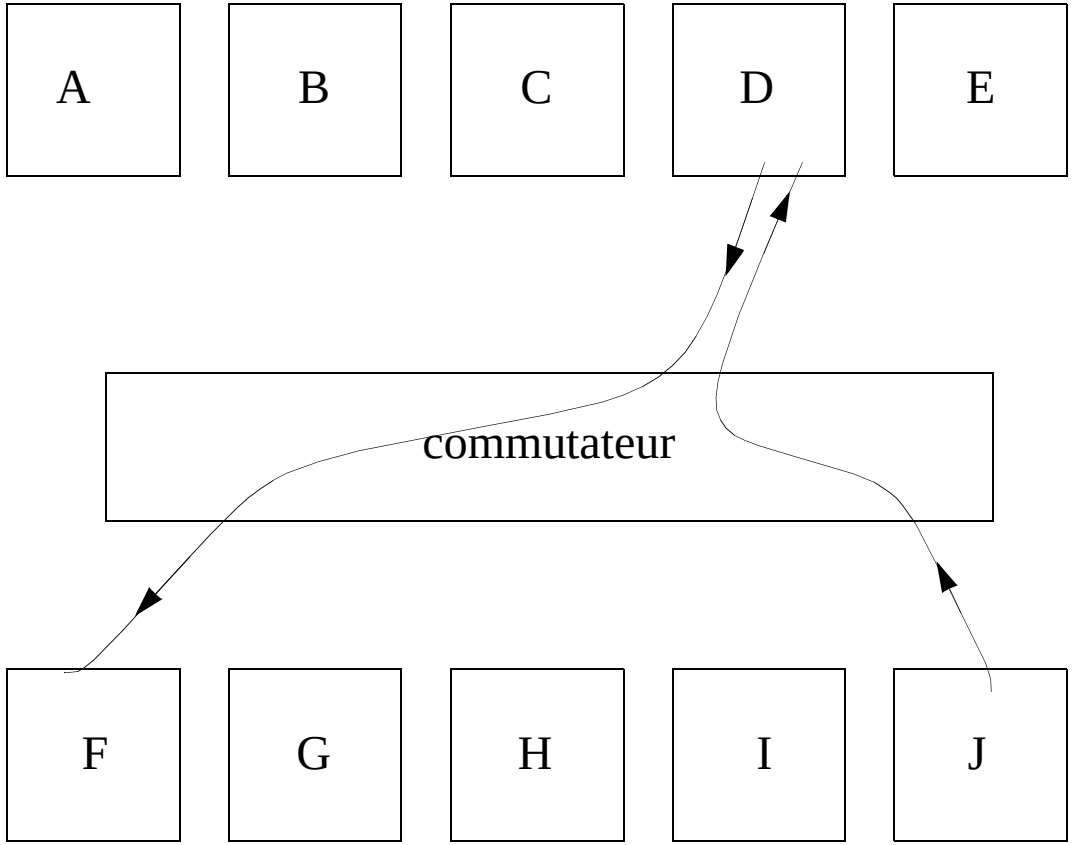
- à la volée
- stockage et retransmission

\+ Switch n'implique pas full duplex (les équipements doivent l'être)

EXEMPLE DE TRAFIC

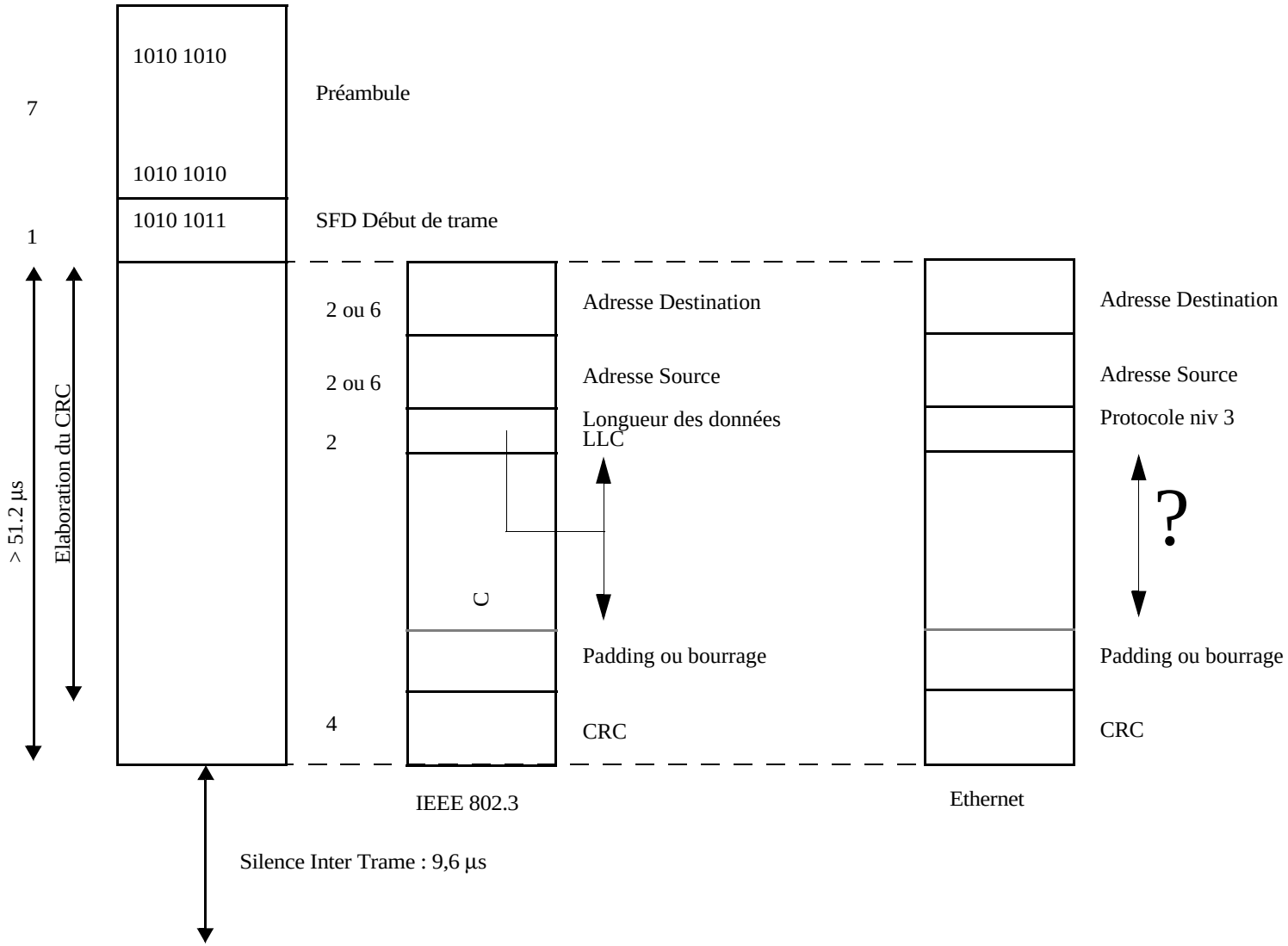


LES COMMUTATEURS (SWITCHS)

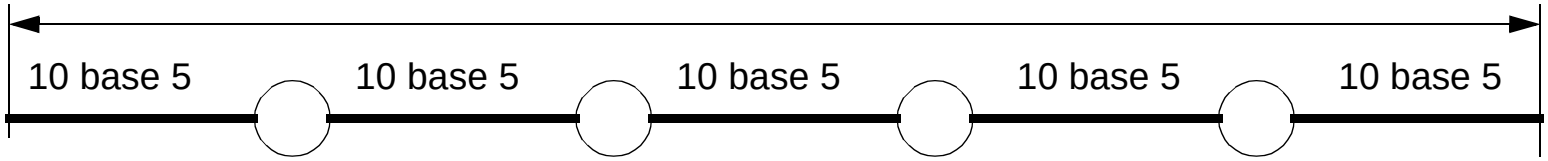


- Problème de contrôle de flux (le CSMA/CD peut être vu comme un algorithme de controle de flux)
- message PAUSE

FORMAT DES TRAMES



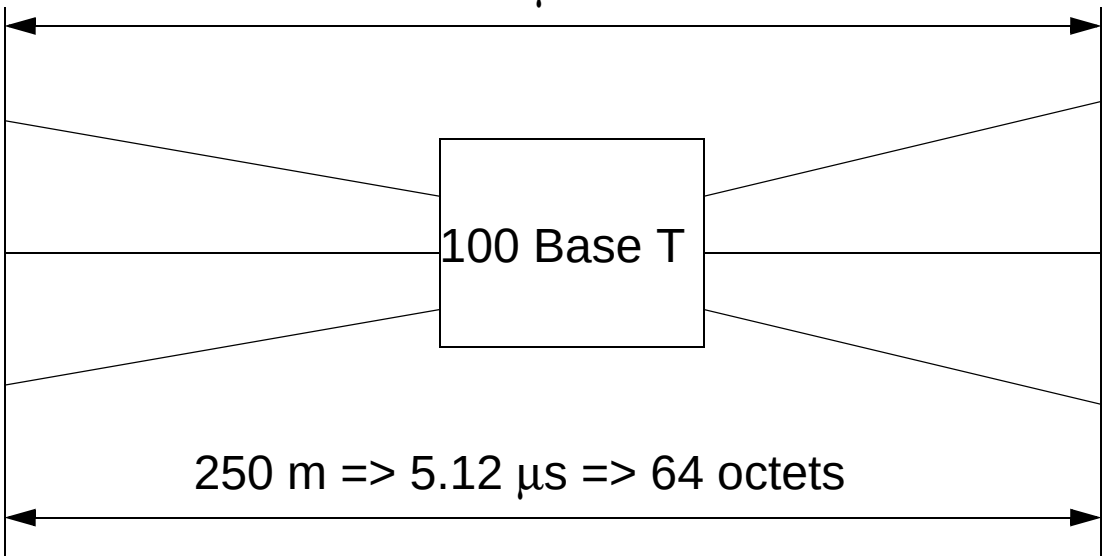
INFLUENCE DE LA VITESSE SUR LE FORMAT DES TRAMES



Solution 1

2.5 Km => 51.2 µs => 640 octets

Interconnexion ???

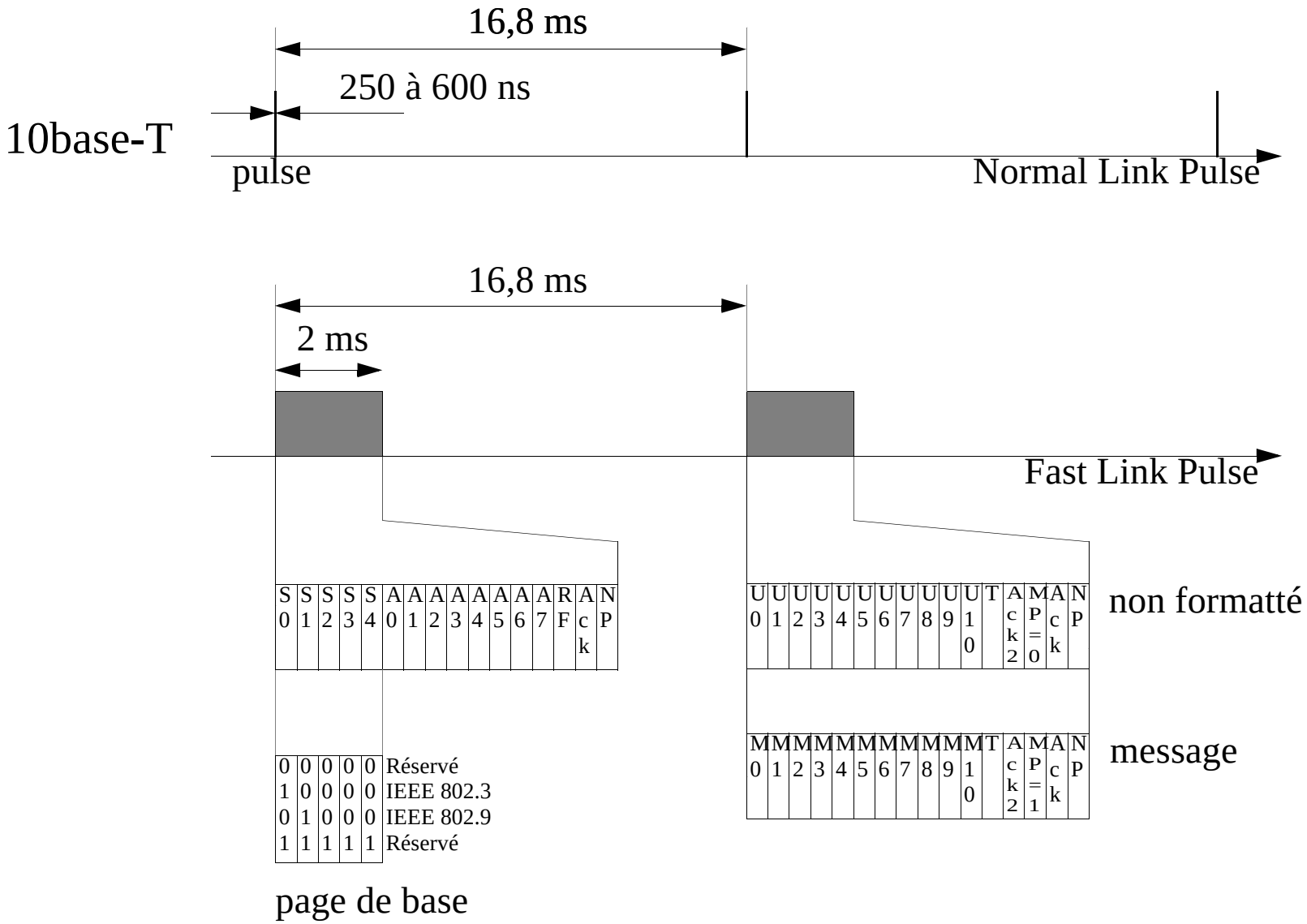


Solution 2

250 m => 5.12 µs => 64 octets

\+ Quelle distance pour Ethernet 1 Giga ?

AUTO-NÉGOCIATION



QUESTIONS

+ Quelle est la distance maximum peut couvrir un réseau 10 base 5 ?

+ Un hub est :
- un répéteur,
-un pont,
- un routeur,
- une passerelle ?

un switch est :

+ Quelle est la différence entre un switch et un pont ?

+ Un répéteur empêche la collision de se propager ?

+ Un commutateur empêche la collision de se propager ?

+ Un pont empêche la collision de se propager ?

+ Un pont peut créer des collisions ?

+ Les contraintes de câblage s'appliquent aussi aux réseaux commutés ?

+ Comment réagit un commutateur sur lequel est connecté des équipements à 10 et 100 Mbit/s

LES PONTS TRANSPARENTS

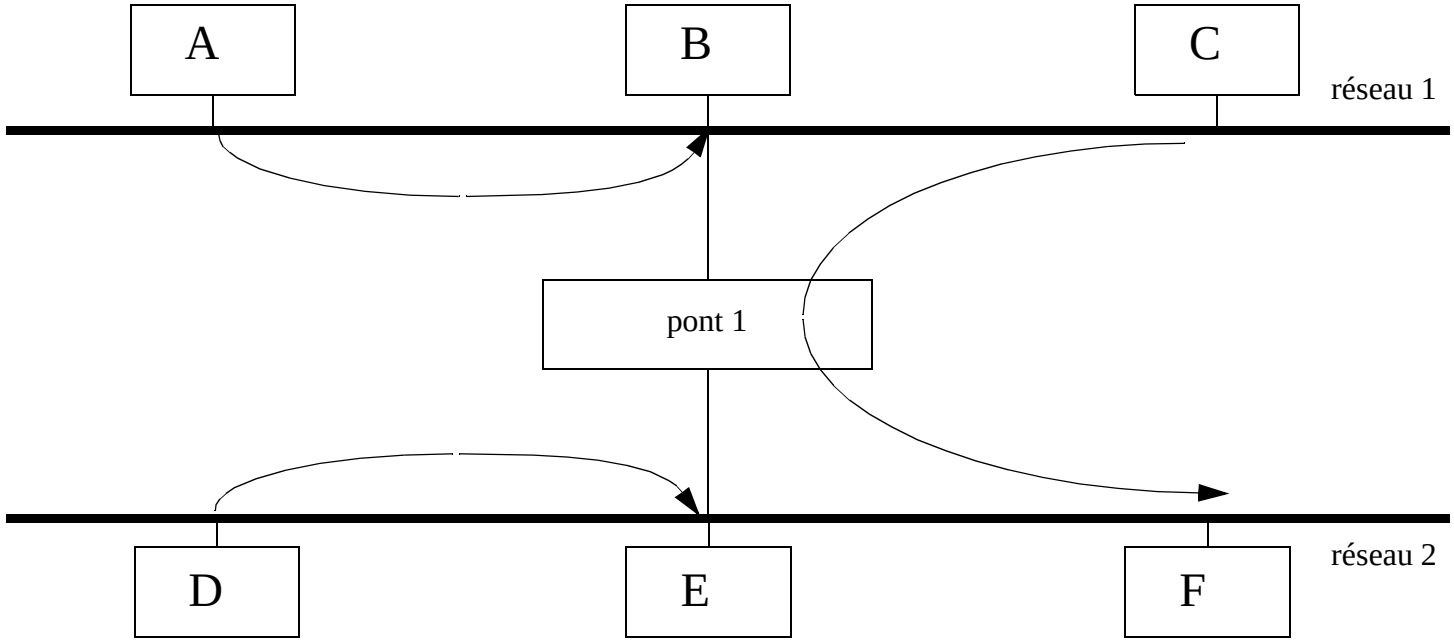
- Principe de fonctionnement :
  - Le pont écoute toute l'activité en mode *Promiscuous*.
  - Il stocke dans sa mémoire les trames.
  - Il retransmet vers le (ou les) autre(s) sous-réseau(x) les messages stockés .
- Un pont n'a pas besoin d'adresse MAC pour fonctionner
- Un pont n'a pas besoin d'être configuré.

QUESTIONS

+ Quelle adresse est mise dans le champ source des trames recopiées ?

+ Quels protocoles de niveau 3 peuvent être utilisés ?

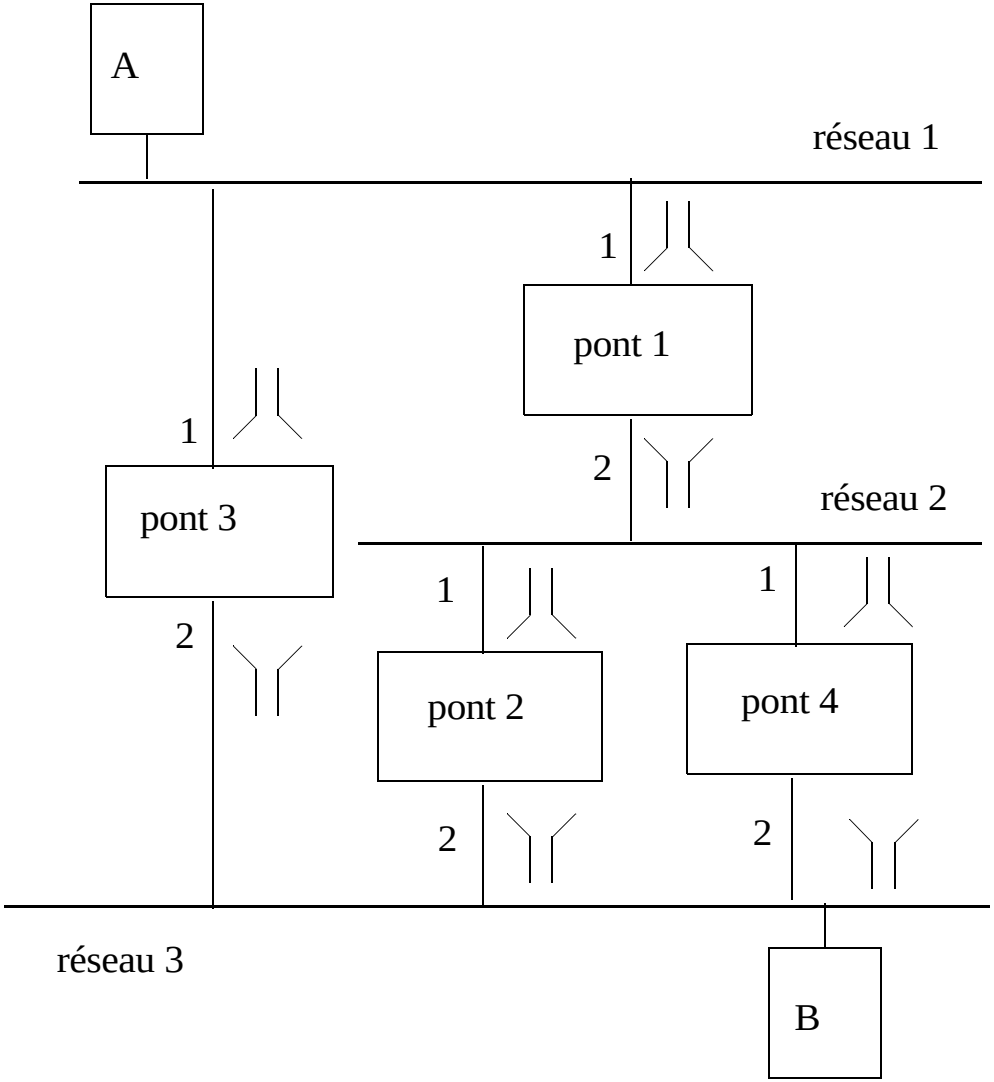
PONTS FILTRANTS



- Les ponts permettent de réduire le trafic global du réseau
- Toujours pas besoin d'adresse MAC ou d'administration

+ Est-ce toujours valable avec 3 réseaux et 2 ponts ?

CAS COMPLEXE



+ Que se passe-t-il quand A émet un message vers B ?

CAS COMPLEXE (SUITE)

Quand plusieurs routes sont possibles, il y a duplication des messages et inondation du réseau

+ IL faut absolument éviter les boucles lors de l'interconnexion des réseaux.

+ Mais :

  - la redondance doit toujours être possible
  - possibilité d'erreur

+ Mettre en œuvre dans les ponts un protocole qui permettra de trouver un arbre couvrant (*Spanning Tree*).

+ Les ponts doivent échanger des messages

ALGORITHME DU SPANNING TREE

Les ponts vont échanger des messages contenant :

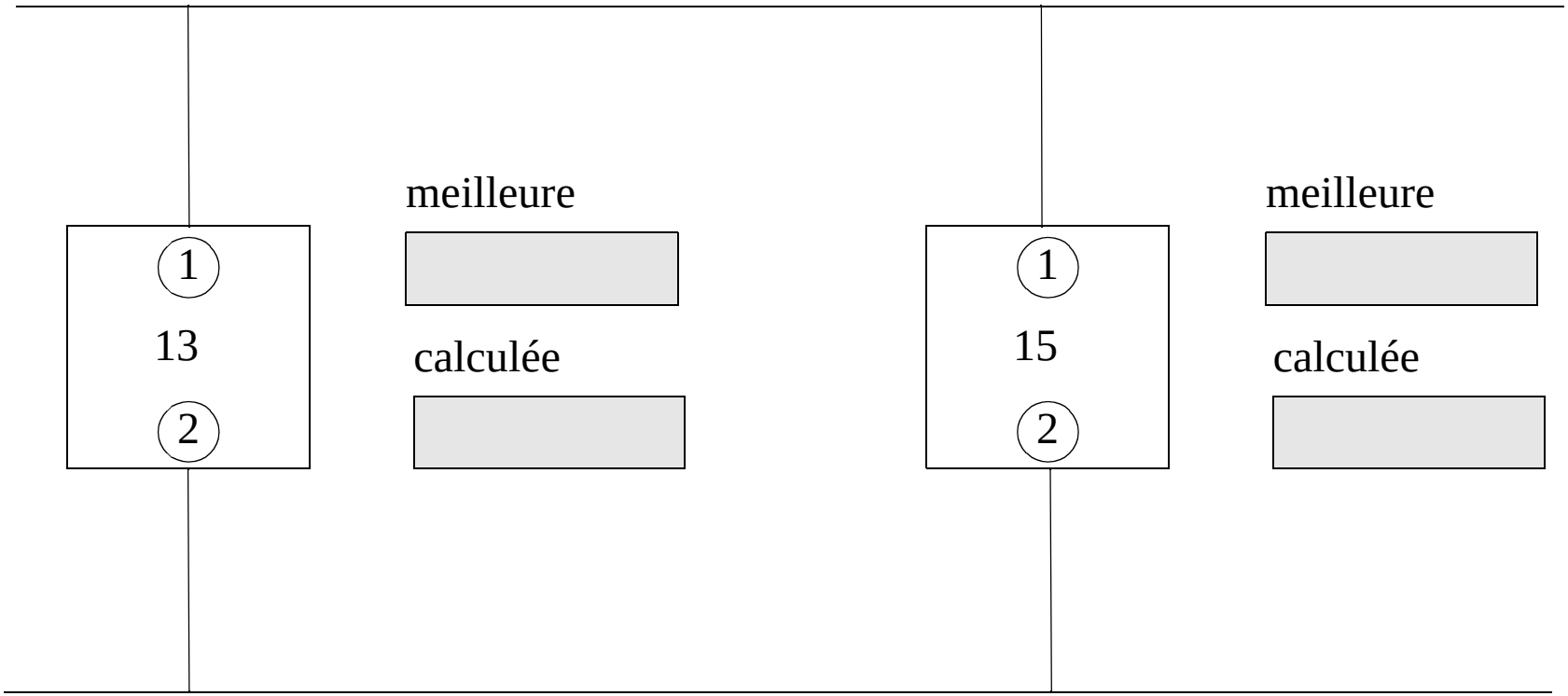
L'identité supposée de la racine. A l'initialisation, ils se supposent racine.
Le coût supposé de la liaison. Pour un pont racine, ce coût est nul.
L'identité de l'émetteur.
Le numéro du port sur lequel le message est émis.

L'algorithme pour chaque pont est le suivant :

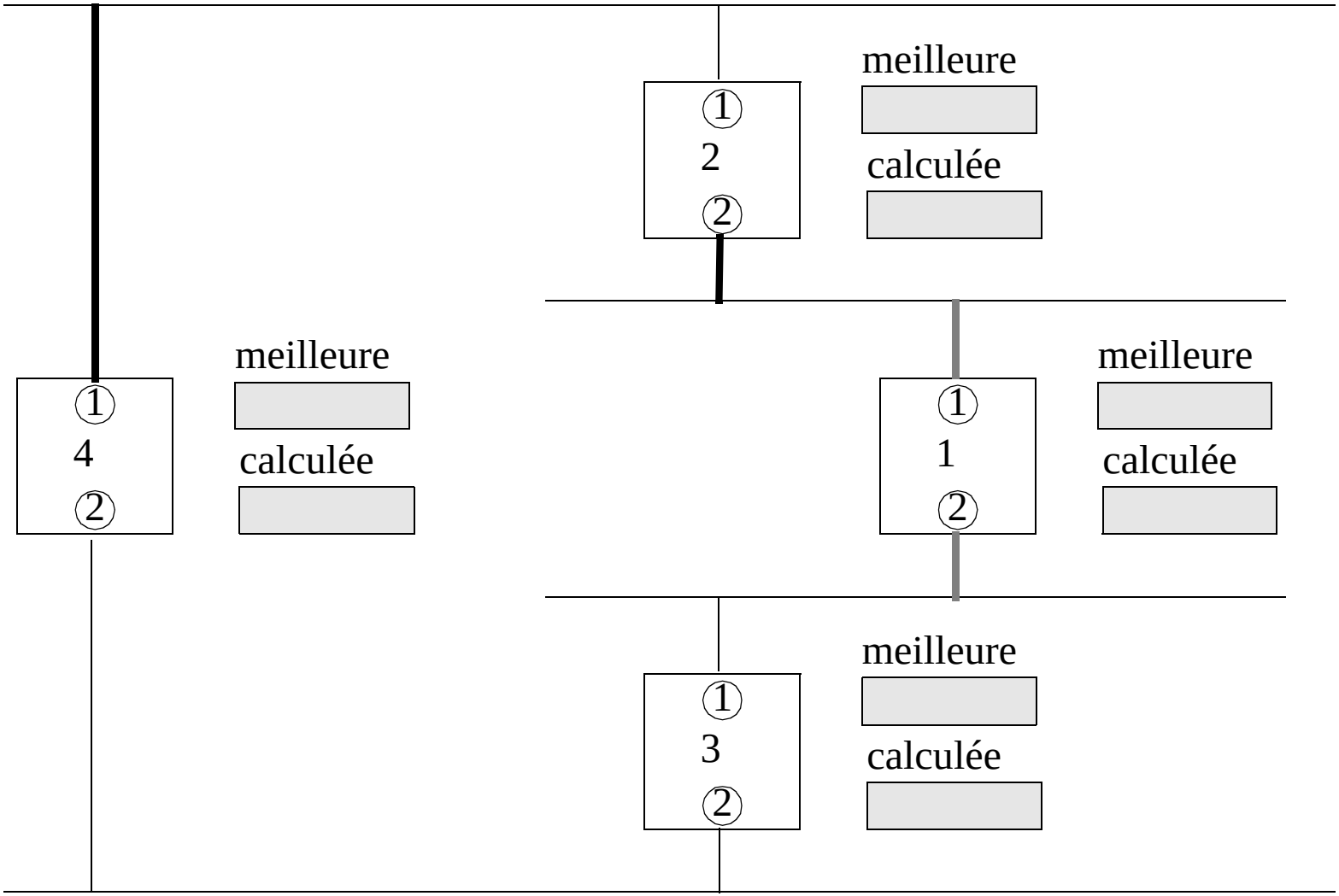
- Recherche du meilleur message sur ses ports.
  Si une des configurations est meilleure que la configuration :
  - Cette voie devient le chemin pour la racine.
    Une nouvelle configuration est calculée. Le coût est augmenté de 1.
  - Les ports qui ont une configuration meilleure que celle nouvellement calculées sont désactivés Les autres ports font partis du *Spanning Tree*.
  - Cette configuration sera émise sur les ports autres que celui qui mène vers la racine

  Si aucun message n'est meilleur que celui émis par le pont, celui-ci se considère comme racine.

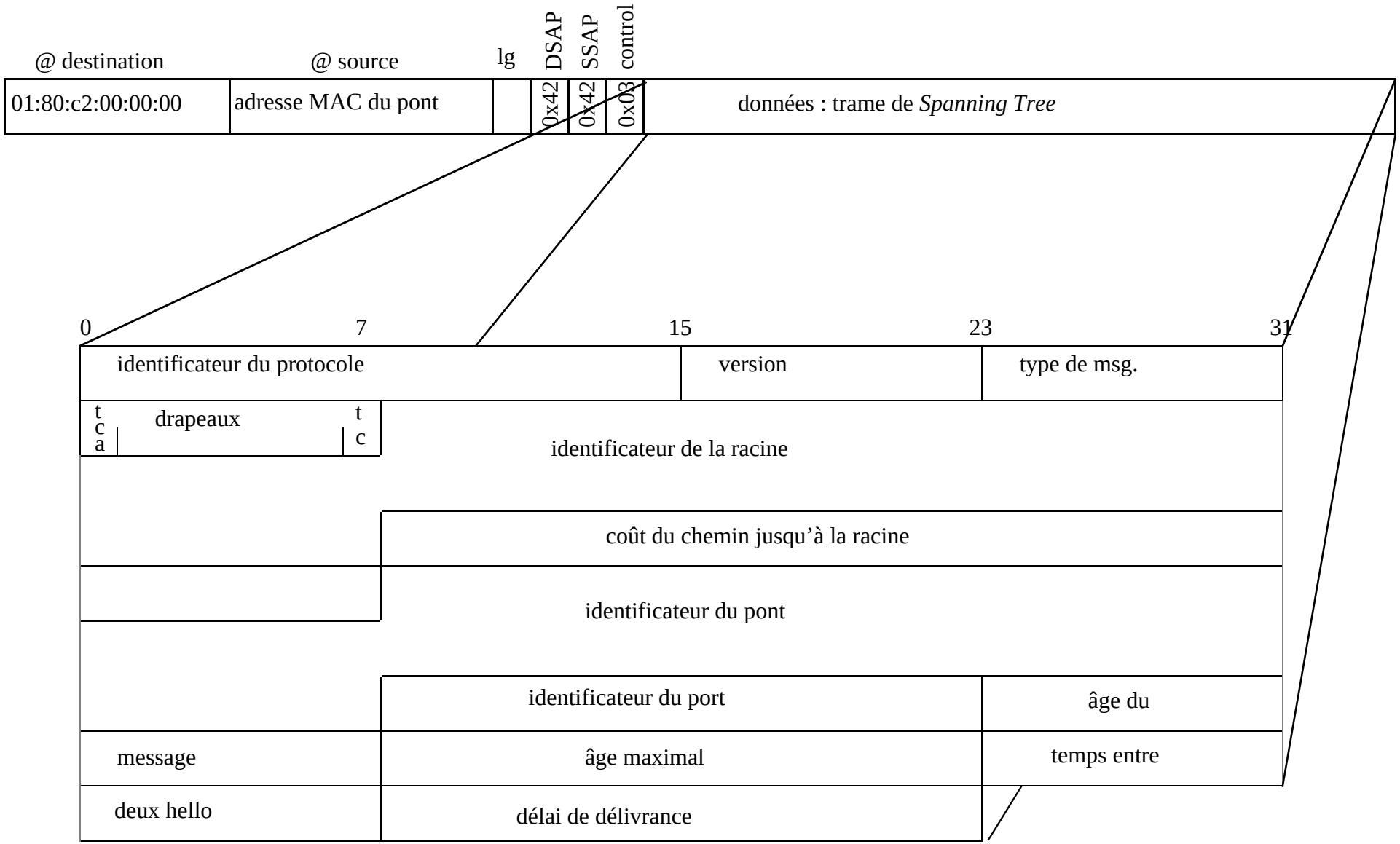
EXEMPLE DE DÉROULEMENT DE L'ALGORITHME DU SPANNING TREE



EXEMPLE DE DÉROULEMENT DE L'ALGORITHME DU SPANNING TREE (SUITE)



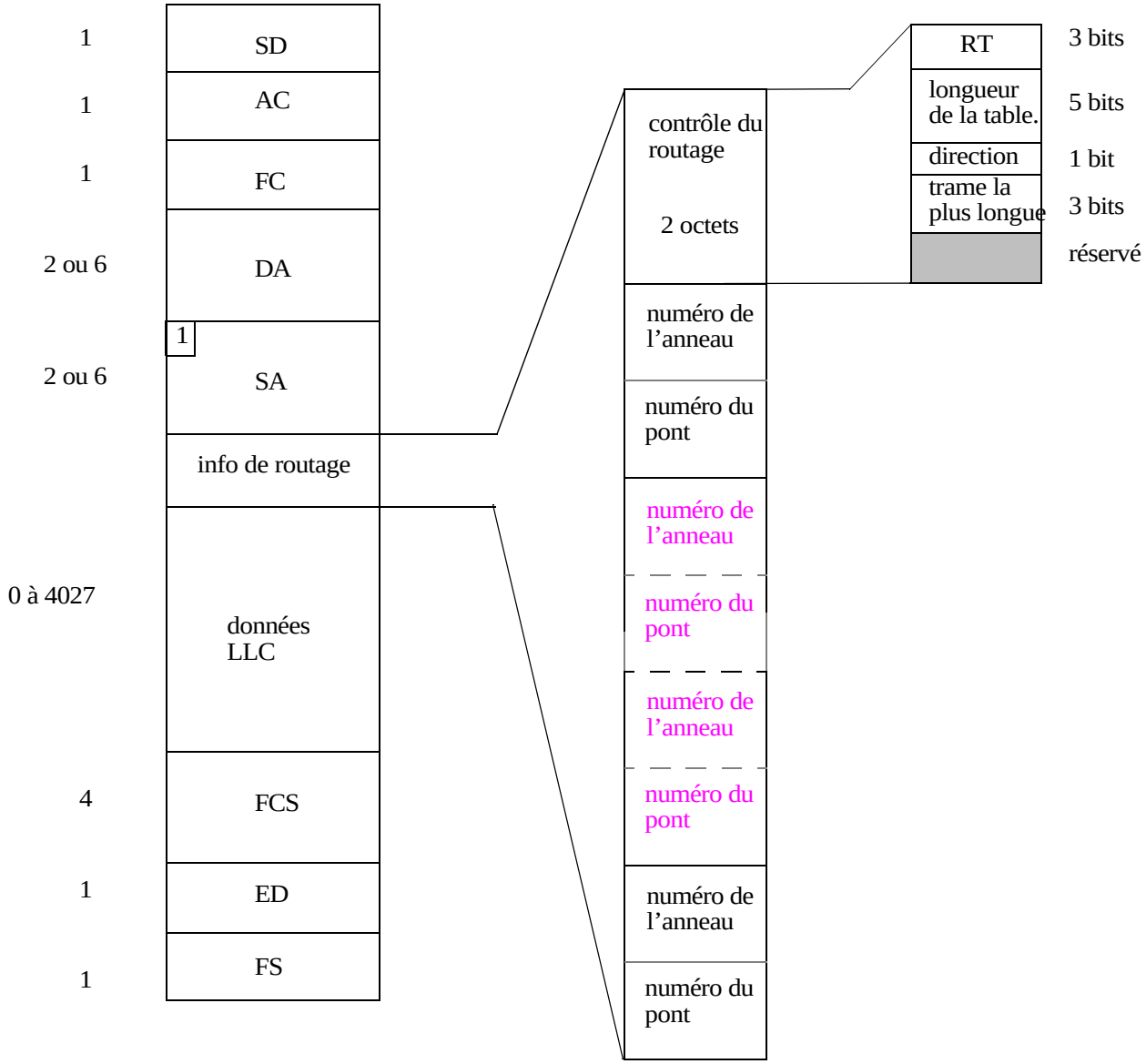
FORMAT DES TRAMES



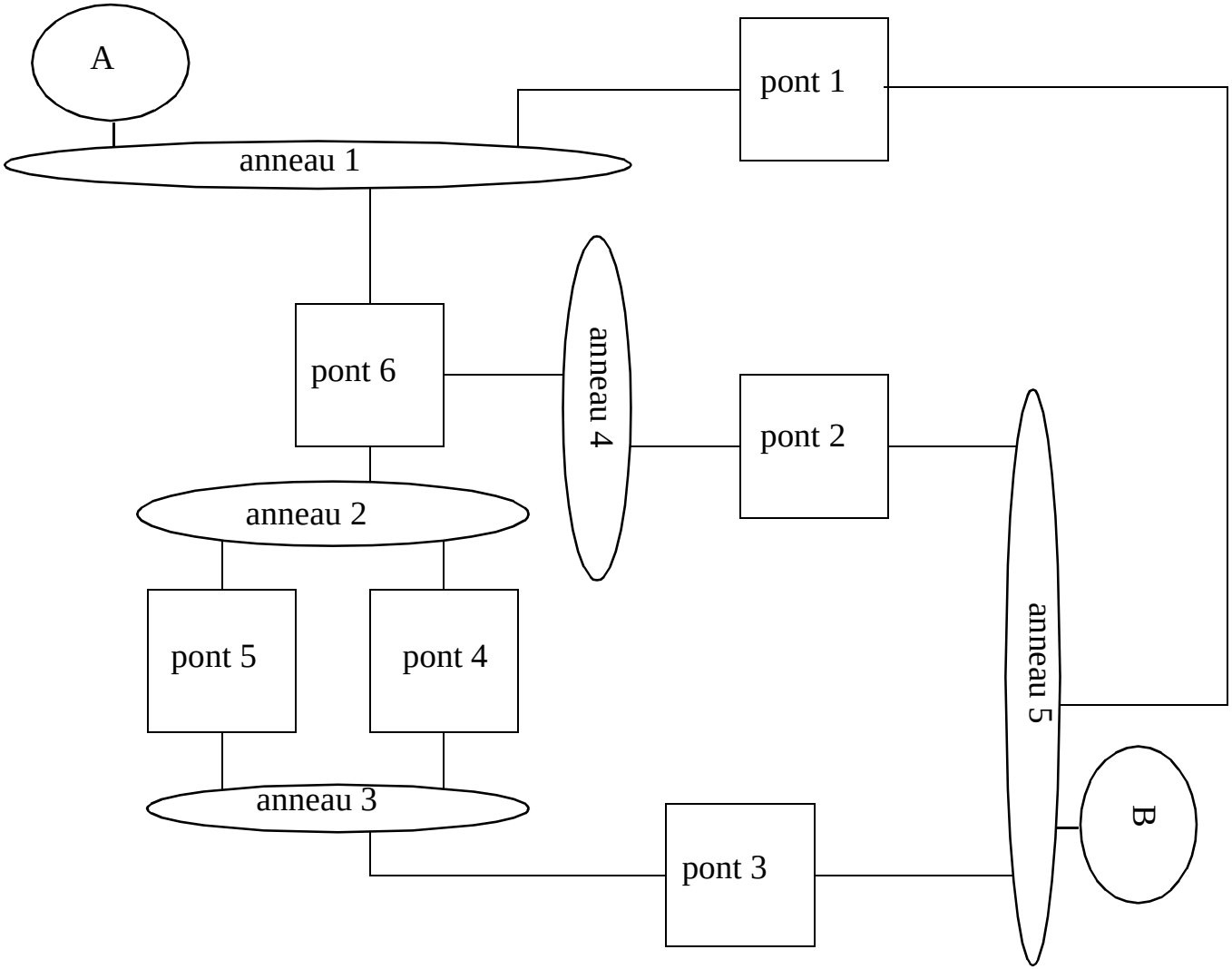
ALGORITHME DU SOURCE ROUTING

- L'intelligence est placée dans les stations.
- Les ponts sont plus simples.
- Compatible avec le Spanning Tree.
- En théorie le Source Routing est meilleur que le Spanning Tree :
  - Il existe des chemins plus direct que ceux trouvés par le Sapnning Tree.
  - Tout le trafic va passer par les mêmes ponts
  - Certains ponts sont inactifs.
- En pratique utilisé uniquement dans les anneaux à jeton

FORMAT DES TRAMES



SOURCE ROUTING PUR.



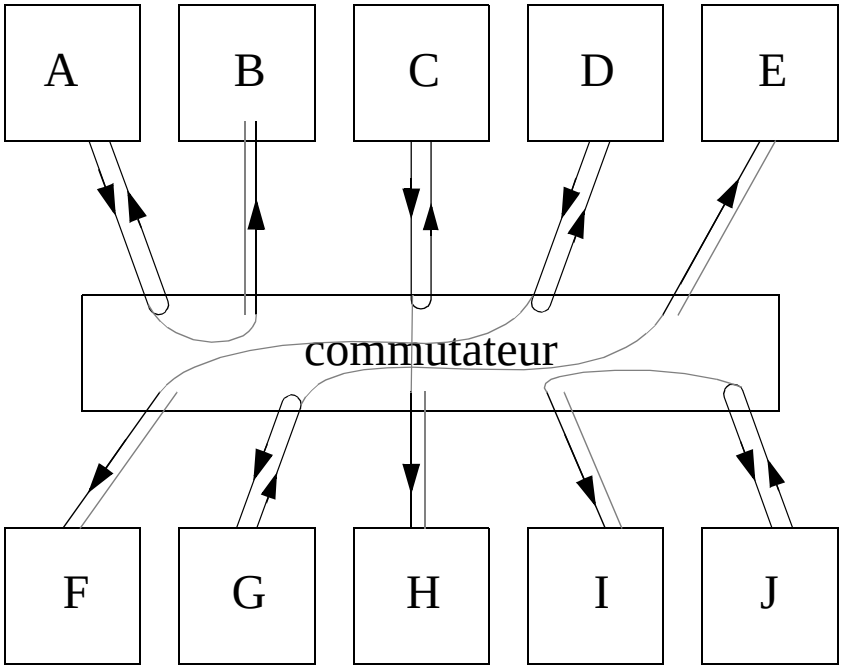
LES RÉSEAUX VIRTUELS

- Garder les avantages des réseaux de niveau 2 :
  - autoconfiguration (adresses, pontage (Spanning Tree), ...)
  - rapidité de transmission grâce à la commutation

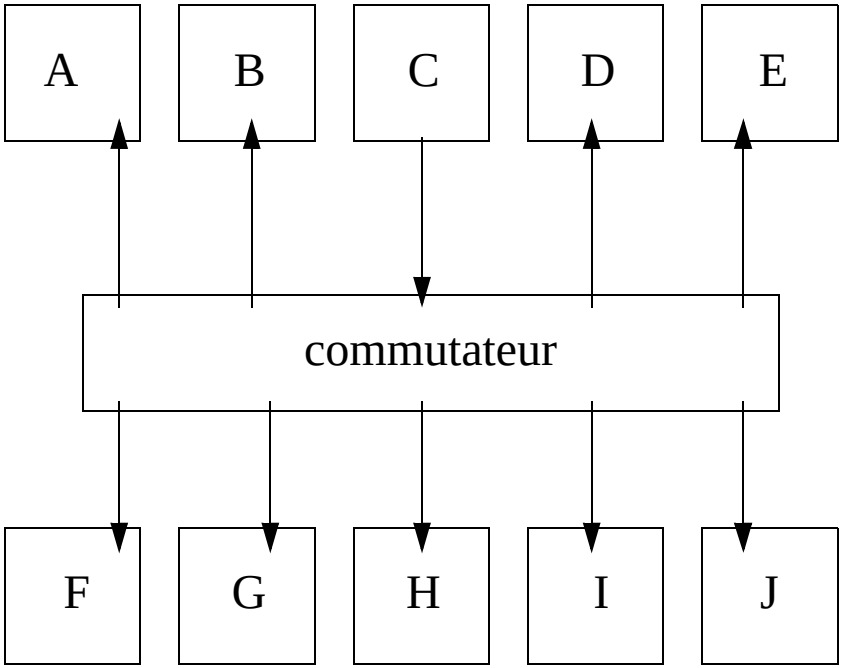
- Offrir des fonctionnalités de réseaux de niveau 3 :
  - extensibilité
  - utilisation sur de grandes distances
  - filtrage du trafic

\+ Réduire l'espace de diffusion

PROBLÈME SUR UN COMMUTATEUR



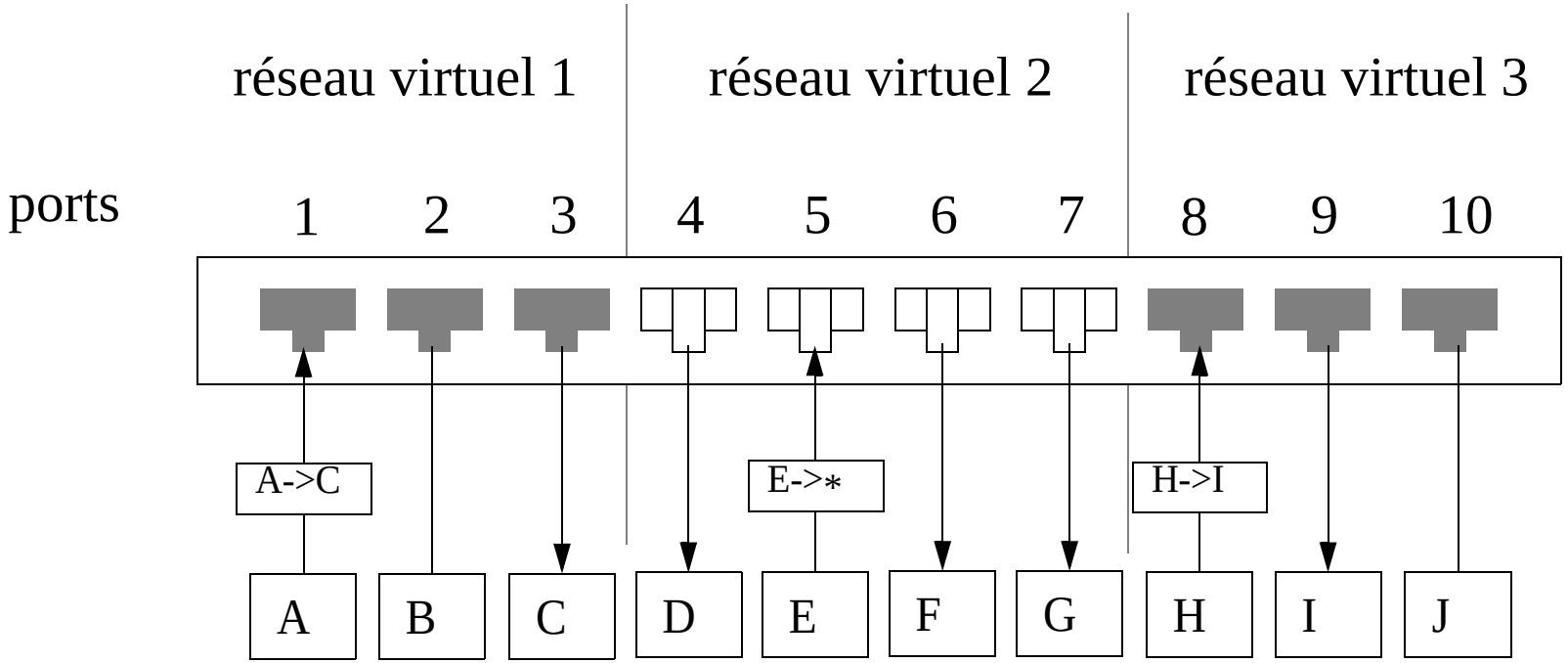
trame point-à-point



trame de broadcast ou de multicast

+ Même problème avec un réseau construit avec des demi-ponts et des LL.

RÉSEAU VIRTUEL = RÉDUIRE L'ESPACE DE DIFFUSION

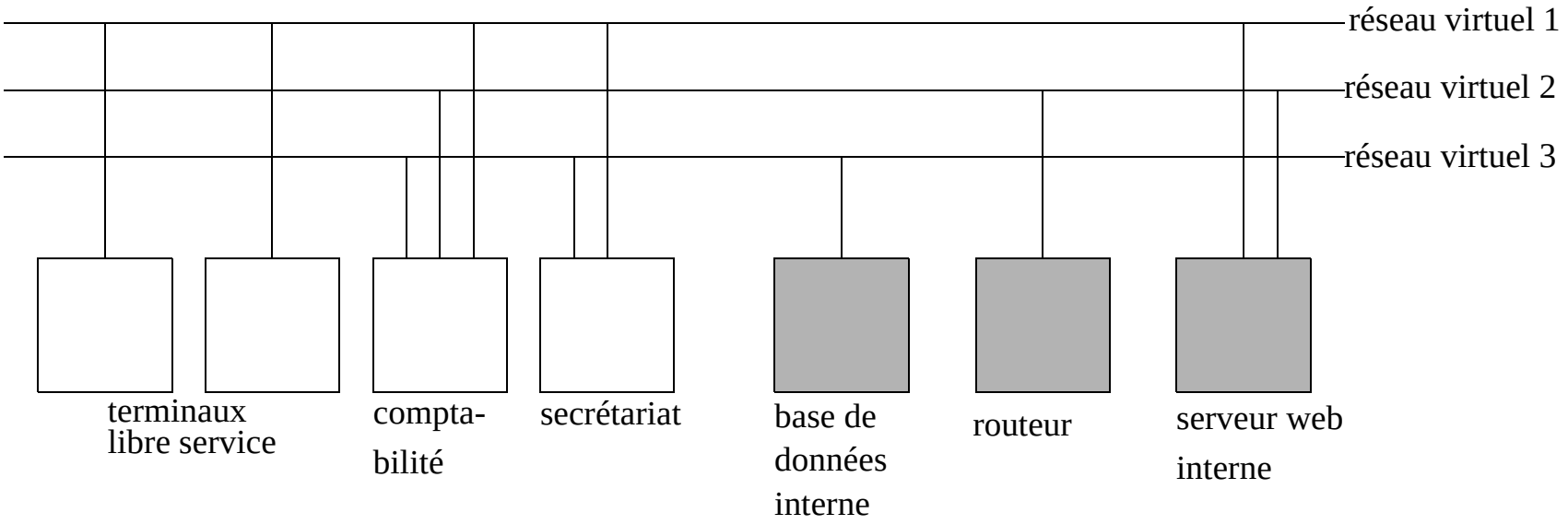


- La communication entre les réseaux virtuels est impossible

=>Utilisation d'un routeur

- Les VLAN fonctionnent sur les ponts et les commutateurs

SÉCURITÉ



+ Un seul réseau physique

COMMENT DÉFINIR L'APPARTENANCE À UN VLAN ?

- Par port :
  - simplicité d'administration
  - très sécurisé
  - manque de souplesse pour les utilisateurs nomades

- Par adresse MAC :
  - souplesse pour les nomades
  - manque de sécurité
  - difficulté d'administration (il faut manipuler les adresses MAC)
  - problèmes avec les protocoles de niveau 3 (par exemple : sous-réseau IP)
    Doit être lié avec des protocoles de configuration automatique (DHCP)

COMMENT DÉFINIR L'APPARTENANCE À UN VLAN ? (SUITE)

- Par protocole de niveau 3 :
  - Permet de limiter les trames en diffusion (Novell : RIP; SAP; IP : ARP)
- Par adresse de niveau 3 :
  - Plus de souplesse pour les nomades
  - Un réseau virtuel par réseau IP
  - Plus lent car analyse des trames/paquets plus complexe
- Combinaison de ces différents modes

COMMENT DÉFINIR DES VLAN SUR PLUSIEURS ÉQUIPEMENTS ?

- Enorme problème de normalisation, solutions constructeurs

- Arrivée prochaine de normes IEEE :
  - IEEE 802.1p (supplément à IEEE 802.1D) + IEEE 802.1Q : numérotation des VLAN, partage d'information entre les équipement.

- Utilisation intensive de plateformes d'administration.

APPARTENANCE À UN VLAN

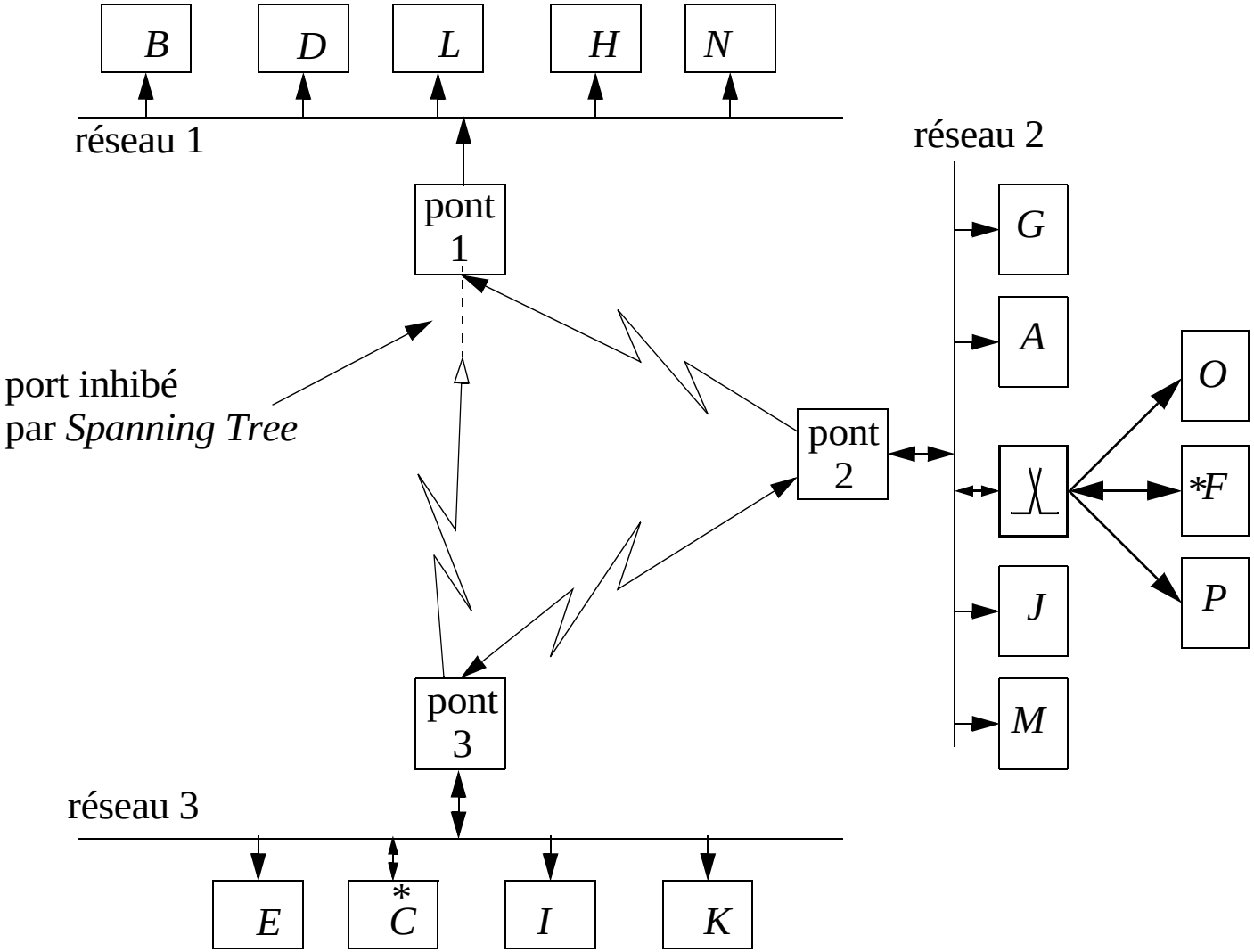
- La machine 12-34-45-67-89 appartient au VLAN 1

- Comment faire pour partager cette information à tous les VLAN ?
  - plateforme d'administration :

    - découvertes des équipements d'interconnexion du réseau
    - découvertes des stations (adresses MAC) utilisées sur le réseau
    - attribution d'un VLAN à une adresse MAC (outil graphique)
    - envoi de cette information à tous les équipements d'interconnexion

  - Protocole d'inondation des équipements d'interconnexion :
    - GVRP : GARP VLAN Registration Protocol
    - GARP : Generic Attribute Registration Protocol

GARP



3 états des stations :

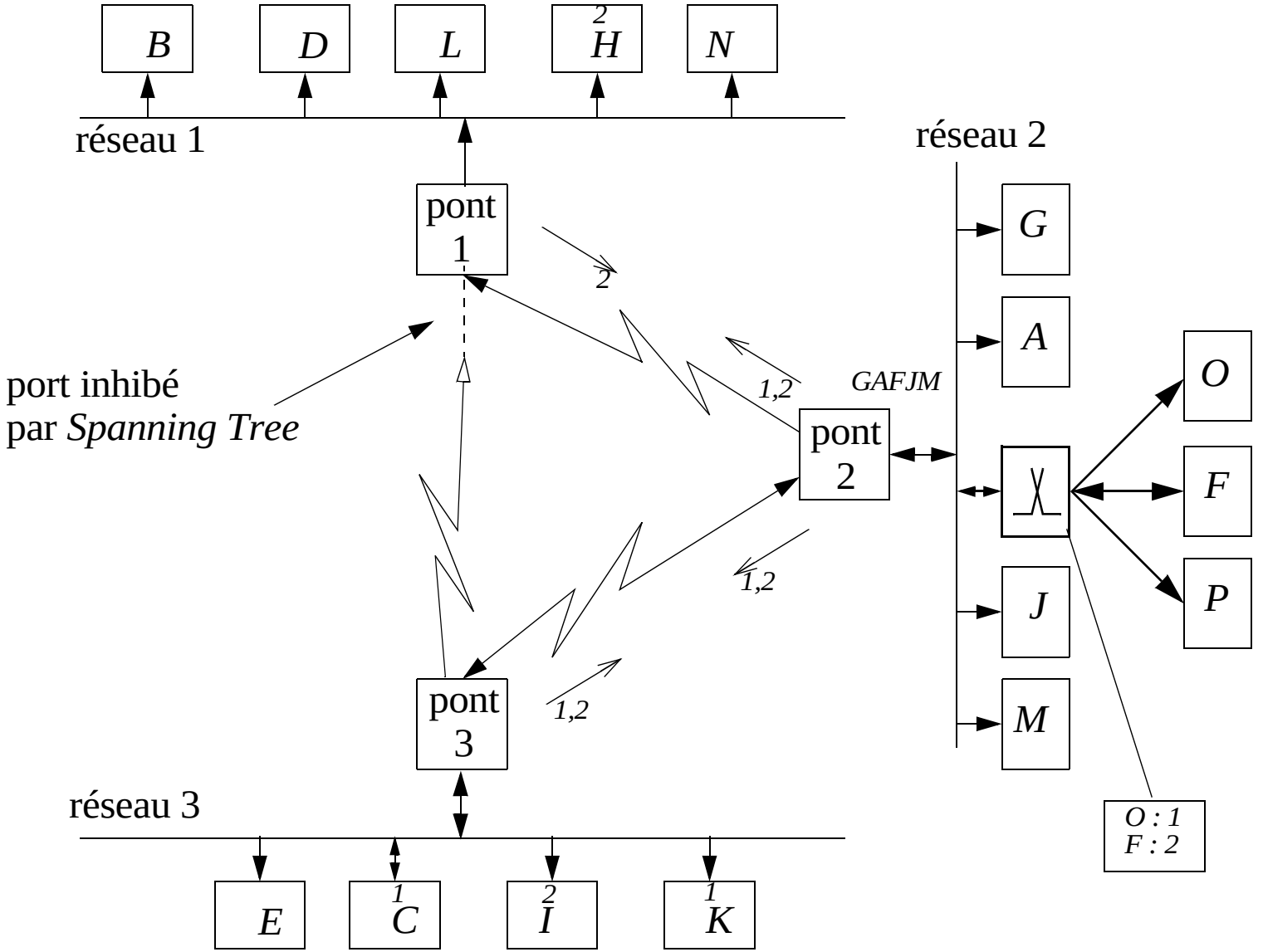
- très anxieuse
- anxieuse
- tranquille

Plusieurs rôles :

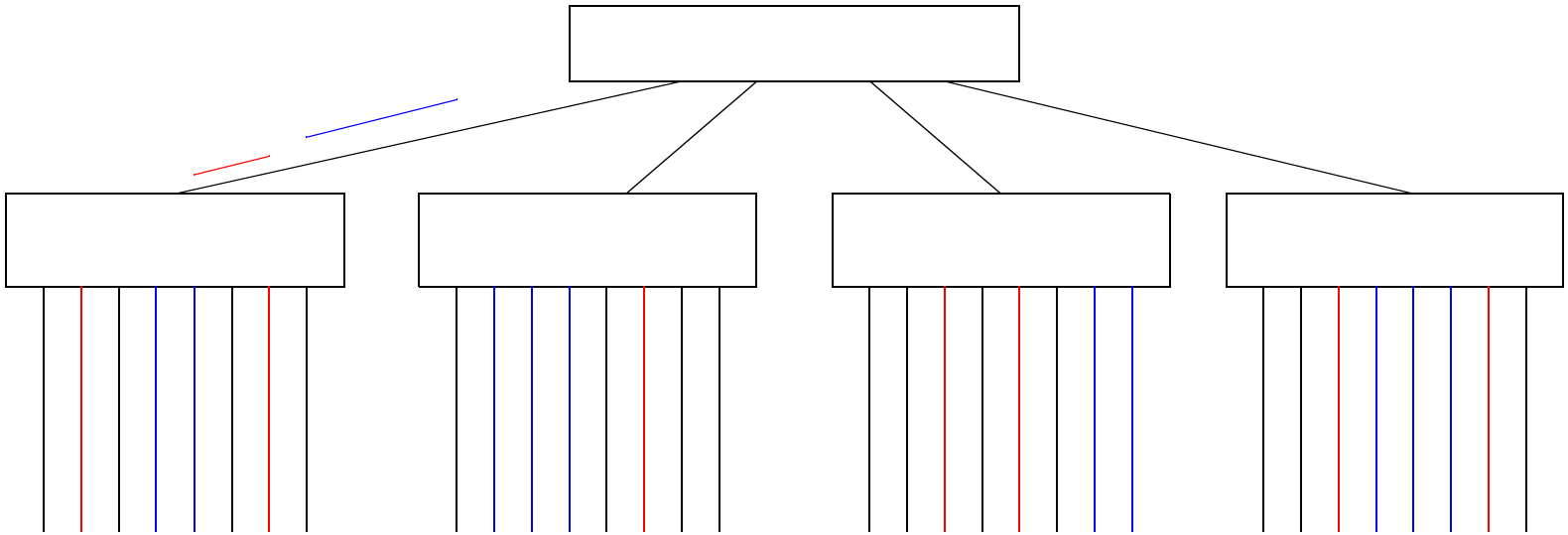
- Observateur
- Membre :
  - actif
  - passif

Périodiquement reconstitution du groupe.

GVRP



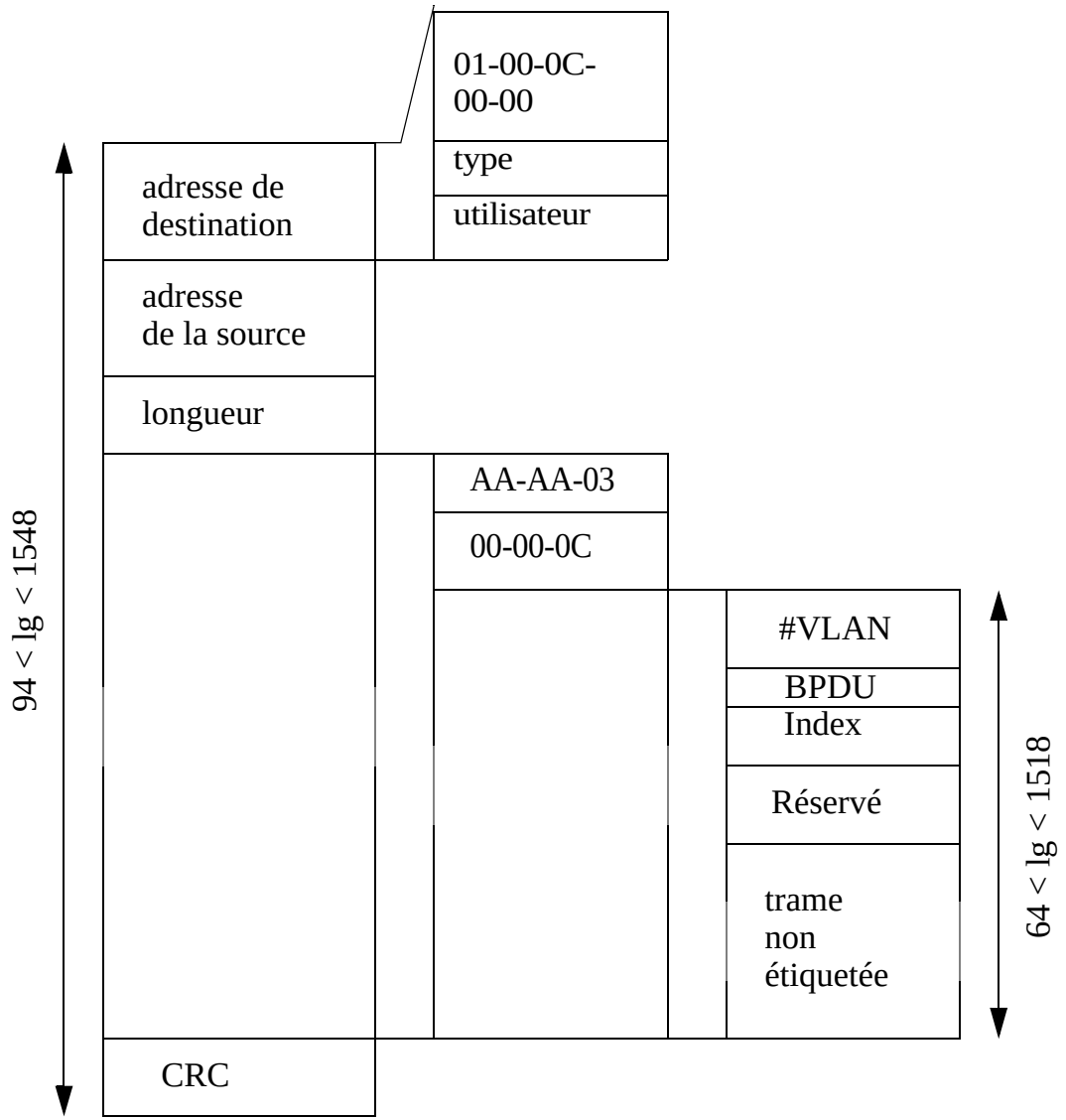
RÉSEAU FÉDÉRATEUR



- Marquage des trames sur le réseau fédérateur
- Trames normales sur le réseau entre le dernier équipement d'interconnexion et la station.

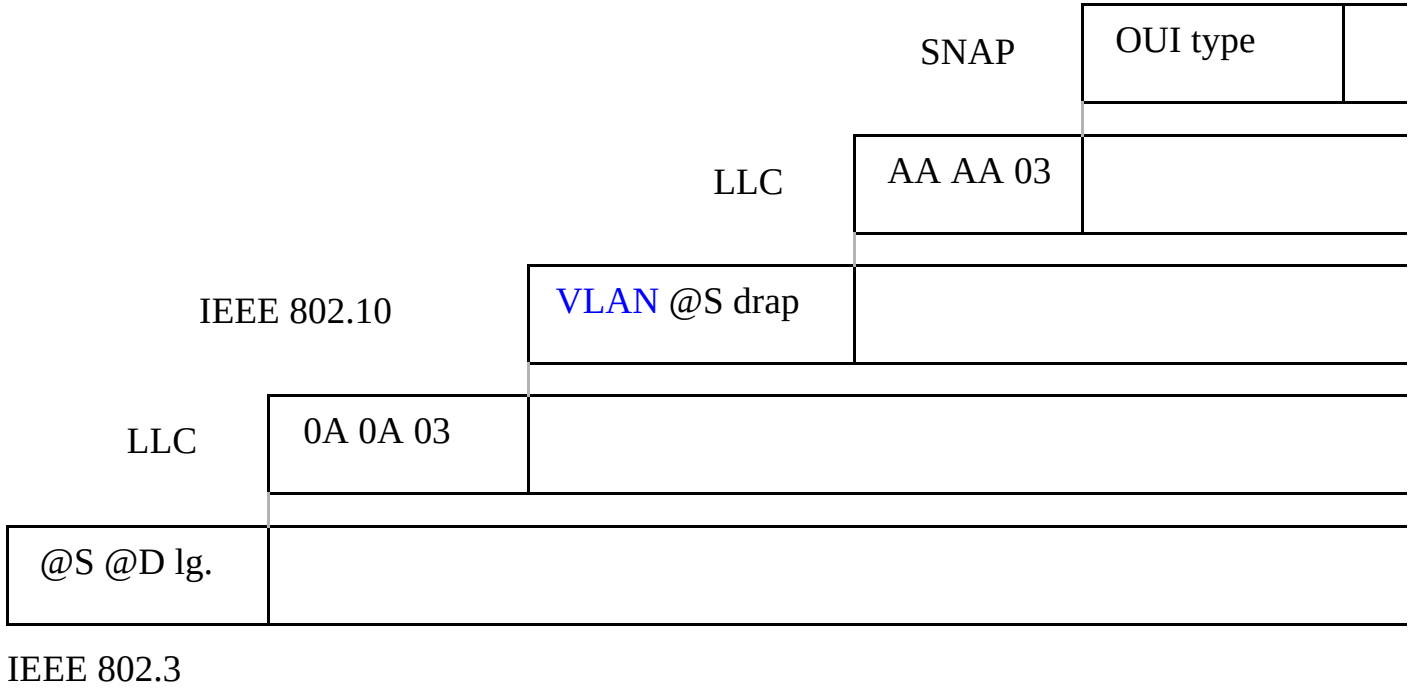
\+ Solution constructeurs (Cisco : ISL, ...)

FORMAT PROPRIÉTAIRE: ISL (CISCO)



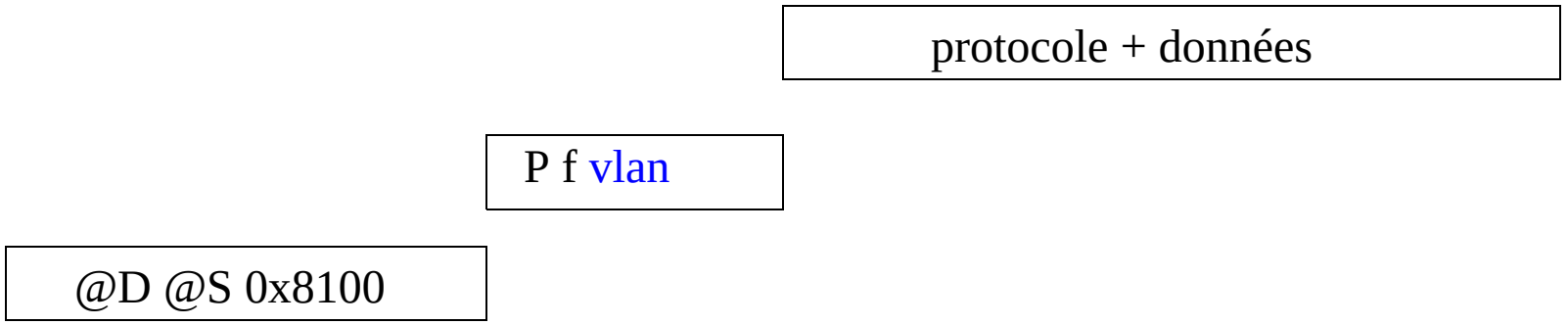
FORMAT PROPRIÉTAIRE : IEEE 802.10

- Norme pour le chiffrement des données
- Réutilisation pour la numérotation des VLAN
- Dans le cas de matériel hétérogène sur un anneau FDDI
- Encapsulation coûteuse.



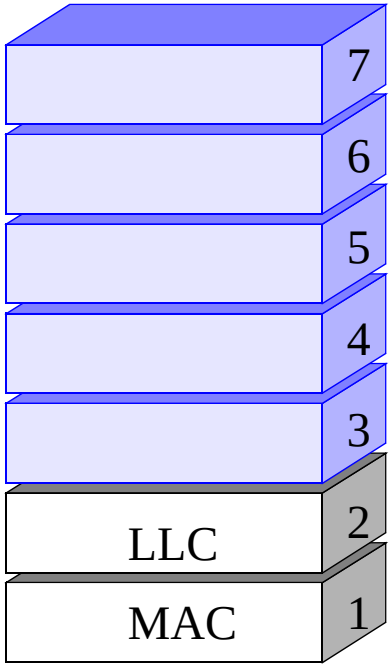
IEEE 802.1P + IEEE 802.1Q

- Nouvelle norme, peu de produit
- Encapsulation réduite



- 12 bits pour numéroter le VLAN
- Taille des trames peut dépasser 1518 octets => problème de compatibilité avec les anciens équipements
- Projet de norme IEEE 802.3ac pour autoriser des longueurs de trames plus grandes

LA COUCHE LLC (LOGICAL LINK CONTROL)



CARTE D'IDENTITÉ DU PROTOCOLE

- Permet de contrôler la communication entre les entités distantes

- Trois modes de fonctionnement :

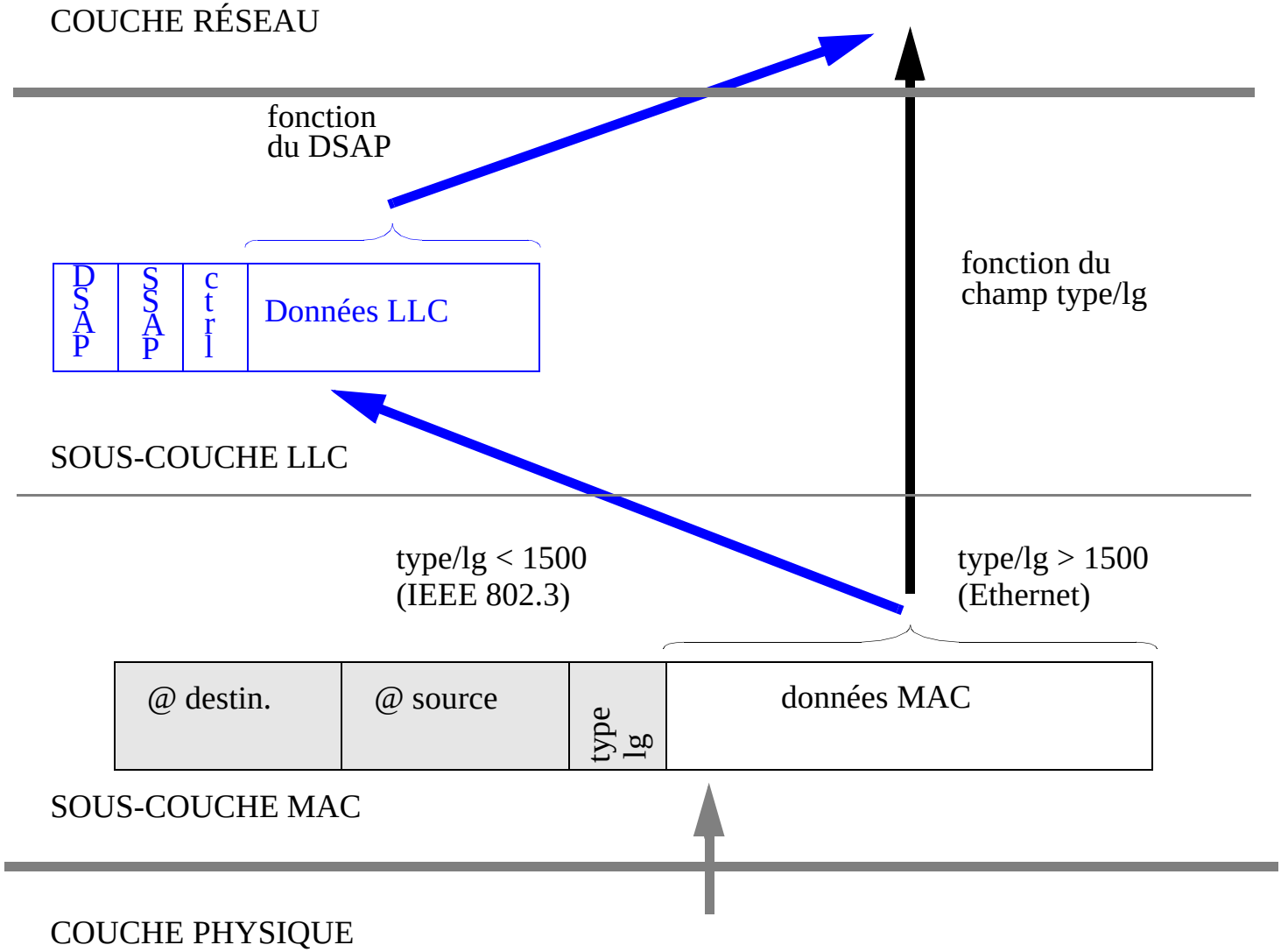
  - LLC type 1 ou mode datagramme.

  - LLC type 2 ou mode transfert sur connexion.

  - LLC type 3 ou mode datagramme acquitté.

- N'est pas utilisé par Ethernet, uniquement par IEEE 802.3 et IEEE 802.5

ETHERNET VS IEEE 802.3



FORMAT DES TRAMES

DSAP	SSAP	Control	information
8 bits	8 bits	8 ou 16 bits	8*M bits

- Le SAP (Service Access Point) permet de désigner la couche supérieure.
  Le format du SAP est :

  AAAAAAUI avec

  - AAAAAA : Numéro du SAP.
  - U : adresse universelle ou locale,
  - I : adresse individuelle ou de groupe.

QUELQUES VALEURS DE SAP

SAP hexa.	SAP décimal	SAP binaire	signification	Equivalent Ethernet
0x00	0	0000 0000	SAP Nul	
0x02	2	0000 0010	Gestion de la couche LLC	
0x06	6	0000 0110	Réseau IP	0x800
0x42	66	0100 0010	Gestion du *Spanning Tree*	
0x7E	126	0111 1110	X.25 niveau 3	0x805
0xAA	170	1010 1010	SNAP	
0xE0	224	1110 0000	IPX: (protocole du réseau NetWare de Novell)	0x8137

CHAMP CONTRÔLE

Trame de type I (Information)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0	N(S)							P/F	N(R)						

Trames de type S (Supervision)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
1	0	0	0					P/F	N(R)							RR
1	0	0	1					P/F	N(R)							REJ
1	0	1	0					P/F	N(R)							RNR

réservé (mis à 0)

Trames de type U (Non Numérotées)

1	2	3	4	5	6	7	8	
1	1	1	1	P	1	1	0	SABME
1	1	0	0	P	0	1	0	DISC
1	1	0	0	F	1	1	0	UA
1	1	1	1	F	0	0	0	DM
1	1	1	0	F	0	0	1	FRMR
1	1	1	1	P/F	1	0	1	XID
1	1	0	0	P/F	1	1	1	TEST
1	1	0	0	F	0	0	0	UI

QUESTIONS

- Combien de SAP peuvent être attribués ?
- Quelle est la longueur de l'en-tête d'un datagramme de type 1 ?
- Comment faire la différence entre une trame Ethernet et IEEE 802.3 ?
- IEEE 802.3 et Ethernet peuvent-ils cohabiter sur un même réseau ?
- Un répéteur peut-il être utilisé pour interconnecter un réseau Ethernet à un réseau IEEE 802.3 ?

+ Poursuivre le désassemblage de la trame IEEE 802.5 :

0 - 10 40 00 00 1D AO CE 51 00 00 1D AO D1 33 AA AA

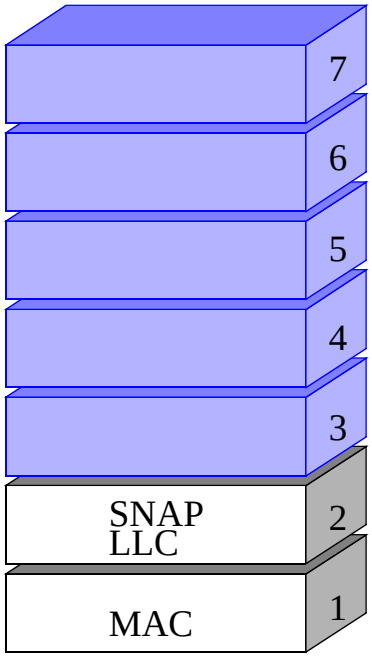
16 - 03 00 00 00 08 00 45 00 00 82 01 01 40 00 FF 11

32 - 60 OB CO 2C 4D 01 CO 2C 4D 04 00 A1 00 A1 00 6E

48 - 00 00 30 64 02 01 00 04 06 70 75 62 6C 69 63 AO

64 - .....

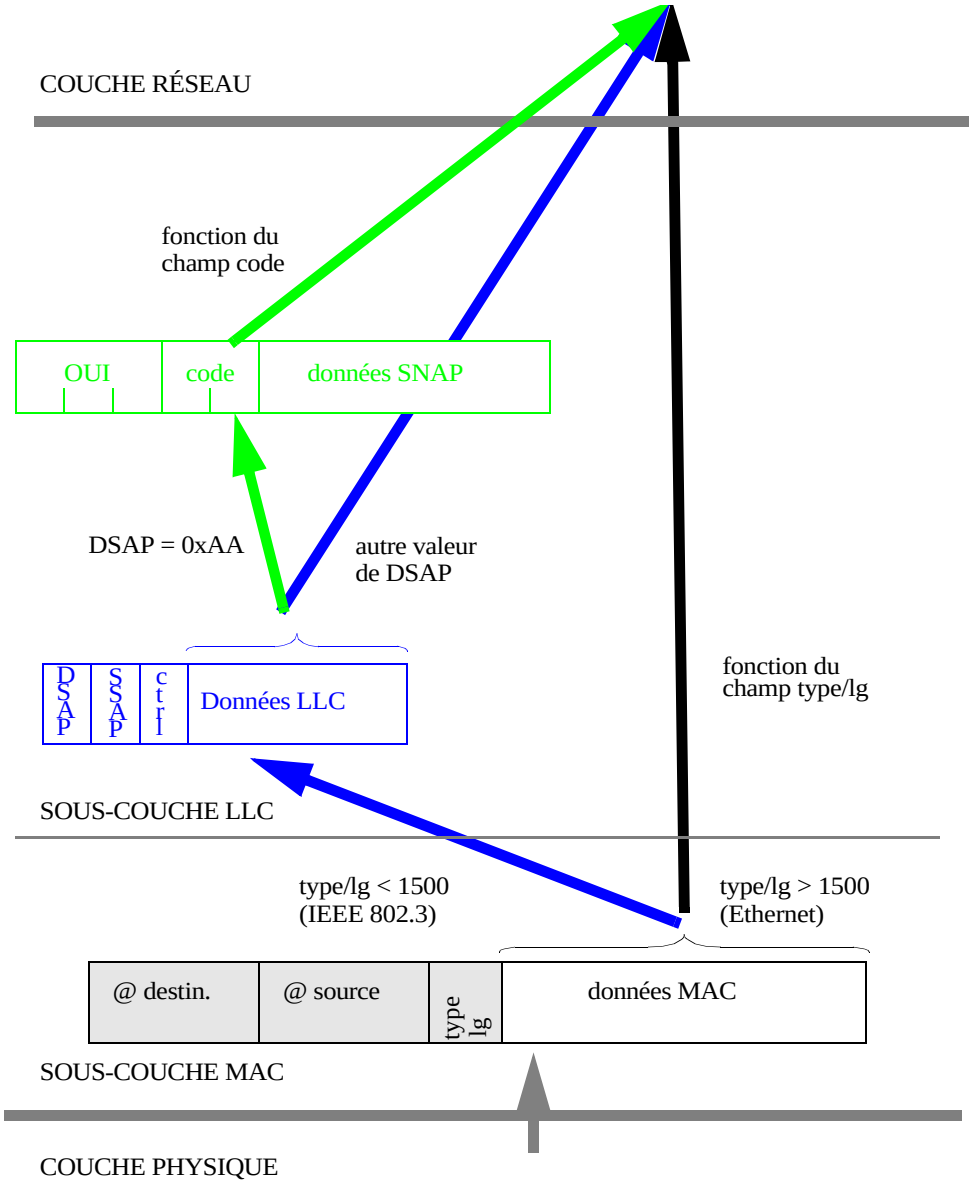
SNAP



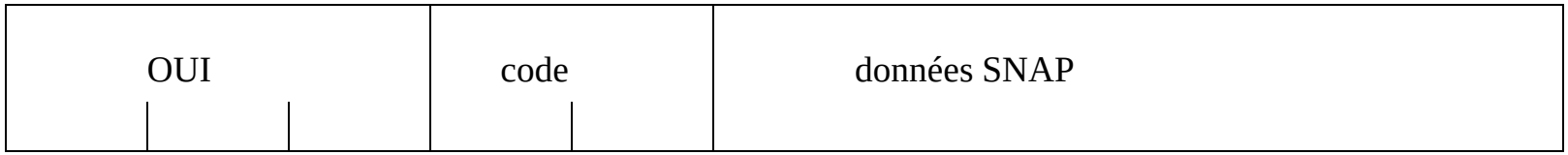
Carte d'identité du protocole

- Permet de combler les lacunes de LLC :
  - Adresse sur un nombre impair de bits : peu performant,
  - Espace de valeurs des SAP limités,
  - Compatibilité avec l'adressage d'Ethernet.
- SNAP propose une encapsulation supplémentaire au dessus de LLC
- Valeur du SAP : 0xAA
- SNAP ne met pas en œuvre de protocole supplémentaire

ETHERNET VS LLC VS SNAP



FORMAT DES TRAMES



-
-
- OUI : Organizational Unit Identifier (3 octets) désigne un vendeur.
  Identique aux 3 premiers octets d'une adresse MAC
- Généralement le OUI est à 0 dans les trames SNAP
- Code : identique aux codes rencontré dans le champ proto des trames Ethernet.

QUESTIONS

+ Poursuivre le désassemblage de la trame IEEE 802.5 :

0 - 10 40 00 00 1D AO CE 51 00 00 1D AO D1 33 AA AA

16 - 03 00 00 00 08 00 45 00 00 82 01 01 40 00 FF 11

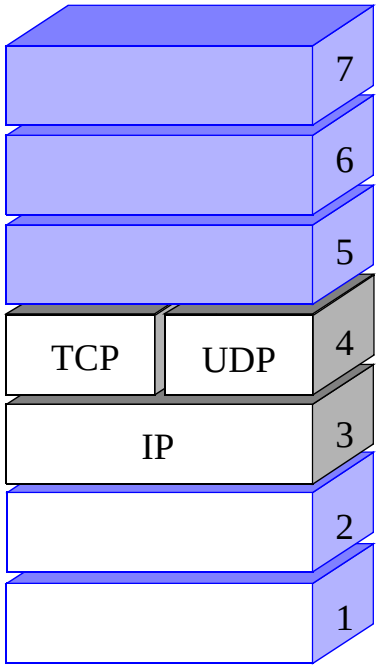
32 - 60 OB CO 2C 4D 01 CO 2C 4D 04 00 A1 00 A1 00 6E

48 - 00 00 30 64 02 01 00 04 06 70 75 62 6C 69 63 AO

64 - .....

+ Quel serait le format pour des trames Ethernet et IEEE 802.3

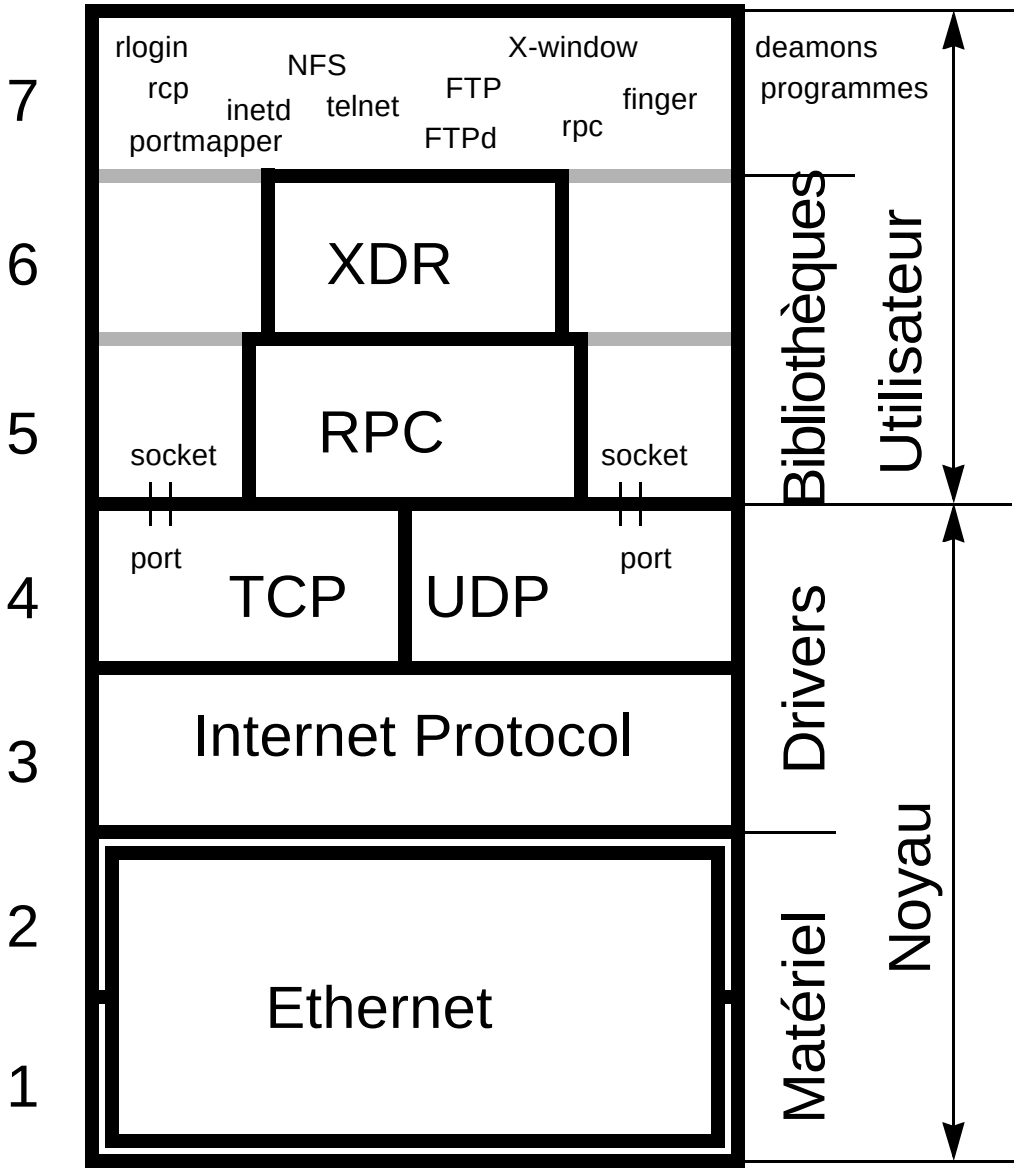
IP - UDP - TCP



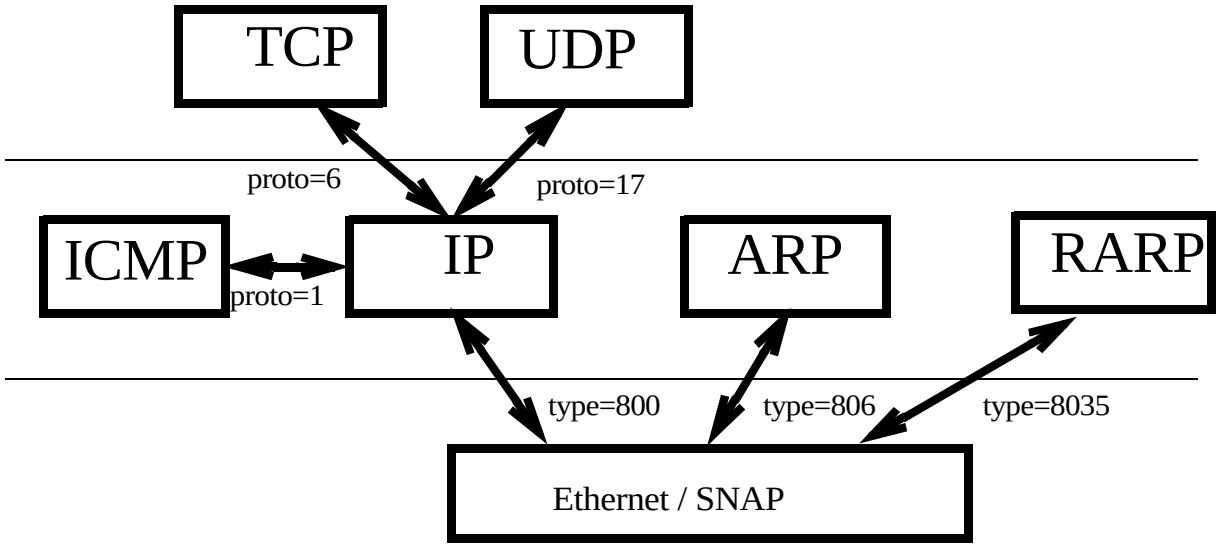
CARTE D'IDENTITÉ DU PROTOCOLE IP

- Protocole en mode datagramme
- La portée des paquets IP est mondiale
- À la base du plus grand réseau informatique mondial : Internet
- Présent dans tous les systèmes UNIX
- Présent de plus en plus dans les PC

ARCHITECTURE D'UN SYSTÈME UNIX



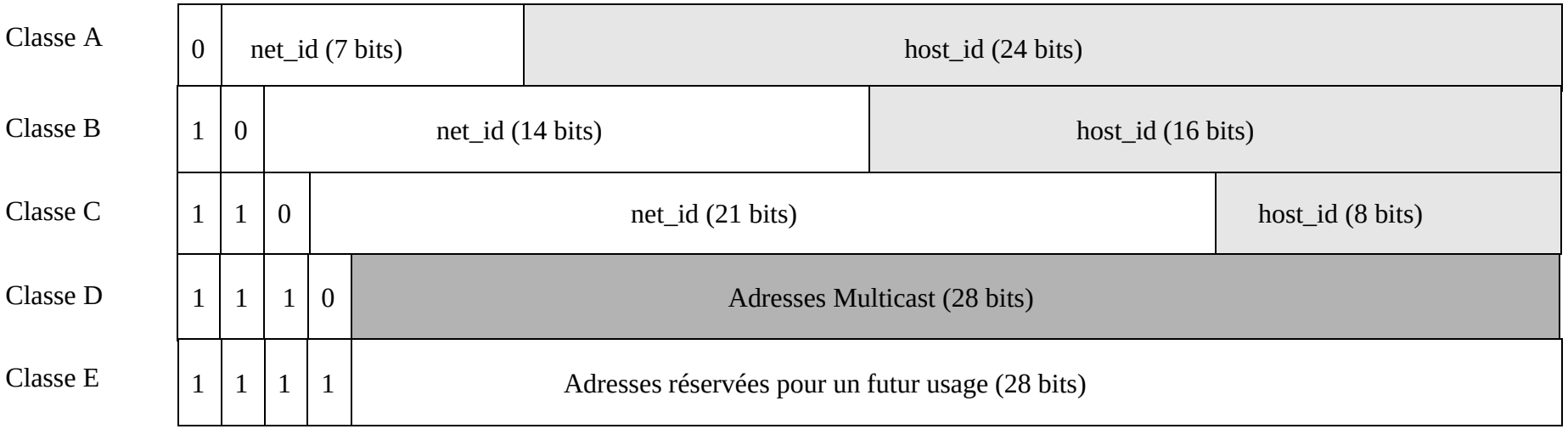
ARCHITECTURE DE LA COUCHE 3



- Protocoles de niveau 3
  - ARP (*Address Resolution Protocol*)
  - RARP (*Reverse Address Resolution Protocol*)
  - ICMP (*Internet Control Message Protocol*)

- Protocoles de niveau 4
  - UDP (*User Datagram Protocol*)
  - TCP (*Transmission Control Protocol*)

L'ADRESSAGE IP

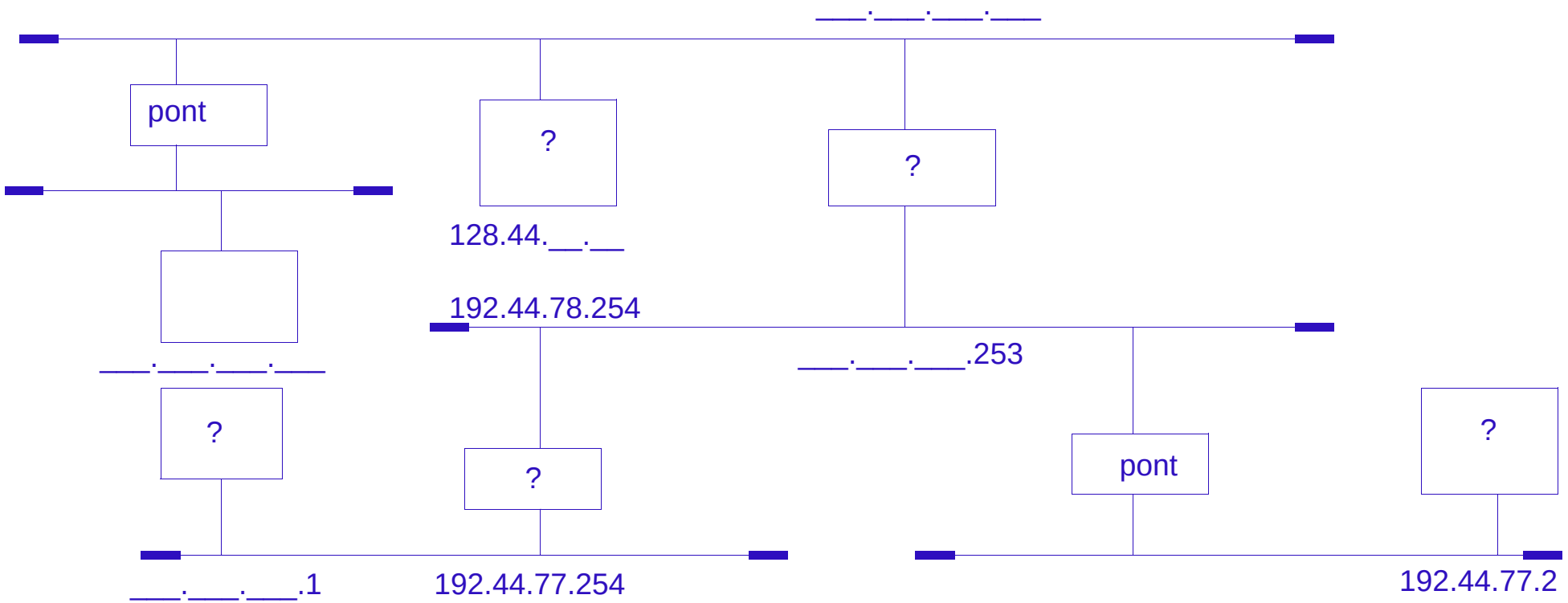


- Les classes A, B et C désignent des stations
- La classe D définit des adresses de groupe
- La classe E est réservée

PLAN D'ADRESSAGE

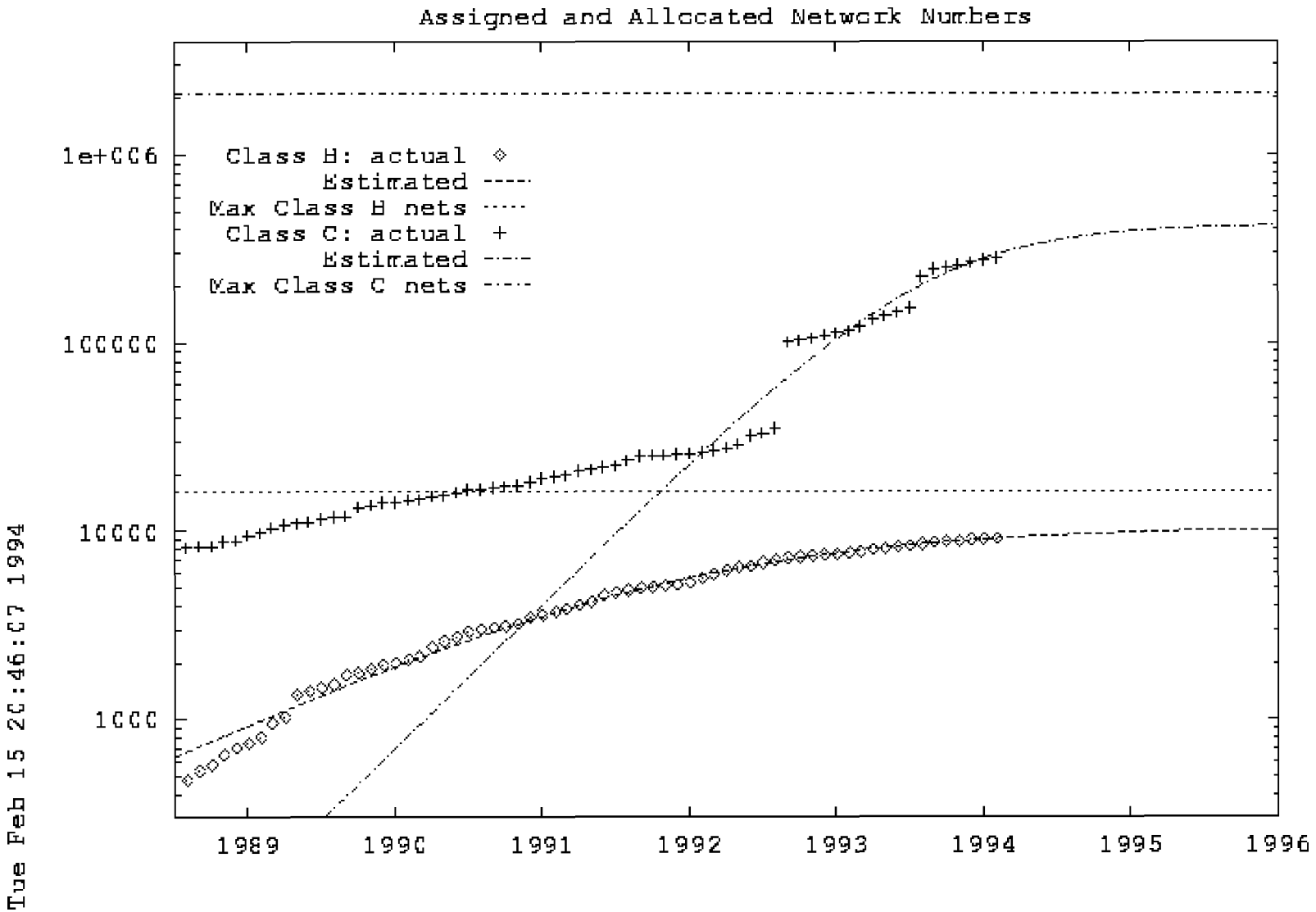
+ Un routeur a une adresse IP par réseau.

+ Le pont est transparent.



+ Est ce que la station A peut avoir cette adresse ?

EVOLUTION DU NOMBRE D'ADRESSES IP

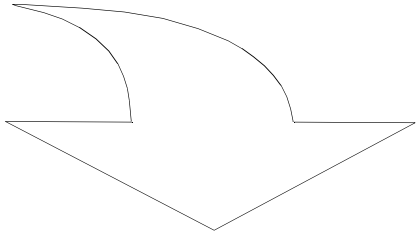


NOTION DE SOUS-RÉSEAU / NETMASK

+ Les numéros attribués (classes B ou C) par le NIC désignent un réseau. La partie restante est à administrer par l'ingénieur système. Elle peut être divisée en réseaux et sous-réseaux.

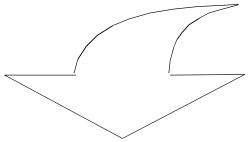
Adresses IP : 192.44.77.79

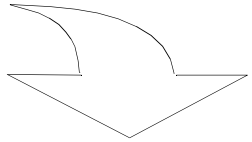
NetMask : 255.255.255.192



1100 0000.0010 1100.0100 1101.0100 1111
1111 1111.1111 1111.1111 1111.1100 0000

réseau machine





1100 0000.0010 1100.0100 1101.0100 0000 1111
192.44.77.64 15

LA TABLE DE ROUTAGE.

- Permet d'indiquer la direction que devront suivre les paquets.
- Contient deux informations :
  - la destination (machine ou réseau IP)
  - le chemin directement accessible (attachement local à IP ou routeur sur le même sous-réseau IP).

Exemple :

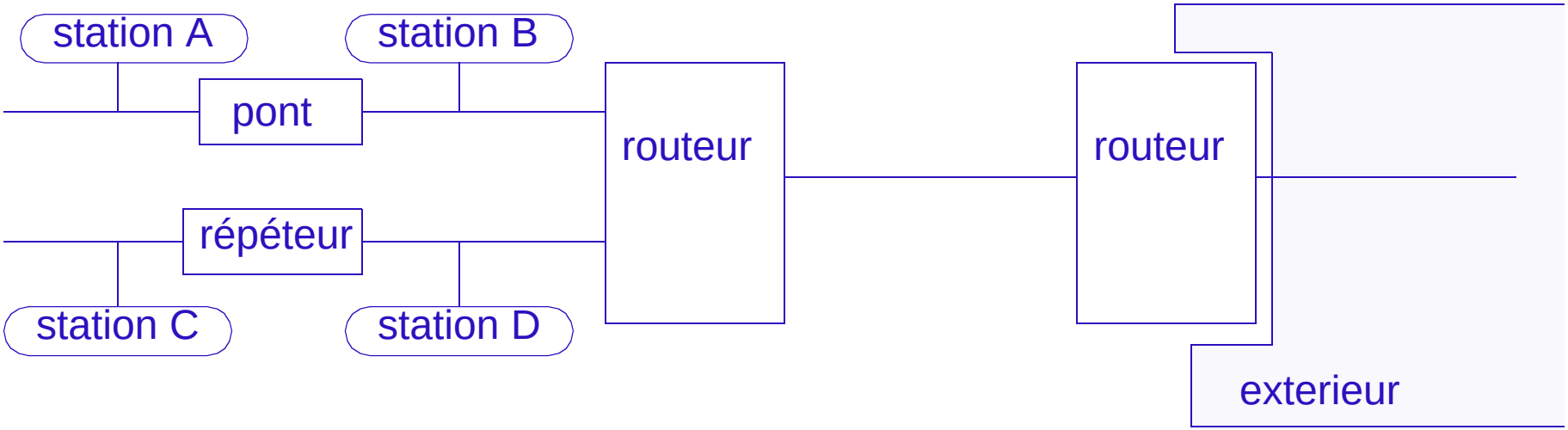
Destination	Gateway
default	mgs-rsm
192.44.77.0	bloodmoney
192.108.119.0	nintendo

+ Cette table provient de la machine bloodmoney :

Proposez une adresse IP pour cette machine,
Proposez une table de routage pour les machines nintendo et msg-rsm.

QUESTIONS

Soit le réseau suivant :



Le NIC vous attribue le numéro de réseau 192.45.67.

\+

Quelle est la classe de cette adresse ?

Définir un netmask pour chacun des réseaux IP.
Donner une adresse pour une station sur chacun des sous-réseaux.
Combien d'adresses IP ont les routeurs ?

CIDR (CLASSLESS INTER-DOMAIN ROUTING)

- Classe B : 65524 adresses,
- Classe C : 254 adresses,
- Pénurie d'adresses ⟹ utilisation d'adresses de classe C,

⟹ Saturation des tables de routages (plusieurs entrées pour le même endroit).

\+ CIDR permet d'annoncer un ensemble d'adresses consécutives.

\+ La notion d'adresses de classes A, B ou C disparaît.

\+ On annonce l'adresse de début et un netmask indiquant la partie variable.

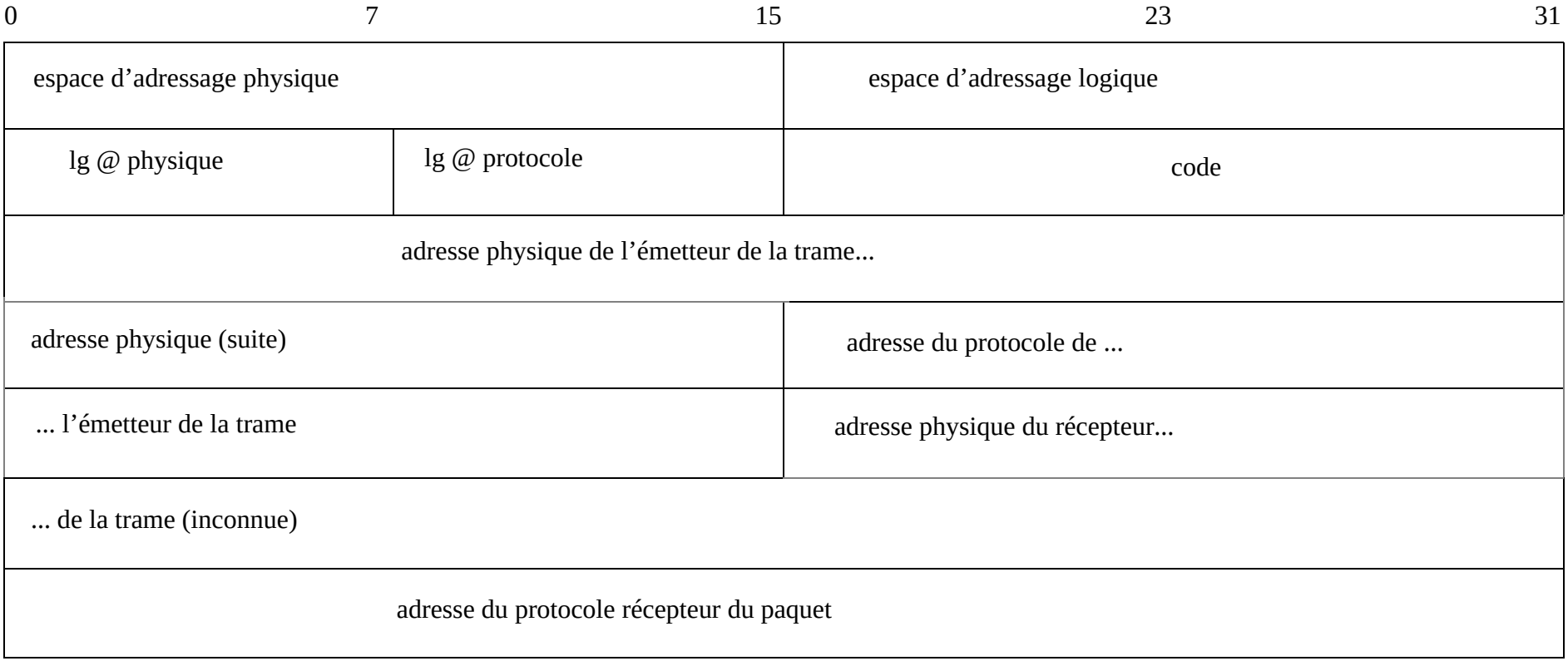
Exemple d'allocation d'adresses

société	nb d'adresses	nb de classe C	adresse de début	adresse de fin	adresse de début	netmask	netmask en binaire
A	< 2048	8	192.24.0	192.24.7	192.24.0	255.255.248.0	255.255. 1111 1000 .0
B	< 4096	16	192.24.16	192.24.31	192.24.16	255.255.240.0	255.255. 1111 0000 .0
C	< 1024	4	192.24.8	192.24.11	192.24.8	255.255.252.0	255.255. 1111 1100 .0
D	< 1024	4	192.24.12	192.24.15	192.24.12	255.255.252.0	255.255. 1111 1100 .0
E	< 512	2	192.24.32	192.24.33	192.24.32	255.255.254.0	255.255. 1111 1110 .0
F	< 512	2	192.24.34	192.24.35	192.24.34	255.255.254.0	255.255. 1111 1110 .0

LE PROTOCOLE ARP / CARTE D'IDENTITÉ

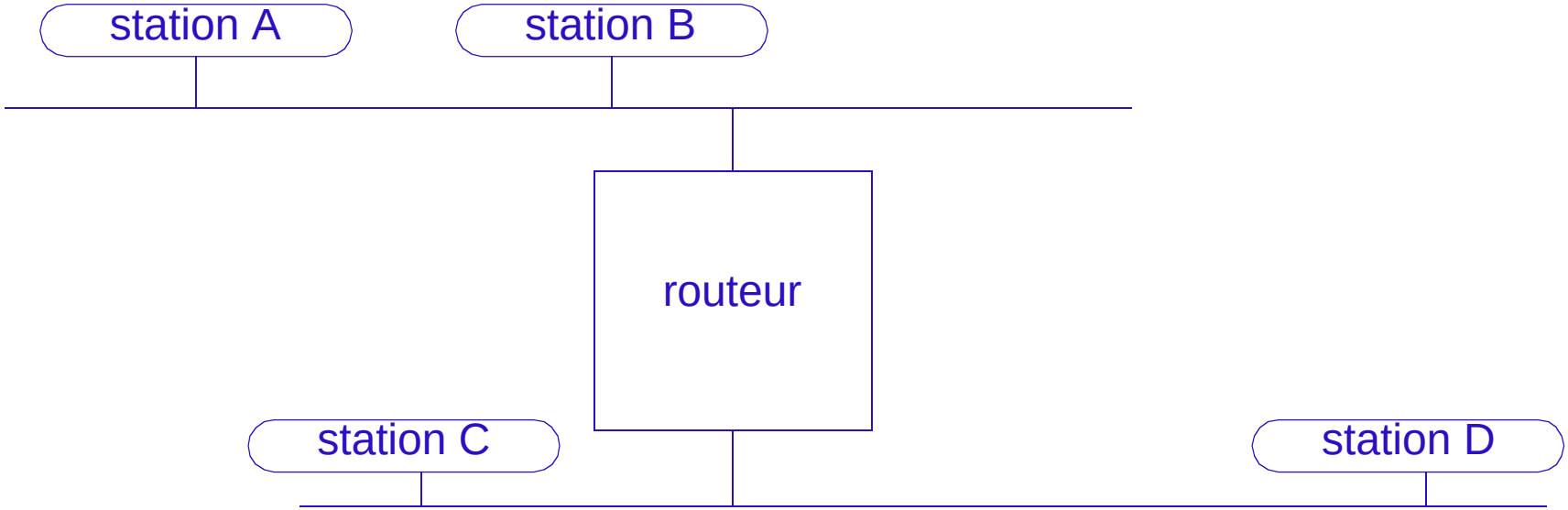
- Identifié par le champ type = 0x806 dans la trame Ethernet ou SNAP.
- Permet de faire la correspondance entre une adresse IP et une adresse MAC.
- Les applications ne manipulent que des adresses IP. Le système doit retrouver l'adresse MAC correspondante.
- Les systèmes construisent une table de correspondance.
- Quand un système ne trouve pas l'adresse MAC, il DIFFUSE une requête ARP et attend une réponse.
- La station qui reconnaît dans la requête ARP son adresse IP répond en émettant une réponse ARP avec son adresse MAC.

FORMAT D'UNE TRAME ARP



QUESTIONS

Soit le réseau suivant :

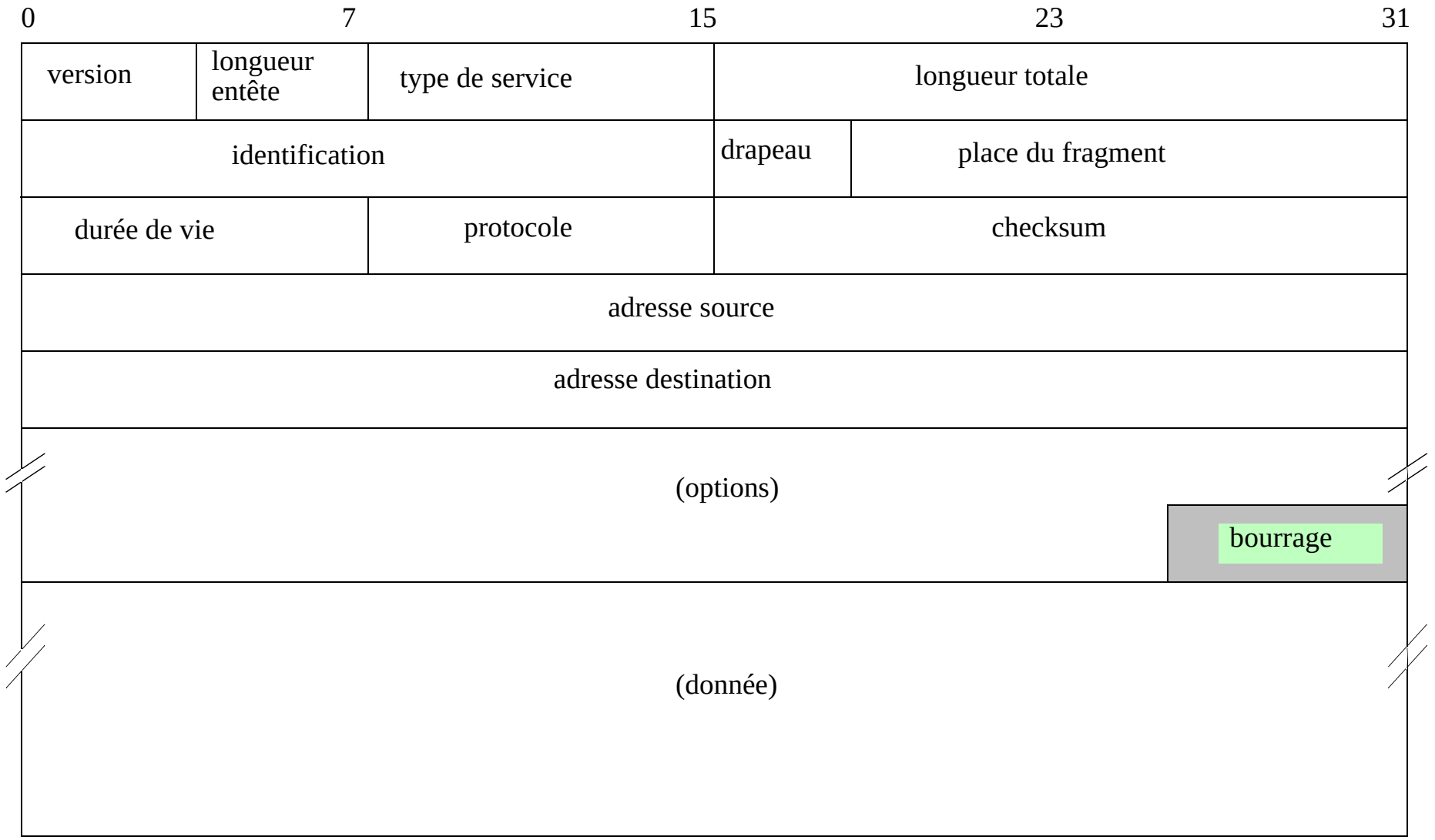


+ Que se passe-t-il quand :

  A veut émettre vers B ?
  A veut émettre vers C ?

+ Quel est le No de SAP pour ARP dans les trames LLC ?

FORMAT DES PAQUETS IP.



QUESTIONS

+ Poursuivre le désassemblage de la trame IEEE 802.5 :

0 - 10 40 00 00 1D AO CE 51 00 00 1D AO D1 33 AA AA

16 - 03 00 00 00 08 00 45 00 00 82 01 01 40 00 FF 11

32 - 60 OB CO 2C 4D 01 CO 2C 4D 04 00 A1 00 A1 00 6E

48 - 00 00 30 64 02 01 00 04 06 70 75 62 6C 69 63 AO

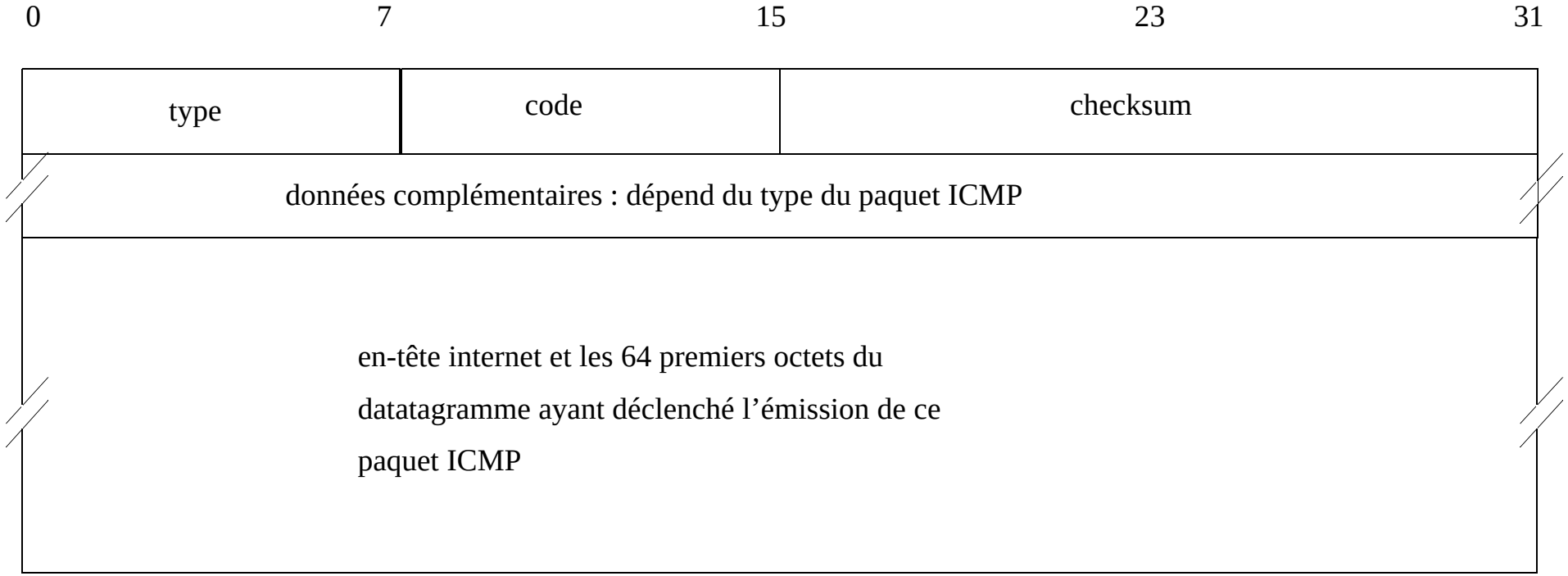
64 - .....

+ Que vaut le champ option ?

+ De quelle classe sont les adresses IP ?

LE PROTOCOLE ICMP / CARTE D'IDENTITÉ

- ICMP (*Internet Control Message Protocol*)
- Encapsulé dans IP (champ proto = 1) mais de niveau 3.
- Sert à contrôler le bon déroulement du protocole IP.



QUELQUES TYPES DE PAQUETS ICMP

- Le message ne peut atteindre sa destination (type = 3)
  - 0 : le réseau ne peut être atteint
  - 1 : la station ne peut être atteinte
  - 2 : le protocole ne peut être atteint
  - 3 : la fragmentation est nécessaire et bit DF=1
  - 4 : port inaccessible
  - 5 : route proposée en option n'est pas valable
- Durée de vie expirée (type = 11)
  - 0. La durée de vie a expiré pendant le transit
  - 1 : La durée de vie a expiré pendant le réassemblage
- Source Quench (type = 4)
- Indication de redirection (type = 5)
- Echo / La commande Ping (type = 8 : demande et type = 0 : réponse)

NOTION DE PORT

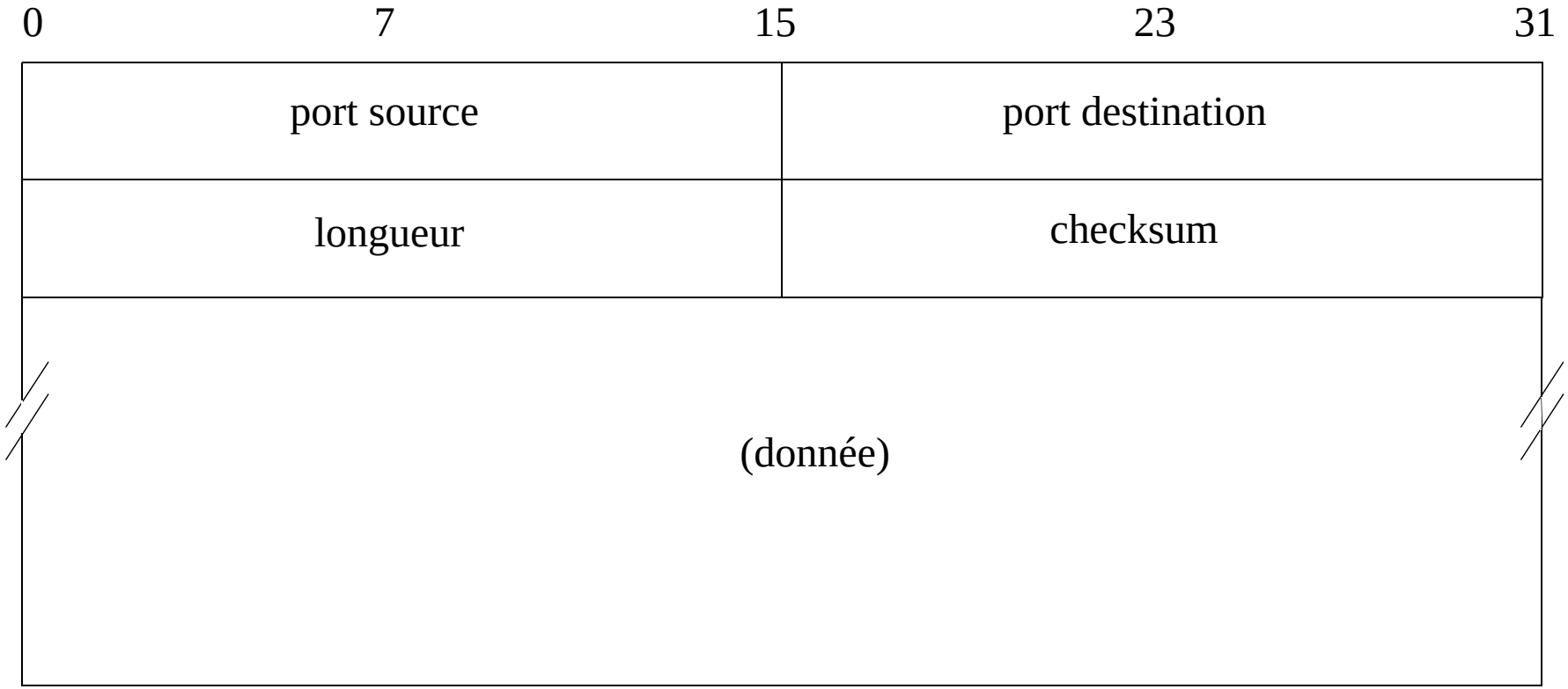
- Les adresses IP permettent de désigner une machine unique dans le monde
- Le port permet de désigner un programme sur cette machine
- 0..1023 : en mode protégé
- 1024...65535 : attribuable à n'importe quelle application.

LE PROTOCOLE UDP / CARTE D'IDENTITÉ

- Identifié par le champ proto = 17 dans le paquet IP
- Protocole en mode datagramme :
  - Aucune sécurité de bout-en-bout
  - Les adresses de classe D sont utilisables uniquement avec ce protocole
- Aucun traitement protocolaire
- Utilisé uniquement pour contenir l'information sur les numéros de port

UDP

- Format du message :

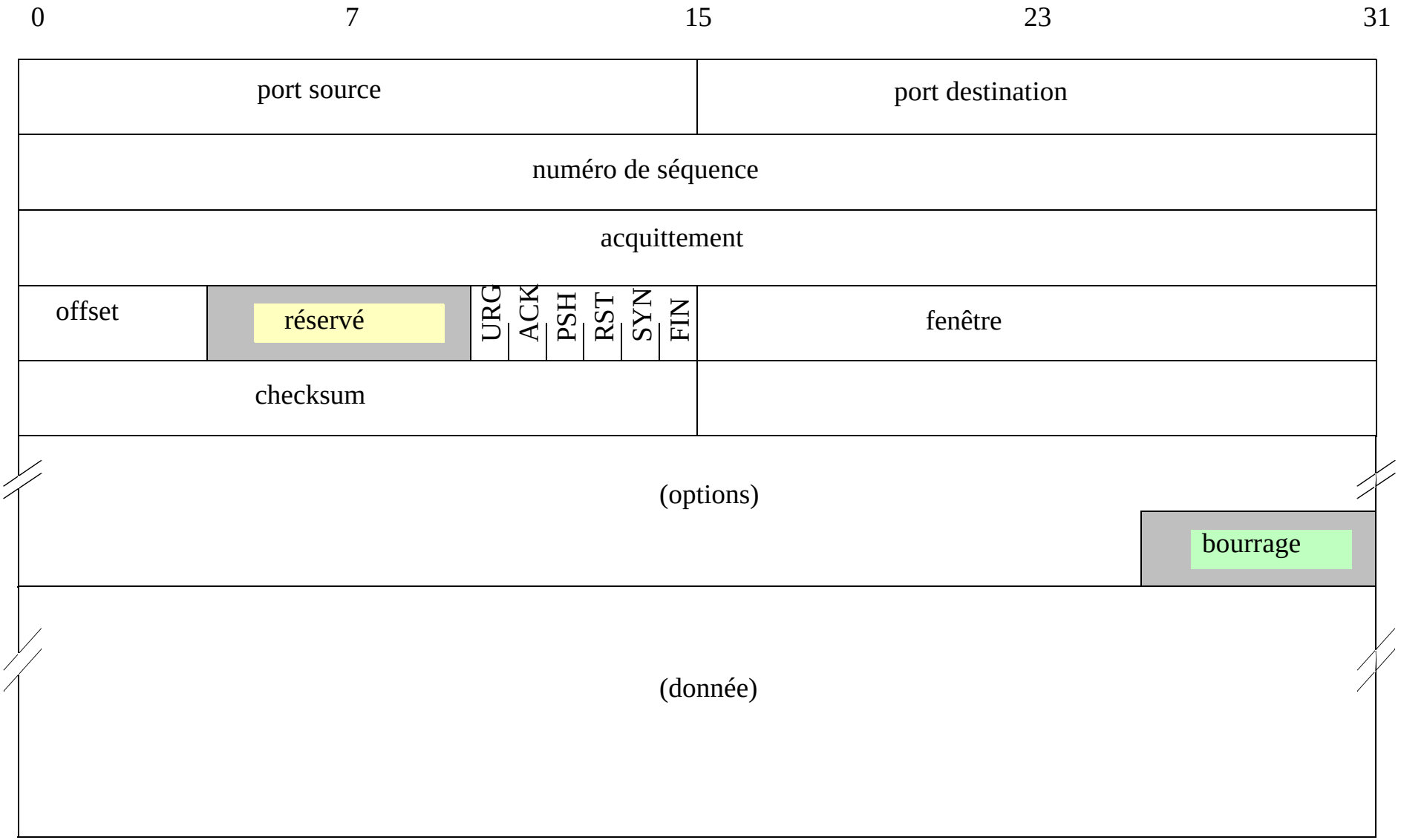


LE PROTOCOLE TCP

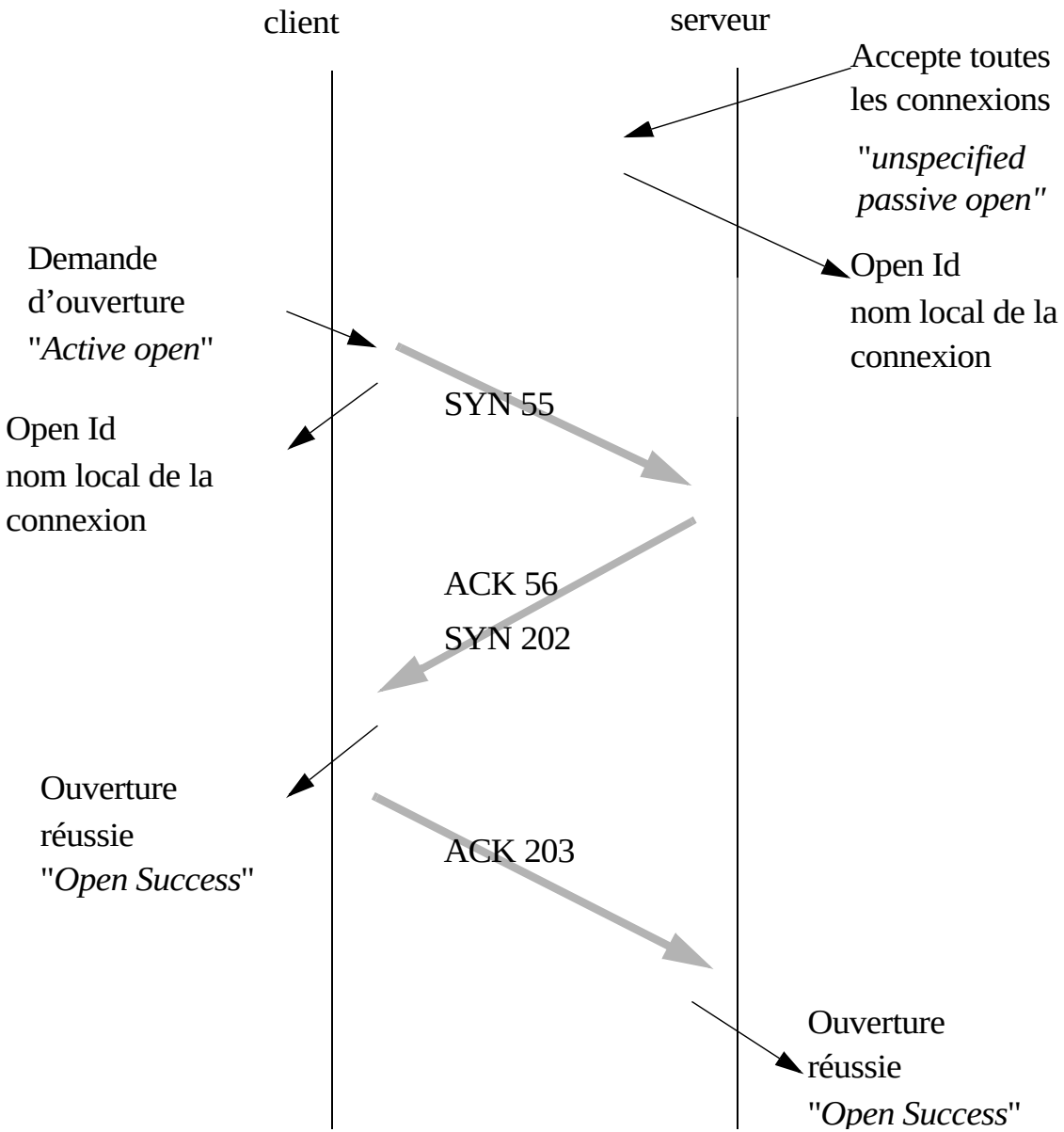
Carte d'identité :

- Identifé par le champ proto = 6 dans le paquet IP
- Protocole en mode connecté :
  - Contrôle de séquencement, de flux
  - Possibilité de données express
  - Uniquement en point-à-point
- Protocole complexe
- Trois phases :
  - Ouverture de la connexion
  - Transfert de données
  - Fermeture de la connexion
- Numérotation des octets transmis

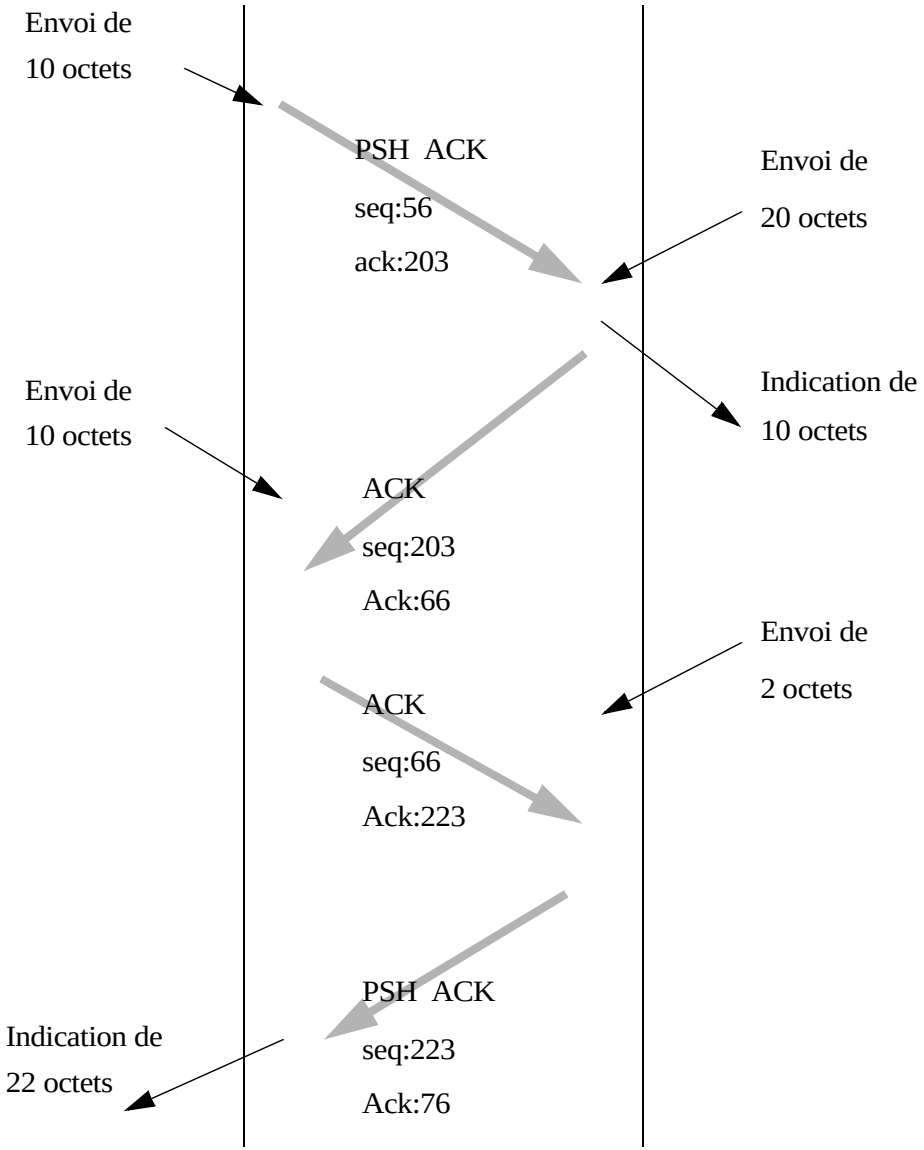
FORMAT DES MESSAGES



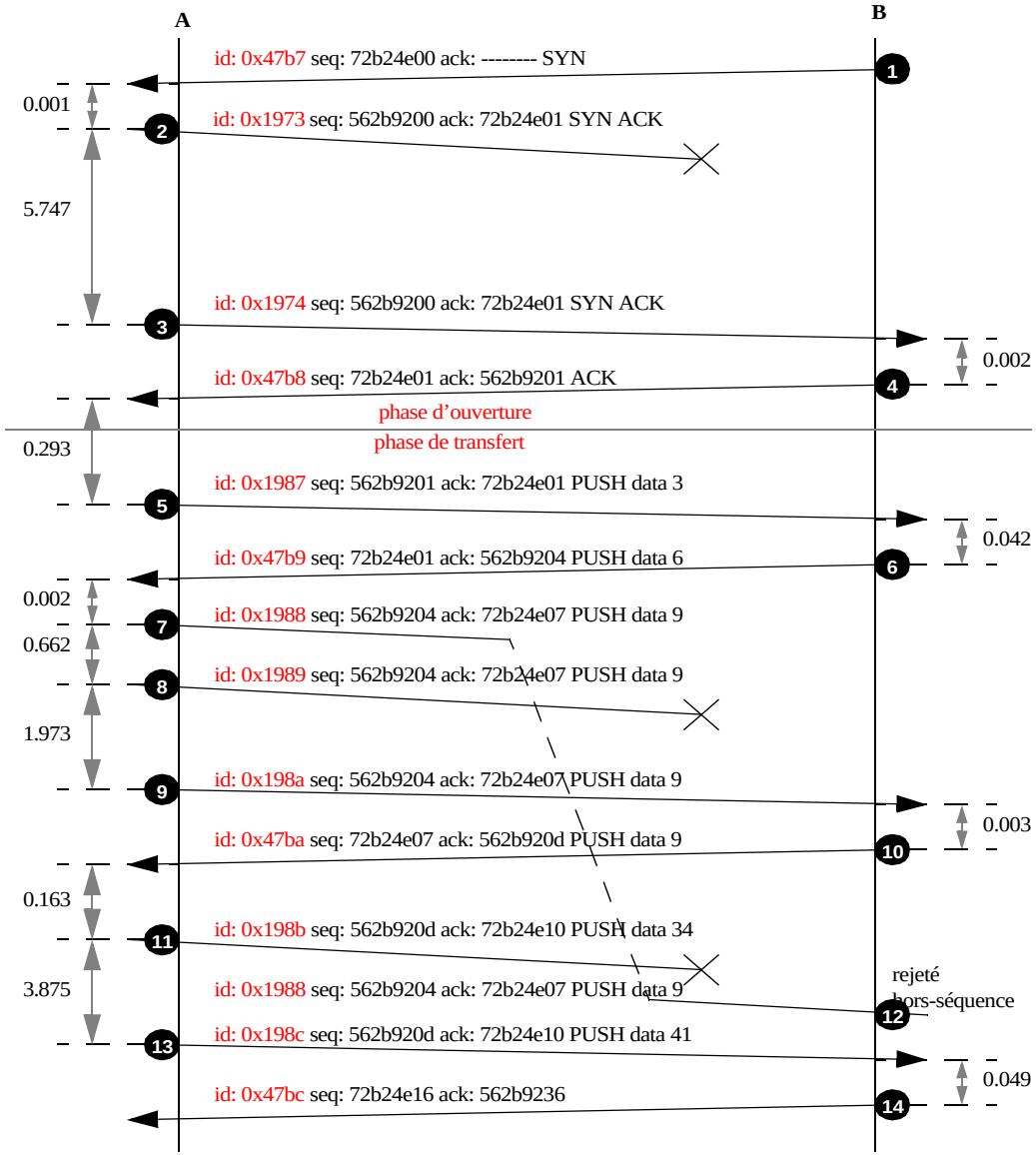
OUVERTURE DE LA CONNEXION



TRANSFERT DE DONNÉES

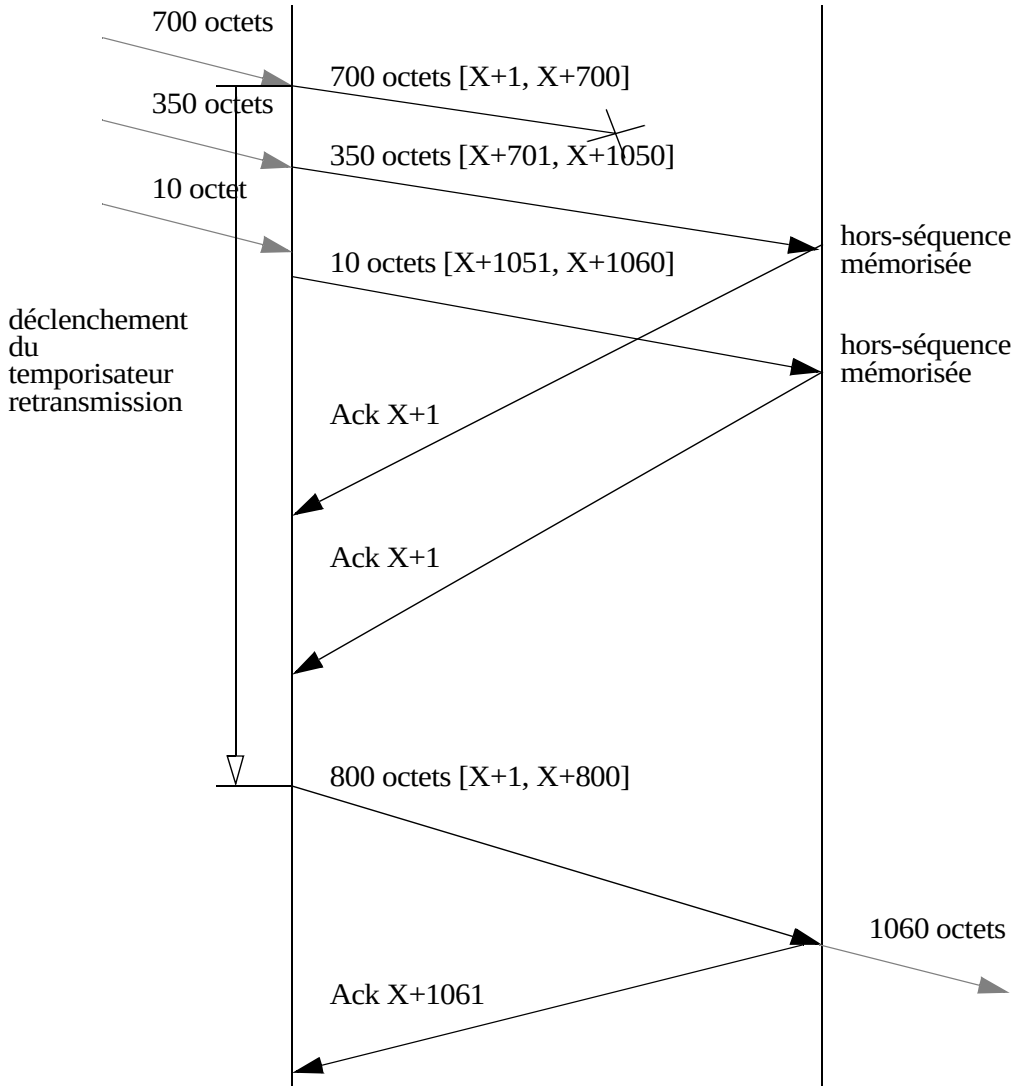


TRANSFERT AVEC ERREURS

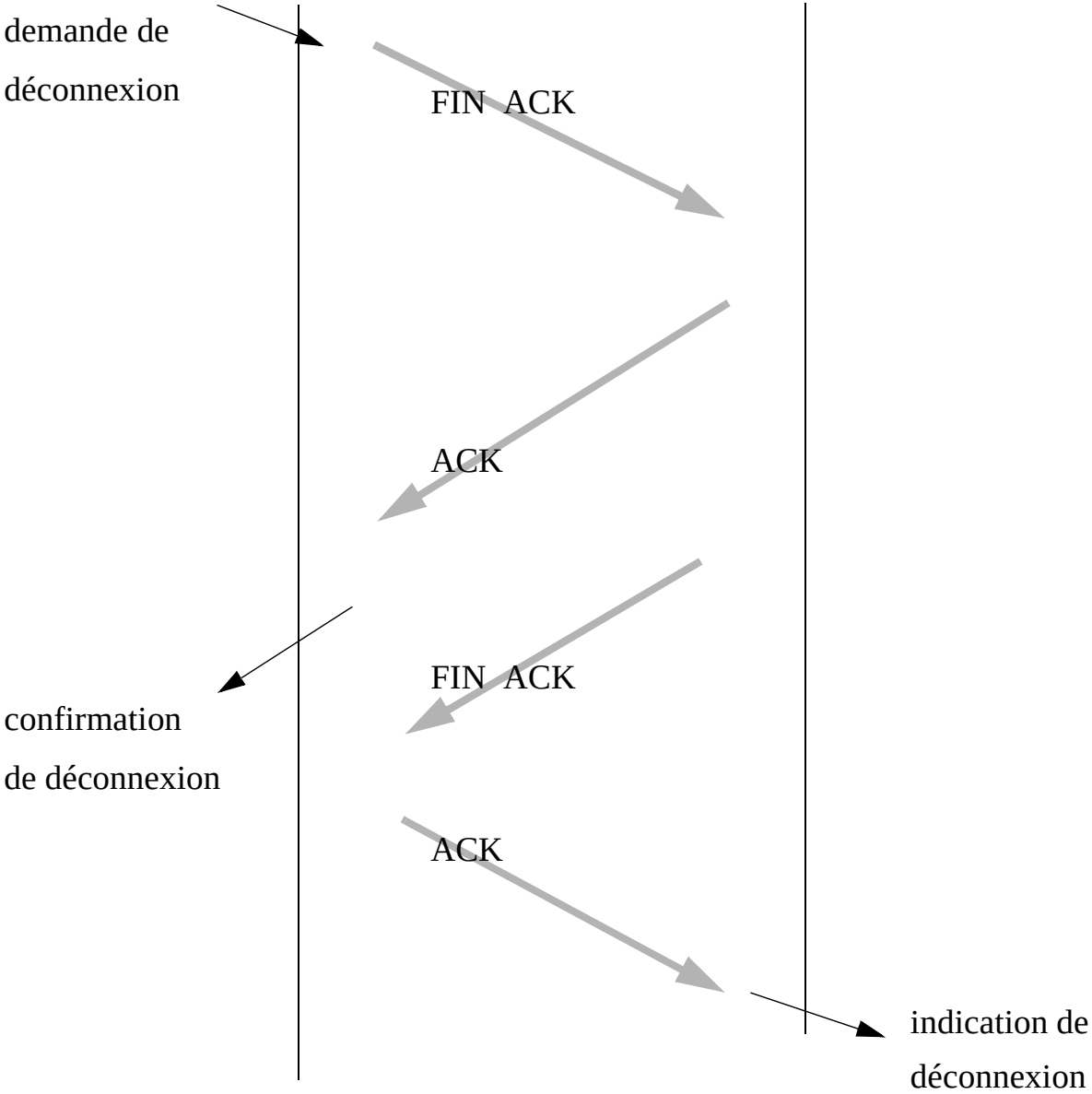


TRANSFERT AVEC ERREUR

MSS = 800 octets
Séquence du dernier octet transmis : X



FERMETURE DE CONNEXION



ADAPTATION À L'ENVIRONNEMENT

- Gestion dynamique des valeurs temporisation
- RTT : *Round Trip Time*. Temps mesuré pour un aller et retour.
- L'estimation du RTT va déterminer le temporisateur de *Retransmission Time Out* (RTO).
- Ne pas prendre en compte le RTT lors de retransmissions (algorithme de Karn)

On a : :

$$erreur = mesure - moyenne$$

$$moyenne = moyenne - \alpha \times erreur$$

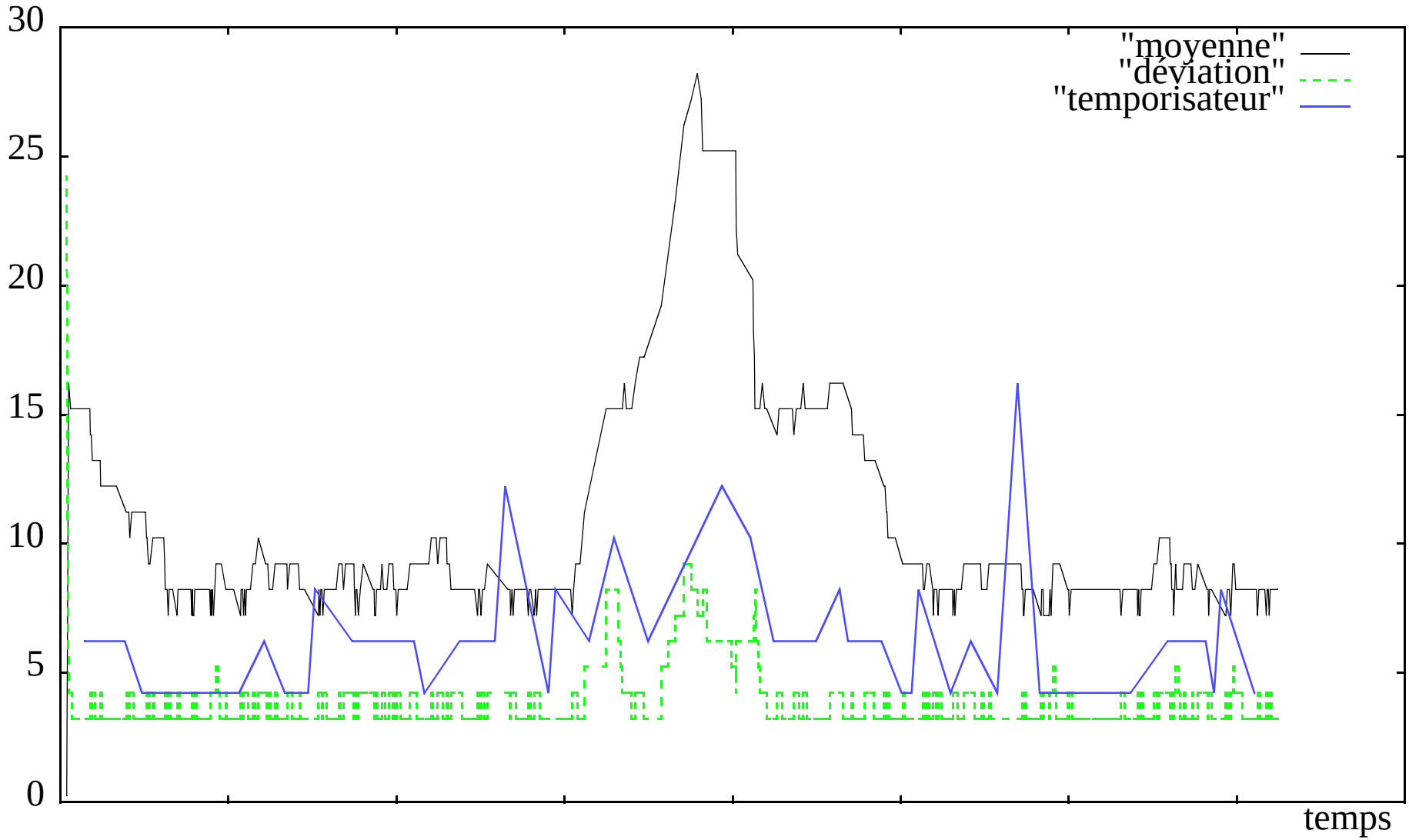
$$deviation = deviation + \beta \times (|erreur| - deviation)$$

et

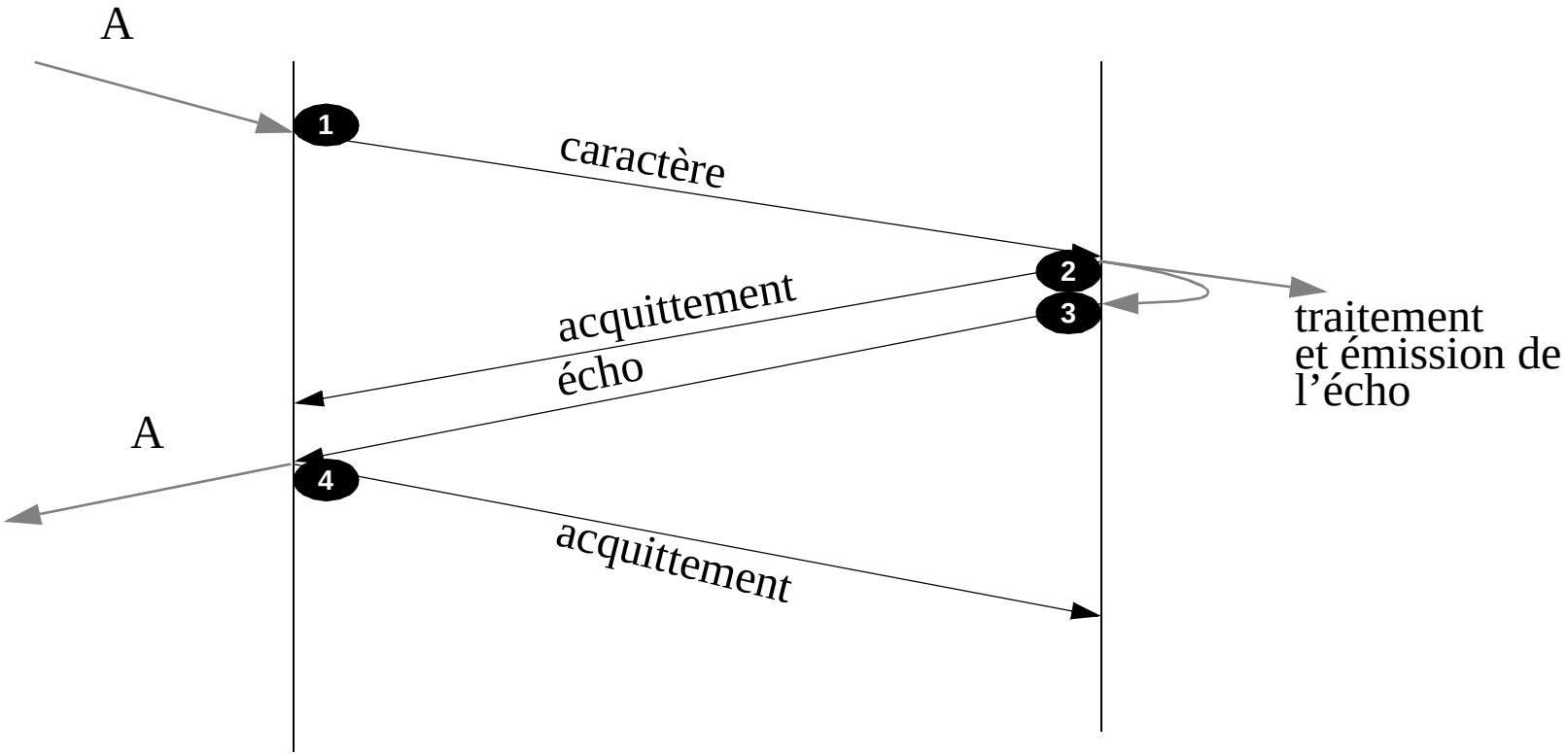
$$RTO = moyenne + 4 \times deviation$$

\+ L'utilisation d'une bonne estimation améliore les performances

EXEMPLE (MESURÉ DANS LE NOYAU SUNOS 4.1.3)



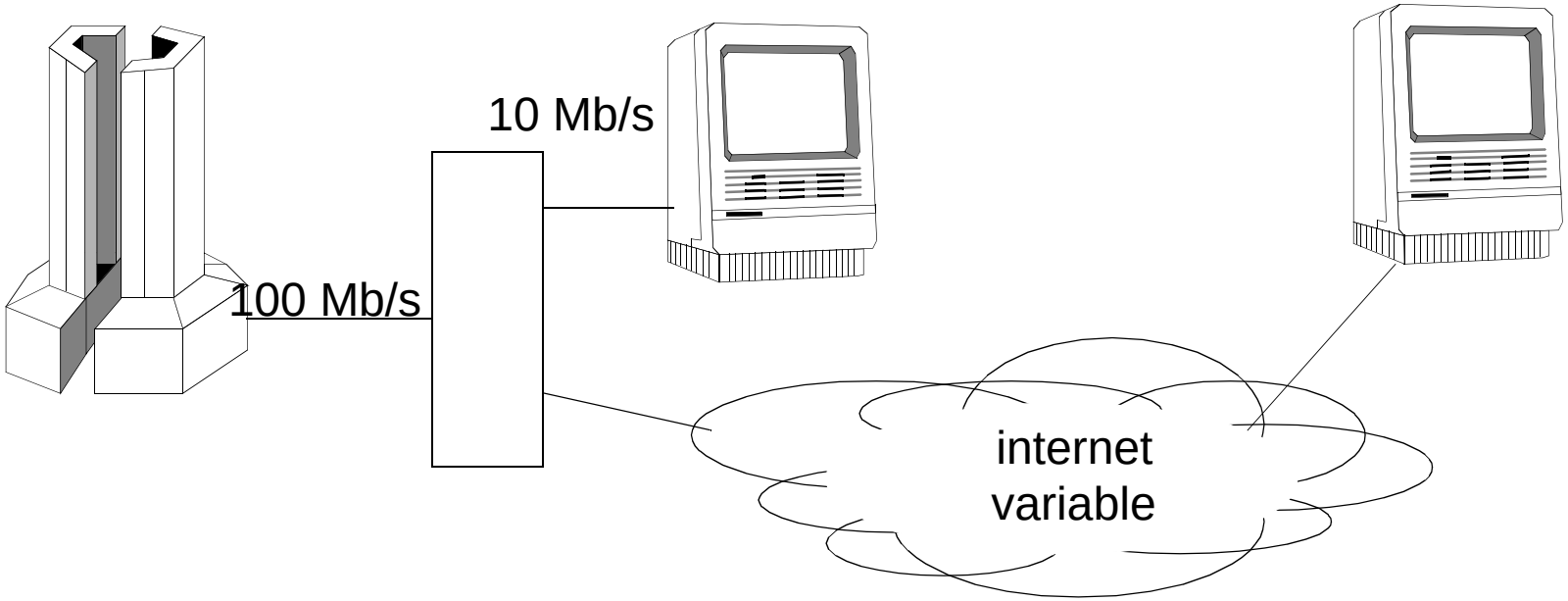
LIMITATION DU TRAFIC



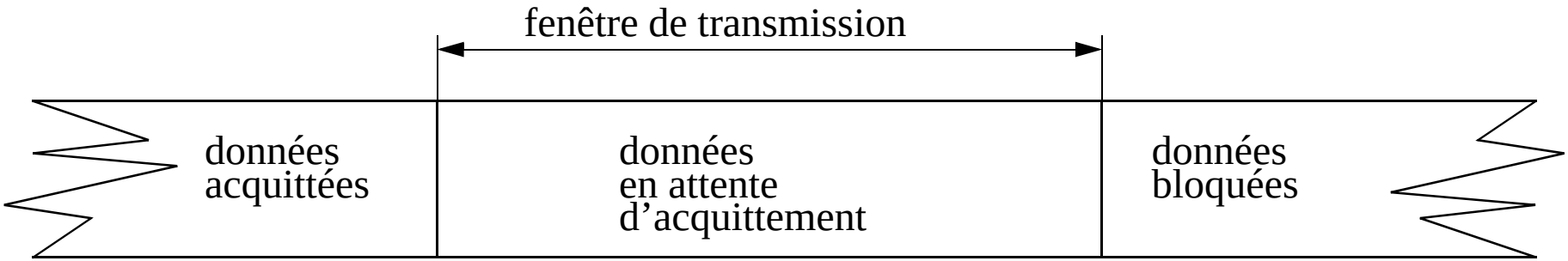
- Retard à l'acquittement de 200 ms pour permettre le "*Piggy Backing*"
- Algorithme de Naggle :
  "*on ne peut émettre sur le réseau qu'un paquet de petite taille non acquité*"

CONTRÔLE DE FLUX

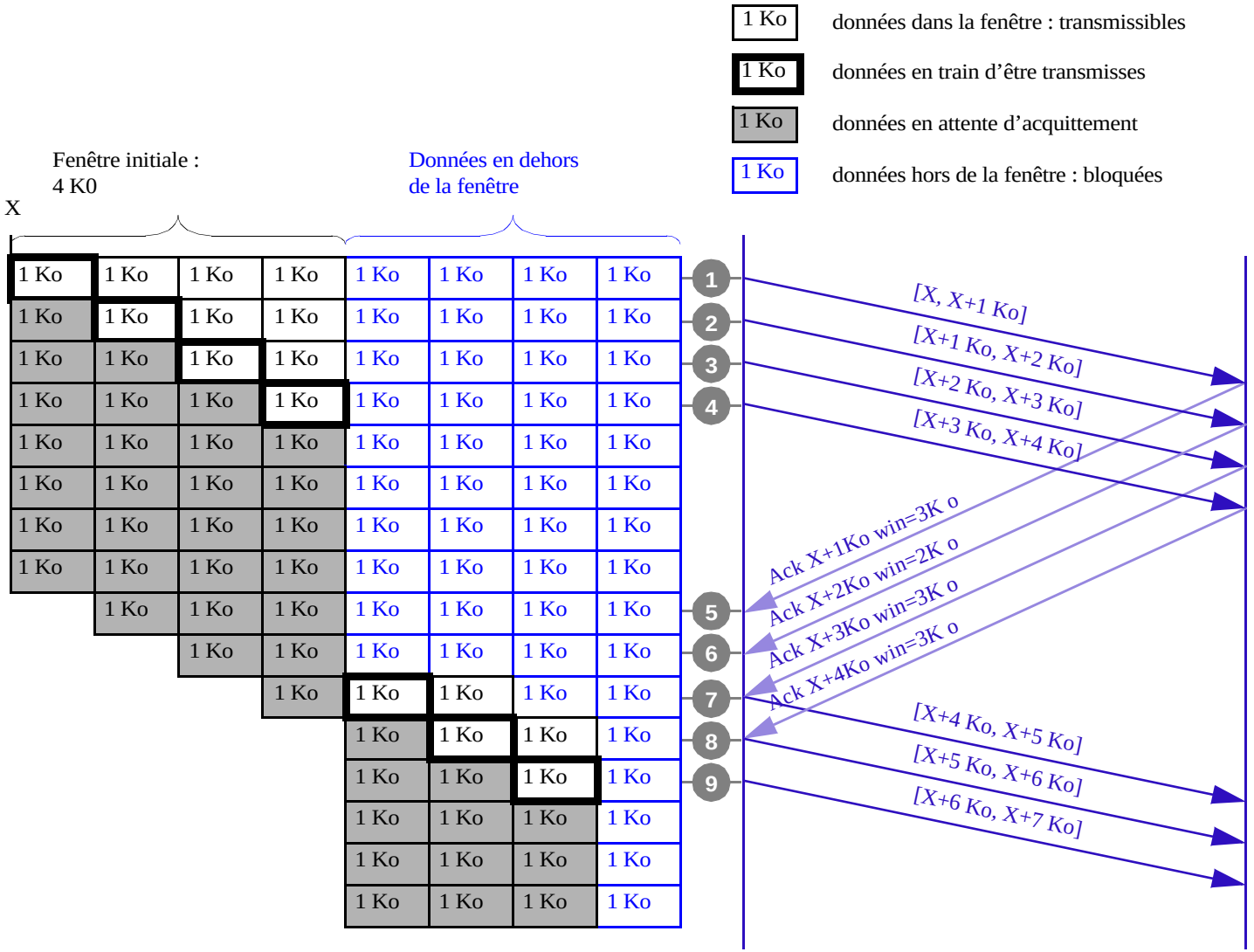
- Adapter dynamiquement le trafic émis à la configuration du réseau et au récepteur.



Basé sur la taille d'une fenêtre de retransmission (nombre d'octets non acquittés qui peuvent être émis)



FENÊTRE D'ANTICIPATION



CONGESTION

Si le débit du réseau varie (dépend des autres utilisateurs), il peut y avoir congestion des routeurs. Règles :

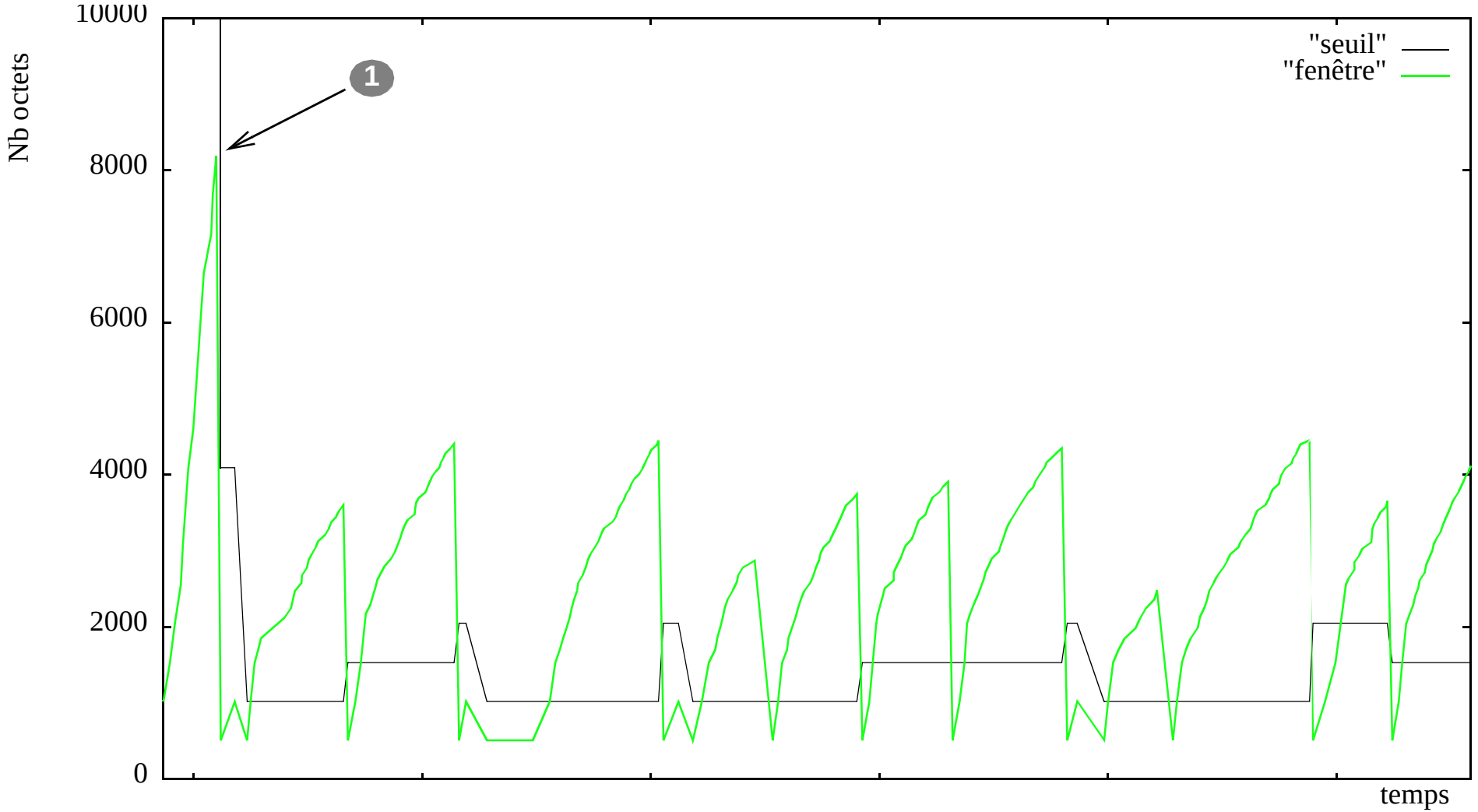
- Les congestions font perdre des paquets, qui seront retransmis, et qui à leur tour feront des congestions, qui feront perdre des paquets et qui ...
- La perte d'un paquet provient plus souvent d'une congestion que d'une erreur de transmission.
- La perte d'un paquet peut se détecter quand l'émetteur reçoit plusieurs acquittements identiques.
- La perte de plusieurs paquets se détecte quand la temporisation de retransmission se déclenche.

\+ Pour s'adapter au débit, faire varier la taille de la fenêtre d'anticipation.

ALGORITHME DU SLOW START

- Soit la variable seuil initialement à 65 535 et la taille de la fenêtre initialement à 1.
- A chaque réception d'un acquittement, ajouter 1 à la taille de la fenêtre.
  La taille de la fenêtre croît exponentiellement.
- Si la taille de la fenêtre atteint le seuil, incrémenter linéairement sa taille.
- Si un paquet est perdu (réception de trois acquittements identiques) alors :
  taille de la fenêtre = taille de la fenêtre / 2
- Si pertes de plusieurs paquets (déclenchement du timer) alors :
  seuil = taille de la fenêtre / 2 (le seuil ne peut pas être inférieur à 2 segments)
  et
  taille de la fenêtre = 1

EXEMPLE



MÉCANISMES DE CONTRÔLE DE FLUX PAR VAN JAKOBSON

TCP

- Estimation du délai :

$$erreur = mesure - moyenne$$

$$moyenne = moyenne - \alpha \times erreur$$

$$dev = dev + \beta \times (|erreur| - dev)$$

$$RTO = moyenne + 4 \times dev$$

- Slow start :

si pertes de plusieurs paquets (déclenchement du timer) alors :

seuil = taille de la fenêtre / 2 (le seuil ne peut pas être inférieur à 2 segments)

et

taille de la fenêtre = 1

UDP

From list-mgr@ISI.EDU Tue Jan 10 22:35:19 1995
Received: by venera.isi.edu (5.65c/5.61+local-20)
id <AA25275>; Wed, 11 Jan 1995 06:31:42 -0800
Received: by rx7.ee.lbl.gov for mbone-na@isi.edu (5.65/1.44r)
id AA01872; Wed, 11 Jan 95 06:35:21 -0800
Message-Id: <9501111435.AA01872@rx7.ee.lbl.gov>
To: mbone-na
Subject: **jpuckett@168.18.130.249 sendig 600kb/s video**
Date: Wed, 11 Jan 95 06:35:19 PST
From: **Van Jacobson** <van@ee.lbl.gov>

Someone named jpuckett in Georgia (domain peachnet.edu) is sending ridiculous amounts of video (400-600kb/s) to the “Live from Antarctica” session (and most other mbone video sessions).
The reverse dns mapping for this address doesn’t work so I haven’t been able to send email to this individual. **Could someone get in touch with them & explain to them that this behavior is antisocial?** Thanks.

- Van

FONCTIONS D'UN ROUTEUR

- Contrairement aux ponts, les routeurs doivent être configurés.
- Ils doivent connaître les adresses des routeurs ou des stations vers lesquels ils envoient les paquets.
- Ces informations peuvent être entrées manuellement (routage statique) ou acquises par dialogue entre routeurs.
- Dans ce dernier cas, les routeurs peuvent détecter la panne ou l'ajout d'un routeur sur le réseau et modifier leur plan de routage.
- Deux grandes familles d'algorithmes existent :
  - Le distance vector et
  - Le Link State.

LES ALGORITHMES DE ROUTAGE

+ Permettent l'échange automatique de tables de routage.

Table de routage :

- pour aller :
  - réseau,
  - sous-réseau,
  - machine.
- passer par :
  - attachement local,
  - routeur.
- avec un coût de (*métrique*) :
  - nombre de sauts,
  - fonction du débit, délai...

CLASSIFICATION DES ALGORITHMES DE ROUTAGE.

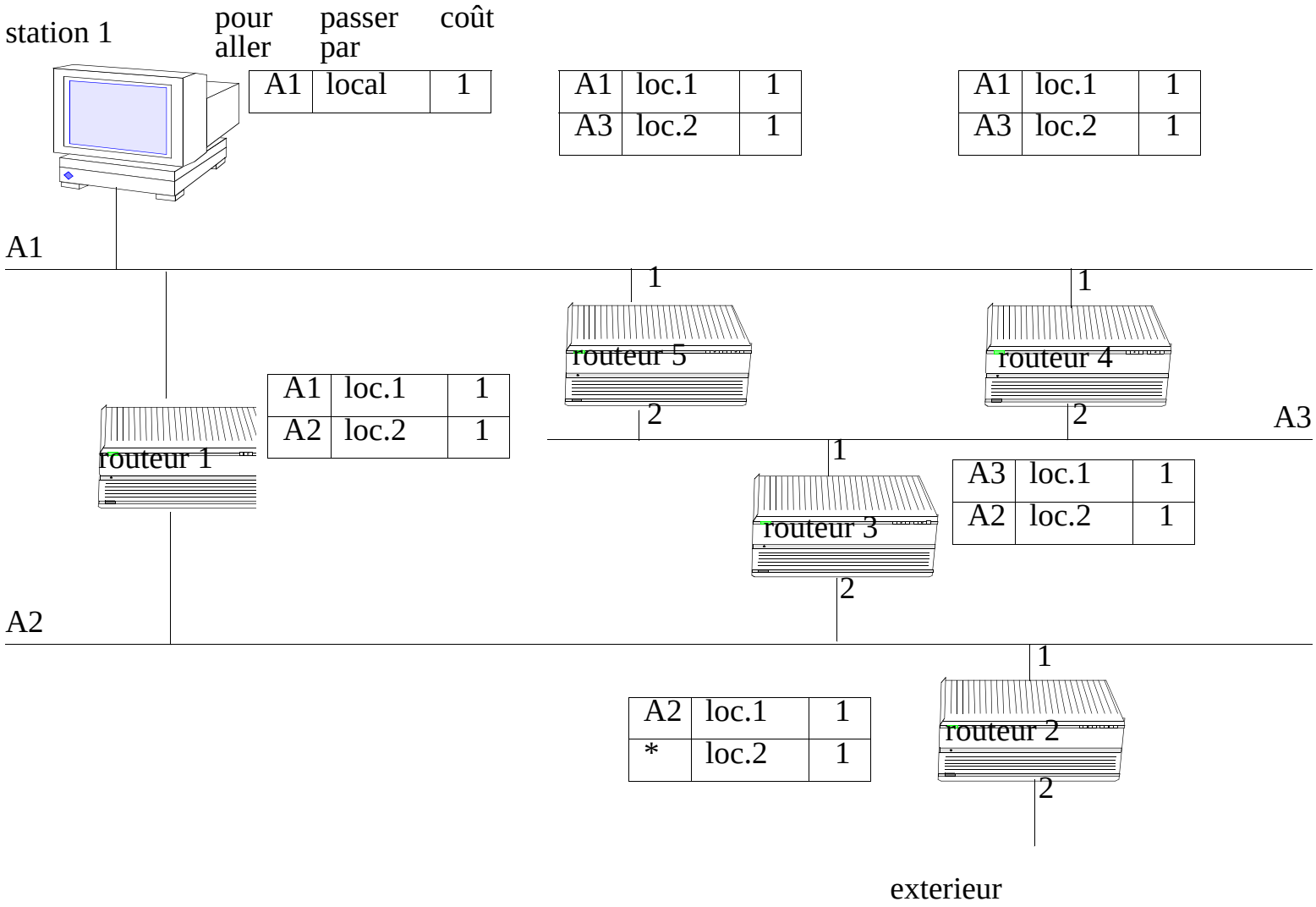
	Monde Internet	Monde ISO
routage intra-domaines (IGP)	distant vector : RIP, RIP-II, IGRP link state :OSPF, DUAL : EIGRP	 IS-IS
routage inter-domaines (EGP)	EGP (obsolète) BGP	 IDRP
Entre routeur et équipement	ICMP redirect	IS-ES

ALGORITHME DU DISTANT VECTOR

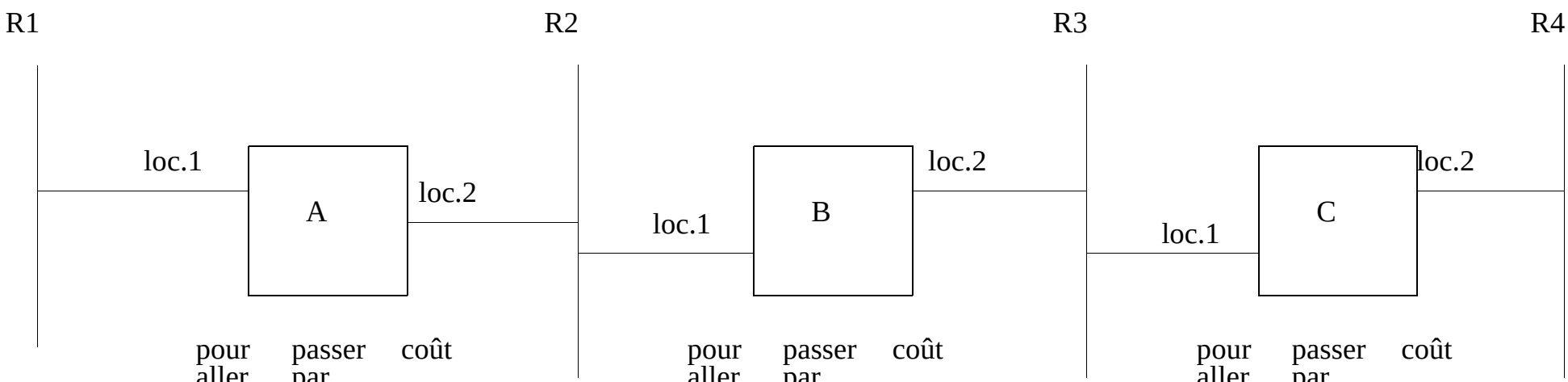
- Echange d'information entre routeurs adjacents.
- Les routeurs diffusent vers les nœuds adjacents leur table de routage rudimentaire constituée de ses différents voisins accessibles et du coût de la liaison.
- Quand un routeur reçoit une nouvelle table, il effectue les traitements suivants pour chaque entrée de la table reçue :
  - Si l'entrée n'est pas dans sa table, il la rajoute.
  - Si le coût de la route proposée par la table plus le coût de la route pour aller jusqu'au routeur (émetteur de la table) est supérieur au coût indiqué dans sa table, sa table de routage est modifiée pour prendre en compte cette nouvelle route.
  - Sinon, il n'y a pas de changement.
- La modification d'une entrée dans la table d'un routeur engendre l'émission de la nouvelle table sur tous les ports du routeur.

\+ Les échanges entres les routeurs continuent jusqu'à ce que l'algorithme converge

EXEMPLE DE DÉROULEMENT DE L'ALGORITHME DU DISTANT VECTOR



PROBLÈME DE CONVERGENCE

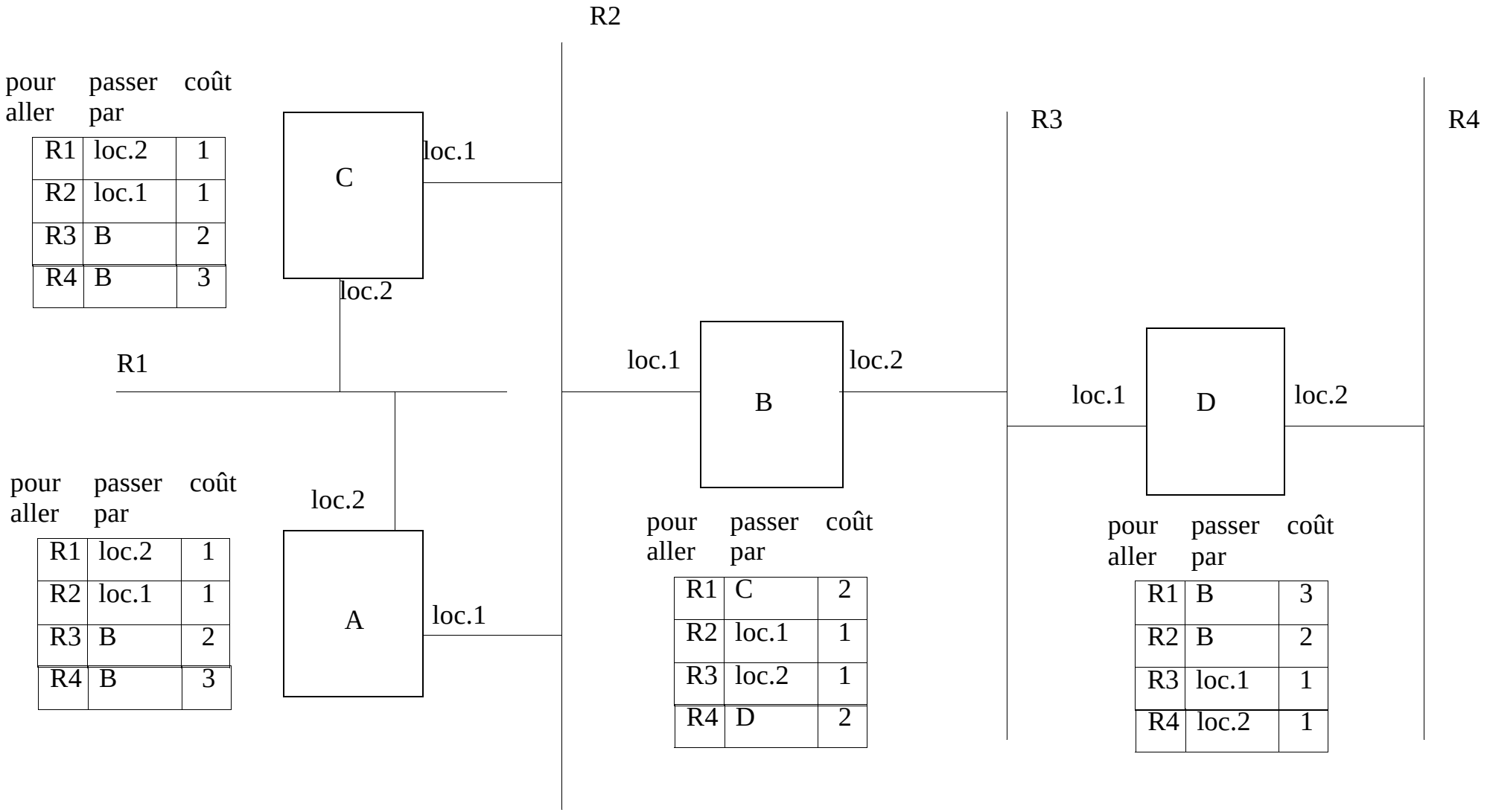


pour aller	passer par	coût
R1	loc.1	1
R2	loc.2	1
R3	B	2
R4	B	3

pour aller	passer par	coût
R1	A	2
R2	loc.1	1
R3	loc.2	1
R4	C	2

pour aller	passer par	coût
R1	B	3
R2	B	2
R3	loc.1	1
R4	loc.2	1

HORIZON COUPÉ



pour aller	passer par	coût
R1	loc.2	1
R2	loc.1	1
R3	B	2
R4	B	3

pour aller	passer par	coût
R1	loc.2	1
R2	loc.1	1
R3	B	2
R4	B	3

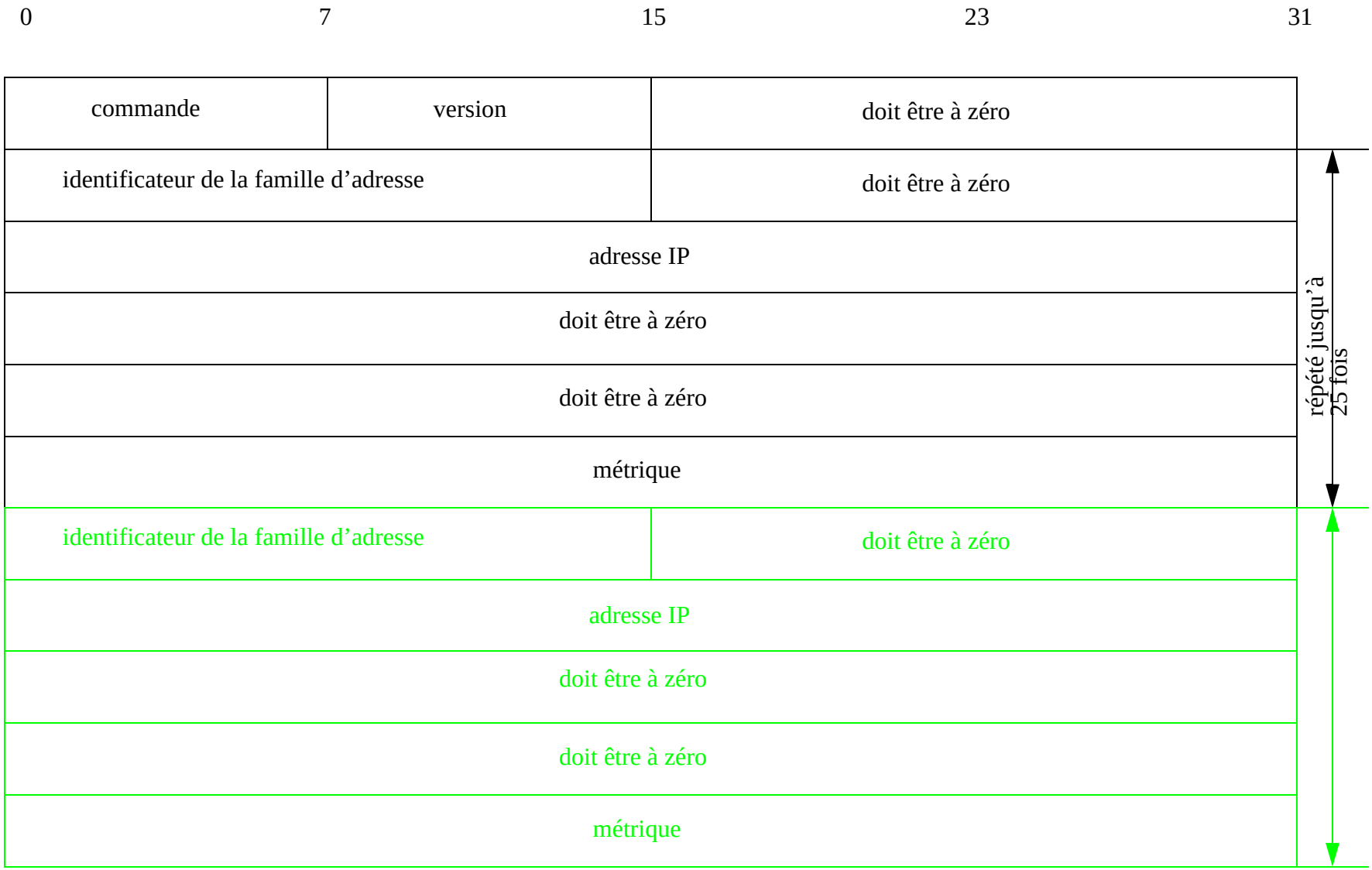
pour aller	passer par	coût
R1	C	2
R2	loc.1	1
R3	loc.2	1
R4	D	2

pour aller	passer par	coût
R1	B	3
R2	B	2
R3	loc.1	1
R4	loc.2	1

RIP

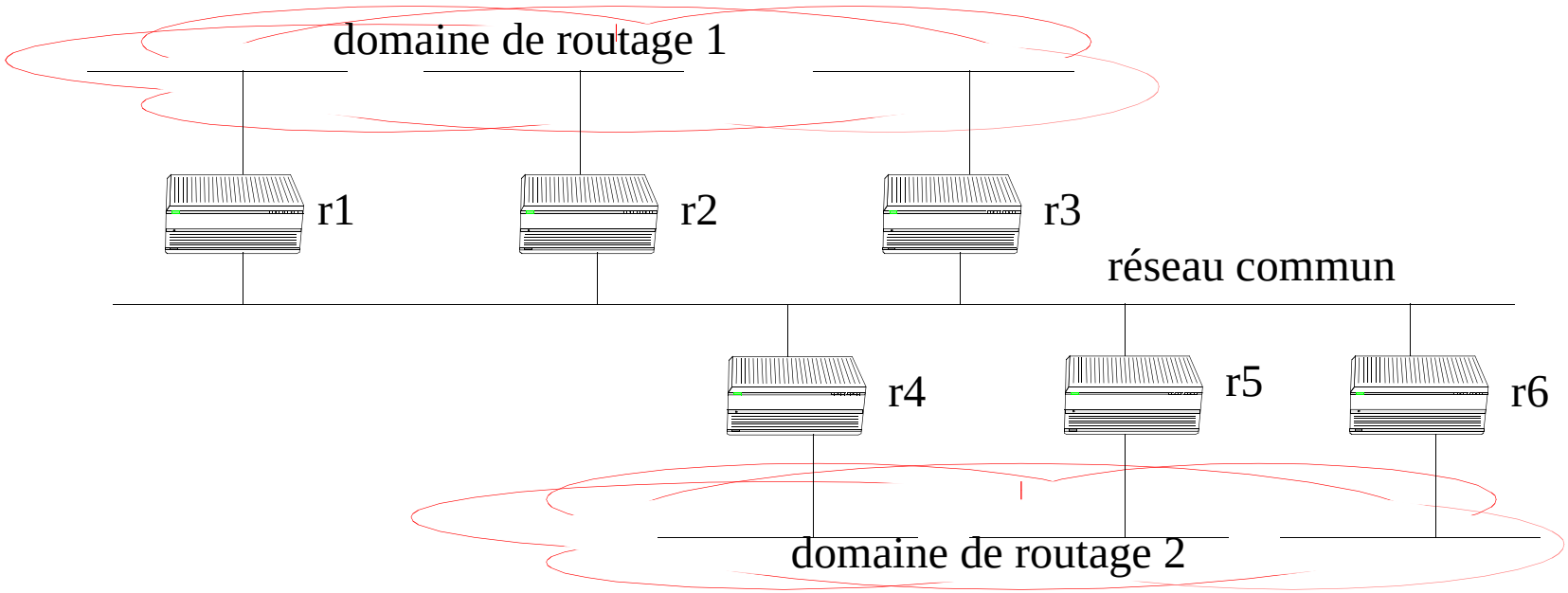
- Algorithme de type Distant Vector
- Implémenté sur la plupart des systèmes (sur BSD : routed)
- Limité à 15 routeurs
- Au dessus de UDP, port 520
- A servi de base pour de nombreux autres protocoles
  - AppleTalk, Novel / IPX, Banyan Vines ...
- Broadcasts toutes les 30 secondes

FORMAT DES PAQUETS



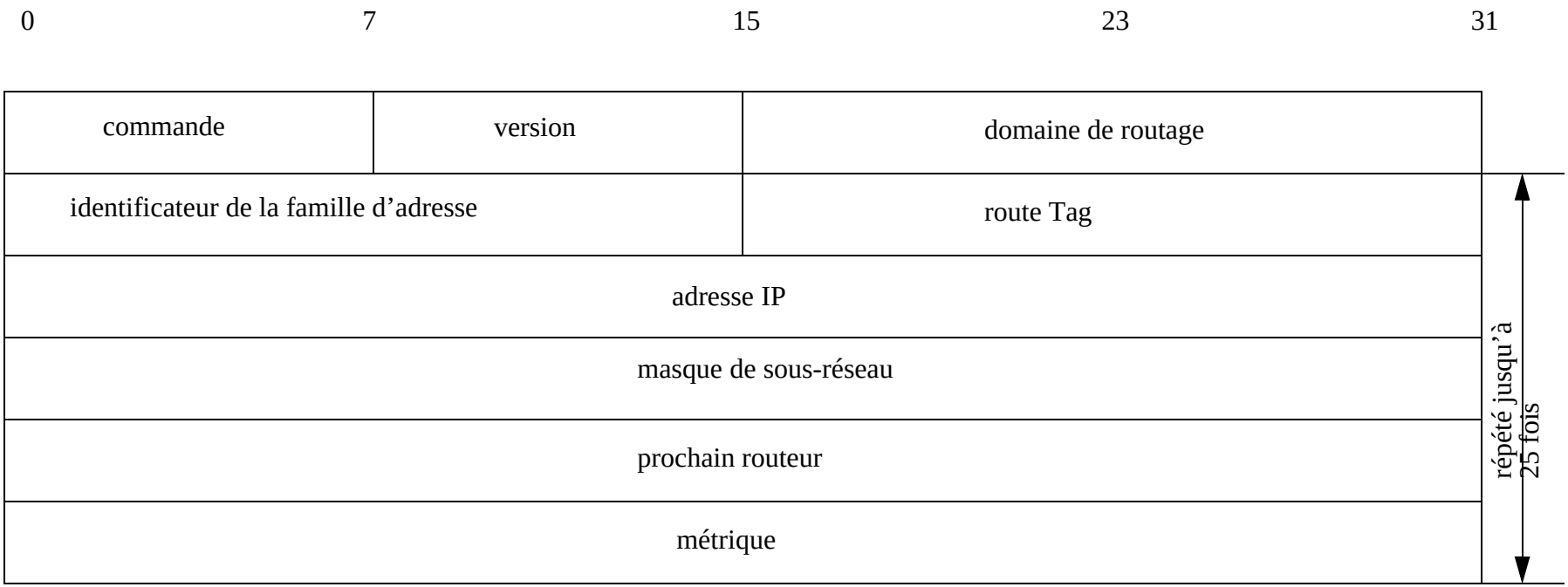
LES EXTENSIONS DE RIP

- RIP-2 (Janvier 1993)
  - Utilisation d'une adresse de Multicast (224.0.0.9)
  - Prise en compte des *netmask*
  - Plusieurs plans de routage (*route tag*)



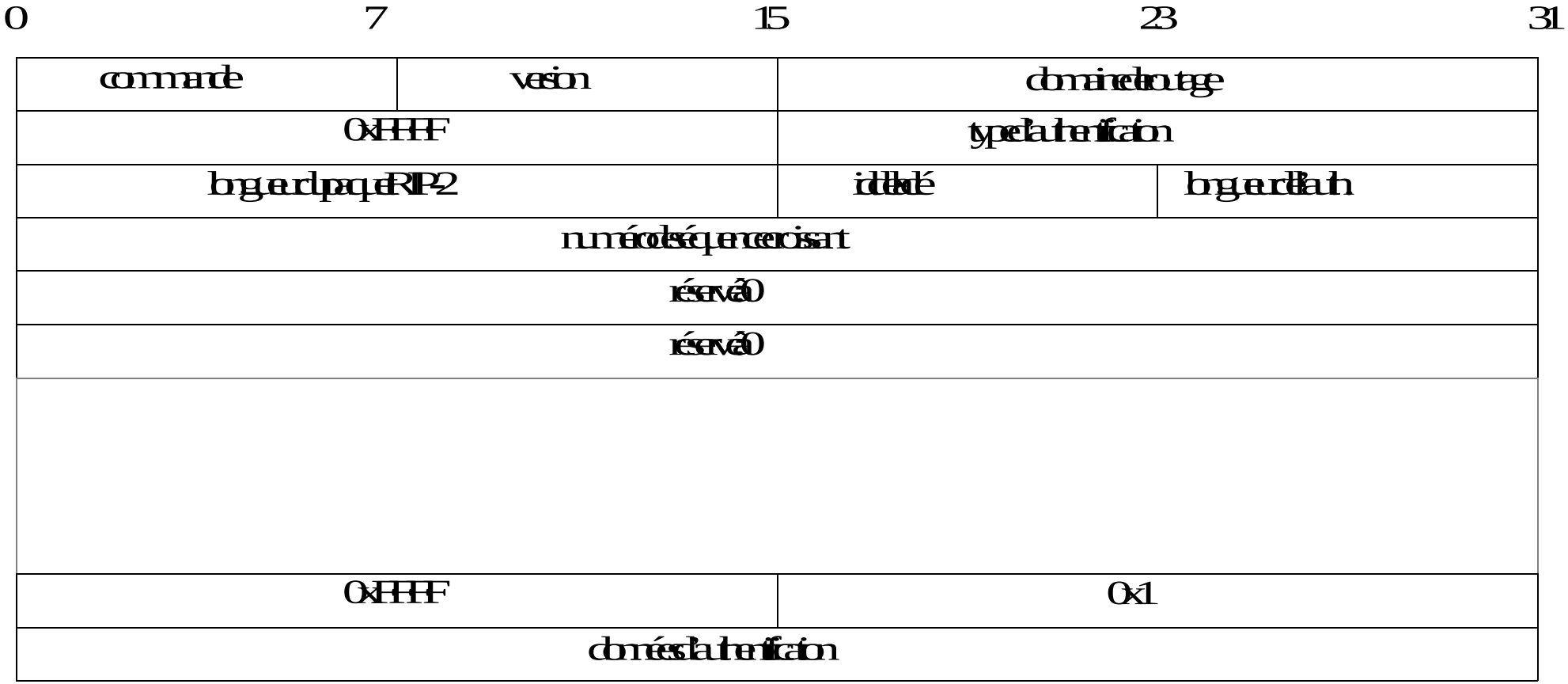
  - Le champ prochain routeur indique la destination qui sera insérée dans la table de routage du destinataire.

RIP-II


0
7
15
23
31
commande
version
domaine de routage
identificateur de la famille d'adresse
route Tag
adresse IP
masque de sous-réseau
prochain routeur
métrique
répété jusqu'à 25 fois

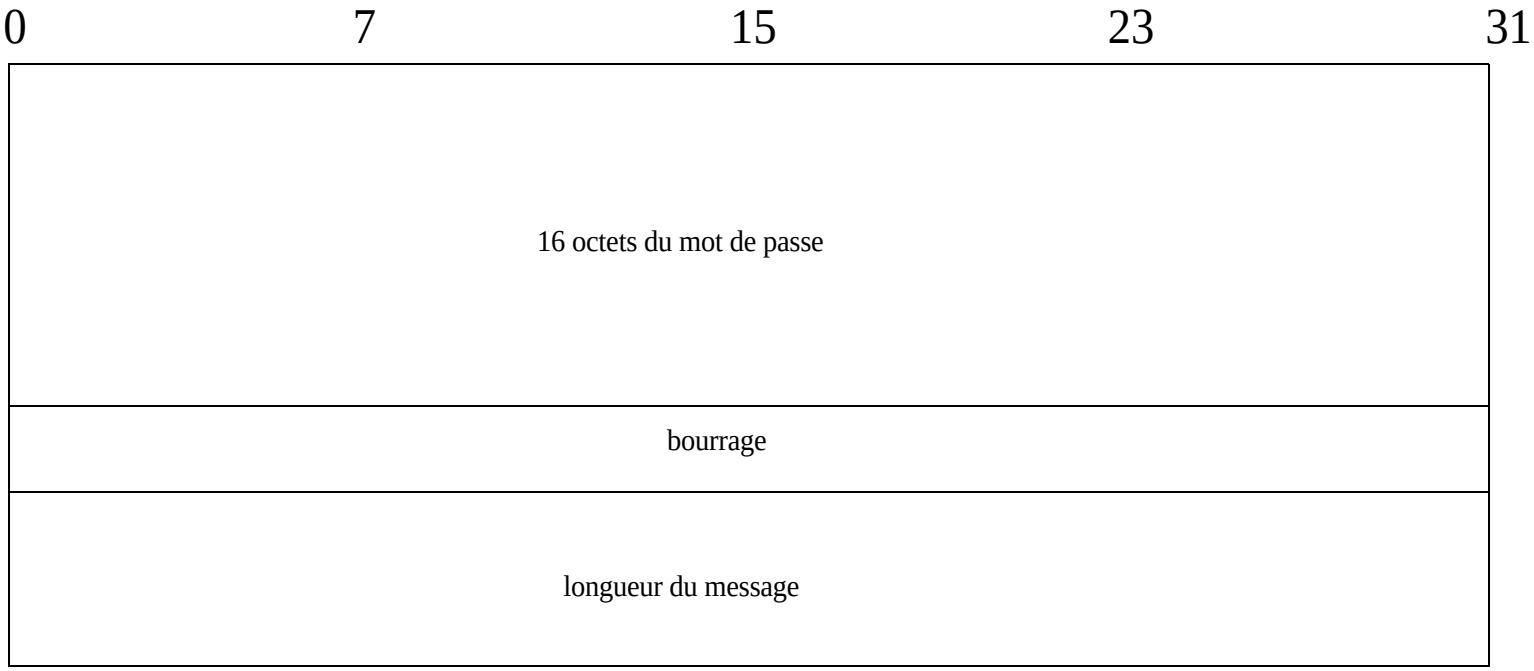
AUTHENTIFICATION

- Authentification simple : mot de passe
  - Peu fiable (découverte par écoute du réseau)
  - Permet surtout de se prémunir contre les équipements mal configurés



AUTHENTIFICATION PAR SECRET COMMUN

- Authentification pas sceau MD5
  - Contre le rejeu
  - Le mot de passe (secret) n'est jamais transmis sur le réseau
  - Les champs suivants sont ajoutés lors du calcul du sceau.



INTERIOR GATEWAY ROUTING PROTOCOL (IGRP)

- Propriétaire : Cisco
- Broadcasts de mise à jour toutes les 90 secondes
- Mesures de protection contre les boucles de routage
- Routage des sous-réseaux
- Routage multi-chemins (secours + load balancing)
  - La table de routage peut avoir plus d'un chemin pour atteindre une même destination
  - Chaque chemin = adresse_prochain_routeur, N°_interface
  - Avantages :
    - secours automatique si un chemin devient indisponible
    - partage du trafic entre les chemins de métriques voisines
- Gestion des routes par défaut (candidates)

IGRP (SUITE)

- La métrique composite d'IGRP
  - la métrique tient compte de :
  - la bande passante (B)
  - le délai de propagation (D)
  - la charge de la liaison (C)
  - et sa fiabilité (F)
    => B et D sont des caractéristiques statiques de la liaison.
    => C et F sont des variables dynamiques calculées par le routeur.
- Avantages :
  - affiner le routage en fonction des besoins, des côuts des ressources, de la redondance ...
  - en faisant varier les composantes de la métrique.
- Evolution :
  - Extended IGRP (E-IGRP)

OSPF (Open Shortest Path First)

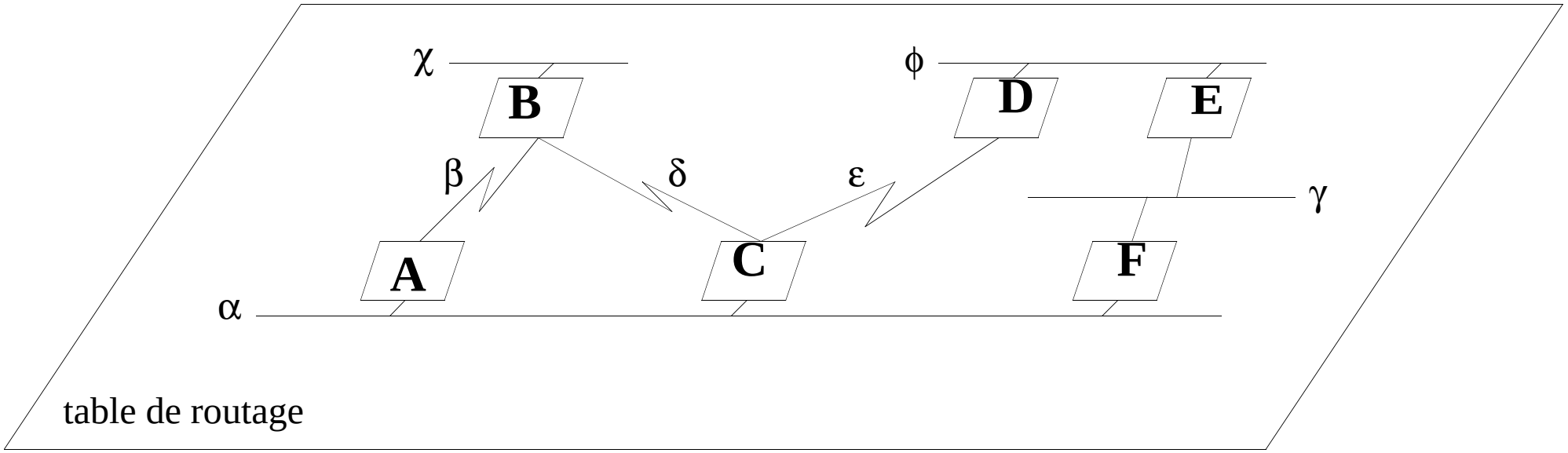
+ Pour de grands réseaux,
+ Force à structurer le réseau,
+ Poids attribuable aux liens, généralement fonction du débit :

$$\text{coût} = \frac{10^8}{\text{bande passante en b/s}}$$

  - Pour un Ethernet à 10 Mbit/s le coût est de 10

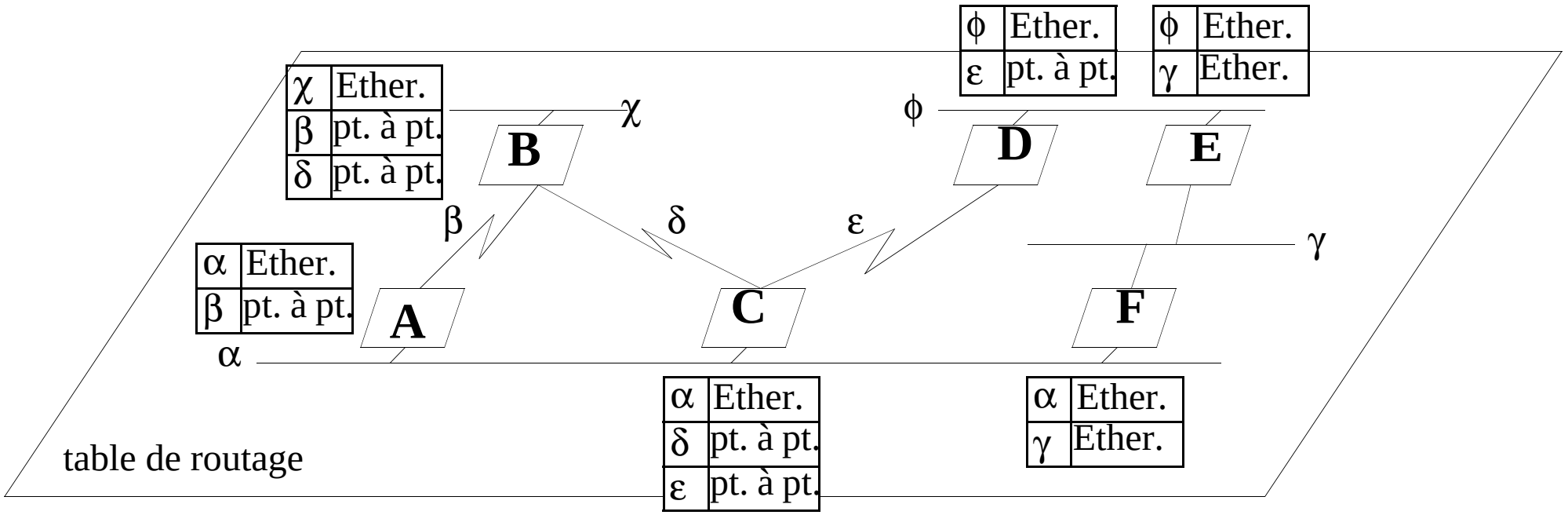
+ Utilise l'algorithme du Link State

EXEMPLE



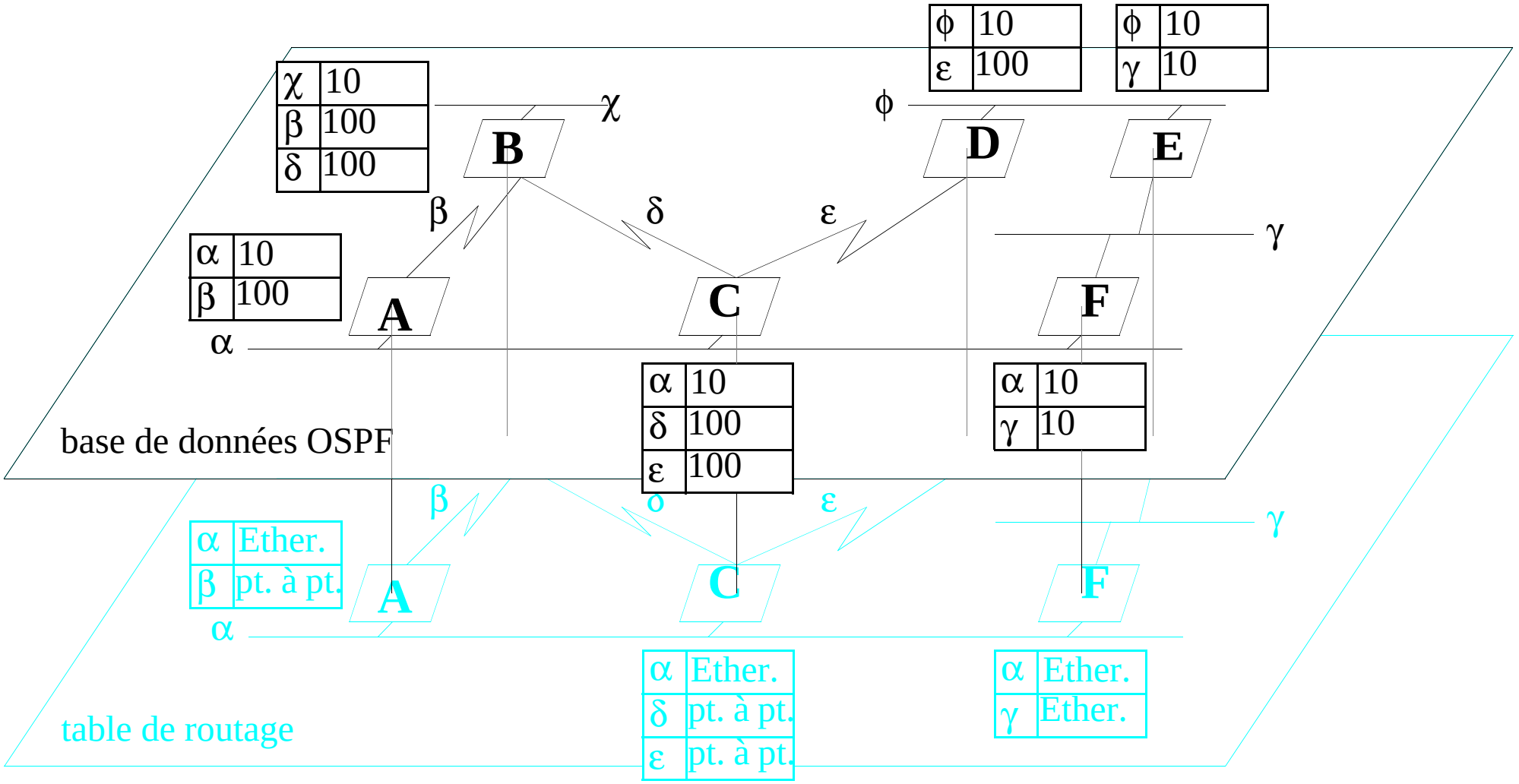
- Chaque routeur connaît uniquement les réseaux auxquel il est connecté

TABLES DE ROUTAGE INITIALES



- Ces informations sont recopiées dans la base de donnée OSPF du routeur avec le coût choisi par l'administrateur

BASE DE DONNÉES OSPF



- Ces informations sont recopiées dans la base de donnée OSPF du routeur avec un coût.

INONDATION

- Un mécanisme d'inondation fiable permet de faire partager à tous les routeurs les informations de chaque routeur
- Après un temps de convergence, chaque routeur dispose des mêmes informations :

A

α	10
β	100

B

χ	10
β	100
δ	100

C

α	10
δ	100
ε	100

D

ϕ	10
ε	100

E

ϕ	10
γ	10

F

α	10
γ	10

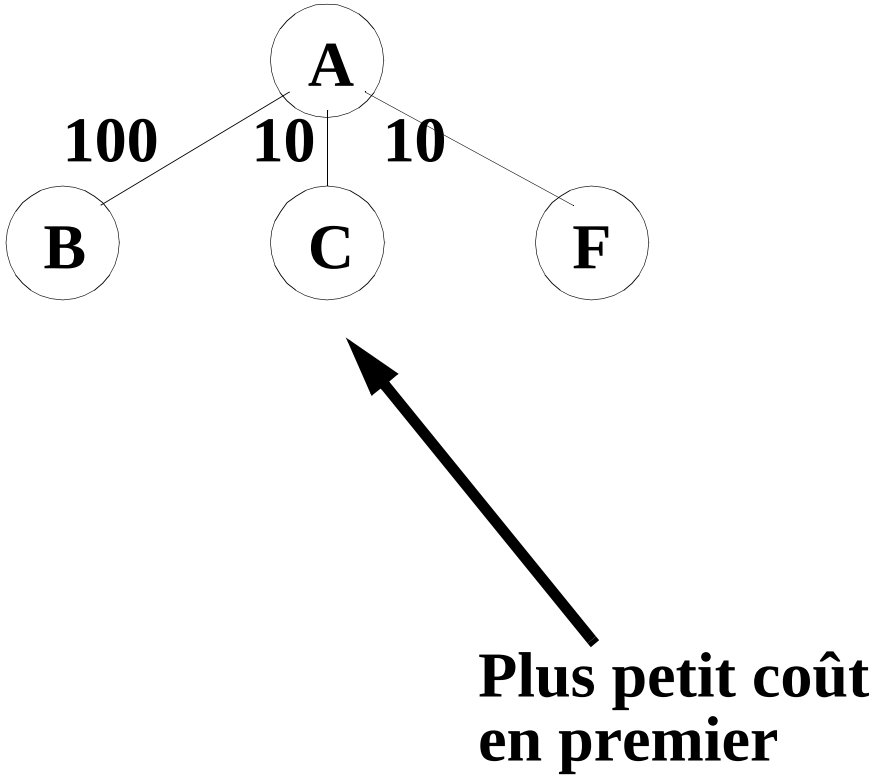
- Chaque routeur exécute l'algorithme du plus court chemin (SPF : Short Path First) pour trouver le plus court chemin pour aller des réseaux de son routeur vers les autres réseaux.

ALGORITHME DU SHORT PATH FIRST

Chaque routeur calcule sa table de routage à la réception de paquets LSP. A l'initialisation mettre le routeur dans la variable PATH puis :

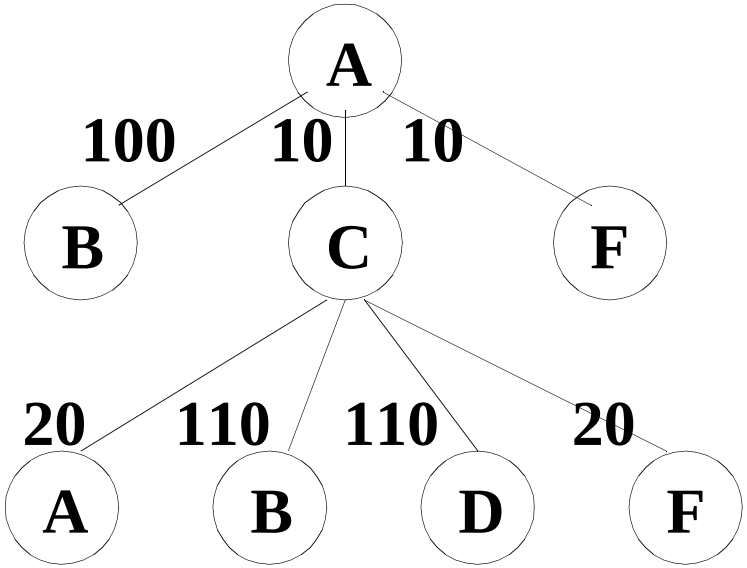
- Pour chaque routeur N contenu dans la structure PATH, examiner les LSP (c'est-à-dire les voisins immédiats) de N.
  - Pour chaque voisin M de N ajouter le coût de la liaison de la racine jusqu'à N au coût de la liaison de N à M.
  - Si M n'est ni dans la structure PATH ni dans la structure TENT avec un meilleur coût, insérer M avec le coût calculé et la direction N dans TENT.
- Si TENT est vide, l'algorithme est terminé, sinon prendre dans TENT l'entrée qui a le coût minimum, la mettre dans PATH et redérouler l'algorithme au début.

EXEMPLE POUR LE ROUTEUR A



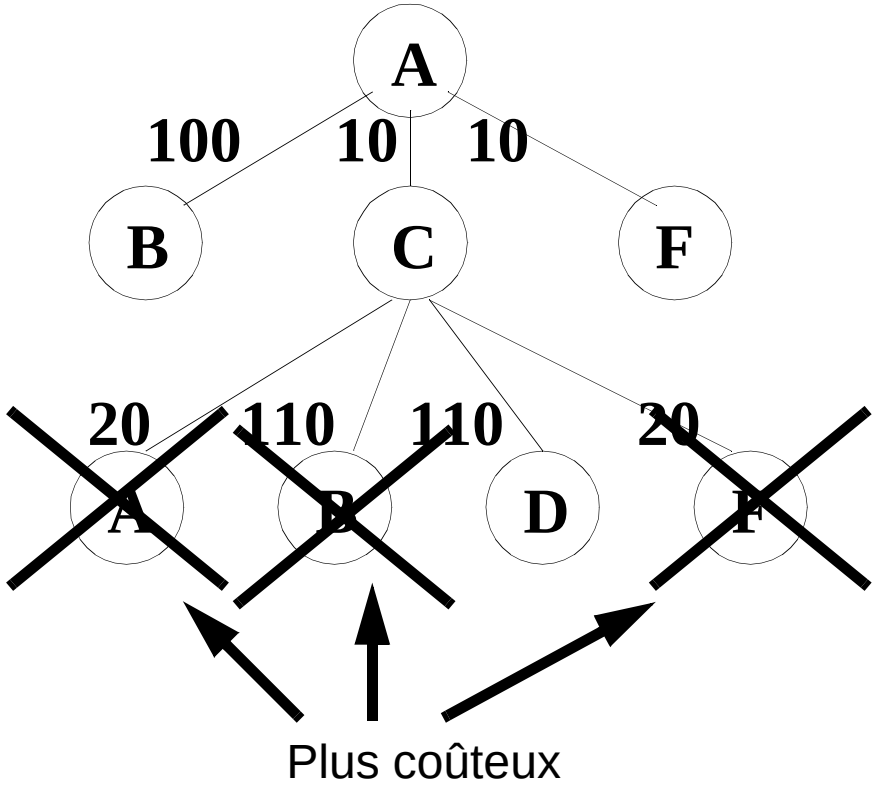
- A regarde les coûts pour ses voisins immédiats

EXEMPLE POUR LE ROUTEUR A

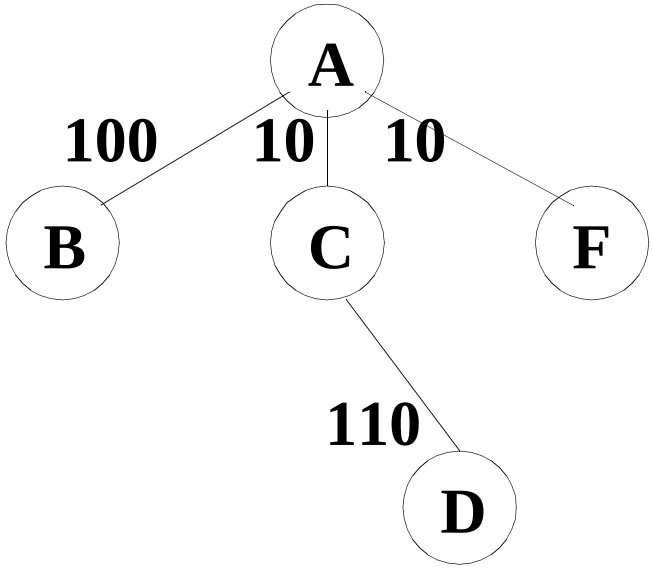


- A regarde les coûts cumulés pour les voisins immédiats de C

EXEMPLE POUR LE ROUTEUR A

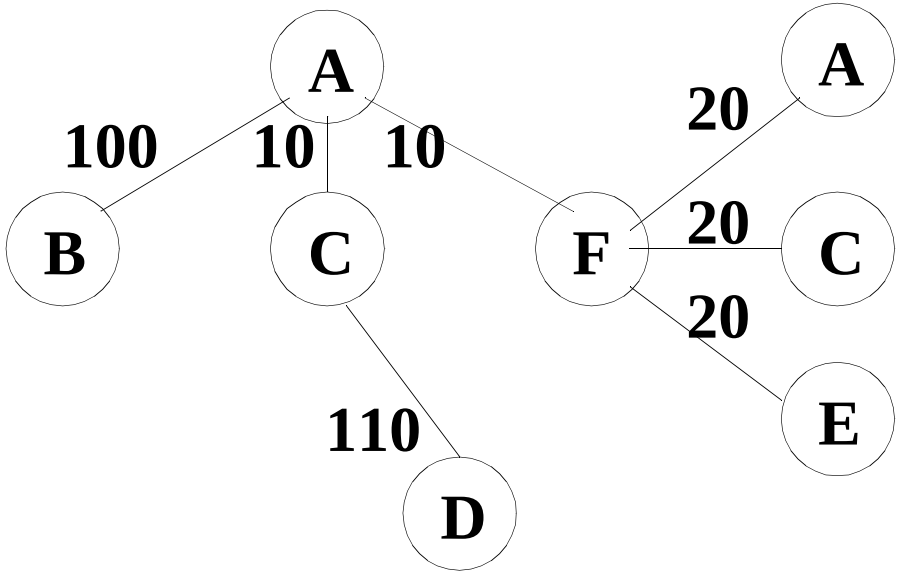


EXEMPLE POUR LE ROUTEUR A

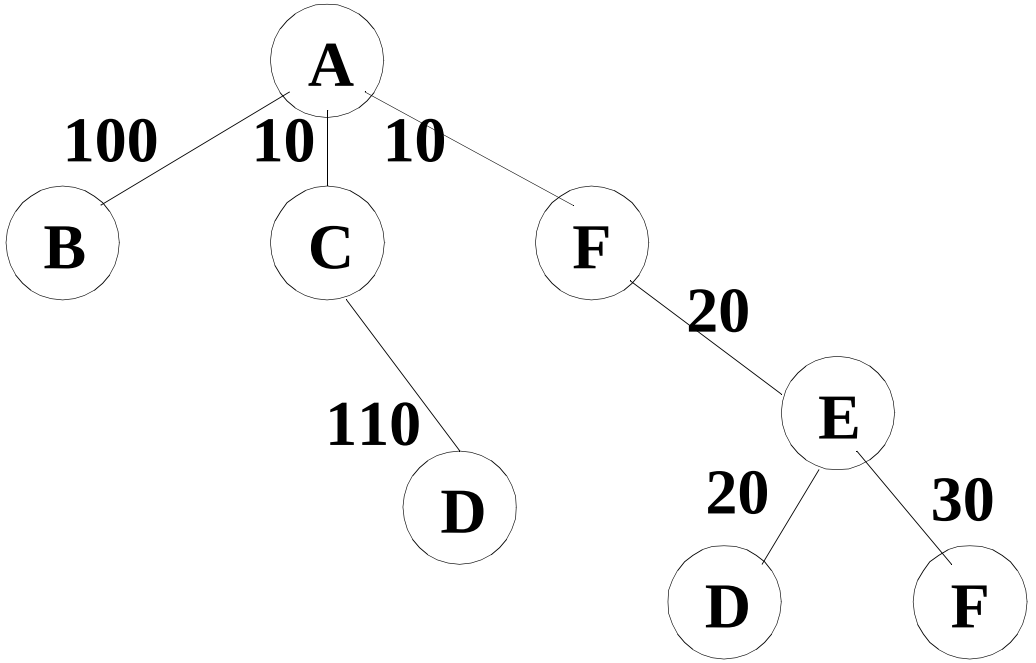


- A s'intéresse maintenant au plus petit chemin de l'arbre non encore exploré
- A calcule des coûts cumulés pour les voisins immédiats de F

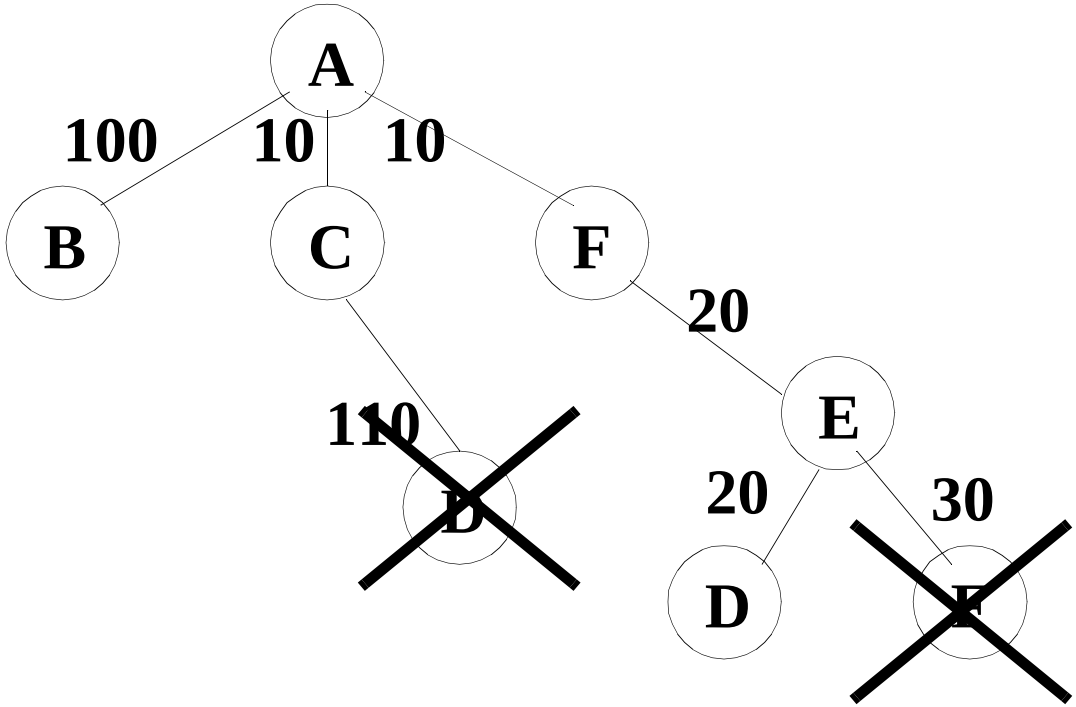
EXEMPLE POUR LE ROUTEUR A



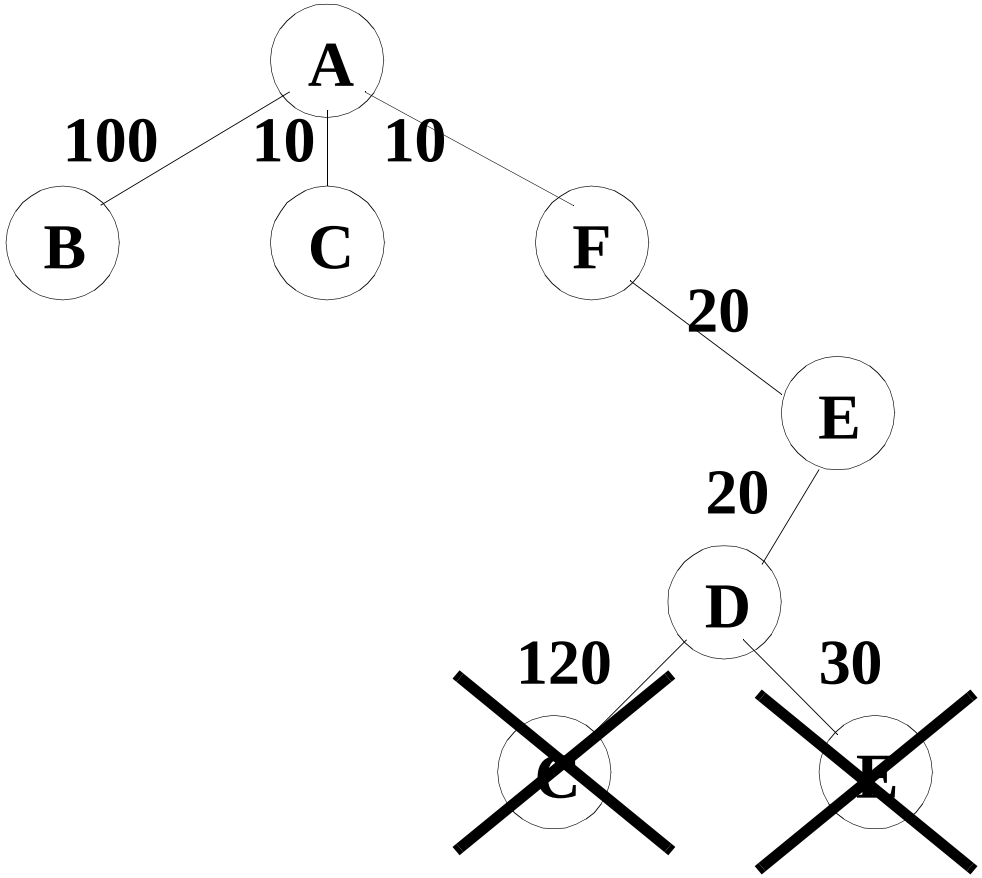
EXEMPLE POUR LE ROUTEUR A



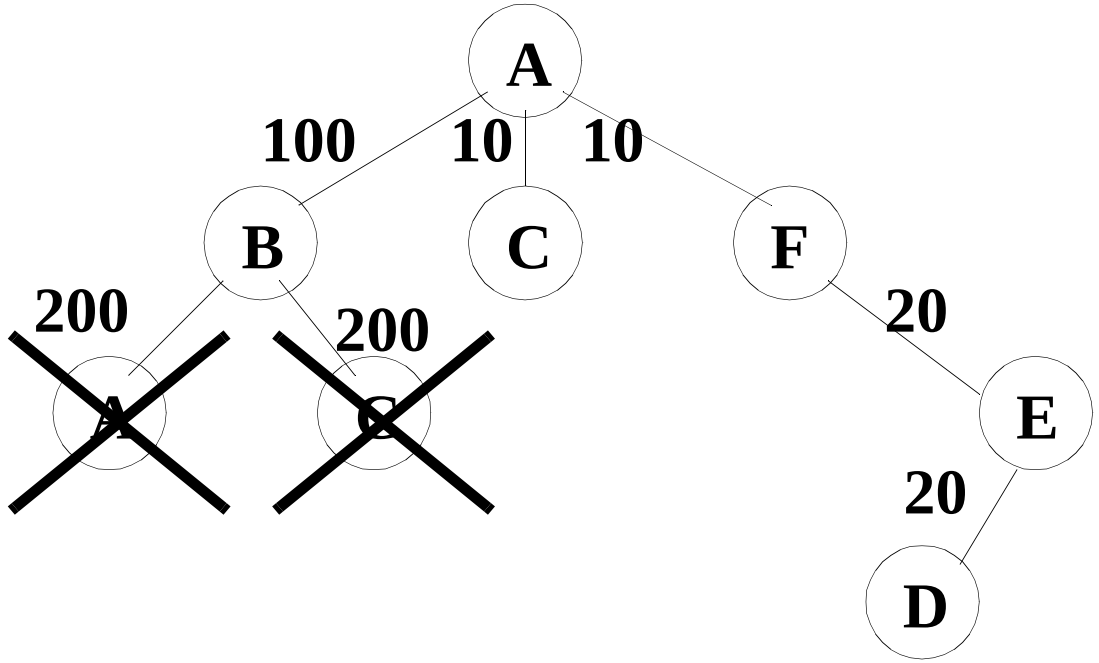
EXEMPLE POUR LE ROUTEUR A



EXEMPLE POUR LE ROUTEUR A

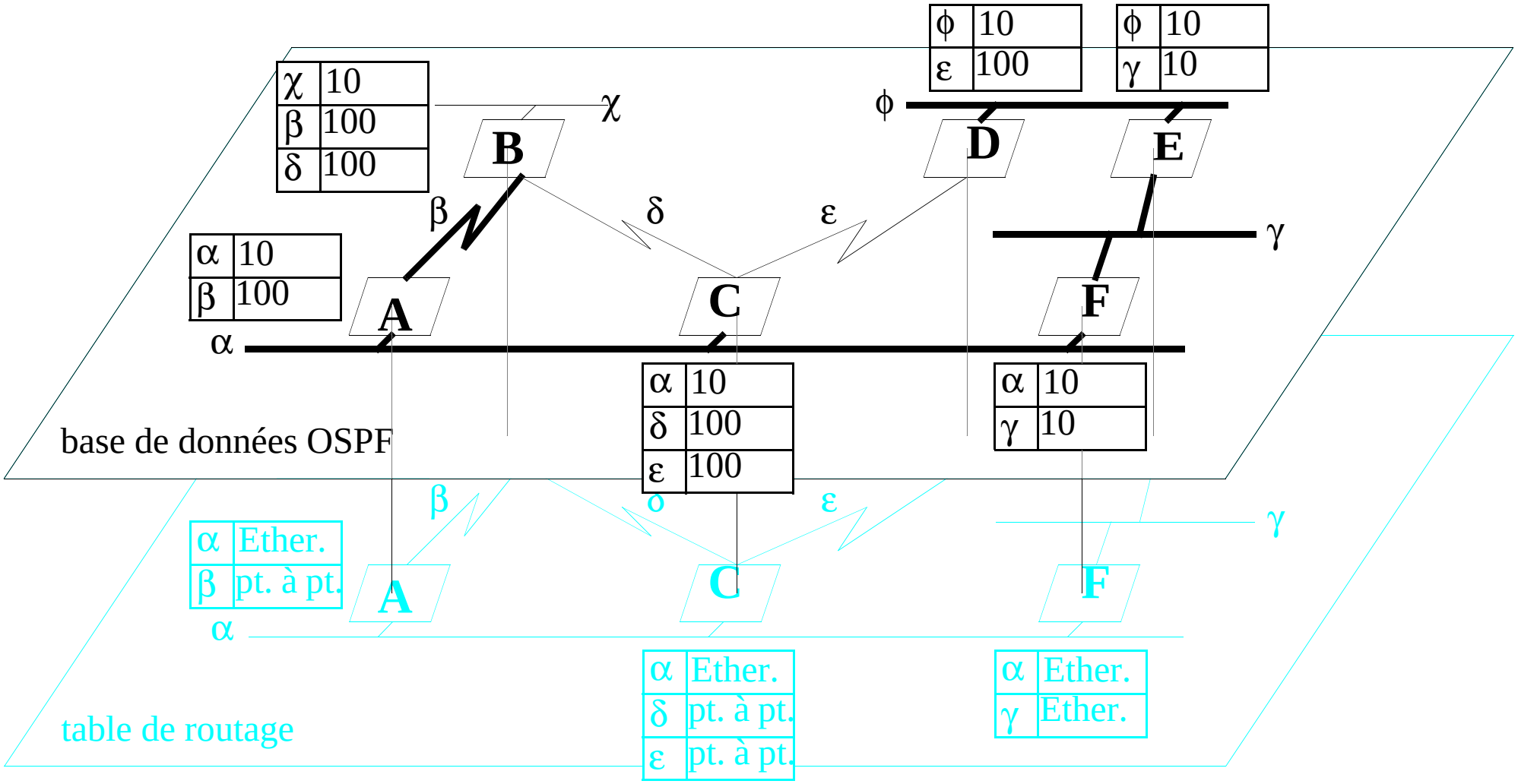


EXEMPLE POUR LE ROUTEUR A



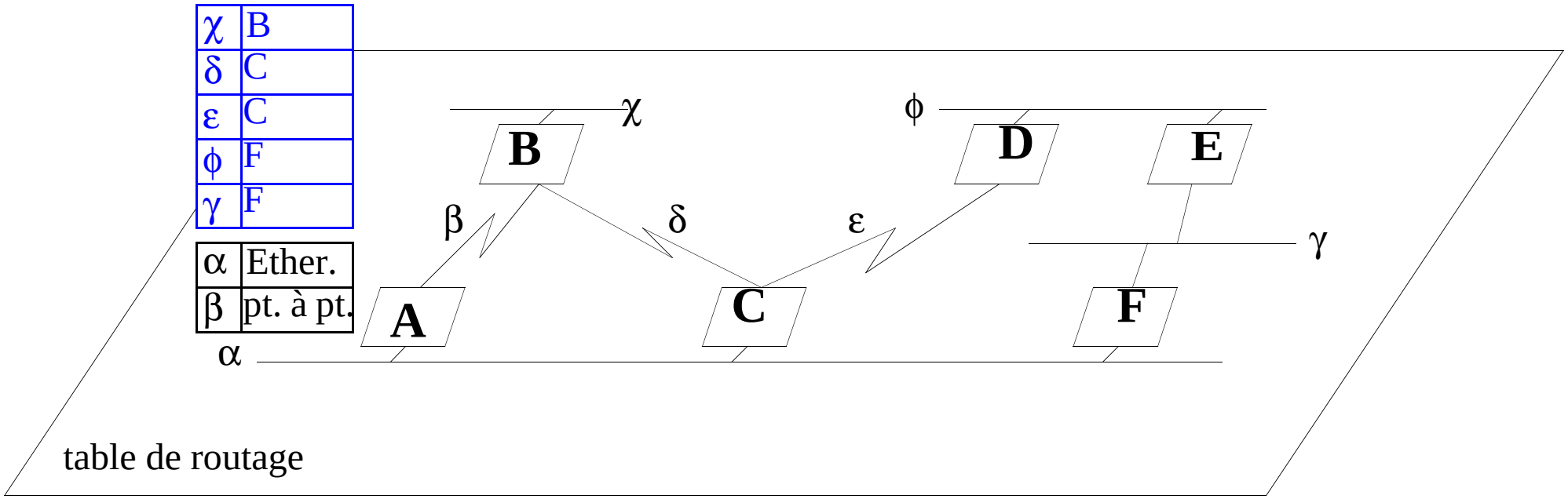
- A en déduit l'arbre des plus courts chemins

ARBRE DES PLUS COURTS CHEMINS POUR A



- A peut donc calculer sa table de routage

MISE À JOUR DE LA TABLE DE ROUTAGE DE A



? Faites la même chose pour le routeur F

STRUTURATION DU RÉSEAU

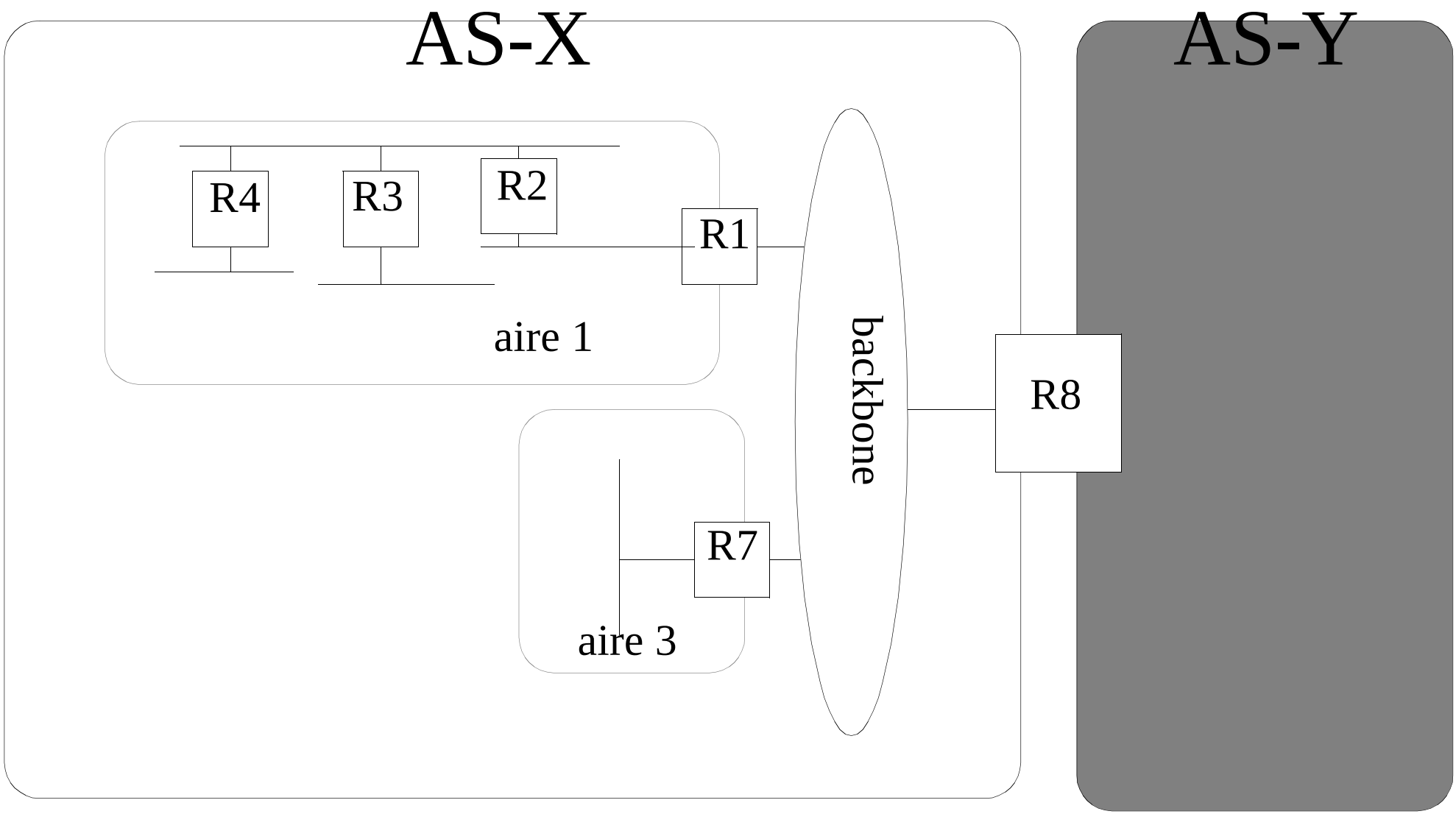
Principes :

- Faire connaître à tous les routeurs les visions locales
  - Protocole d'inondation évitant les boucles,
  - Permettre les mises à jours quand l'état d'un routeur évolue.
- Calculer une route en fonction de cette vue générale
  - Algorithme du link-state
  - Algorithme coûteux en temps de calcul
  - Doit être relancé après chaque modification d'un routeur

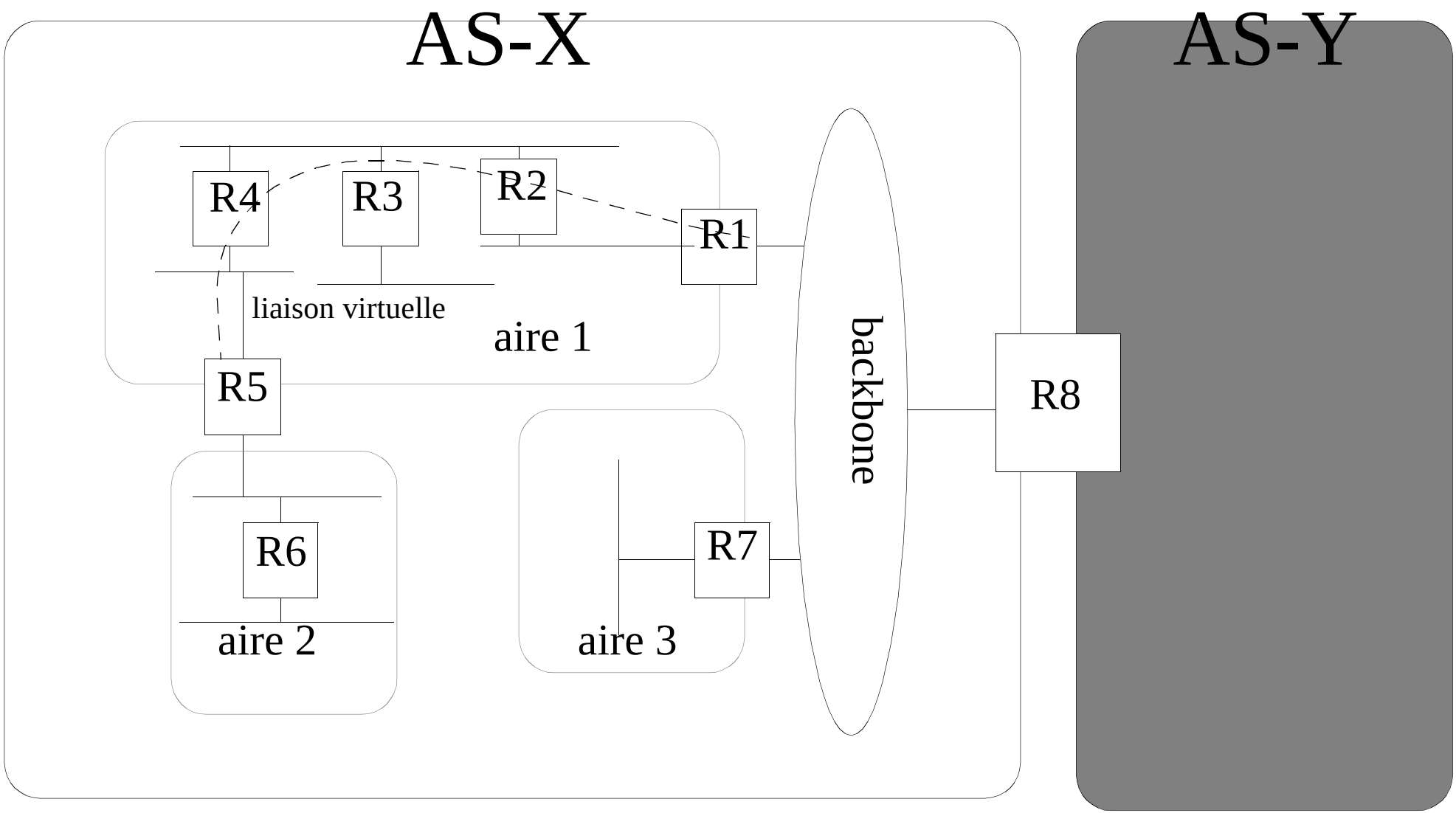
\+ Pour éviter des instabilités et un trop grand échange d'informations

- Division en aires : moins d'instabilités, moins d'échange de routes
- Définition d'un plan d'adressage conforme : possibilité d'agrégation

DIVISION EN AIRE

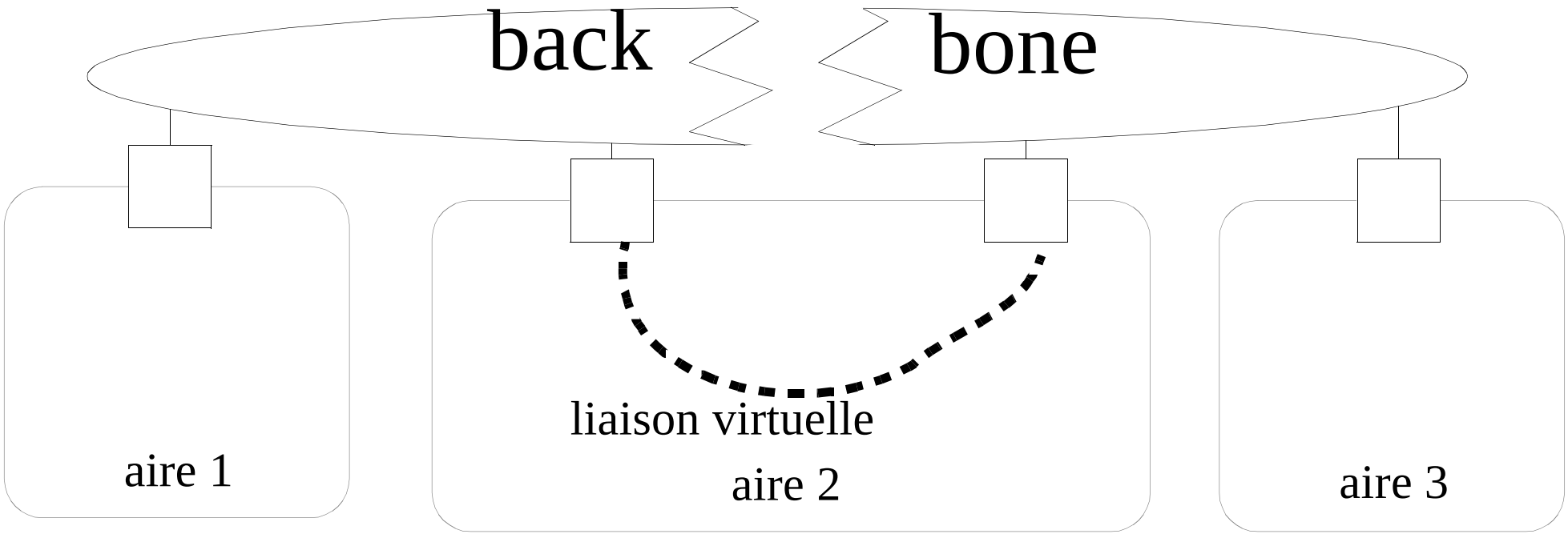


LIAISON VIRTUELLE



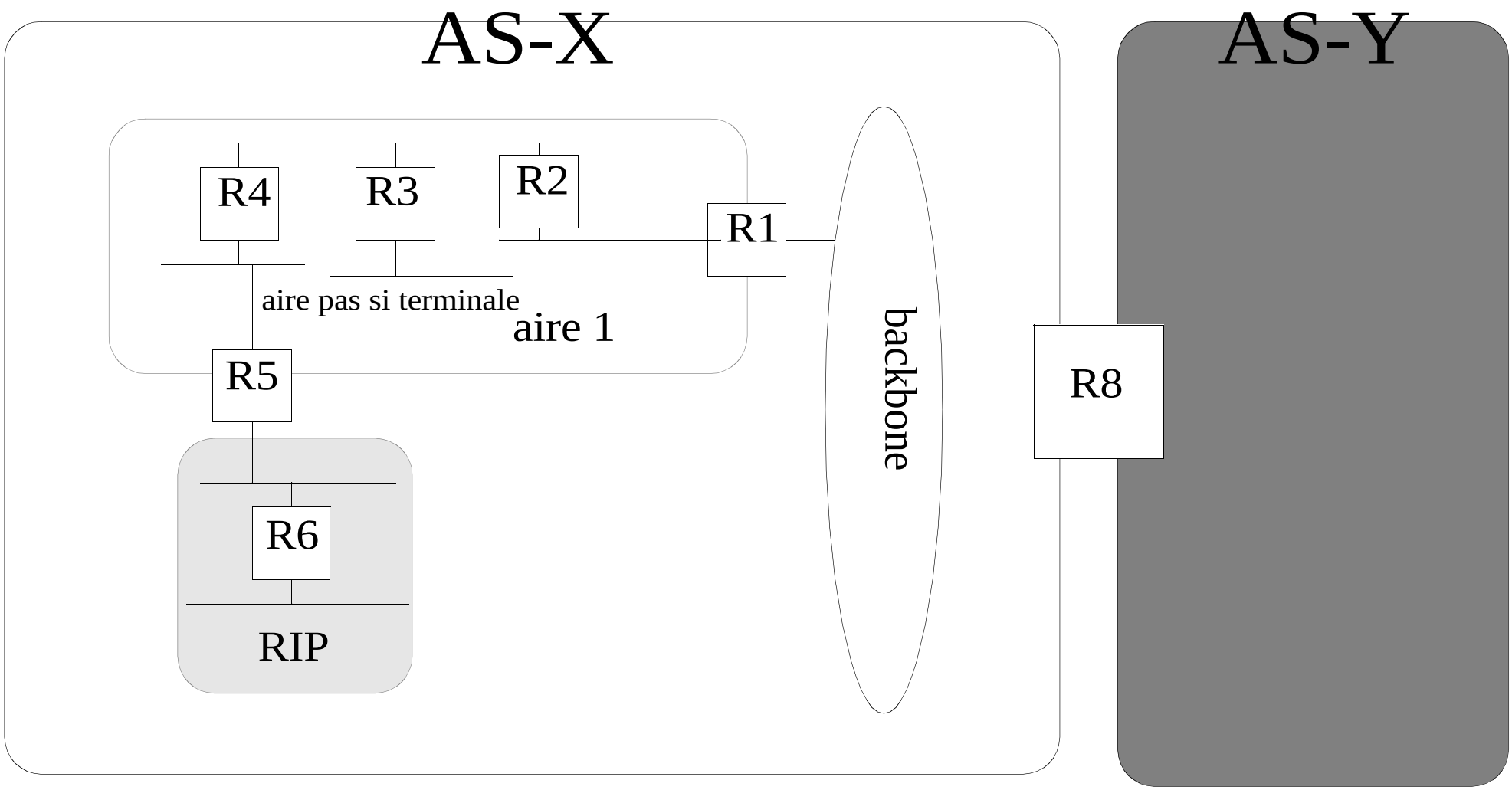
- A déconseiller à cause du manque de stabilité des routes

LIAISON VIRTUELLE



- Le backbone doit toujours être connexe
- Le lien virtuel peut servir à le réparer

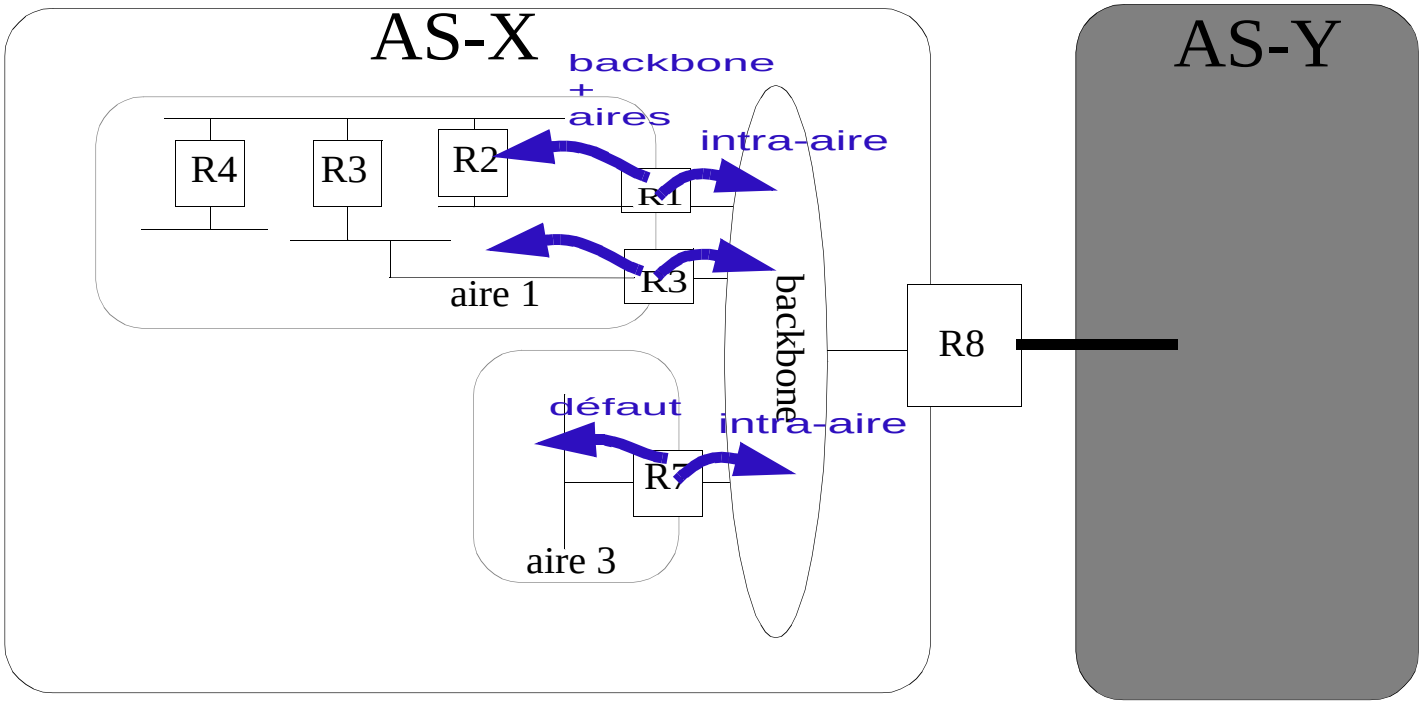
AIRE PAS SI TERMINALE



- Mais qui utilise encore RIP dans ce genre de configuration ?

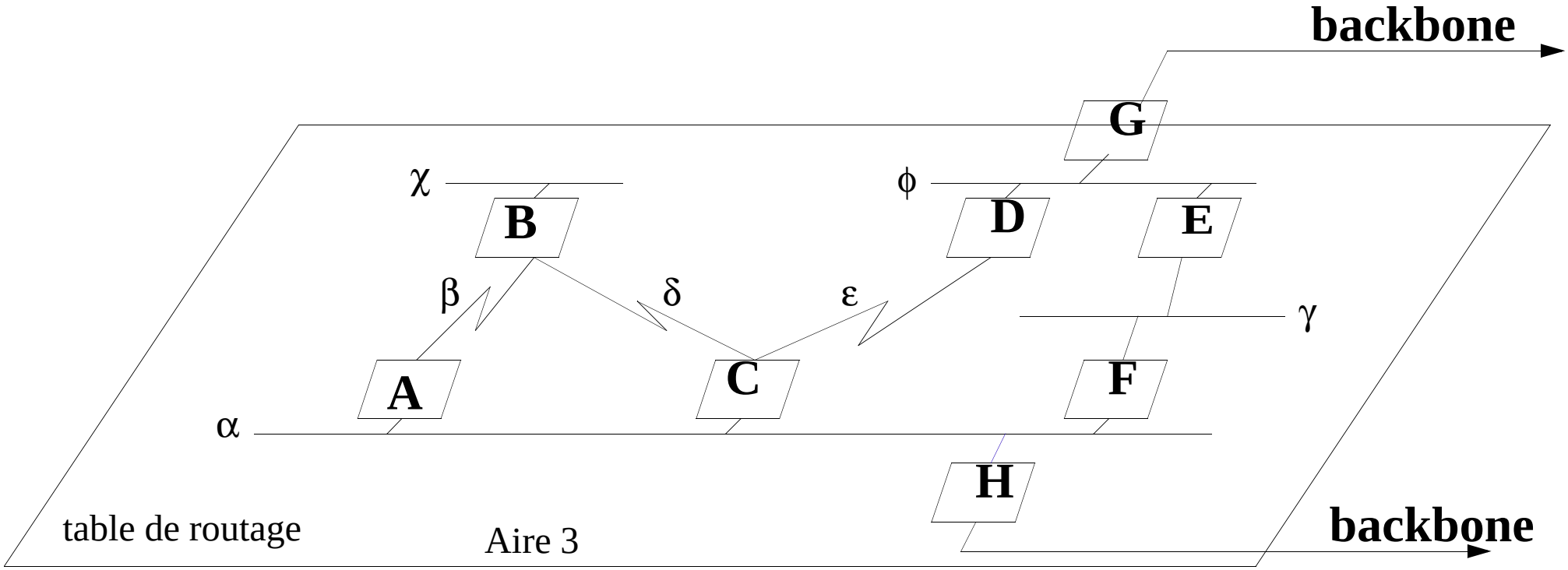
ROUTES ENTRE LES AIRES

- Des réseaux dans les autres aires :
  - Contient la liste de tous les réseaux contenus dans une aire.
  - Produit uniquement par le routeur en frontière d'Aire
  - Les routeurs à l'intérieur de l'aire ajoutent ces informations à leur table de routage



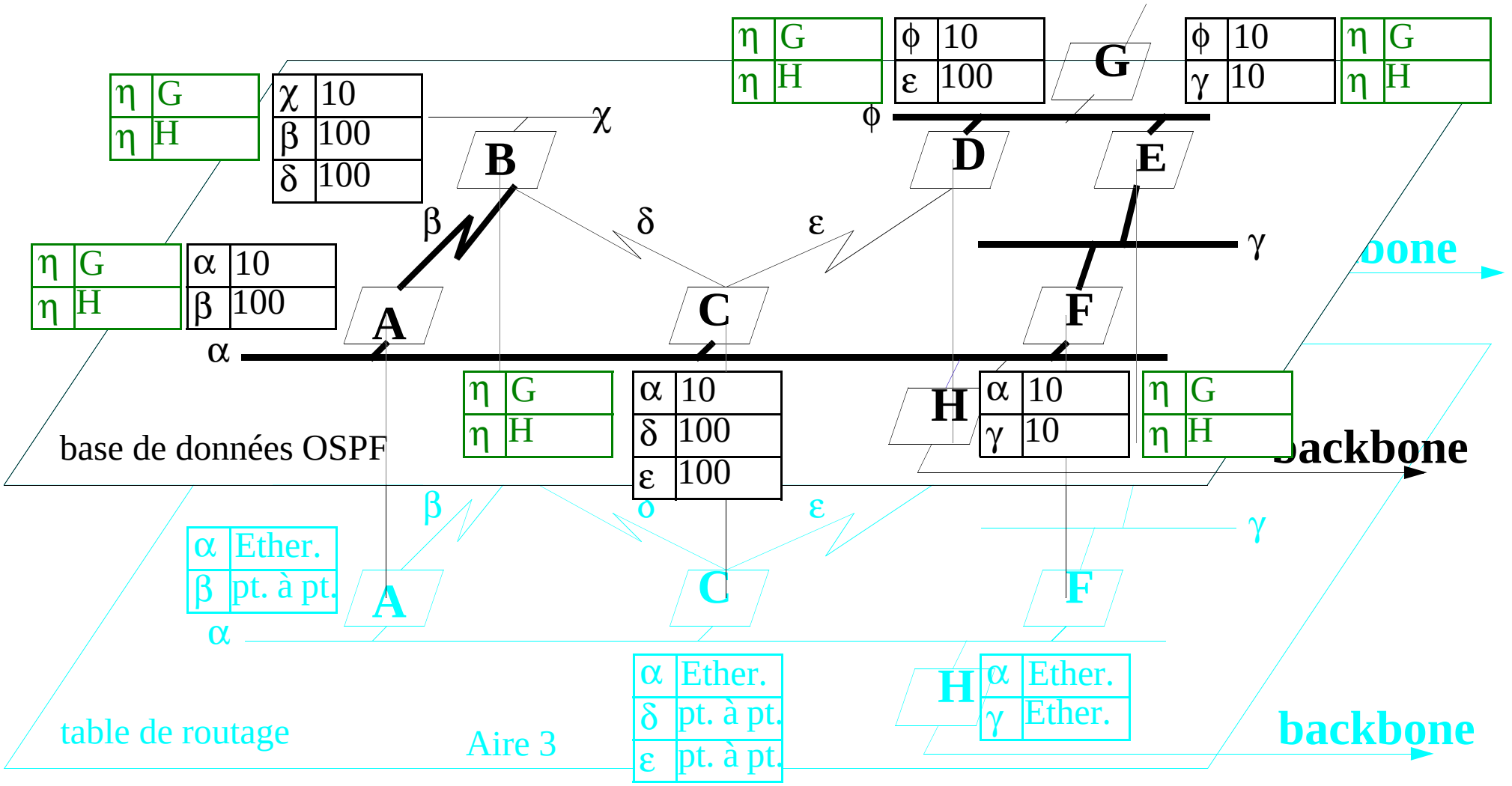
EXEMPLE

η : backbone



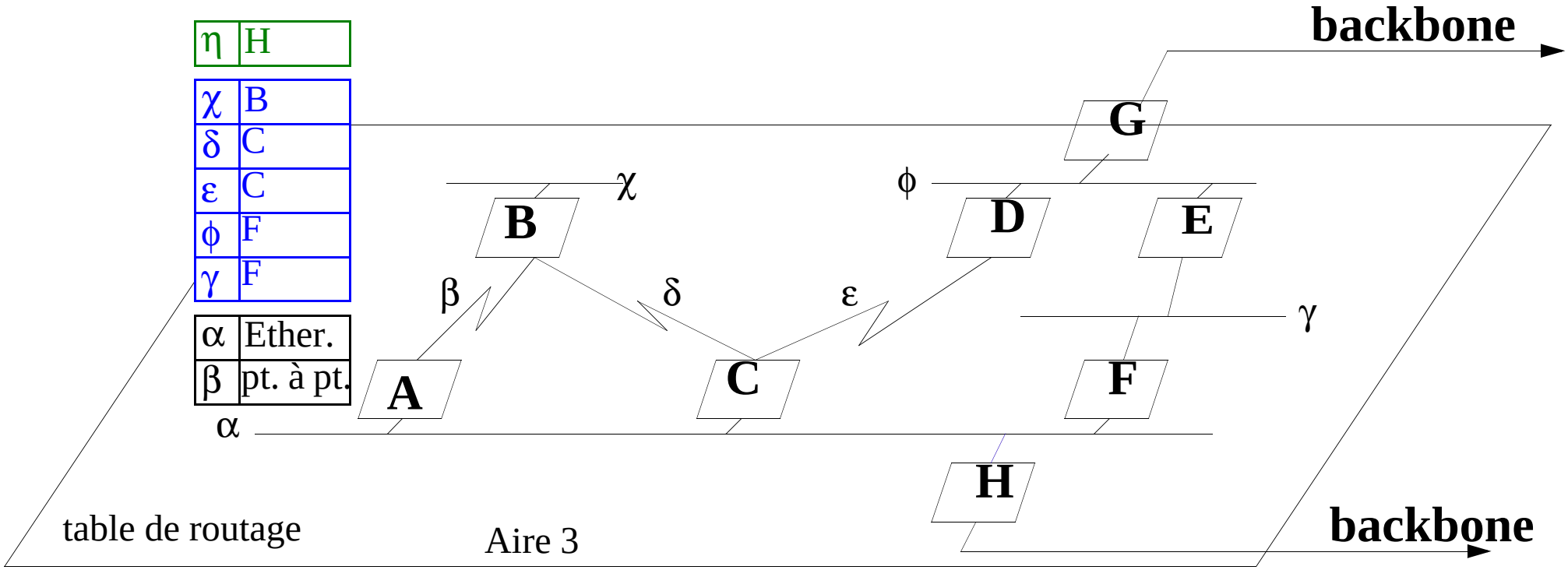
- Les routeurs G et H appartiennent au backbone, ils connaissent le réseau η.

DIFFUSION DES ROUTES EXTERIEURES À L'AIRE

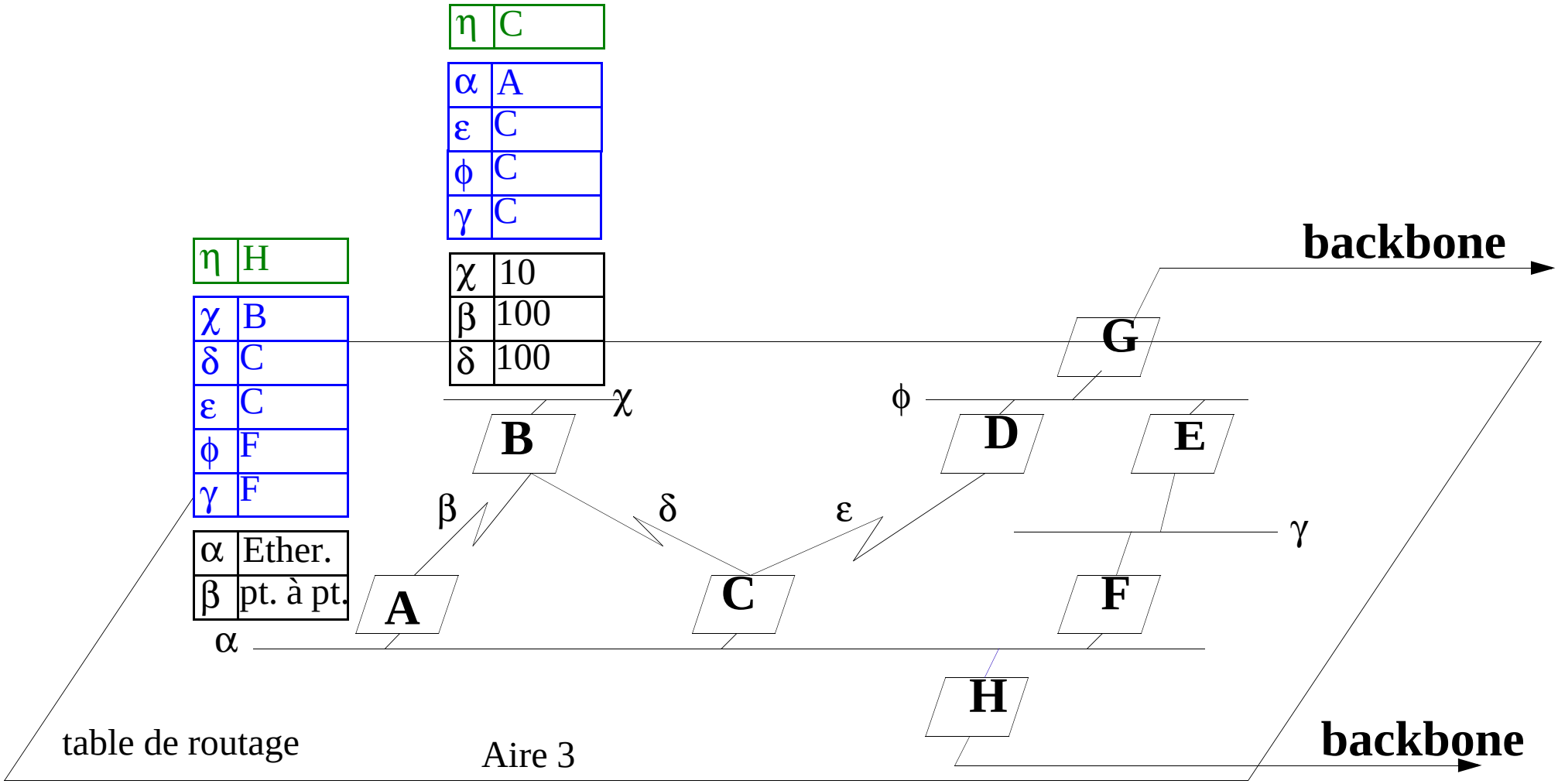


- Les routeurs calculent le meilleur chemin pour joindre les routeurs annonçant

MISE À JOUR DE LA TABLE DE ROUTAGE DE A

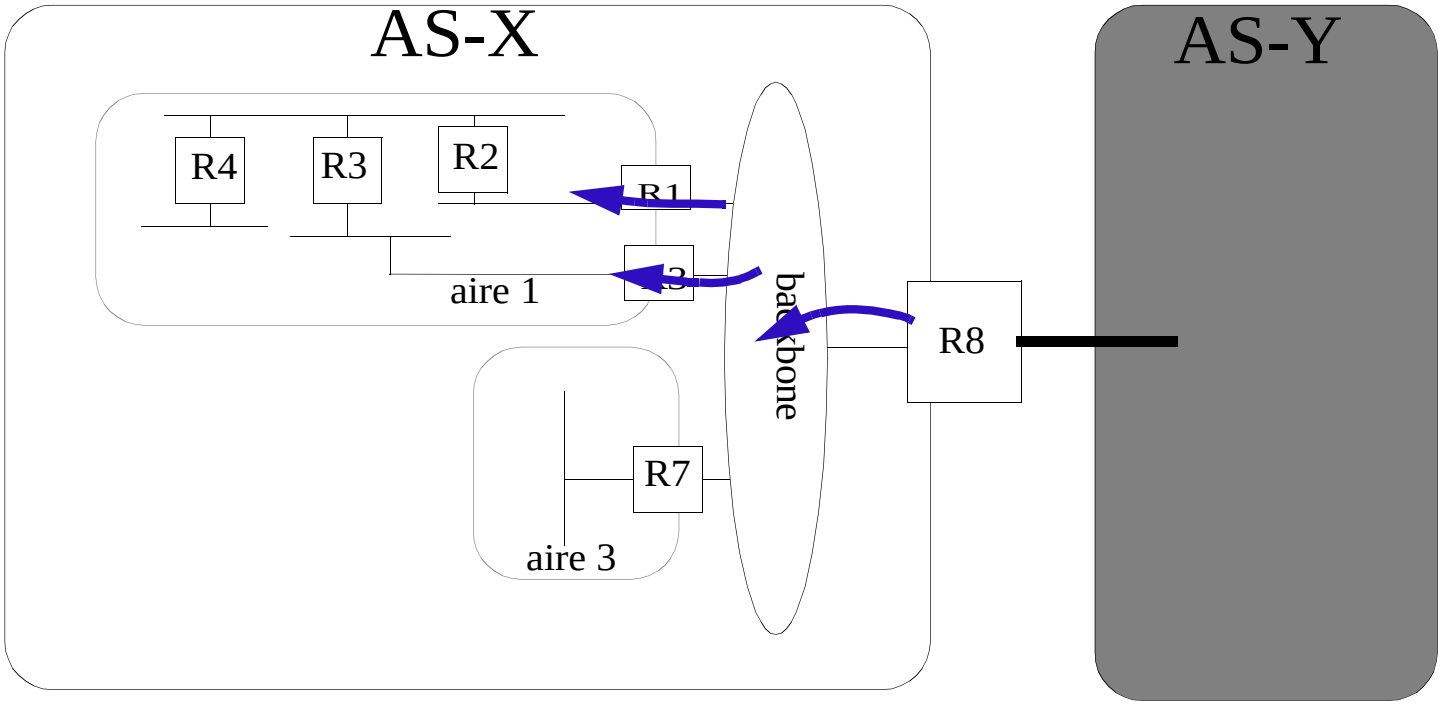


MISE À JOUR DE LA TABLE DE ROUTAGE DE B



ROUTES EXTERNES

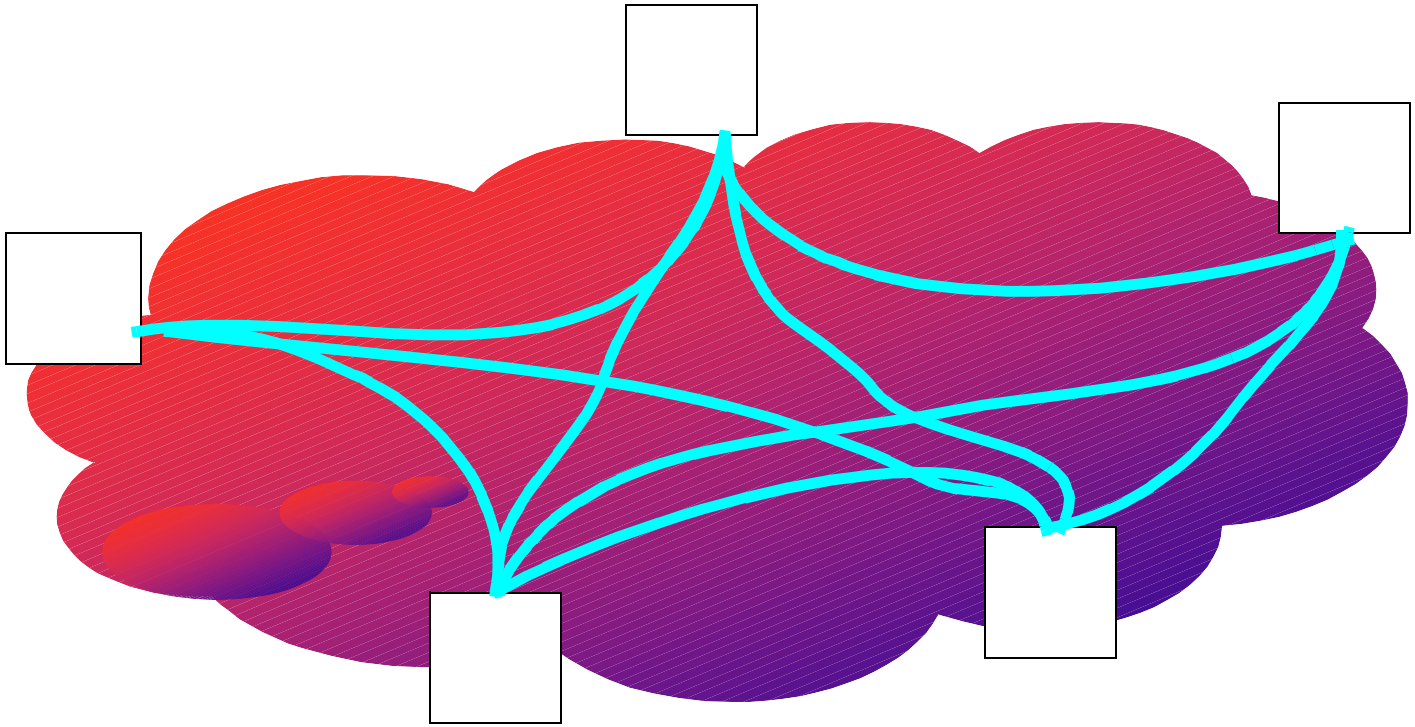
- Marquées différemment pour ne pas les réexporter
- Importées par les routeurs en frontière d'aire



- Deux types de routes :
  - type 1 (préféré) : coût = interne + externe
  - type 2 : coût = externe

RÉCAPITULATIF SUR LES BASES DE DONNÉES D'OSPF

- Des routeurs du lien : pour les réseaux NBMA (Frame Relay, X.25,...)
  - permet de se synchroniser pour connaître tous les routeurs qui forment le lien,
  - détermine le routeur désigné,
  - doit connaître au moins un des éléments par configuration statique,



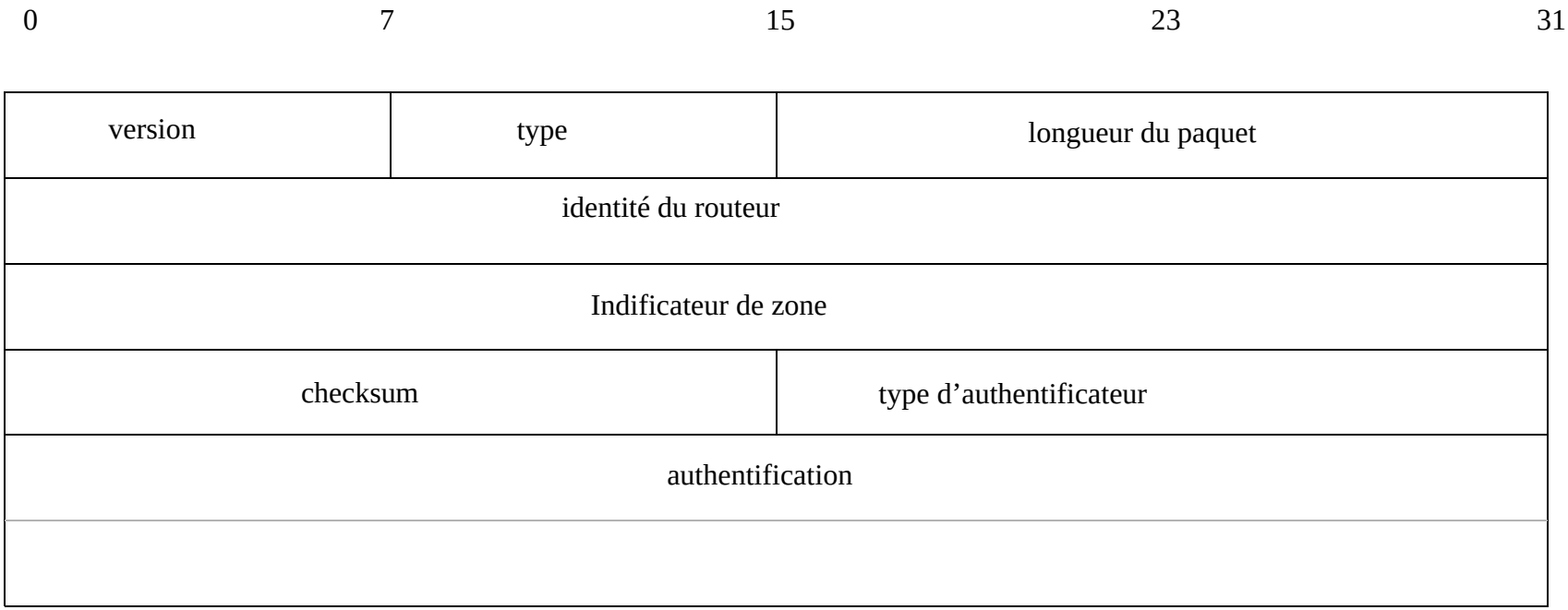
LE PROTOCOLE OSPF

- Meilleur que RIP
- Au dessus d'IP (proto = 89)
- Utilise les adresses de Multicast
  - 224.0.0.5 tous les routeurs du lien
  - 224.0.0.6 le routeur désigné et celui en secours
- Nombre illimité de routeurs
- Faible trafic (seules les modifications sont transmises)
- Authentification

Mais

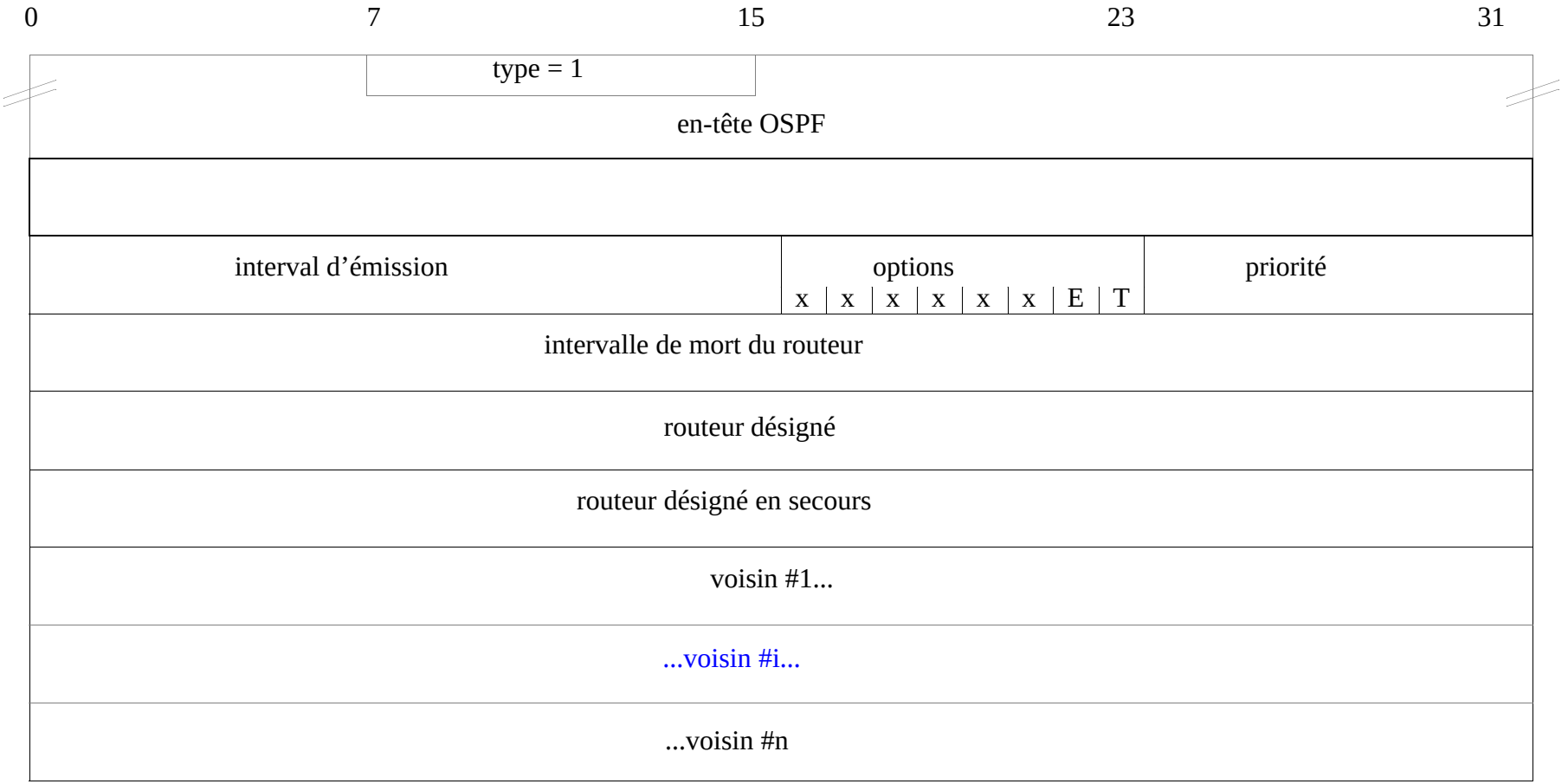
- Demande des routeurs puissants
- Difficile à configurer (nombreux paramètres)

FORMAT DES EN-TÊTE OSPF



- Paquet Hello : émis périodiquement pour identification des voisins
- Paquet description de la base de donnée : contient l'information sur la topologie du réseau
- Paquet demande, mise à jour, acquittement de l'état de la liaison : permettent de mettre à jour les tables de routeurs

PAQUET D'HELLO

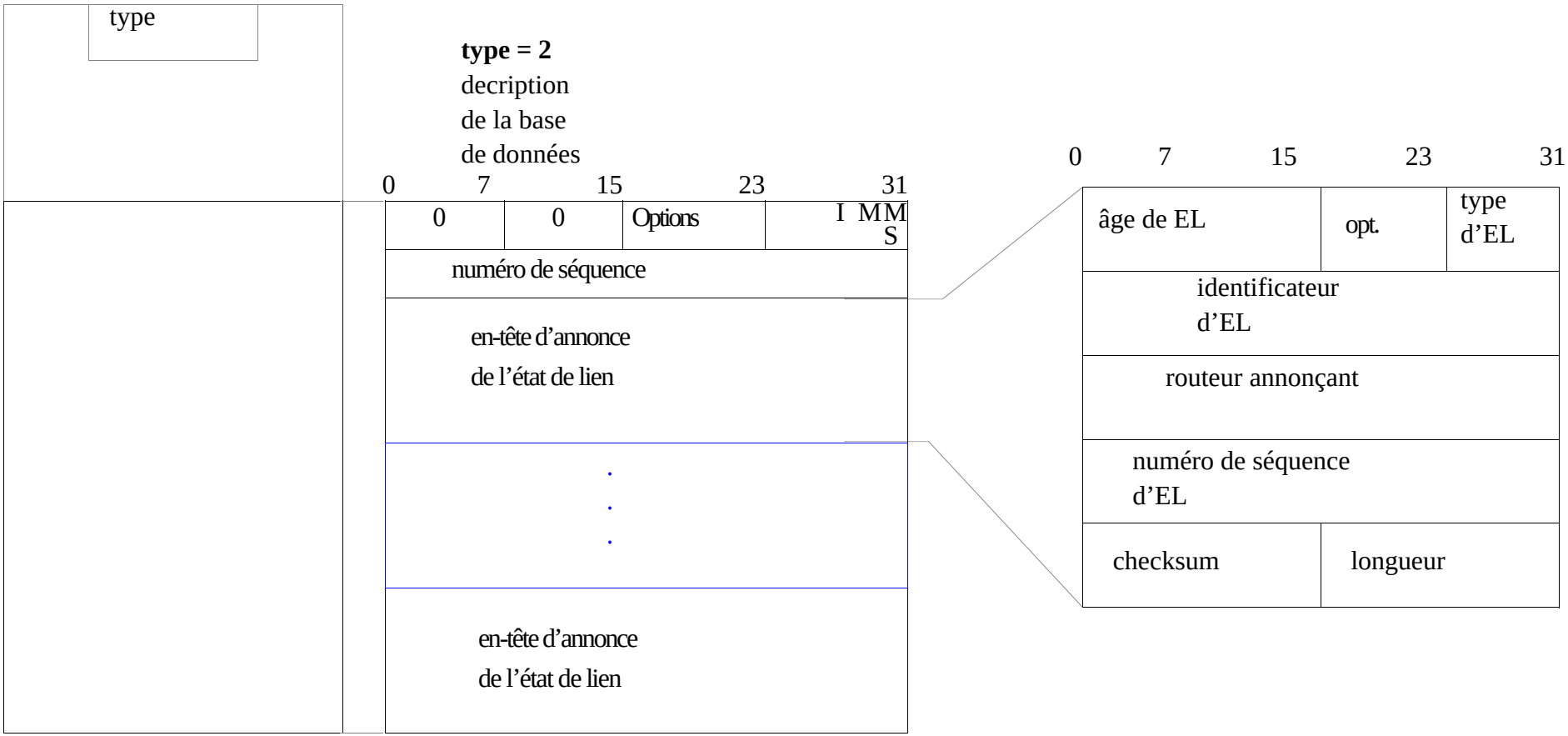


ALGORITHME DU LINK STATE

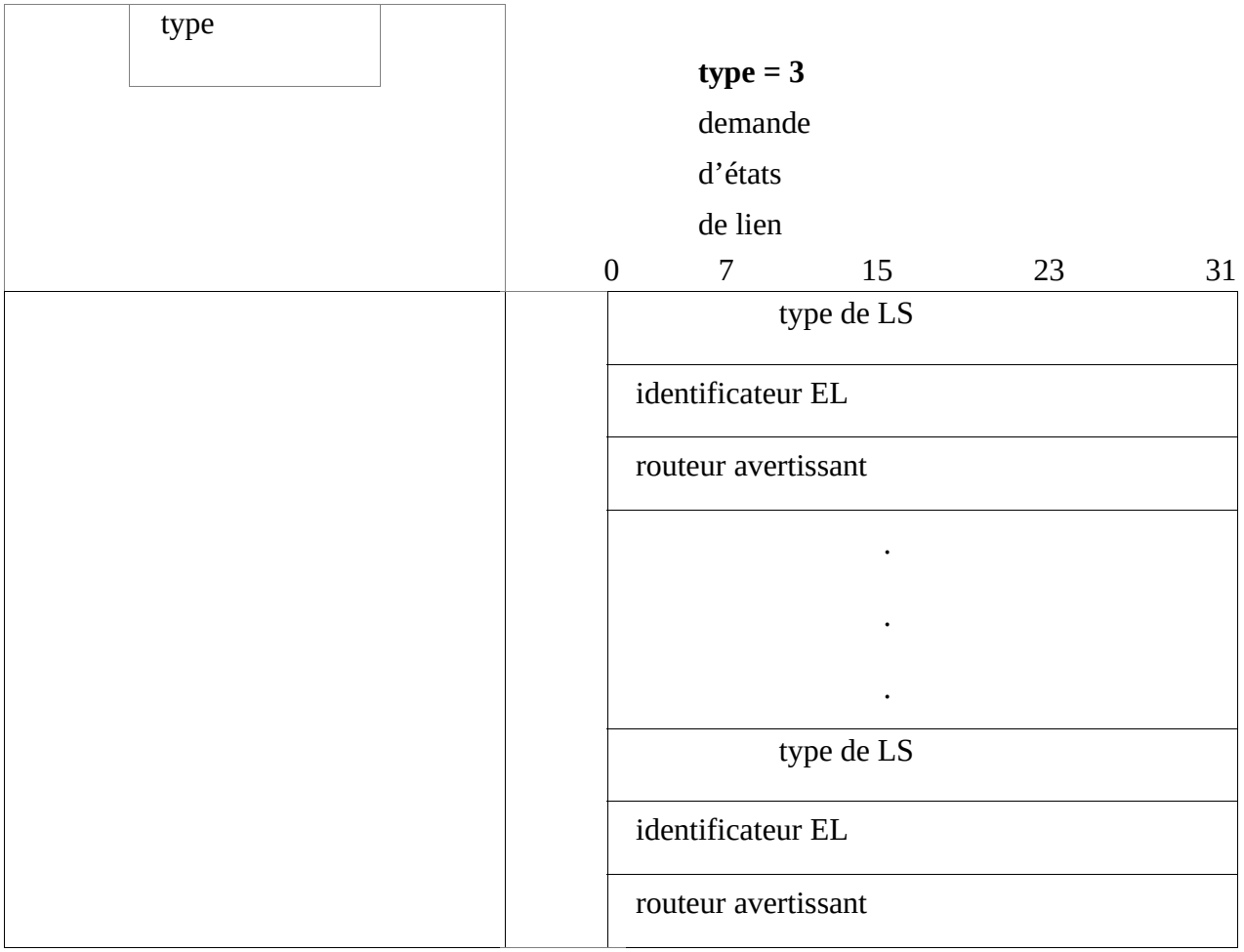
L'algorithme du *Link State* est le suivant :

- chaque routeur identifie ses voisins immédiats.
- un routeur principal (*designated router*) et de secours (*backup designated router*) sont désignés sur chaque réseau grâce à un mécanisme d'élection.
- le routeur acquière la base de donnée du routeur désigné.
- chaque routeur construit un message contenant la liste de ses voisins immédiats ainsi que le coût associé à la liaison. Ce message sera appelé LSP pour *Link State Packet*.
- ce paquet est transmis à tous les autres routeurs du réseau avec un mécanisme de diffusion qui limite la propagation des messages et évite les boucles.
- chaque routeur met à jour sa base de donnée ce qui lui donne une vision globale du réseau et il peut en déduire ses tables de routage.

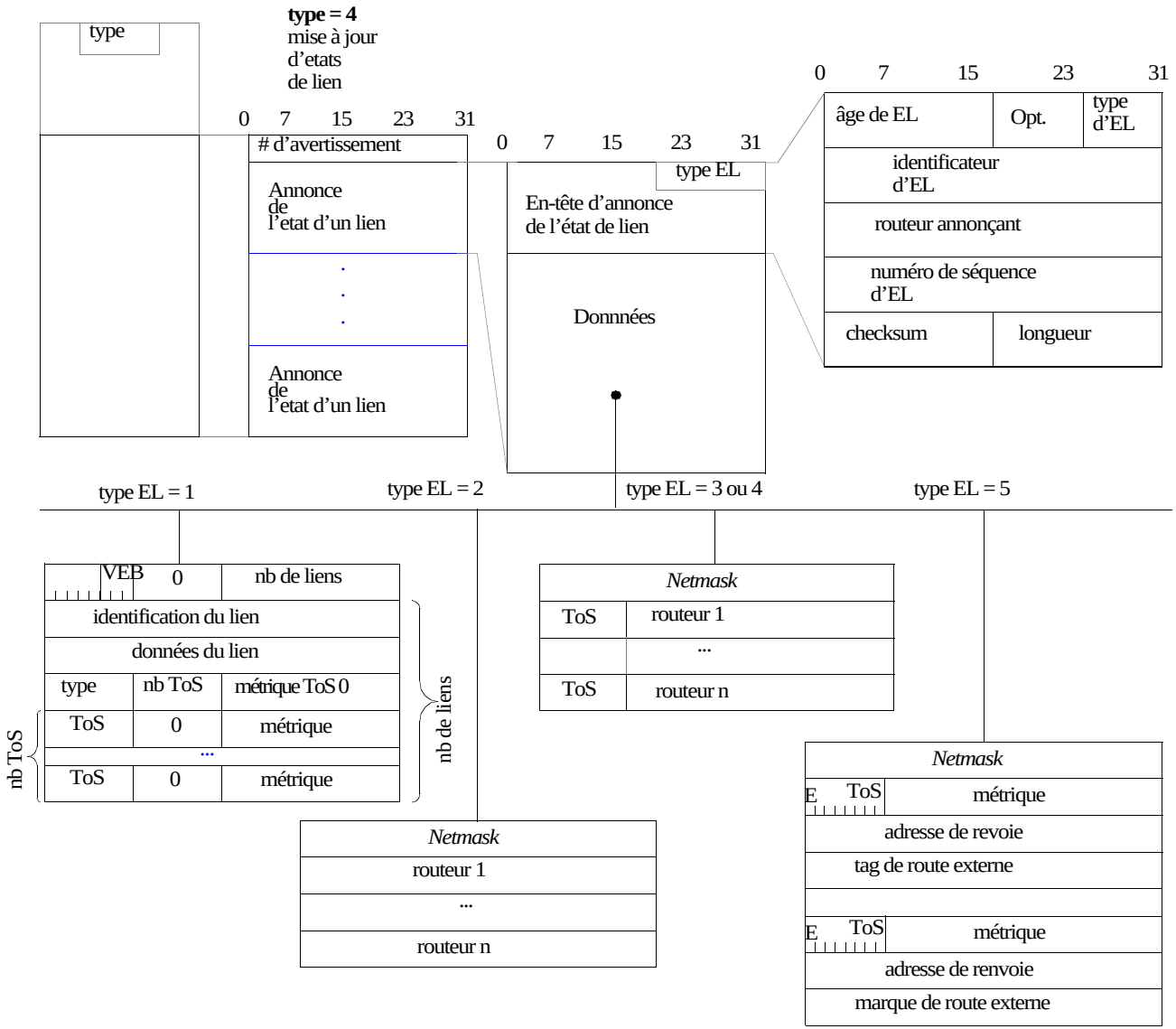
PAQUET DE DESCRIPTION DE LA BASE DE DONNÉES



PAQUET DE DEMANDE D'ÉTAT DE LIEN



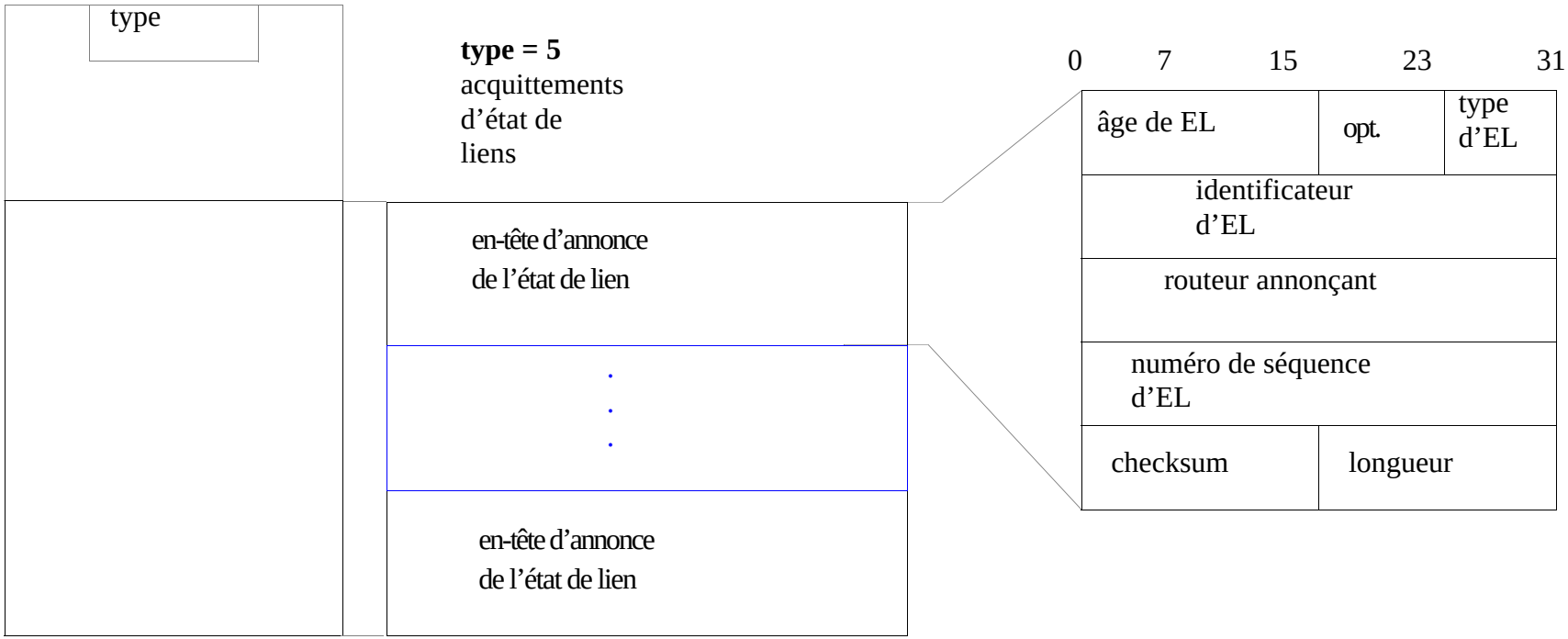
MISE À JOUR DE L'ÉTAT DE LIEN



+ces paquets inondent le réseau

+la transmission doit être fiable

ACQUITTEMENT.

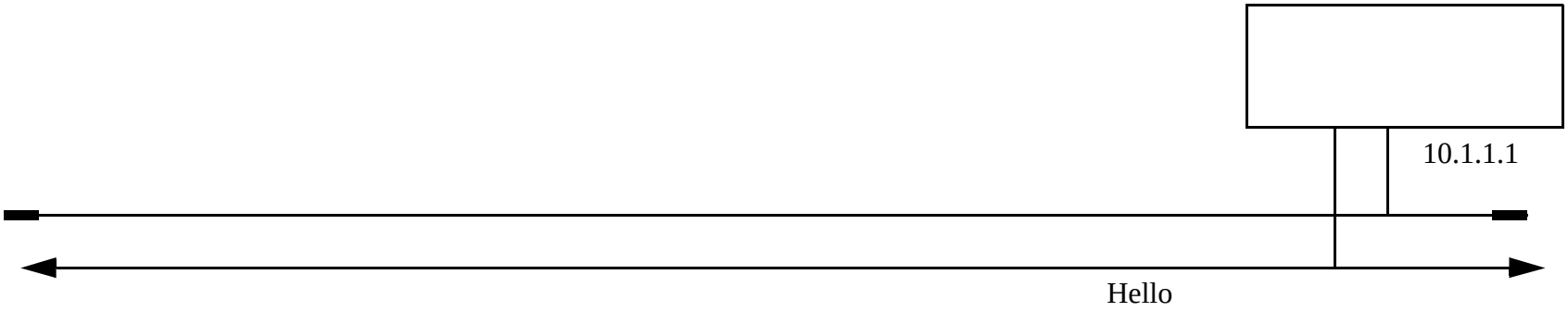


EXEMPLE

?

- On ne sait rien du réseau,
- Découverte de la topologie en écoutant le trafic OSPF

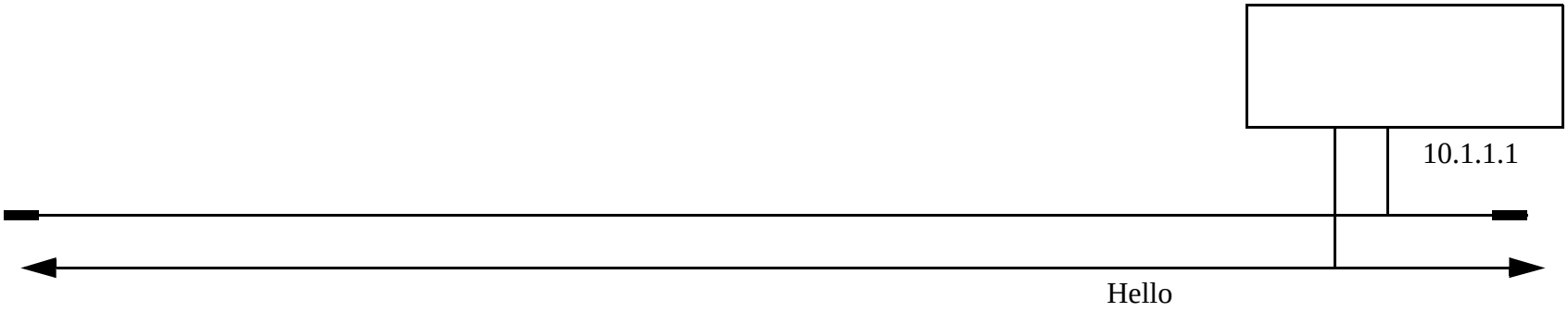
EXEMPLE : HELLO (ÉQUIPEMENT ISOLÉ)



10.1.1.1 > 224.0.0.5: OSPFv2-hello 44:
area 0.0.0.1 E mask 255.255.255.0 int 10 pri 5 dead 40 dr 10.1.1.1
nbrs

- Un routeur présent (10.1.1.1) de priorité 5
- Le préfixe est 10.1.1.0/24
- On se trouve dans l'aire 1
- Ce routeur ne connaît par d'autres voisins sur ce lien

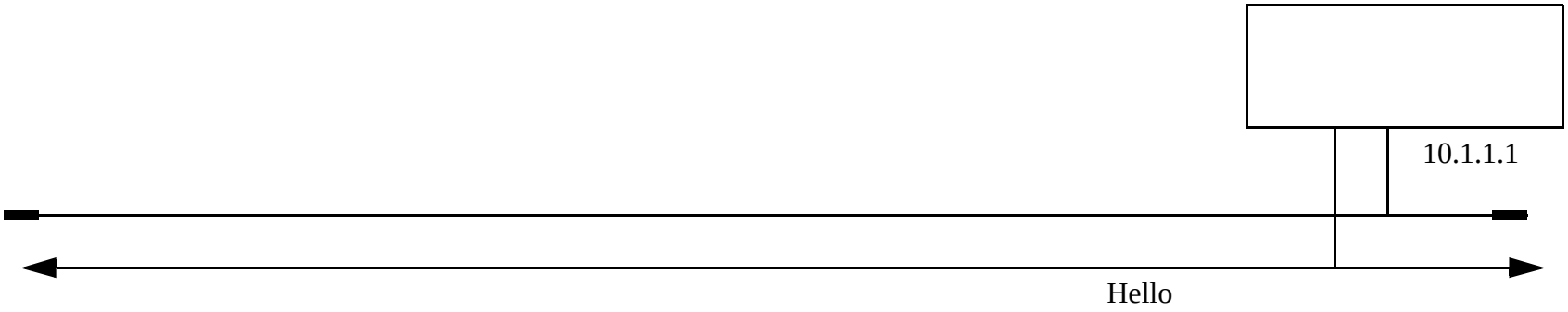
EXEMPLE HELLO (ÉQUIPEMENT ISOLÉ)



10.1.1.1 > 224.0.0.5: OSPFv2-hello 44:
area 0.0.0.1 E mask 255.255.255.0 int 10 pri 5 dead 40 dr 10.1.1.1
nbrs

- La période d'annonce des Hello est de 10 secondes
- Si on ne reçoit pas de message Hello pendant 40 seconde, le routeur est considéré comme mort.
- Priorité 5 utilisée pendant l'élection du routeur désigné

EXEMPLE : HELLO (ÉQUIPEMENT ISOLÉ)



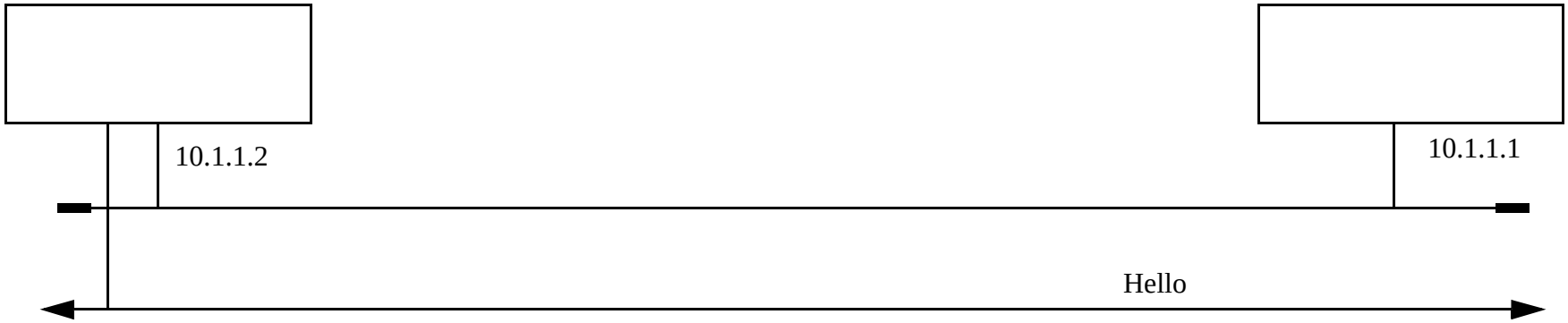
10.1.1.1 > 224.0.0.5: OSPFv2-hello 44:
area 0.0.0.1 E mask 255.255.255.0 int 10 pri 5 dead 40 dr 10.1.1.1
nbrs

10.1.1.1 > 224.0.0.5: OSPFv2-hello 44:
area 0.0.0.1 E mask 255.255.255.0 int 10 pri 5 dead 40 dr 10.1.1.1
nbrs

10.1.1.1 > 224.0.0.5: OSPFv2-hello 44:
area 0.0.0.1 E mask 255.255.255.0 int 10 pri 5 dead 40 dr 10.1.1.1
nbrs

....

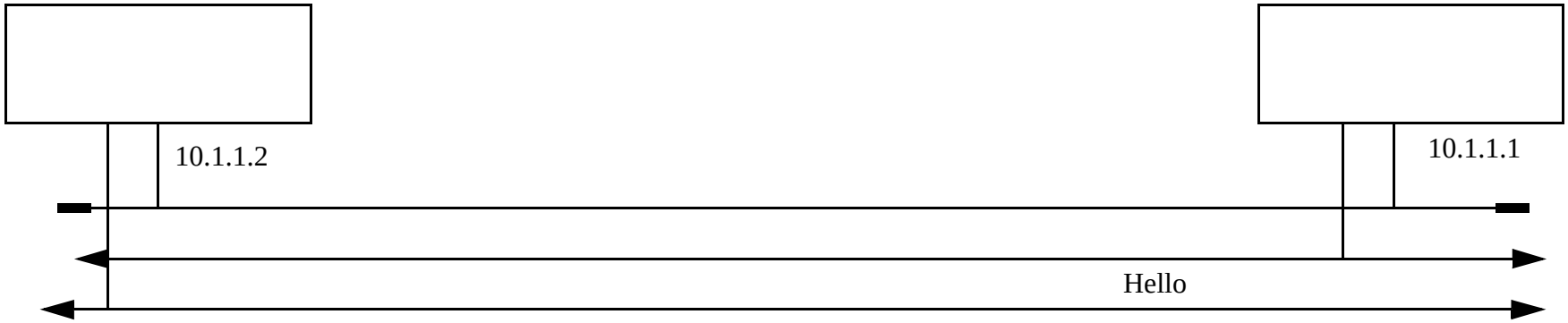
EXEMPLE : ECHANGE DE HELLO



10.1.1.2 > 224.0.0.5: OSPFv2-hello 44:
area 0.0.0.1 E mask 255.255.255.0 int 10 pri 27 dead 40
nbrs

- Un deuxième routeur émet un message Hello

EXEMPLE : ECHANGE DE HELLO



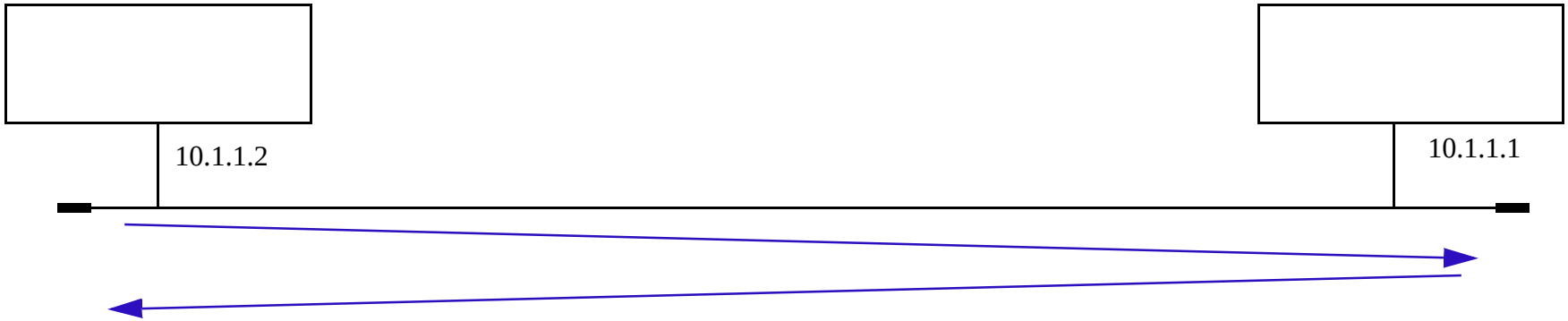
10.1.1.1 > 224.0.0.5: OSPFv2-hello 48:
area 0.0.0.1 E mask 255.255.255.0 int 10 pri 5 dead 40
dr 10.1.1.1 nbrs 10.1.1.2

- Le routeur 10.1.1.1 devient routeur désigné
- Le routeur 10.1.1.2 apparaît dans sa liste de voisins

EXEMPLE HELLO

- La connectivité est bi-directionnelle
- Le routeur 10.1.1.2 peut se synchroniser avec le routeur désigné.

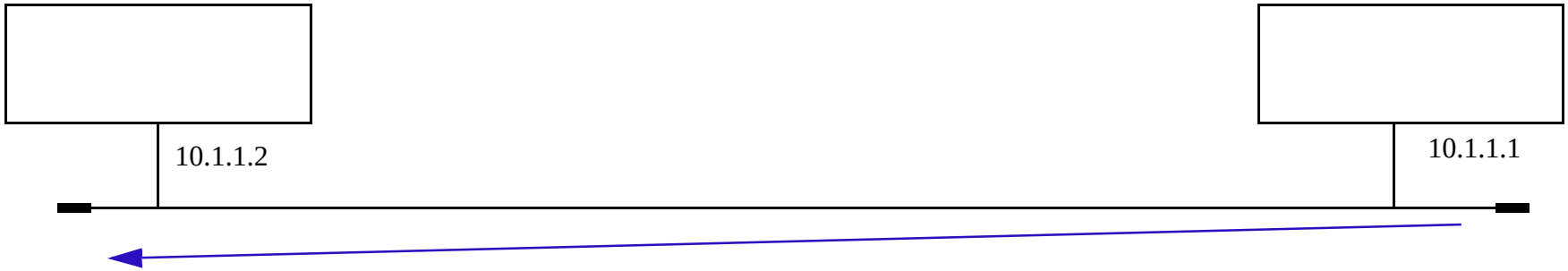
EXEMPLE : DESCRIPTION DE LA BASSE DE DONNÉES



10.1.1.2 > 10.1.1.1: OSPFv2-dd 32: area 0.0.0.1 E I/M/MS S **B**
10.1.1.1 > 10.1.1.2: OSPFv2-dd 32: area 0.0.0.1 E I/M/MS S **1973**

- 10.1.1.2 a comme numéro de séquence B
- 10.1.1.1 a comme numéro de séquence 1973
- I : premier paquet, M : d'autres paquets vont suivre, MS : maître.
- 10.1.1.2 va gagner (n° de séquence plus petit)

EXEMPLE : DESCRIPTION DE LA BASE DE DONNÉES

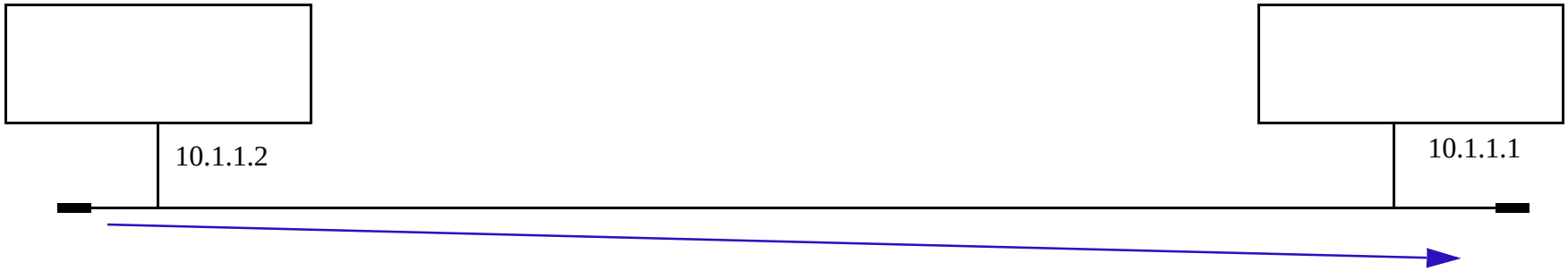


10.1.1.1 > 10.1.1.2: OSPFv2-dd 112: area 0.0.0.1 E M S **B**

{ E S 80000002 age 3:09 rtr 10.1.1.1 }	TYPE 1
{ E S 80000001 age 2:49 sum 10.1.2.0 abr 10.1.1.1 }	TYPE 3
{ E S 80000003 age 2:44 sum 10.1.100.0 abr 10.1.1.1 }	TYPE 3
{ E S 80000001 age 2:59 abr 10.1.1.1 rtr 10.1.1.1 }	TYPE 4

- 10.1.1.1 décrit sa base de donnée :
  - TYPE 1 : liaison que connaît le routeur, TYPE 3 : réseaux présents à l'exterieur de l'aire, TYPE 4 : routeurs faisant l'agrégation pour cette aire.

EXEMPLE : DESCRIPTION DE LA BASE DE DONNÉES



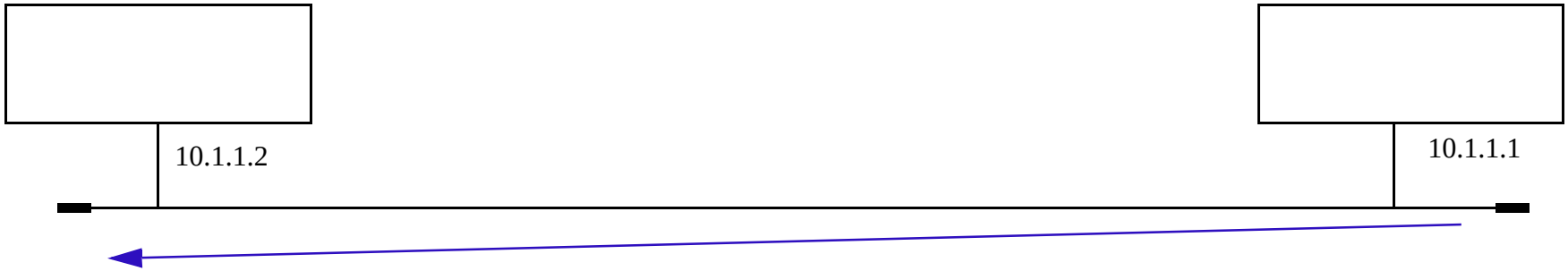
10.1.1.2 > 10.1.1.1: OSPFv2-dd 52: area 0.0.0.1 E MS S **C**
{ S 80000002 age 5 rtr 10.1.1.2 } TYPE 1

- 10.1.1.2 acquitte le message précédent en incrémentant le champ séquence et décrit ses connaissances :
  - •• les réseaux auxquel il est attaché.
- plus d'autres informations (le bit M n'est plus présent)

EXEMPLE : DESCRIPTION DE LA BASE DE DONNÉES

{ S 80000002 age 5 rtr 10.1.1.2 }

{ E S 80000002 age 3:09 rtr 10.1.1.1 }
{ E S 80000001 age 2:49 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:44 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000001 age 2:59 abr 10.1.1.1 rtr 10.1.1.1 }



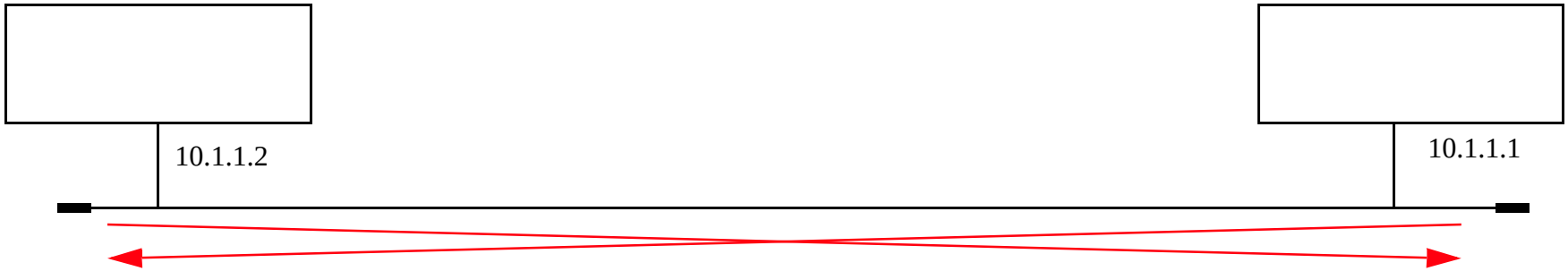
15:14:50.472224 10.1.1.1 > 10.1.1.2: OSPFv2-dd 32: area 0.0.0.1 E S C

- 10.1.1.1 n'a plus rien à décrire,
- fin de la phase de descrition de la base de données.
- chaque routeur sait ce que connaît l'autre.

EXEMPLE : DEMANDE D'INFORMATION

{ S 80000002 age 5 rtr 10.1.1.2 }

{ E S 80000002 age 3:09 rtr 10.1.1.1 }
{ E S 80000001 age 2:49 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:44 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000001 age 2:59 abr 10.1.1.1 rtr 10.1.1.1 }



10.1.1.2 > 10.1.1.1: OSPFv2-ls_req 72: area 0.0.0.1
{ rtr 10.1.1.1 } { sum 10.1.2.0 abr 10.1.1.1 } { sum 10.1.100.0 abr 10.1.1.1 }
{ abr 10.1.1.1 rtr 10.1.1.1 }

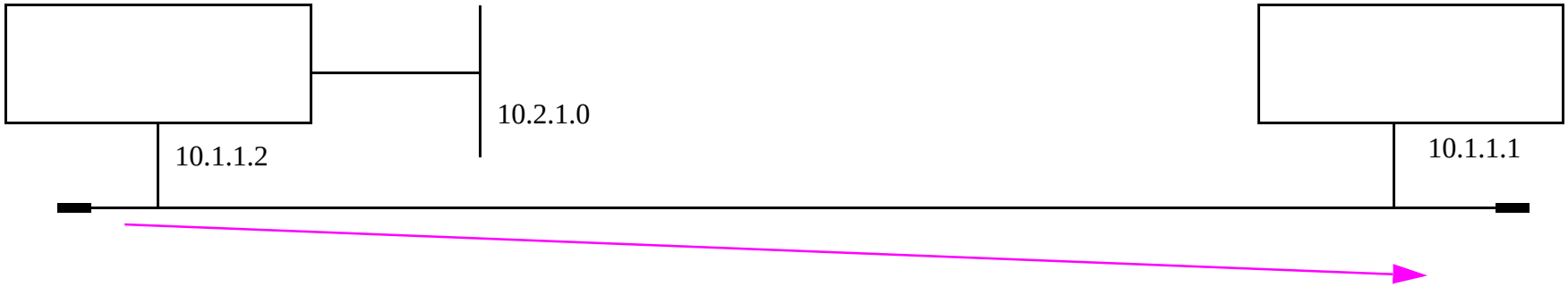
10.1.1.1 > 10.1.1.2: OSPFv2-ls_req 36: area 0.0.0.1
{ rtr 10.1.1.2 }

- Les routeurs veulent les informations qui :
  - leur manquent ou
  - qui sont plus récentes (pas dans cet exemple)

EXEMPLE : MISE À JOUR DE LA BASE DE DONNÉES

{ E S 80000002 age 3:09 rtr 10.1.1.1 }
{ E S 80000001 age 2:49 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:44 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000001 age 2:59 abr 10.1.1.1 rtr 10.1.1.1 }
{ S 80000002 age 6 rtr 10.1.1.2 }

{ S 80000002 age 5 rtr 10.1.1.2 }



10.1.1.2 > 10.1.1.1: OSPFv2-ls_upd 76: area 0.0.0.1
{ S 80000002 age 6 rtr 10.1.1.2
{ net 10.1.1.0 mask 255.255.255.0 tos 0 metric 1 }
{ net 10.2.1.0 mask 255.255.255.0 tos 0 metric 1 } }

- Enregistrement complet du type 1 décrivant les attachements locaux du routeur
- L'âge de l'enregistrement a évolué

EXEMPLE : MISE À JOUR DE LA BASE DE DONNÉES

{ S 80000002 age 5 rtr 10.1.1.2 }
{ E S 80000002 age 3:10 rtr 10.1.1.1 }
{E S 80000001 age 2:50 sum 10.1.2.0 abr 10.1.1.1}
{E S 80000003 age 2:45 sum 10.1.100.0 abr 10.1.1.1 }
{E S 80000001 age 3:01 abr 10.1.1.1 rtr 10.1.1.1}

{ E S 80000002 age 3:09 rtr 10.1.1.1 }
{ E S 80000001 age 2:49 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:44 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000001 age 2:59 abr 10.1.1.1 rtr 10.1.1.1 }
{ S 80000002 age 6 rtr 10.1.1.2 }

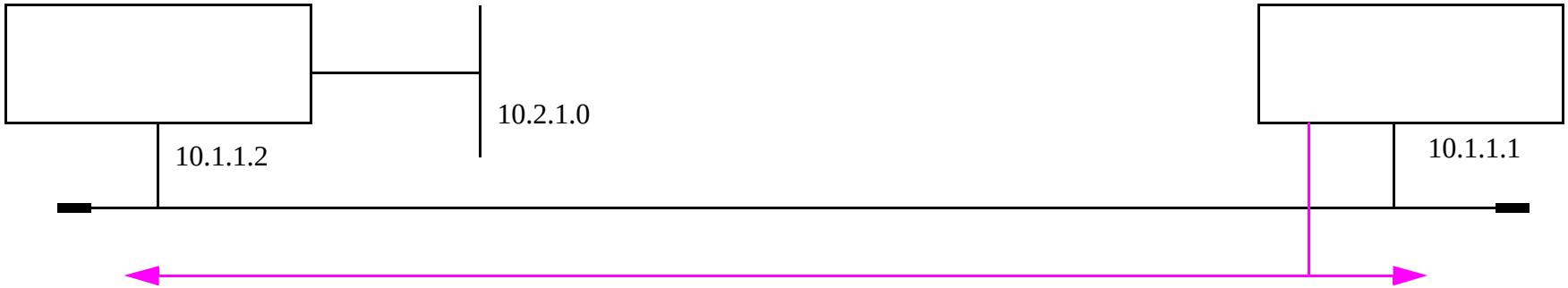
10.1.1.1 > 10.1.1.2: OSPFv2-ls_upd 148: area 0.0.0.1
{ E S 80000002 age 3:10 rtr 10.1.1.1 B
{ net 10.1.1.0 mask 255.255.255.0 tos 0 metric 10 } }
{ E S 80000001 age 2:50 sum 10.1.2.0 abr 10.1.1.1 mask 255.255.255.0 tos 0 metric 20 }
{ E S 80000003 age 2:45 sum 10.1.100.0 abr 10.1.1.1 mask 255.255.255.0 tos 0 metric 10 }
{ E S 80000001 age 3:01 abr 10.1.1.1 rtr 10.1.1.1 tos 0 metric 16777215 }

- Informations du routeur 10.1.1.1 :
  - Un seul attachement dans l'aire 1 et deux réseaux (10.1.2.0/24 et 10.1.100.0/24) à l'extérieur de l'aire.

EXEMPLE : MISE À JOUR DE LA BASE DE DONNÉES

{ S 80000002 age 5 rtr 10.1.1.2 }
{ E S 80000002 age 3:10 rtr 10.1.1.1 }
{ E S 80000001 age 2:50 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:45 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000001 age 3:01 abr 10.1.1.1 rtr 10.1.1.1 }

{ E S 80000002 age 3:09 rtr 10.1.1.1 }
{ E S 80000001 age 2:49 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:44 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000001 age **1:00:00** abr 10.1.1.1 rtr 10.1.1.1 }
{ S 80000002 age 6 rtr 10.1.1.2 }



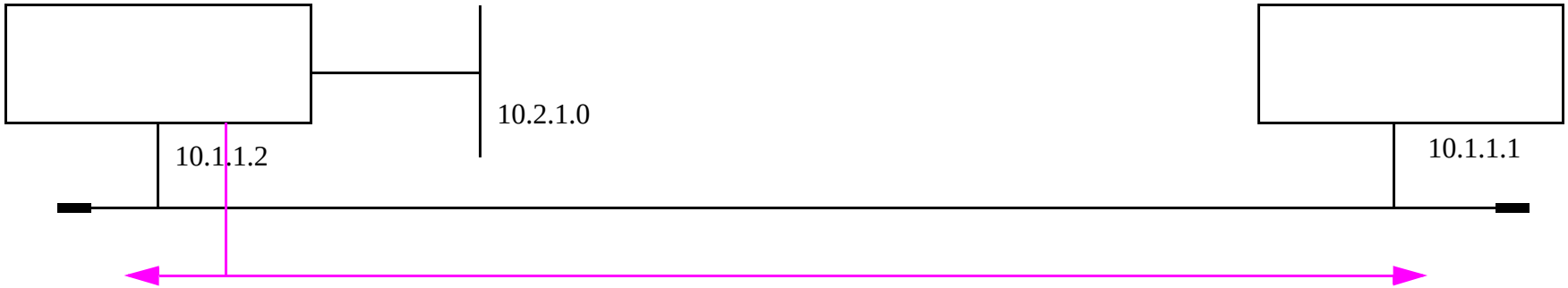
10.1.1.1 > 224.0.0.5: OSPFv2-ls_upd 56: area 0.0.0.1
{ E S 80000002 age 1:00:00 abr 10.1.1.1 rtr 10.1.1.1 tos 0 metric 16777215 }

- 10.1.1.1 rend l'enregistrement obsolète

EXEMPLE : MISE À JOUR DE LA BASE DE DONNÉES

{ S 80000003 age 1 rtr 10.1.1.2 }
{ E S 80000002 age 3:10 rtr 10.1.1.1 }
{ E S 80000001 age 2:50 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:45 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000001 age 3:01 abr 10.1.1.1 rtr 10.1.1.1 }

{ E S 80000002 age 3:09 rtr 10.1.1.1 }
{ E S 80000001 age 2:49 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:44 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000001 age **1:00:00** abr 10.1.1.1 rtr 10.1.1.1 }
{ S 80000002 age 6 rtr 10.1.1.2 }



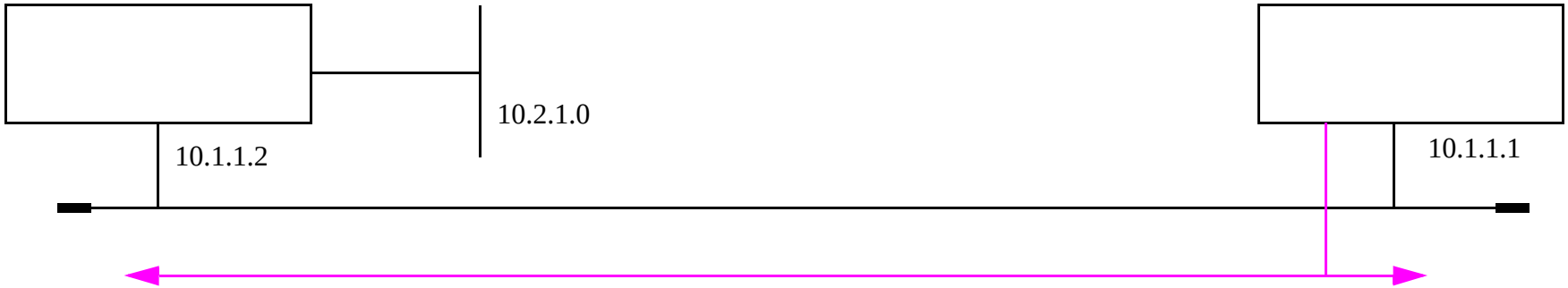
10.1.1.2 > 224.0.0.5: OSPFv2-ls_upd 76: area 0.0.0.1
{ S 80000003 age 1 rtr 10.1.1.2
{ dr 10.1.1.1 if 10.1.1.2 tos 0 metric 1 }
{ net 10.2.1.0 mask 255.255.255.0 tos 0 metric 1 } }

- Le routeur émet une nouvelle version de son enregistrement

EXEMPLE : MISE À JOUR DE LA BASE DE DONNÉES

{ S 80000003 age 1 rtr 10.1.1.2 }
{ E S 80000002 age 3:10 rtr 10.1.1.1 }
{ E S 80000001 age 2:50 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:45 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000001 age 3:01 abr 10.1.1.1 rtr 10.1.1.1 }

{ E S 80000003 age 1 rtr 10.1.1.1 }
{ E S 80000001 age 2:49 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:44 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000001 age 1:00:00 abr 10.1.1.1 rtr 10.1.1.1 }
{ S 80000002 age 6 rtr 10.1.1.2 }
{ E S 80000001 age 1 net dr 10.1.1.1 }



10.1.1.1 > 224.0.0.5: OSPFv2-ls_upd 64: area 0.0.0.1
{ E S 80000003 age 1 rtr 10.1.1.1 B
{ dr 10.1.1.1 if 10.1.1.1 tos 0 metric 10 } }

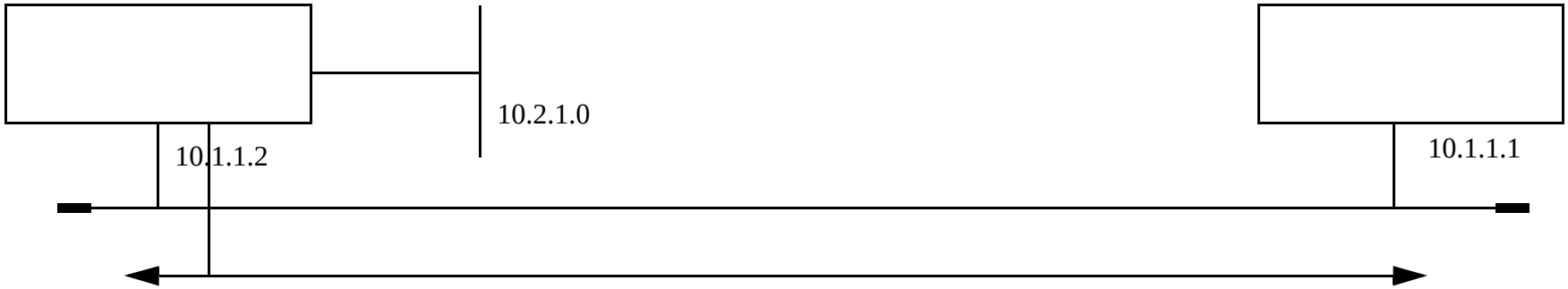
10.1.1.1 > 224.0.0.5: OSPFv2-ls_upd 60: area 0.0.0.1 TYPE 2
{ E S 80000001 age 1 net dr 10.1.1.1 if 10.1.1.1 mask 255.255.255.0 rtrs 10.1.1.1 10.1.1.2 }

- Remet à jour l'enregistrement sur les liaisons du routeur.
- Emet un nouvel enregistrement de type 2.

EXEMPLE

{ S 80000003 age 1 rtr 10.1.1.2 }
{ E S 80000002 age 3:10 rtr 10.1.1.1 }
{ E S 80000001 age 2:50 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:45 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000001 age 3:01 abr 10.1.1.1 rtr 10.1.1.1 }

{ E S 80000003 age 1 rtr 10.1.1.1 }
{ E S 80000001 age 2:49 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:44 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000001 age 1:00:00 abr 10.1.1.1 rtr 10.1.1.1 }
{ S 80000002 age 6 rtr 10.1.1.2 }
{ E S 80000001 age 1 net dr 10.1.1.1 }

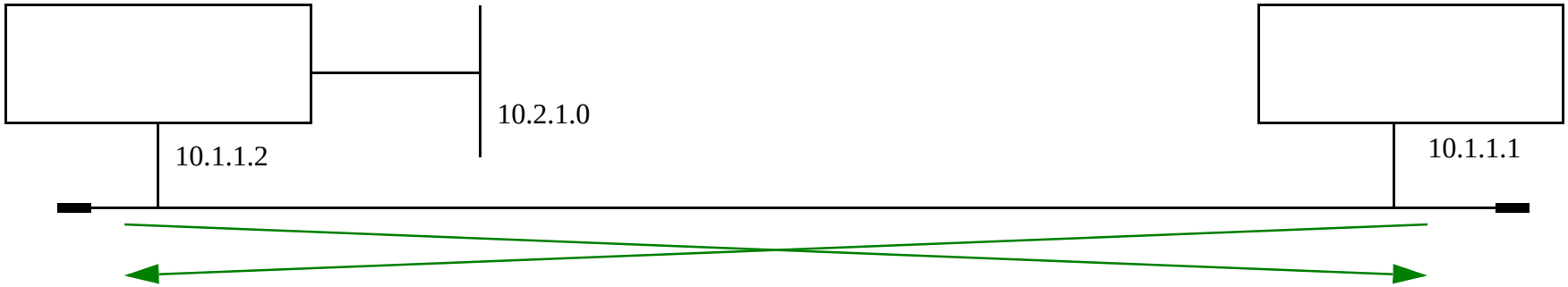


10.1.1.2 > 224.0.0.5: OSPFv2-hello 48:
area 0.0.0.1 E mask 255.255.255.0 int 10 pri 27 dead 40
dr 10.1.1.1 bdr 10.1.1.2 nbrs 10.1.1.1

EXEMPLE : ACQUITTEMENT

{ S 80000003 age 1 rtr 10.1.1.2 }
{ E S 80000002 age 3:10 rtr 10.1.1.1 }
{ E S 80000001 age 2:50 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:45 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000001 age 3:01 abr 10.1.1.1 rtr 10.1.1.1 }

{ E S 80000003 age 1 rtr 10.1.1.1 }
{ E S 80000001 age 2:49 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:44 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000001 age 1:00:00 abr 10.1.1.1 rtr 10.1.1.1 }
{ S 80000002 age 6 rtr 10.1.1.2 }
{ E S 80000001 age 1 net dr 10.1.1.1 }



10.1.1.2 > 224.0.0.5: OSPFv2-ls_ack 44: area 0.0.0.1
{ E S 80000001 age 1 net dr 10.1.1.1 if 10.1.1.1 }

10.1.1.1 > 224.0.0.5: OSPFv2-ls_ack 44: area 0.0.0.1
{ S 80000002 age 6 rtr 10.1.1.2 }

EXEMPLE

{ S 80000003 age 1 rtr 10.1.1.2 }
{ E S 80000002 age 3:10 rtr 10.1.1.1 }
{ E S 80000001 age 2:50 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:45 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000001 age 3:01 abr 10.1.1.1 rtr 10.1.1.1 }

{ E S 80000003 age **5** rtr 10.1.1.1 }
{ E S 80000001 age 2:49 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:44 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 8000000**2** age 1:00:00 abr 10.1.1.1 rtr 10.1.1.1 }
{ S 80000002 age 6 rtr 10.1.1.2 }
{ E S 80000001 age 1 net dr 10.1.1.1 }

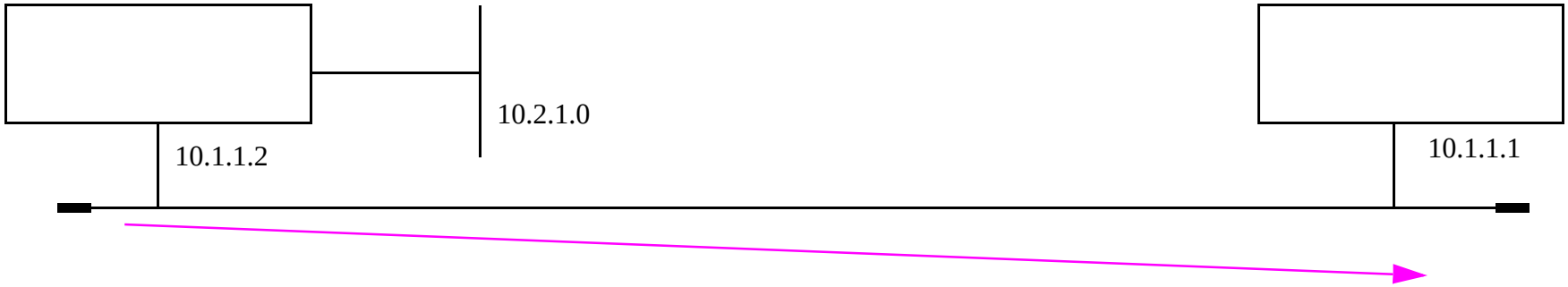
10.1.1.1 > 10.1.1.2: OSPFv2-ls_upd 92: area 0.0.0.1
{ E S 80000002 age 1:00:00 abr 10.1.1.1 rtr 10.1.1.1 tos 0 metric 16777215 }
{ E S 80000003 age 5 rtr 10.1.1.1 B
{ dr 10.1.1.1 if 10.1.1.1 tos 0 metric 10 } }

- Expiration de l'enregistrement sur le routeur d'agrégation

EXEMPLE

{ S 80000003 age **6** rtr 10.1.1.2 }
{ E S 80000002 age 3:10 rtr 10.1.1.1 }
{ E S 80000001 age 2:50 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:45 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000001 age 3:01 abr 10.1.1.1 rtr 10.1.1.1 }

{ E S 80000003 age 5 rtr 10.1.1.1 }
{ E S 80000001 age 2:49 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:44 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000002 age 1:00:00 abr 10.1.1.1 rtr 10.1.1.1 }
{ S 80000002 age 6 rtr 10.1.1.2 }
{ E S 80000001 age 1 net dr 10.1.1.1 }

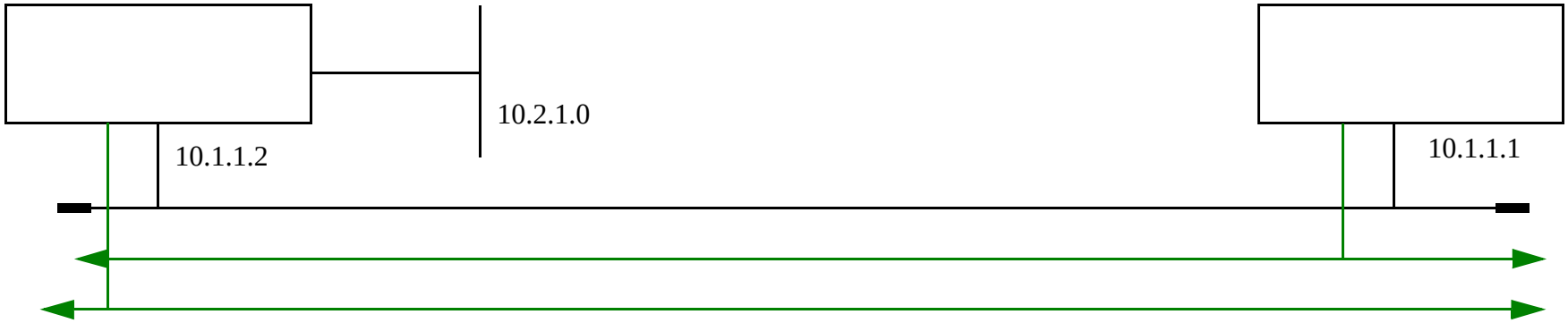


10.1.1.2 > 10.1.1.1: OSPFv2-ls_upd 76: area 0.0.0.1
{ S 80000003 age 6 rtr 10.1.1.2
{ dr 10.1.1.1 if 10.1.1.2 tos 0 metric 1 }
{ net 10.2.1.0 mask 255.255.255.0 tos 0 metric 1 } }

EXEMPLE

{ S 80000003 age 6 rtr 10.1.1.2 }
{ E S 80000002 age 3:10 rtr 10.1.1.1 }
{ E S 80000001 age 2:50 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:45 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000001 age 3:01 abr 10.1.1.1 rtr 10.1.1.1 }

{ E S 80000003 age 5 rtr 10.1.1.1 }
{ E S 80000001 age 2:49 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:44 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000002 age 1:00:00 abr 10.1.1.1 rtr 10.1.1.1 }
{ S 80000002 age 6 rtr 10.1.1.2 }
{ E S 80000001 age 1 net dr 10.1.1.1 }



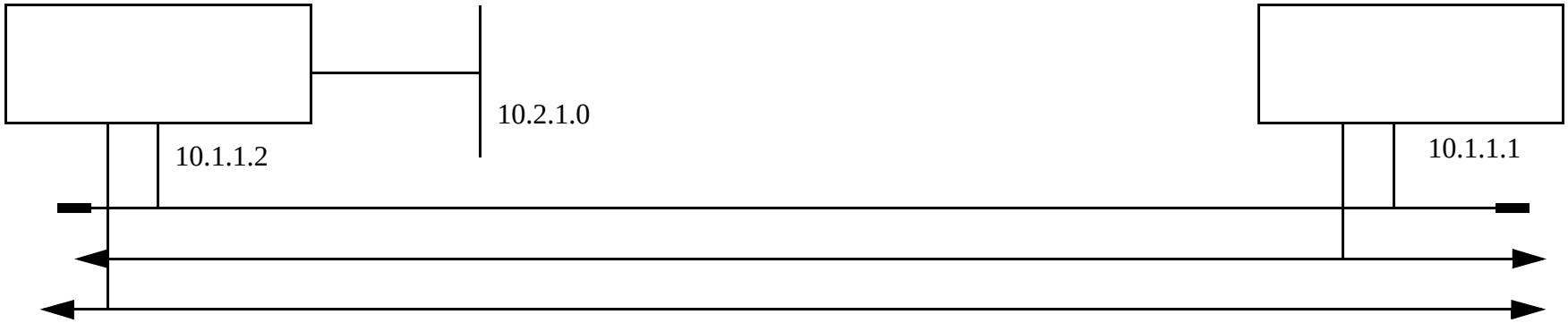
10.1.1.2 > 224.0.0.5: OSPFv2-ls_ack 64: area 0.0.0.1
{ E S 80000003 age 5 rtr 10.1.1.1 }
{ E S 80000002 age 1:00:00 abr 10.1.1.1 rtr 10.1.1.1 }

10.1.1.1 > 224.0.0.5: OSPFv2-ls_ack 44: area 0.0.0.1
{ S 80000003 age 6 rtr 10.1.1.2 }

EXEMPLE

{ S 80000003 age 6 rtr 10.1.1.2 }
{ E S 80000002 age 3:10 rtr 10.1.1.1 }
{ E S 80000001 age 2:50 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:45 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000001 age 3:01 abr 10.1.1.1 rtr 10.1.1.1 }

{ E S 80000003 age 5 rtr 10.1.1.1 }
{ E S 80000001 age 2:49 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:44 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000002 age 1:00:00 abr 10.1.1.1 rtr 10.1.1.1 }
{ S 80000002 age 6 rtr 10.1.1.2 }
{ E S 80000001 age 1 net dr 10.1.1.1 }



10.1.1.1 > 224.0.0.5: OSPFv2-hello 48: area 0.0.0.1 E mask 255.255.255.0 int 10 pri 5 dead 40
dr 10.1.1.1 bdr 10.1.1.2 nbrs 10.1.1.2

10.1.1.2 > 224.0.0.5: OSPFv2-hello 48: area 0.0.0.1 E mask 255.255.255.0 int 10 pri 27 dead 40
dr 10.1.1.1 bdr 10.1.1.2 nbrs 10.1.1.1 [ttl 1] (id 58902)

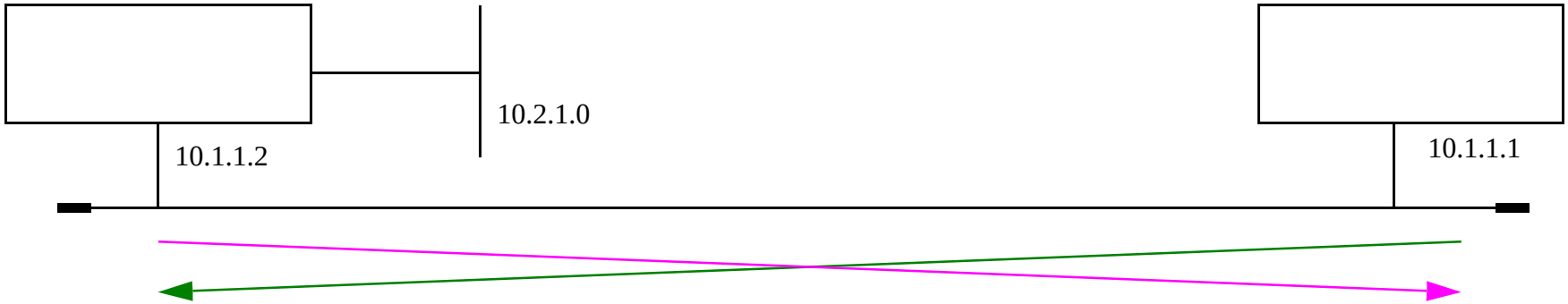
10.1.1.1 > 224.0.0.5: OSPFv2-hello 48: area 0.0.0.1 E mask 255.255.255.0 int 10 pri 5 dead 40
dr 10.1.1.1 bdr 10.1.1.2 nbrs 10.1.1.2 [tos 0xc0] [ttl 1] (id 130)

....

EXEMPLE

{ S 80000004 age 1 rtr 10.1.1.2 }
{ E S 80000002 age 3:10 rtr 10.1.1.1 }
{ E S 80000001 age 2:50 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:45 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000001 age 3:01 abr 10.1.1.1 rtr 10.1.1.1 }

{ E S 80000003 age 5 rtr 10.1.1.1 }
{ E S 80000001 age 2:49 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 2:44 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000002 age 1:00:00 abr 10.1.1.1 rtr 10.1.1.1 }
{ S 80000002 age 6 rtr 10.1.1.2 }
{ E S 80000001 age 1 net dr 10.1.1.1 }

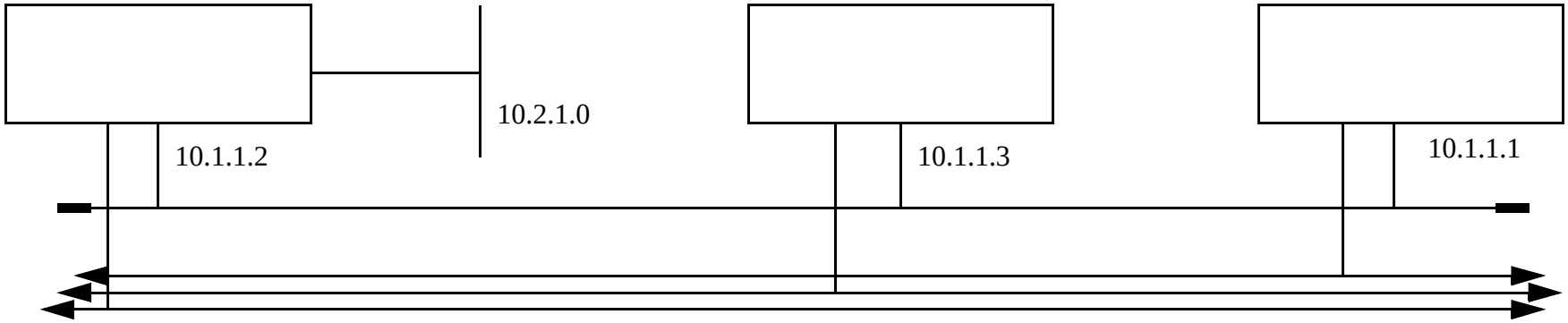


10.1.1.2 > 224.0.0.5: OSPFv2-ls_upd 76: area 0.0.0.1
{ S 80000004 age 1 rtr 10.1.1.2
{ dr 10.1.1.1 if 10.1.1.2 tos 0 metric 1 }
{ net 10.2.1.0 mask 255.255.255.0 tos 0 metric 1 } }

10.1.1.1 > 224.0.0.5: OSPFv2-hello 48: area 0.0.0.1 E mask 255.255.255.0 int 10 pri 5 dead 40
dr 10.1.1.1 bdr 10.1.1.2 nbrs 10.1.1.2 [tos 0xc0] [ttl 1] (id 132)

10.1.1.1 > 224.0.0.5: OSPFv2-ls_ack 44: area 0.0.0.1
{ S 80000004 age 1 rtr 10.1.1.2 }

EXEMPLE



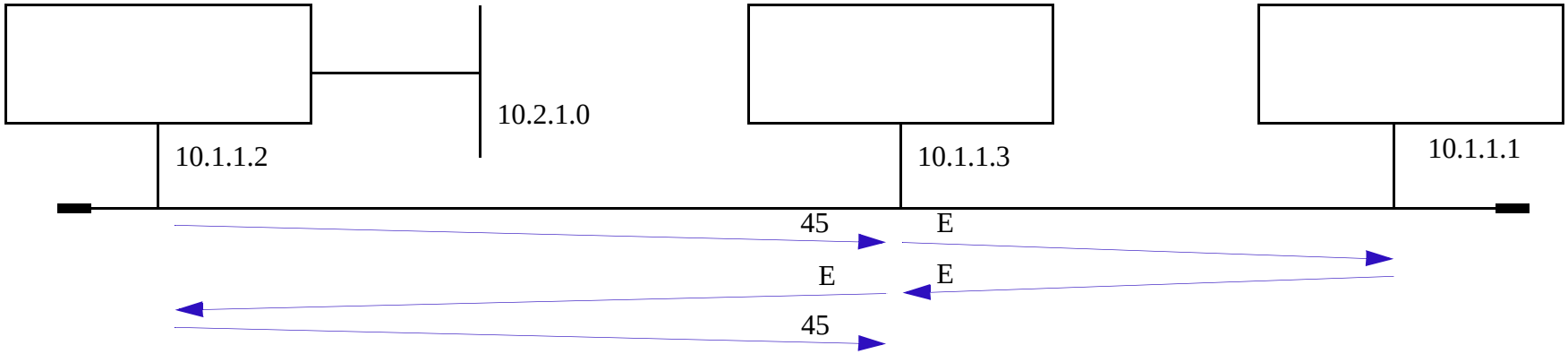
10.1.1.3 > 224.0.0.5: OSPFv2-hello 44: area 0.0.0.1 E mask 255.255.255.0 int 10 pri 17 dead 40 nbrs

10.1.1.1 > 224.0.0.5: OSPFv2-hello 48: area 0.0.0.1 E mask 255.255.255.0 int 10 pri 5 dead 40
dr 10.1.1.1 bdr 10.1.1.2 nbrs 10.1.1.2 [tos 0xc0] [ttl 1] (id 139)

10.1.1.2 > 224.0.0.5: OSPFv2-hello 48: area 0.0.0.1 E mask 255.255.255.0 int 10 pri 27 dead 40
dr 10.1.1.1 bdr 10.1.1.2 nbrs 10.1.1.1 [ttl 1] (id 58925)

10.1.1.3 > 224.0.0.5: OSPFv2-hello 52: area 0.0.0.1 E mask 255.255.255.0 int 10 pri 17 dead 40
nbrs 10.1.1.1 10.1.1.2

EXEMPLE



10.1.1.2 > 10.1.1.3: OSPFv2-dd 32: area 0.0.0.1 E I/M/MS S **45**

10.1.1.1 > 224.0.0.5: OSPFv2-hello 52: area 0.0.0.1 E mask 255.255.255.0 int 10 pri 5 dead 40
dr 10.1.1.1 bdr 10.1.1.2 nbrs 10.1.1.3 10.1.1.2

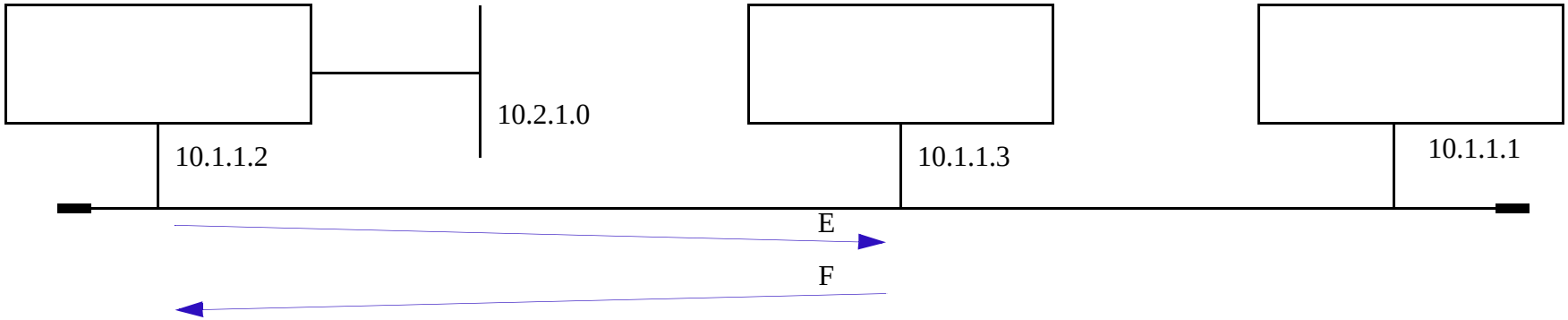
10.1.1.2 > 224.0.0.5: OSPFv2-hello 52: area 0.0.0.1 E mask 255.255.255.0 int 10 pri 27 dead 40
dr 10.1.1.1 bdr 10.1.1.2 nbrs 10.1.1.1 10.1.1.3 [ttl 1] (id 58930)

10.1.1.3 > 10.1.1.1: OSPFv2-dd 32: area 0.0.0.1 E I/M/MS S **E**

10.1.1.3 > 10.1.1.2: OSPFv2-dd 32: area 0.0.0.1 E I/M/MS S **E**

10.1.1.2 > 10.1.1.3: OSPFv2-dd 32: area 0.0.0.1 E I/M/MS S **45**

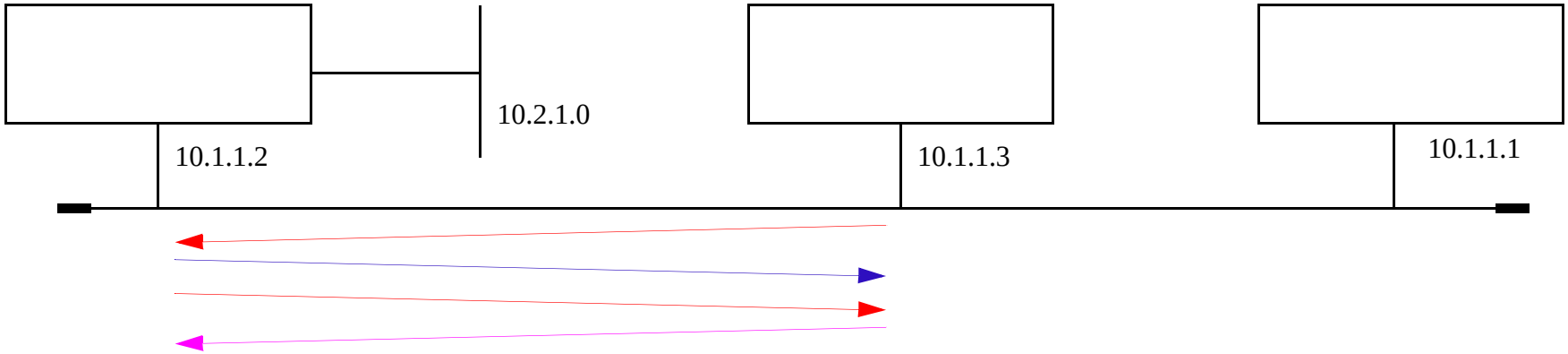
EXEMPLE



10.1.1.2 > 10.1.1.3: OSPFv2-dd 132: area 0.0.0.1 E S E
{ S 80000005 age 25 rtr 10.1.1.2 }
{ E S 80000003 age 1:00 rtr 10.1.1.1 }
{ E S 80000001 age 1:01 net dr 10.1.1.1 if 10.1.1.1 }
{ E S 80000003 age 3:45 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000001 age 3:50 sum 10.1.2.0 abr 10.1.1.1 } [ttl 1] (id 58932)

10.1.1.3 > 10.1.1.2: OSPFv2-dd 52: area 0.0.0.1 E MS S F
{ S 80000002 age 7 rtr 10.1.1.3 } [ttl 1] (id 2762)

EXEMPLE



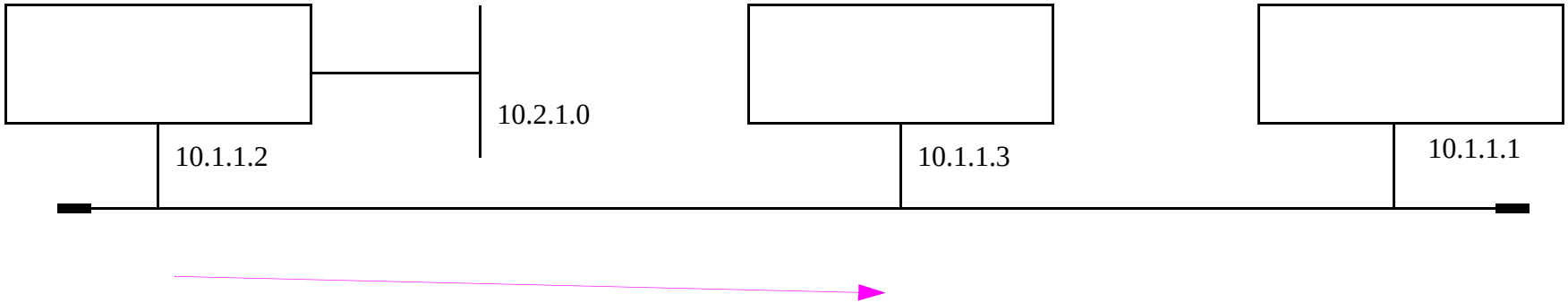
10.1.1.3 > 10.1.1.2: OSPFv2-ls_req 84: area 0.0.0.1
{ rtr 10.1.1.2 } { rtr 10.1.1.1 } { net dr 10.1.1.1 if 10.1.1.1 }
{ sum 10.1.100.0 abr 10.1.1.1 } { sum 10.1.2.0 abr 10.1.1.1 }

10.1.1.2 > 10.1.1.3: OSPFv2-dd 32: area 0.0.0.1 E S F

10.1.1.2 > 10.1.1.3: OSPFv2-ls_req 36: area 0.0.0.1
{ rtr 10.1.1.3 } [ttl 1] (id 58934)

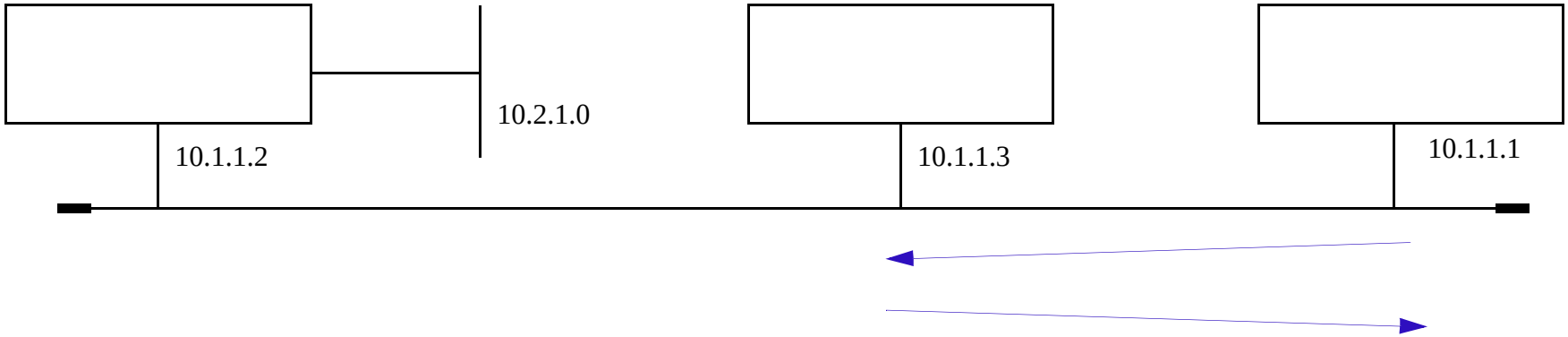
10.1.1.3 > 10.1.1.2: OSPFv2-ls_upd 64: area 0.0.0.1
{ S 80000002 age 8 rtr 10.1.1.3
{ net 10.1.1.0 mask 255.255.255.0 tos 0 metric 1 } }

EXEMPLE



110.1.1.2 > 10.1.1.3: OSPFv2-ls_upd 200: area 0.0.0.1
{ S 80000005 age 26 rtr 10.1.1.2
{ dr 10.1.1.1 if 10.1.1.2 tos 0 metric 1 }
{ net 10.2.1.0 mask 255.255.255.0 tos 0 metric 1 } }
{ E S 80000003 age 1:01 rtr 10.1.1.1 B
{ dr 10.1.1.1 if 10.1.1.1 tos 0 metric 10 } }
{ E S 80000001 age 1:02 net dr 10.1.1.1 if 10.1.1.1 mask 255.255.255.0 rtrs 10.1.1.1 10.1.1.2 }
{ E S 80000003 age 3:46 sum 10.1.100.0 abr 10.1.1.1 mask 255.255.255.0 tos 0 metric 10 }
{ E S 80000001 age 3:51 sum 10.1.2.0 abr 10.1.1.1 mask 255.255.255.0 tos 0 metric 20 }

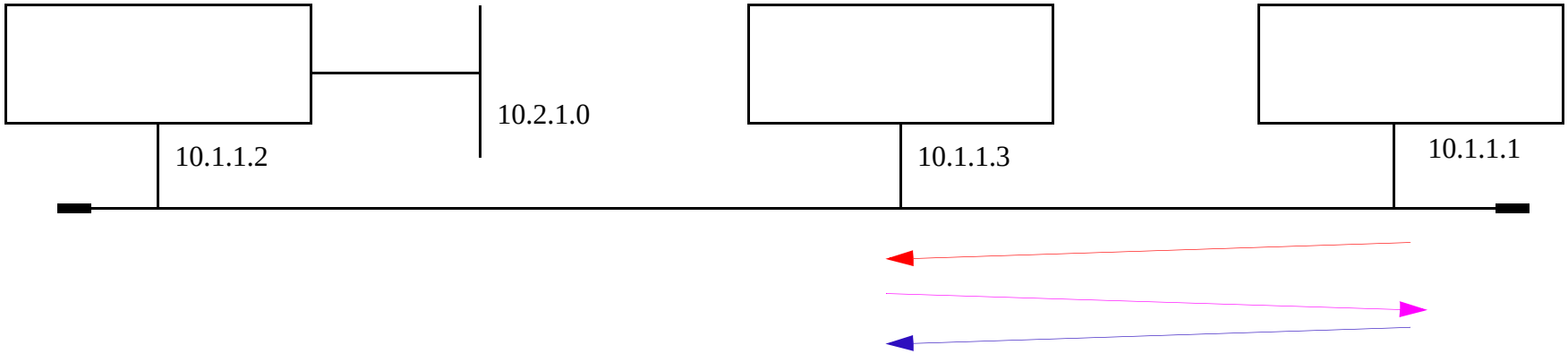
EXEMPLE



10.1.1.1 > 10.1.1.3: OSPFv2-dd 132: area 0.0.0.1 E M S E
{ E S 80000003 age 1:00 rtr 10.1.1.1 } { S 80000005 age 26 rtr 10.1.1.2 }
{ E S 80000001 age 1:00 net dr 10.1.1.1 if 10.1.1.1 }
{ E S 80000001 age 3:50 sum 10.1.2.0 abr 10.1.1.1 }
{ E S 80000003 age 3:45 sum 10.1.100.0 abr 10.1.1.1 }

10.1.1.3 > 10.1.1.1: OSPFv2-dd 152: area 0.0.0.1 E MS S F
{ S 80000005 age 26 rtr 10.1.1.2 } { S 80000002 age 7 rtr 10.1.1.3 }
{ E S 80000003 age 1:01 rtr 10.1.1.1 }
{ E S 80000001 age 1:02 net dr 10.1.1.1 if 10.1.1.1 }
{ E S 80000003 age 3:46 sum 10.1.100.0 abr 10.1.1.1 }
{ E S 80000001 age 3:51 sum 10.1.2.0 abr 10.1.1.1 } [ttl 1] (id 2765)

EXEMPLE

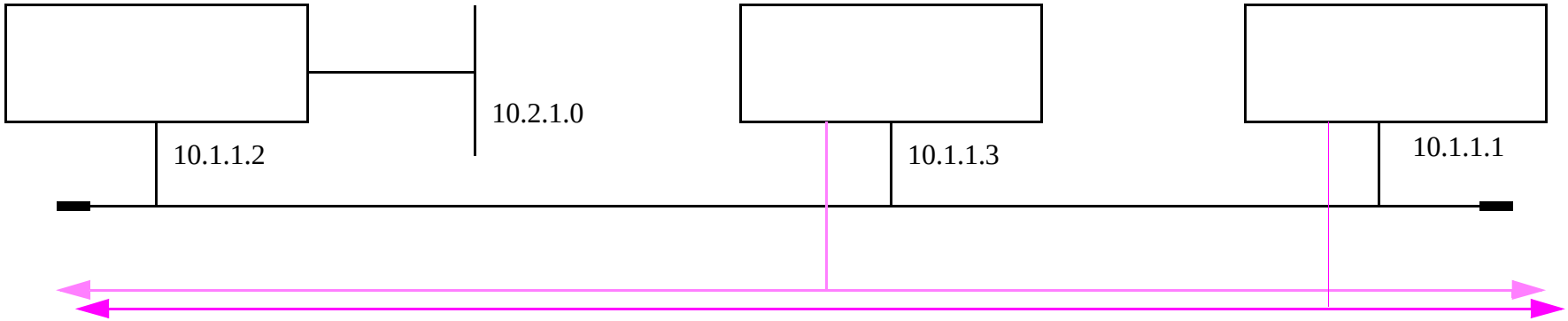


10.1.1.1 > 10.1.1.3: OSPFv2-ls_req 36: area 0.0.0.1 { rtr 10.1.1.3 }

10.1.1.3 > 10.1.1.1: OSPFv2-ls_upd 64: area 0.0.0.1
{ S 80000002 age 8 rtr 10.1.1.3
{ net 10.1.1.0 mask 255.255.255.0 tos 0 metric 1 } }

15:15:51.264608 10.1.1.1 > 10.1.1.3: OSPFv2-dd 32: area 0.0.0.1 E S F

EXEMPLE

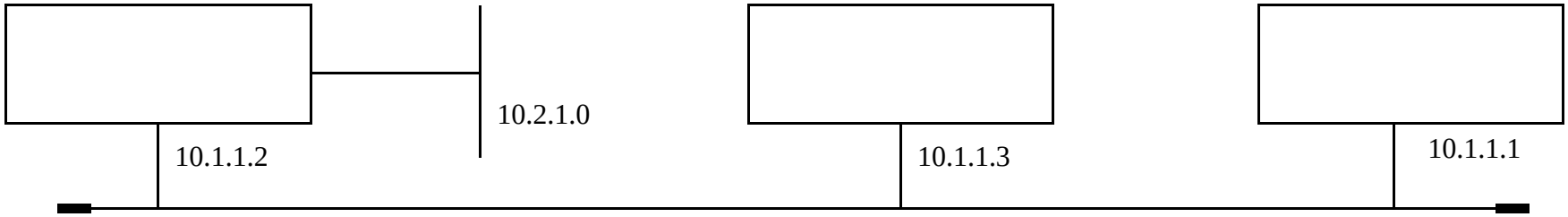


10.1.1.3 > 224.0.0.6: OSPFv2-ls_upd 64: area 0.0.0.1
{ S 80000003 age 1 rtr 10.1.1.3
{ dr 10.1.1.1 if 10.1.1.3 tos 0 metric 1 } }

10.1.1.1 > 224.0.0.5: OSPFv2-ls_upd 64: area 0.0.0.1
{ S 80000002 age 9 rtr 10.1.1.3
{ net 10.1.1.0 mask 255.255.255.0 tos 0 metric 1 } }

10.1.1.1 > 224.0.0.5: OSPFv2-ls_upd 64: area 0.0.0.1
{ E S 80000002 age 1 net dr 10.1.1.1 if 10.1.1.1 mask 255.255.255.0
rtrs 10.1.1.1 10.1.1.3 10.1.1.2 }

EXEMPLE



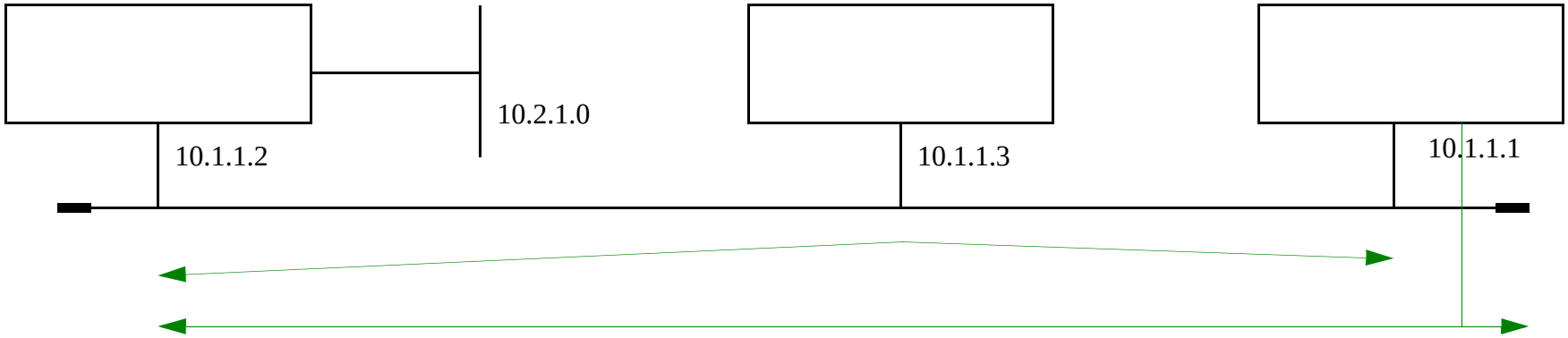
10.1.1.2 > 224.0.0.5: OSPFv2-ls_ack 84: area 0.0.0.1
{ E S 80000002 age 1 net dr 10.1.1.1 if 10.1.1.1 } { S 80000002 age 9 rtr 10.1.1.3 }
{ S 80000002 age 8 rtr 10.1.1.3 }

10.1.1.1 > 10.1.1.3: OSPFv2-ls_upd 64: area 0.0.0.1
{ E S 80000002 age 6 net dr 10.1.1.1 if 10.1.1.1 mask 255.255.255.0 rtrs 10.1.1.1
10.1.1.3 10.1.1.2 }

10.1.1.3 > 224.0.0.6: OSPFv2-ls_ack 44: area 0.0.0.1 { E S 80000002 age 6 net dr 10.1.1.1 if 10.1.1.1 }

10.1.1.3 > 224.0.0.5: OSPFv2-hello 52: area 0.0.0.1 E mask 255.255.255.0 int 10 pri 17 dead 40
dr 10.1.1.1 bdr 10.1.1.2 nbrs 10.1.1.1 10.1.1.2

EXEMPLE

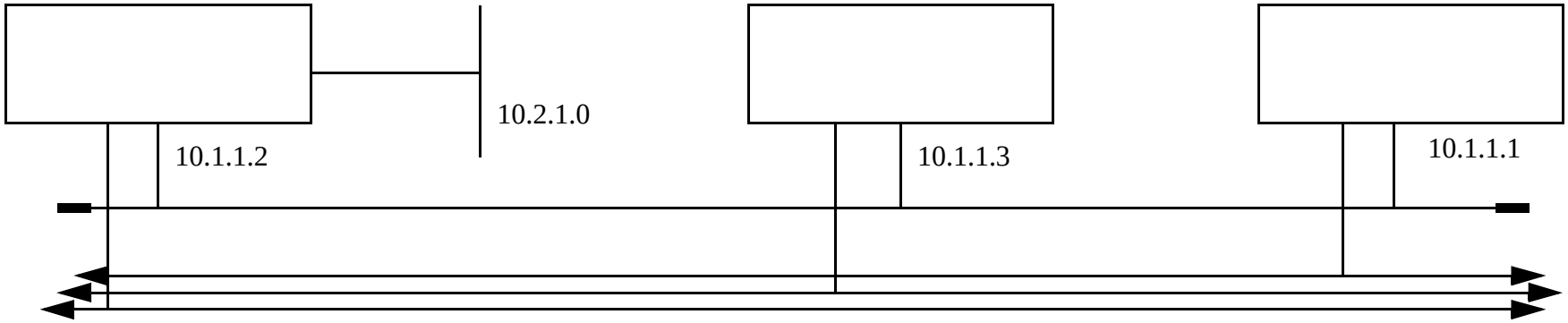


10.1.1.3 > 10.1.1.1: OSPFv2-ls_upd 64: area 0.0.0.1
{ S 80000003 age 9 rtr 10.1.1.3
{ dr 10.1.1.1 if 10.1.1.3 tos 0 metric 1 } }

10.1.1.3 > 10.1.1.2: OSPFv2-ls_upd 64: area 0.0.0.1
{ S 80000003 age 9 rtr 10.1.1.3
{ dr 10.1.1.1 if 10.1.1.3 tos 0 metric 1 } }

10.1.1.1 > 224.0.0.5: OSPFv2-ls_upd 64: area 0.0.0.1
{ S 80000003 age 10 rtr 10.1.1.3
{ dr 10.1.1.1 if 10.1.1.3 tos 0 metric 1 } }

EXEMPLE



10.1.1.1 > 224.0.0.5: OSPFv2-hello 52: area 0.0.0.1 E mask 255.255.255.0 int 10 pri 5 dead 40
dr 10.1.1.1 bdr 10.1.1.2 nbrs 10.1.1.3 10.1.1.2

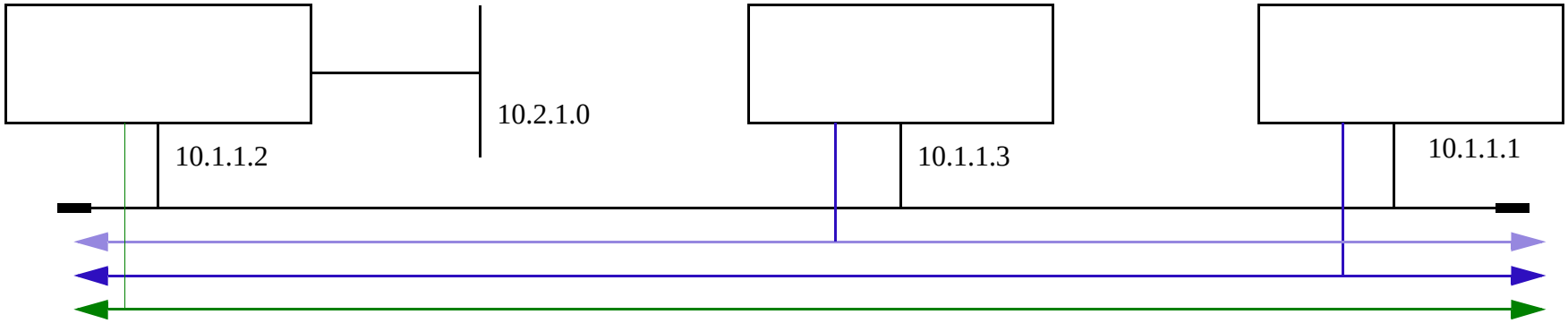
110.1.1.2 > 224.0.0.5: OSPFv2-ls_ack 64: area 0.0.0.1
{ S 80000003 age 9 rtr 10.1.1.3 } { S 80000003 age 9 rtr 10.1.1.3 }

10.1.1.2 > 224.0.0.5: OSPFv2-hello 52: area 0.0.0.1 E mask 255.255.255.0 int 10 pri 27 dead 40
dr 10.1.1.1 bdr 10.1.1.2 nbrs 10.1.1.1 10.1.1.3

10.1.1.3 > 224.0.0.5: OSPFv2-hello 52: area 0.0.0.1 E mask 255.255.255.0 int 10 pri 17 dead 40
dr 10.1.1.1 bdr 10.1.1.2 nbrs 10.1.1.1 10.1.1.2 [ttl 1] (id 2773)

...

EXEMPLE



10.1.1.3 > 224.0.0.6: OSPFv2-ls_upd 64: area 0.0.0.1
{ S 80000004 age 1 rtr 10.1.1.3
{ dr 10.1.1.1 if 10.1.1.3 tos 0 metric 1 } }

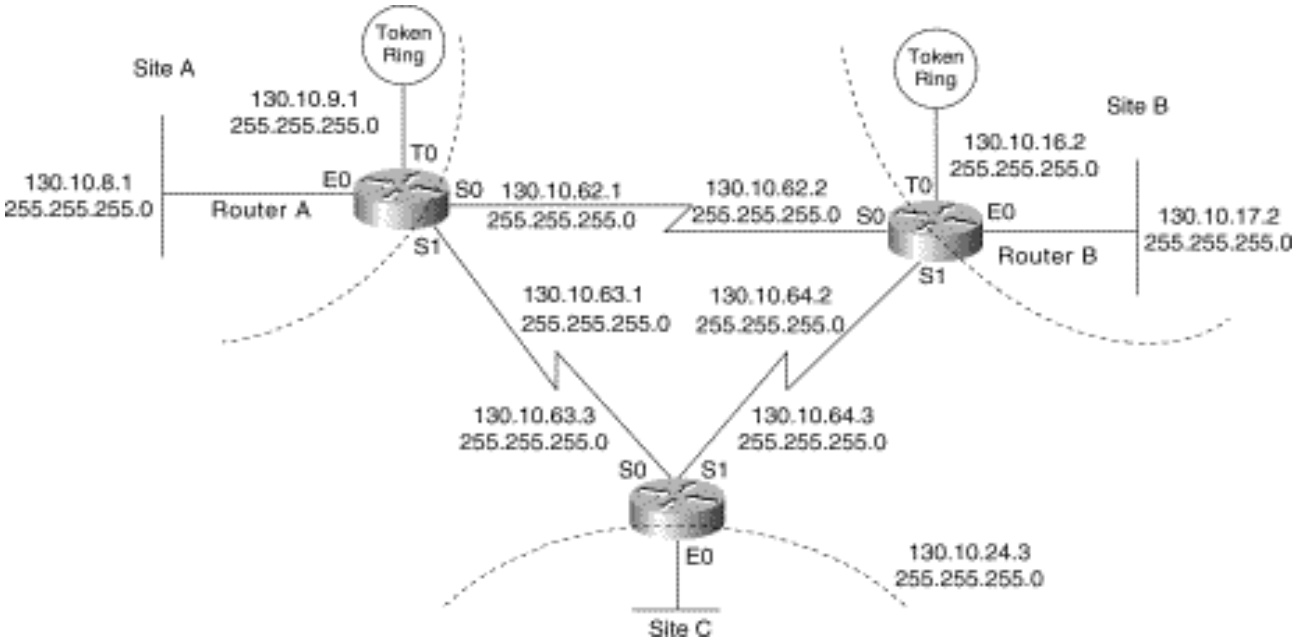
15:16:16.136172 10.1.1.1 > 224.0.0.5: OSPFv2-ls_upd 64: area 0.0.0.1
{ S 80000004 age 2 rtr 10.1.1.3
{ dr 10.1.1.1 if 10.1.1.3 tos 0 metric 1 } }

10.1.1.2 > 224.0.0.5: OSPFv2-ls_ack 64: area 0.0.0.1
{ S 80000004 age 1 rtr 10.1.1.3 }
{ S 80000004 age 1 rtr 10.1.1.3 }

EXEMPLE DE CONFIGURATION OSPF

```
Router#sh conf
Using 508 out of 32762 bytes
!
version 11.3
no service password-encryption
!
hostname Router
!
enable secret 5 $1$lB2a$hXkclsFbROv7zvnIa8s6A.
enable password root
!
!
!
interface Ethernet0
 ip address 192.108.119.137 255.255.255.0
 ip ospf cost 4
!
interface Serial0
 no ip address
 no ip mroute-cache
 shutdown
!
interface Serial1
 no ip address
 shutdown
!
router ospf 1
 network 192.108.119.0 0.0.0.255 area 0.0.0.0
!
ip classless
!
line con 0
line aux 0
line vty 0 4
 password root
 login
!
end
```

EXEMPLE PASSAGE DE RIP À OSPF



Router A :

interface serial 0
  ip address 130.10.62.1 255.255.255.0

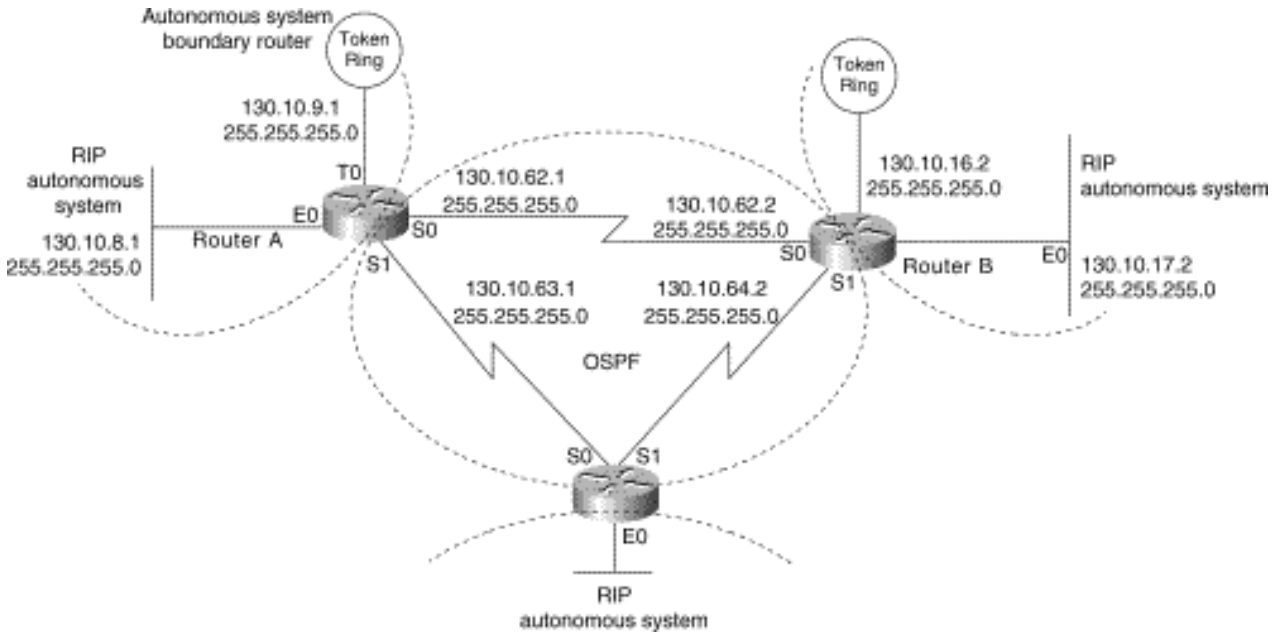
interface serial 1
  ip address 130.10.63.1 255.255.255.0

interface ethernet 0
  ip address 130.10.8.1 255.255.255.0

interface tokenring 0
  ip address 130.10.9.1 255.255.255.0

router rip
  network 130.10.0.0

EXEMPLE PASSAGE DE RIP À OSPF



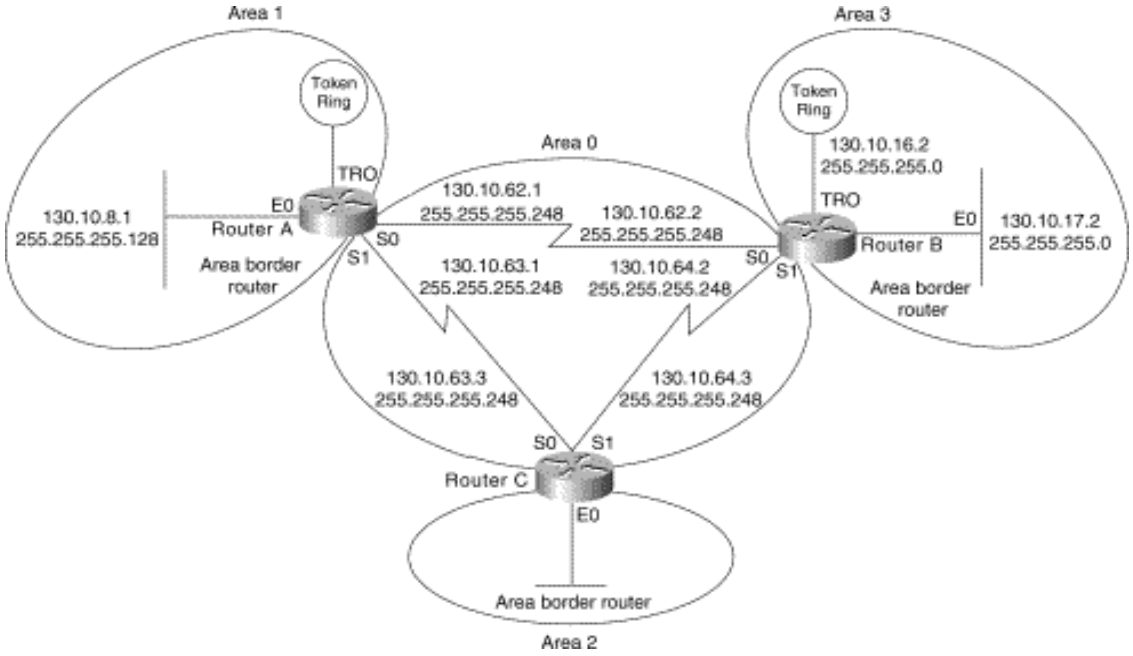
Router A :

! *même définition des interfaces*

router rip
  default-metric 10
  network 130.10.0.0
  passive-interface serial 0
  passive-interface serial 1
  redistribute ospf 109 match internal external 1 external 2
!
!

router ospf 109
  network 130.10.62.0 0.0.0.255 area 0
  network 130.10.63.0 0.0.0.255 area 0
  redistribute rip subnets
  distribute-list 11 out rip
!
access-list 11 permit 130.10.8.0 0.0.7.255
access-list 11 deny 0.0.0.0 255.255.255.255

EXEMPLE PASSAGE DE RIP À OSPF



Router A :

interface serial 0
 ip address 130.10.62.1 255.255.255.248

interface serial 1
 ip address 130.10.63.1 255.255.255.248

interface ethernet 0
 ip address 130.10.8.1 255.255.255.0
 ip irdp

interface tokenring 0
 ip address 130.10.9.1 255.255.255.0
 ip irdp

router ospf 109
 network 130.10.62.0 0.0.0.255 area 0
 network 130.10.63.0 0.0.0.255 area 0
 network 130.10.8.0 0.0.7.255 area 1
 area 1 range 130.10.8.0 255.255.248.0

IS-IS

- Intermediate System to Intermediate System
- Protocole défini par l'ISO
- Utilise :
  - CLNP pour router ISO CLNP
  - CLNP pour router ISO CLNP et IP
  - IP pour router IP
- Utilisé à l'origine par les opérateurs car :
  - permet la gestion des réseaux ISO et IP avec le même protocole
  - opérationnel avant OSPF
- Opérateurs actuels utilisent plutôt OSPF
- Groupe de travail ISIS à l'IETF :
  - permet d'ajouter plus facilement un nouveau protocole (par exemple : IPv6)
- Principe très similaire à OSPF (deux niveaux de réseau : niveau 1 : équivalent aux aires d'OSPF et niveau 2 : équivalent au backbone d'OSPF)

SYSTÈMES AUTONOMES

- Impossible de gérer le réseau Internet de manière centralisée.
- Un système autonome dispose d'une autonomie de gestion
  - numérotation des machines
  - plan de routage interne (manuel ou IGP)
  - ...
- Un système autonome est représenté par un entier sur 16 bits
- Alloué par l'IANA
- Il faut des points de centralisation de l'information.

LES CENTRES DE COORDINATION DE L'INTERNET

- IANA (ou ICANN)
- InterNIC (USA) : Monde
- RIPE-NCC : Europe
- APNIC : Asie Pacifique

LES BASES DE DONNÉES DE RÉFÉRENCE

- InterNIC (USA) : rs.internic.net
- ARIN (Amérique du Nord) : whois.arin.net
- RIPE-NCC (Pays Bas) : whois.ripe.net
- APNIC (Japon) : whois.nic.ad.jp

LES AUTRES BASES DE DONNÉES

- Merit (USA) : http://www.ra.net
- MCI (USA) : http://www.mci.net
- Ca'net (Canada) : http://www.canet.ca
- ...

BASES RÉGIONALES

- AFNIC (France) : http://www.nic.fr
- Allemagne : http://www.nic.de
- ...

LES OBJETS RÉPÉRTORIÉS

+ Pour RIPE la définition est donnée dans des documents RIPE-xxx

  - Adresse de réseau : appartenance (RIPE 141 et 142)
  - Systèmes Autonomes : politiques de routage (RIPE 147, 140 et 181)
  - Noms de domaine : serveur de noms, réseaux (RIPE 49)
  - Personnes : coordonnées des responsables (RIPE 119)
  - Routes : ASN, agrégats (RIPE 181)
  - ...

+ Pour d'autres organismes, la structuration est plus libre

LA BASE RIPE

La base RIPE n'est pas directement utilisée pour le routage, mais :

- permet d'avoir une vision centralisée du réseau
- permet de trouver les incohérences dans les règles de routage distribué

La base RIPE est maintenue par les gestionnaires du système autonome.

La base RIPE est gérée par les réseaux qui ont plusieurs attachements à plusieurs systèmes autonomes

Par exemple, RENATER gère les informations des systèmes autonomes qui lui sont connectés.

+ La base contient la liste des AS avec lesquels des informations de routage sont échangées.

LA BASE RIPE

>whois 192.44.77.0

inetnum: 192.44.77.0
netname: FNET-ENSTB-3
descr: ENST Bretagne - antenne de Rennes
descr: BP 78, 35512 Cesson Sevigne CEDEX, France
country: FR
admin-c: Pierre Rolin
tech-c: Marc Fradin
remarks: OR
changed: <Annie.Renard@inria.fr> 930908
source: RIPE

route: 192.44.77.0/24
descr: FNET-ENSTB-3
origin: AS1717
remarks: OR
mnt-by: AS1717-MNT
changed: ripe-dbm@ripe.net 941121
source: RIPE

person: Pierre Rolin
address: ENST Bretagne - antenne de Rennes

address: BP 78, F-35512 Cesson Sevigne CEDEX, France
phone: +33 99 12 70 21
e-mail: Pierre.Rolin@rennes.enst-bretagne.fr
changed: Annie.Renard@inria.fr 940810
source: RIPE

person: Marc Fradin
address: ENST Bretagne - antenne de Rennes
address: BP 78, F-35512 Cesson Sevigne CEDEX, France
phone: +33 99 12 70 24
e-mail: Marc.Fradin@rennes.enst-bretagne.fr
changed: Annie.Renard@inria.fr 940810
source: RIPE

AUX US, LA BASE ARIN

>**whois -h whois.arin.net 18.0.0.0**
Massachusetts Institute of Technology (NET-MIT-TEMP)
1 Amherst Street Cambridge, MA 02139-1986

Netname: MIT
Netblock: 18.0.0.0 - 18.255.255.255

Coordinator:
Schiller, Jeffrey I (JIS-ARIN) jis@MIT.EDU
+1 617 253-8400 (FAX) +1 617 258-8736

Domain System inverse mapping provided by:

STRAWB.MIT.EDU 18.71.0.151
W20NS.MIT.EDU 18.70.0.160
BITSY.MIT.EDU 18.72.0.3

Record last updated on 26-Sep-98.
Database last updated on 12-Feb-99 16:12:44 EDT.

The ARIN Registration Services Host contains ONLY Internet Network Information: Networks, ASN's, and related POC's. Please use the whois server at rs.internic.net for DOMAIN related Information and nic.mil for NIPRNET Information.

LA COMMANDE WHOIS

>**whois AS1717**

aut-num: AS1717
descr: RENATER
descr: Reseau National de telecommunications pour la Technologie
descr: l'Enseignement et la Recherche
descr: FR
as-in: from AS1755 100 accept ANY
as-in: from AS5511 100 accept ANY
as-in: from AS2470 100 accept AS2470
as-in: from AS789 50 accept AS789
as-in: from AS786 50 accept AS786
as-in: from AS1899 50 accept AS1899
as-in: from AS2917 50 accept AS2917
as-in: from AS3215 50 accept AS3215
as-out: to AS5511 announce AS1717
as-out: to AS2470 announce AS1717
as-out: to AS1755 announce AS1717 AS2470
as-out: to AS786 announce AS1717
as-out: to AS789 announce AS1717
as-out: to AS1899 announce AS1717
as-out: to AS2917 announce AS1717
as-out: to AS3215 announce AS1717

default: **AS1755 10**
admin-c: Michel Lartail
tech-c: Isabelle Morel
tech-c: Marie-Helene Guilmin
mnt-by: AS1717-MNT
changed: rensvp@renater.fr 960423
source: RIPE

NOM DES AS

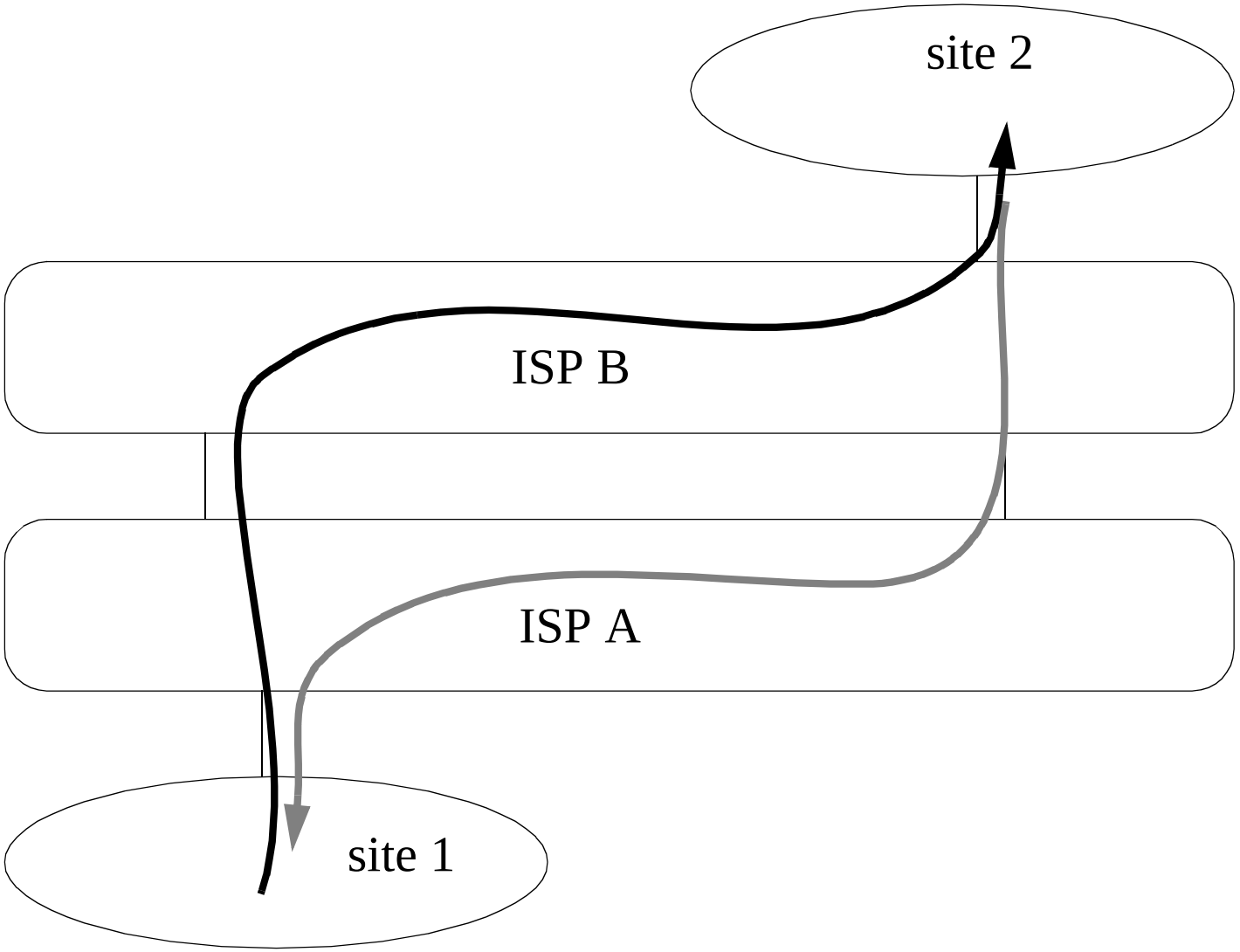
AS	Nom
AS1755	EBONE
AS5511	TRANSPAC
AS2470	Ile de la Réunion Département d'Outre-Mer
AS789	IN2P3
AS786	The JANET IP Service
AS1899	Fnet, EUnet-France
AS2917	OLEANE - PIPEX International

AS	Nom
AS3215	RAIN (Réseau d'Acces àl'INternet TRANSPAC)

OUTILS : TRACEROUTE

```
>./traceroute rki.kbs.co.kr
traceroute to rki.kbs.co.kr (210.115.193.23): 1-30 hops, 38 byte packets
 1  mgs-in.rennes.enst-bretagne.fr (193.52.74.2)  3.11 ms  2.74 ms  2.31 ms
 2  ft-renater.rennes.enst-bretagne.fr (193.51.128.81)  3.21 ms  3.16 ms  3.2 ms
 3  193.48.78.41 (193.48.78.41)  4.21 ms  93.6 ms  38.2 ms
 4  rennes.or-br.ft.net (193.48.78.25)  4.47 ms  4.50 ms  4.60 ms
 5  rennes.renater.ft.net (193.55.253.170)  4.77 ms  9.50 ms  4.90 ms
 6  stamand2.renater.ft.net (195.220.180.153)  10.3 ms  12.9 ms  10.6 ms
 7  rbs2.renater.ft.net (195.220.180.34)  10.3 ms  10.4 ms  10.3 ms
 8  paii.renater.ft.net (195.220.180.29)  13.1 ms  12.1 ms  12.1 ms
 9  relay-pos-6.opentransit.net (193.55.152.70)  87.1 ms  87.9 ms  86.9 ms
10  sl-bb11-rly-0-1.sprintlink.net (144.232.8.209)  127 ms  96.3 ms  88.0 ms
11  sl-bb2-dc-4-0-0.sprintlink.net (144.232.7.142)  101 ms  90.2 ms  88.5 ms
12  core7-hssi0-0-0.Washington.cw.net (206.157.77.33)  89.9 ms  89.3 ms  93.2 ms
13  bordercore2.SanFrancisco.cw.net (166.48.14.1)  160 ms (ttl=240!)  172 ms (ttl=240!)  159 ms (ttl=240!)
14  dacom.SanFrancisco.cw.net (166.48.15.246)  192 ms (ttl=239!)  164 ms (ttl=239!)  161 ms (ttl=239!)
15  gateway.bora.net (203.233.35.249)  298 ms (ttl=238!)  300 ms (ttl=238!)  321 ms (ttl=238!)
16  210.120.128.4 (210.120.128.4)  312 ms (ttl=237!)  311 ms (ttl=237!)  324 ms
17  203.233.37.146 (203.233.37.146)  317 ms (ttl=236!)  336 ms (ttl=236!)  324 ms (ttl=236!)
18  rki.kbs.co.kr (210.115.193.23)  346 ms (ttl=108!)  375 ms (ttl=108!)  329 ms  (ttl=108!)
```

SYMÉTRIE DU ROUTAGE ???



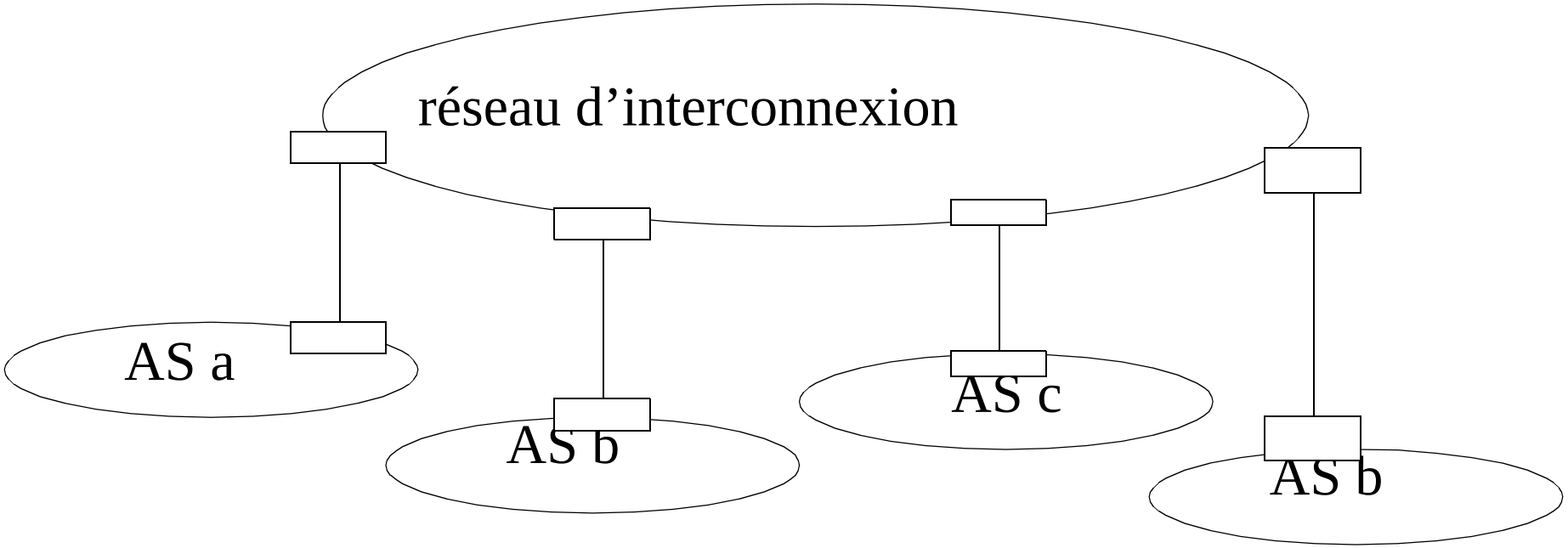
TRACEROUTE AVEC SYSTEME AUTONOME

```
>./traceroute -A -h whois.ripe.net -n rki.kbs.co.kr
traceroute to rki.kbs.co.kr (210.115.193.23): 1-30 hops, 38 byte packets
 1  193.52.74.2 [AS1717 - FR-OR-ENSTRE]  2.95 ms  2.31 ms  2.29 ms
 2  193.51.128.81 [AS1717 - RENATER]  3.7 ms  3.23 ms  3.13 ms
 3  193.48.78.41 [AS1717 - proxy aggregate done by ENONE of RENATER (AS1717) prefixes] 4.23 ms 4.21 ms 4.12 ms
 4  193.48.78.25 [AS1717 - proxy aggregate done by ENONE of RENATER (AS1717) prefixes] 4.45 ms 4.63 ms 4.48 ms
 5  193.55.253.170 [AS1717 - RENATER]  4.57 ms  4.64 ms  4.79 ms
 6  195.220.180.153 [AS1717 - FR-RENATER-RNI]  11.1 ms  10.3 ms  10.5 ms
 7  195.220.180.34 [AS1717 - FR-RENATER-RNI]  12.7 ms  15.5 ms  11.2 ms
 8  195.220.180.29 [AS1717 - FR-RENATER-RNI]  13.3 ms  12.3 ms  11.6 ms
 9  193.55.152.70 [AS5511 - OPENTRANSIT.NET]  87.4 ms  87.6 ms  89.7 ms
10  144.232.8.209  88.7 ms  88.0 ms  87.6 ms
11  144.232.7.142  88.2 ms  88.9 ms  88.2 ms
12  206.157.77.33  90.2 ms  89.8 ms  98.0 ms
13  166.48.14.1  159 ms (ttl=240!)  160 ms (ttl=240!)  160 ms (ttl=240!)
14  166.48.15.246  163 ms (ttl=239!)  160 ms (ttl=239!)  159 ms (ttl=239!)
15  203.233.35.249  332 ms (ttl=238!)  328 ms (ttl=238!)  329 ms (ttl=238!)
16  210.120.128.4  327 ms (ttl=237!)  325 ms (ttl=237!)  326 ms (ttl=237!)
17  203.233.37.146  374 ms (ttl=236!)  347 ms (ttl=236!)  340 ms (ttl=236!)
18  210.115.193.23  345 ms (ttl=108!)  333 ms (ttl=108!)  353 ms (ttl=108!)
```

LE PROTOCOLE EGP

- Protocole très simple
- Obsolète (sauf cas très particuliers)
- Ne prend pas en compte une hiérarchie de réseaux
- Basé sur un réseau d'interconnexion unique (topologie d'origine de l'Internet)
- Plutôt un protocole d'annonce
- Symétrie : accord pour envoyer des annonces & accord pour les recevoir

EXEMPLE



ANNONCES

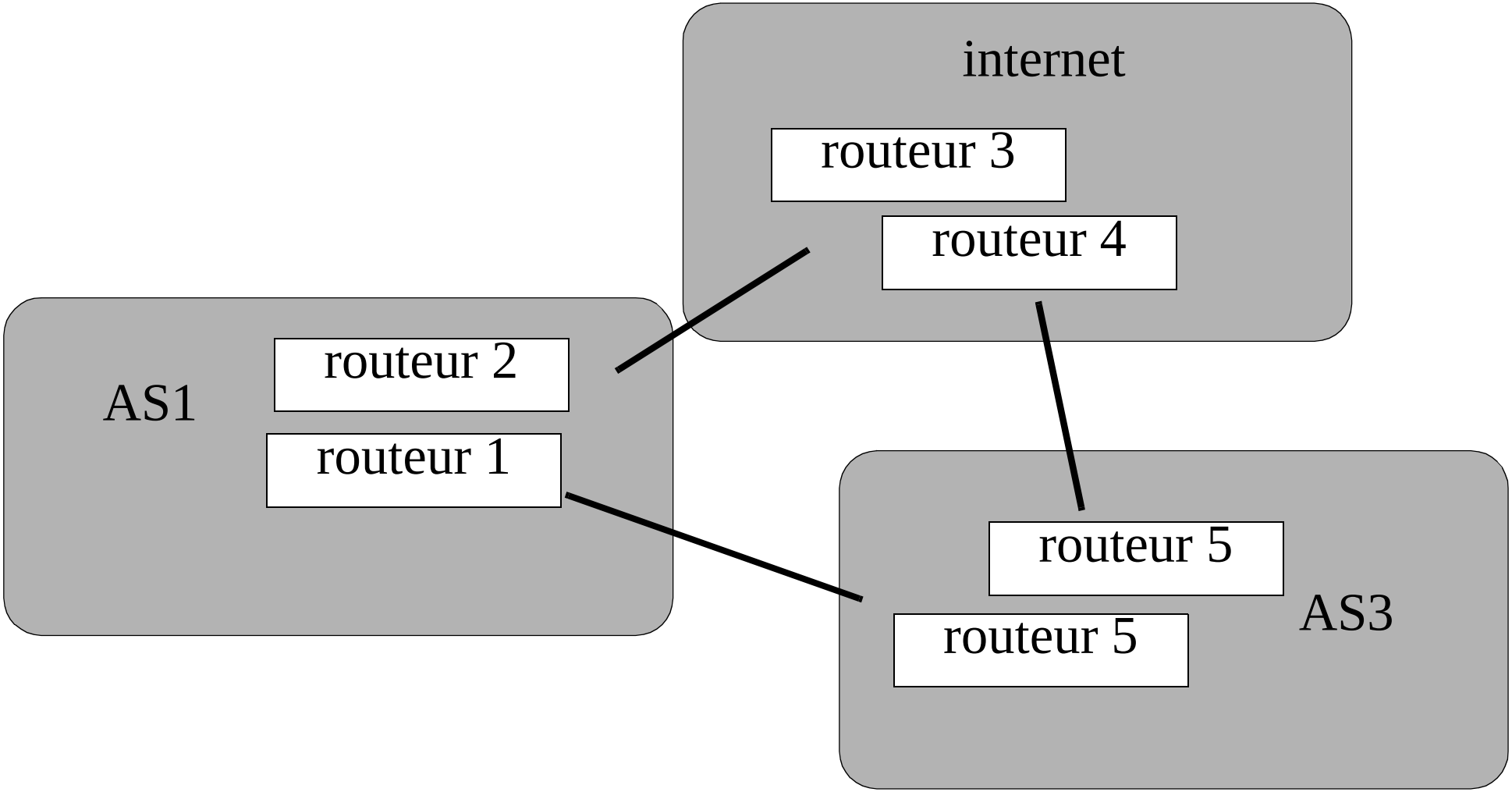
Problèmes :

- Le domaine A ne veut pas servir de réseau de transit pour B
  - Ne pas annoncer vers l'extérieur les routes pour B

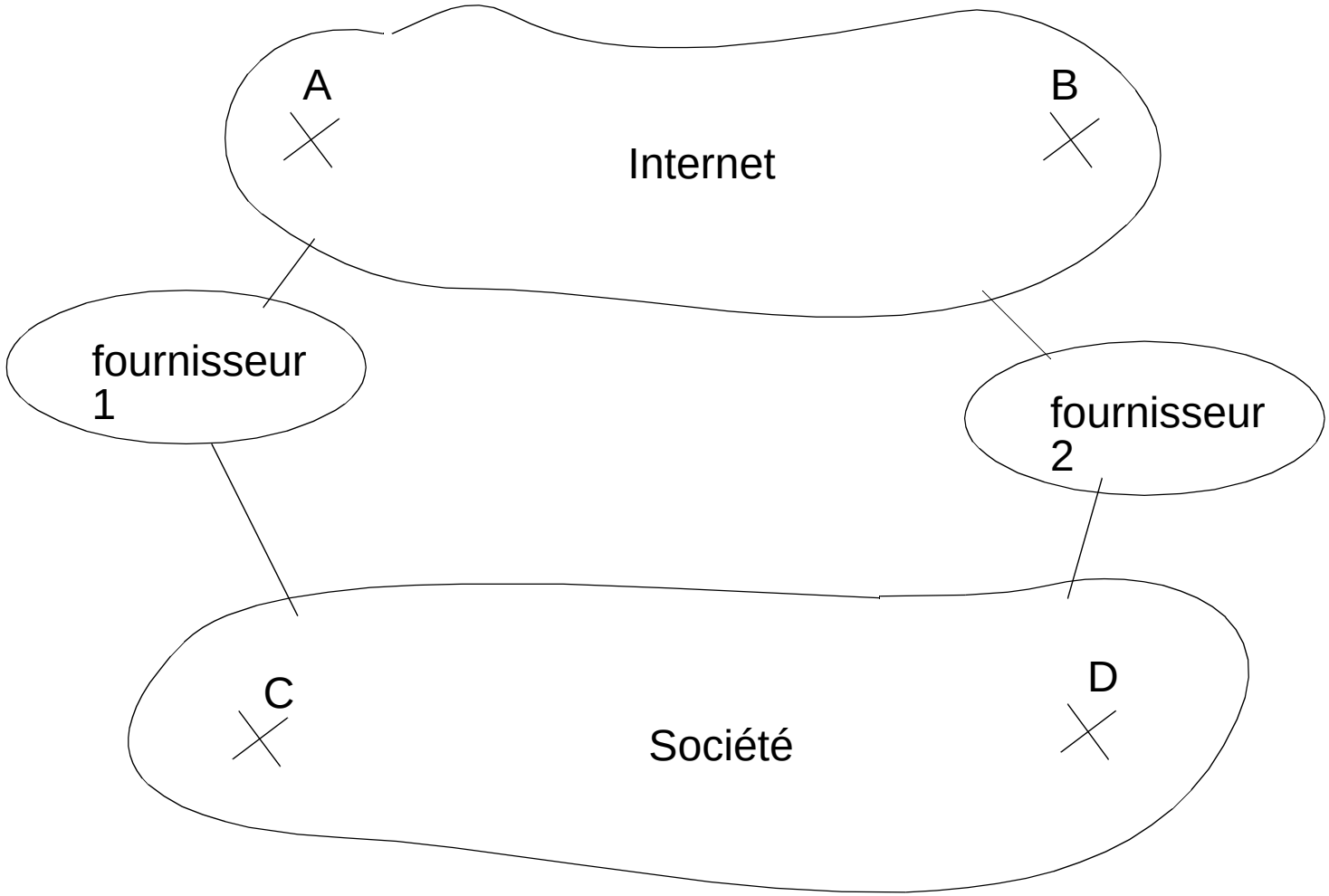
- Les stations du domaine A veulent utiliser la liaison en coulisse pour joindre le domaine B
  - Le domaine B doit annoncer les routes vers ses machines
  - Le domaine A ne doit pas les réannoncer sur le backbone

- En cas de panne du routeur en frontière du domaine A, le domaine A veut pouvoir être accessible via le domaine B et la liaison en coulisse.
  - Le domaine B doit annoncer les routes apprises du domaine A mais avec une préférence moins bonne.

LIMITE D'EGP


internet
routeur 3
routeur 4
AS1
routeur 2
routeur 1
routeur 5
routeur 5
AS3

CAS PLUS COMPLEXE



ANNONCES

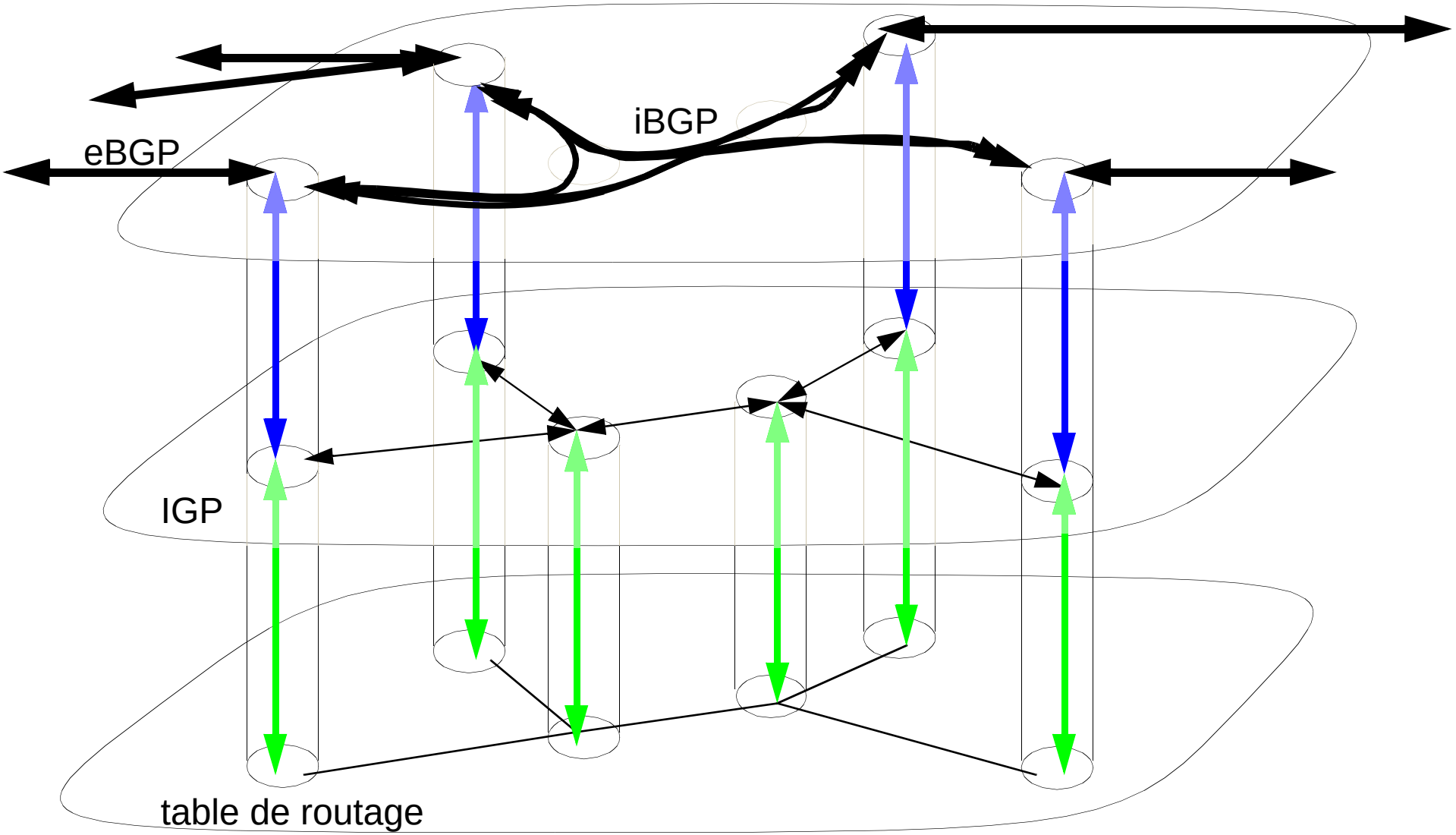
- La machine C veut joindre la machine B
  - Si utilisation de routes par défaut, le paquet sort au plus vite du réseau pour aller via le fournisseur 1 vers l'Internet
  - Si les annonces du fournisseur 2 sont diffusées dans le réseau de la société, le paquet passera par le fournisseur 2 (si la métrique est meilleure).
- La machine A veut joindre la machine C
  - Si la société annonce que les routes "proches" du fournisseur 1, le paquet transitera via l'internet.
  - Si la société annonce toutes ses routes, le paquet passera par le fournisseur 1.

- Si l'utilisateur de A utilise l'option de Source Routing pour joindre B
  - il peut utiliser les ressources privées de la société
  - filtrer les options de Source Routing en frontière de zone

BGP

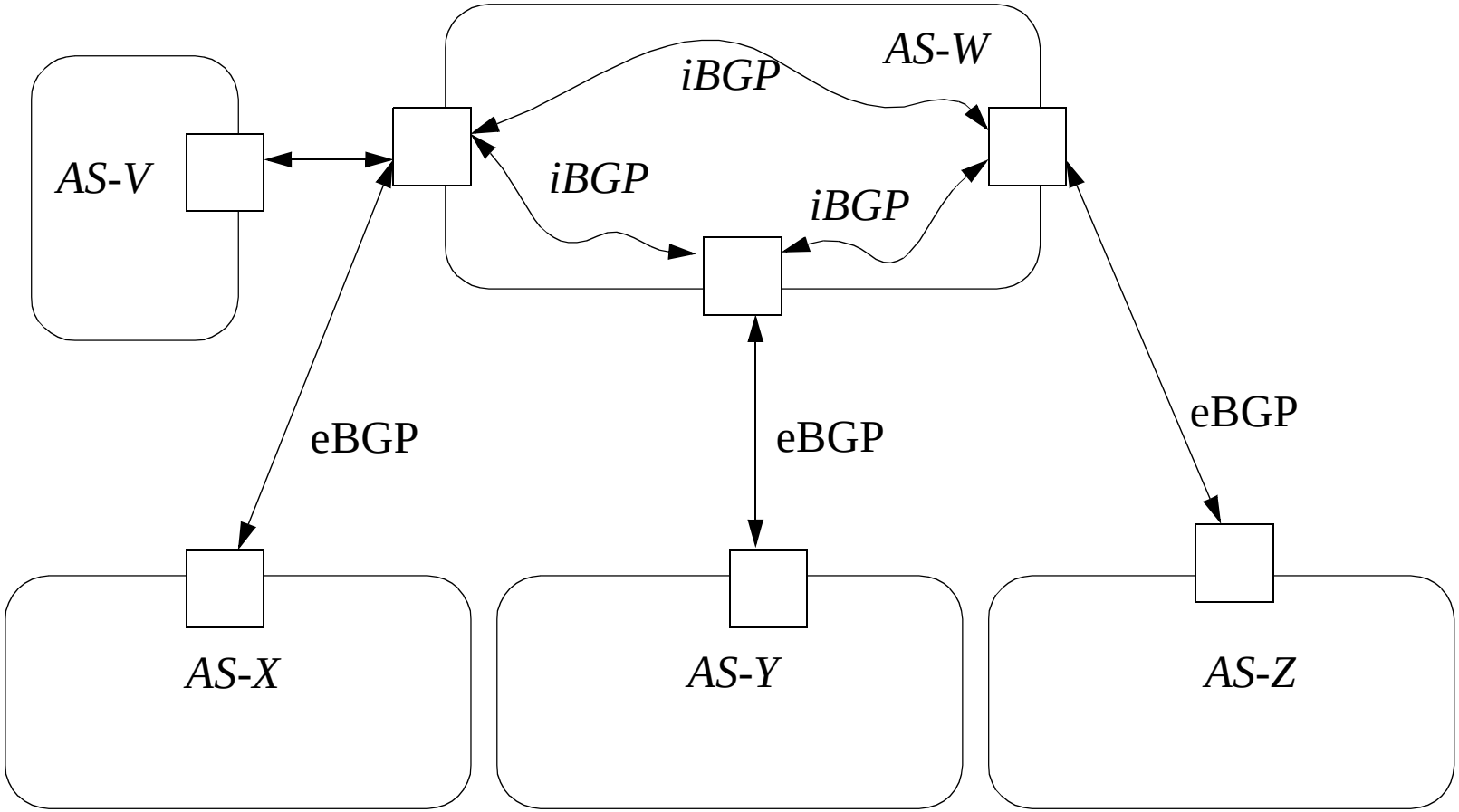
- Propagation des annonces de routes
- Evite des boucles par maintien du chemin des AS traversés (semblable au Source Routing de l'anneau à jeton).
- BGP 4 permet l'adressage CIDR
- La future version sera le protocole IDRP de l'ISO
- Utilisation de TCP (EGP utilise UDP) port 179 pour faciliter :
  - ouverture
  - mise à jour (échange d'information de routage)
  - notification (détection d'une erreur de protocole)
  - sonde (vérifier la vie de la connexion)

PRINCIPES



IBGP

- Protocole de routage Externe (mais utilisé en interne)
- Synchronisation des bases de données
- Pose des problèmes de synchronisation avec les protocoles de routage internes.

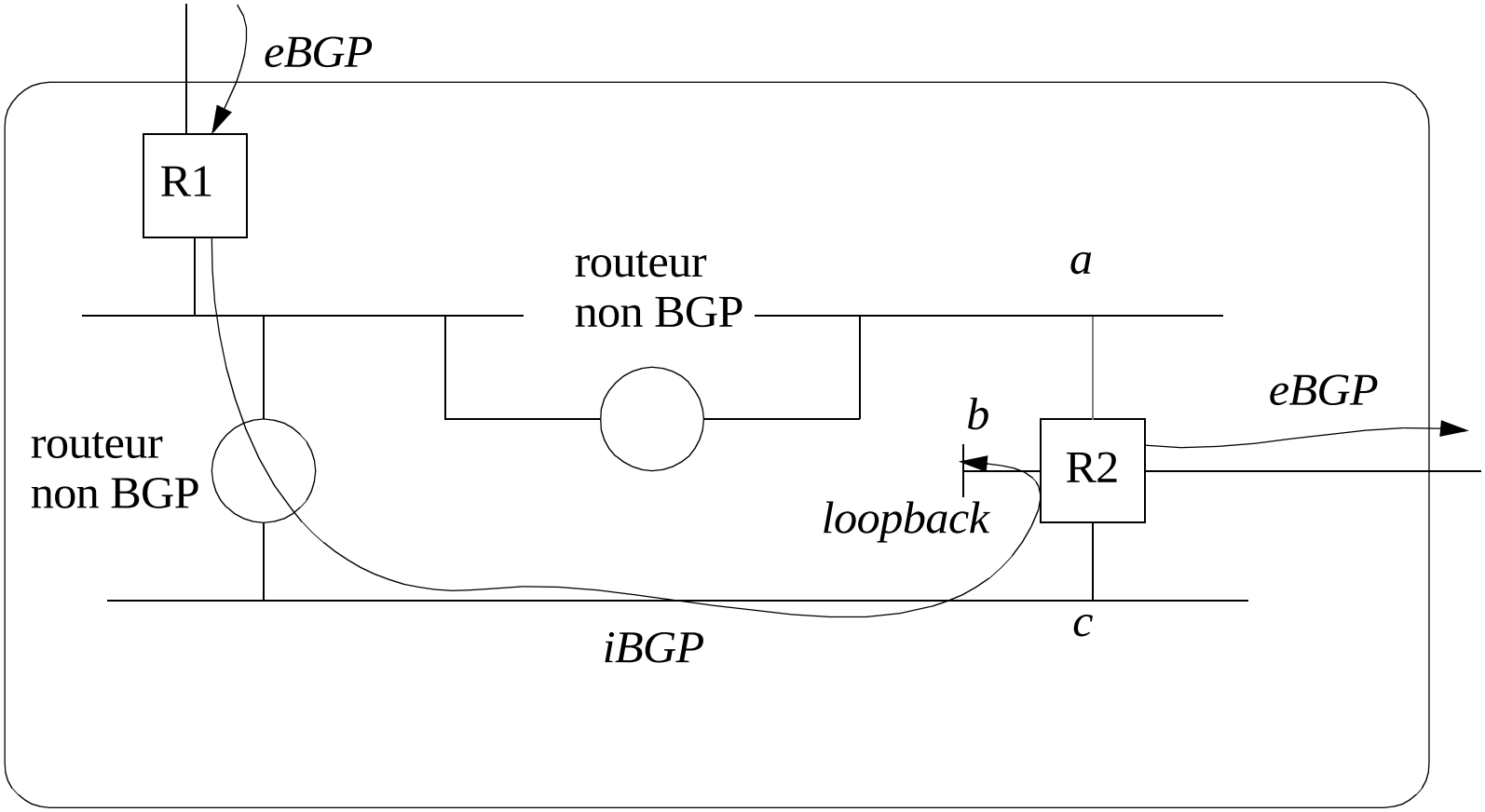


ATTRIBUTS.

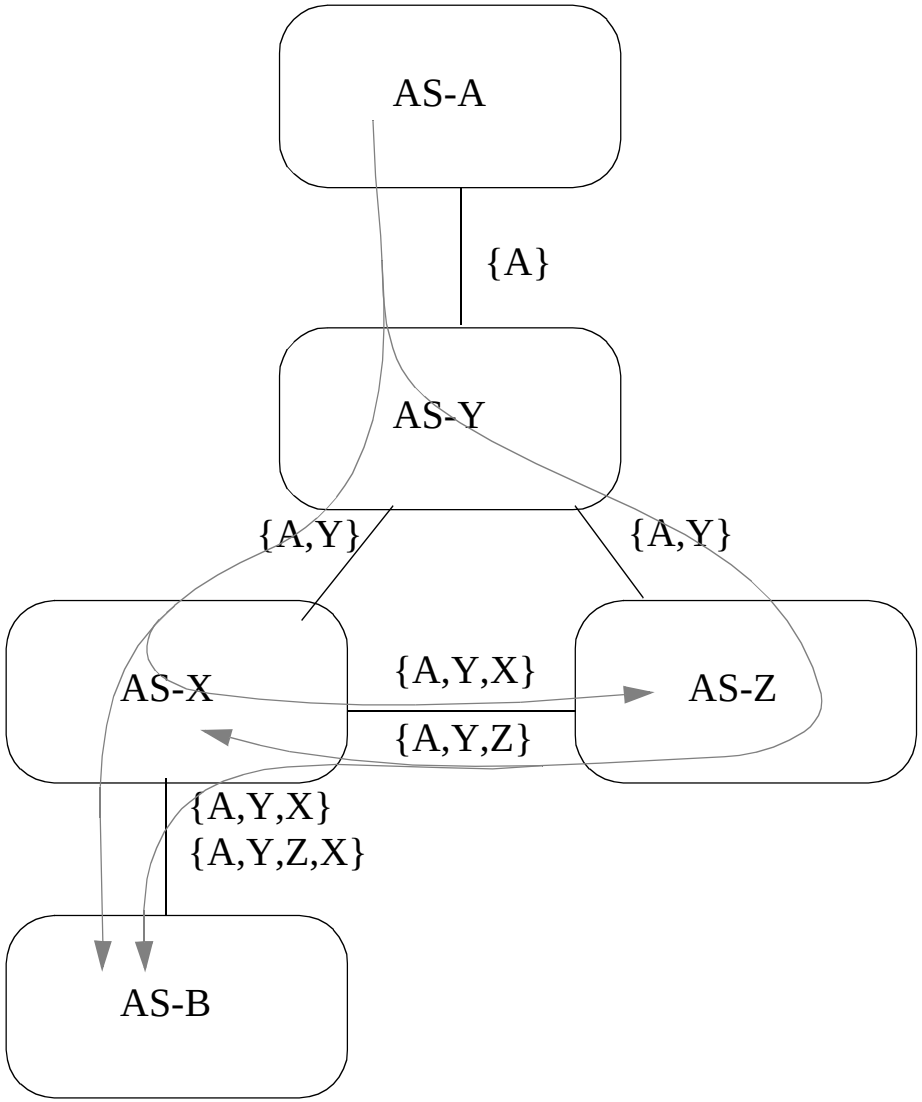
drapeau	code	Attribut
bien connu obligatoire	1	**ORIGIN**
bien connu obligatoire	2	**AS_PATH**
bien connu obligatoire	3	**NEXT_HOP**.
optionnel, non transitif	4	**MULTI_EXT_DISCR** (BGP-4) ou **INTER-AS** (BGP-3).
bien connu facultatif	5	**LOCAL_PREF**
bien connu facultatif	6	**ATOMIC_AGGREGATE.**
optionnel transitif	7	**AGGREGATOR.**
optionnel transitif	8	**COMMUNITY.**t.
optionnel non-transitif	9	**ORIGINATOR_ID.**
optionnel non-transitif	10	**CLUSTER_LIST**
	11	**DPA**
optionnel non-transitif	12	**ADVERTISER**
optionnel non-transitif	13	**RCID_PATH / CLUSTER_ID**

- bien connu : mis en œuvre dans tous les routeurs
- transitif : peut sortir de l'AS

UTILISATION D'UNE ADRESSE *LOOPBACK*



AS_PATH



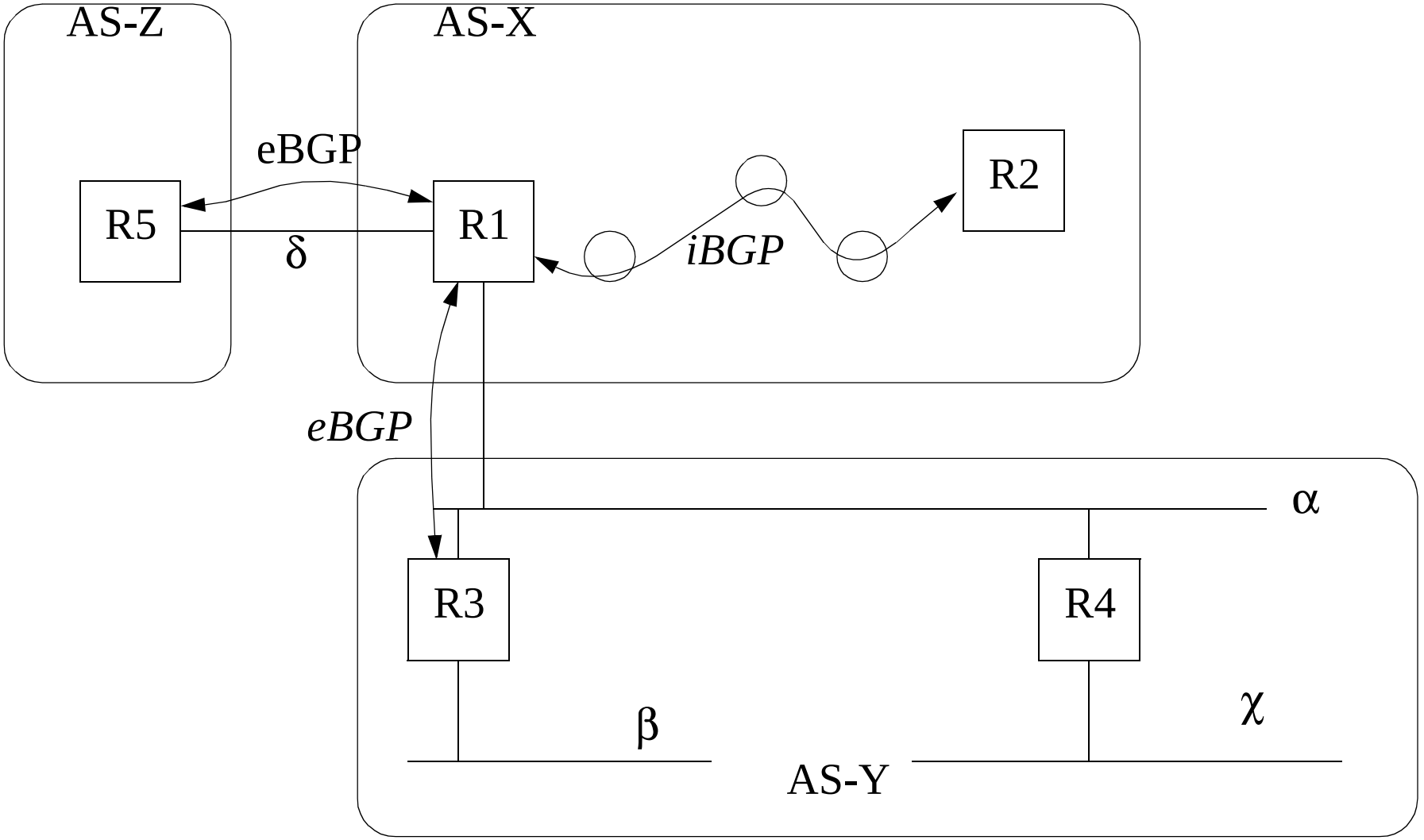
Evite les boucles dans les annonces

Permet de choisir le chemin le plus court

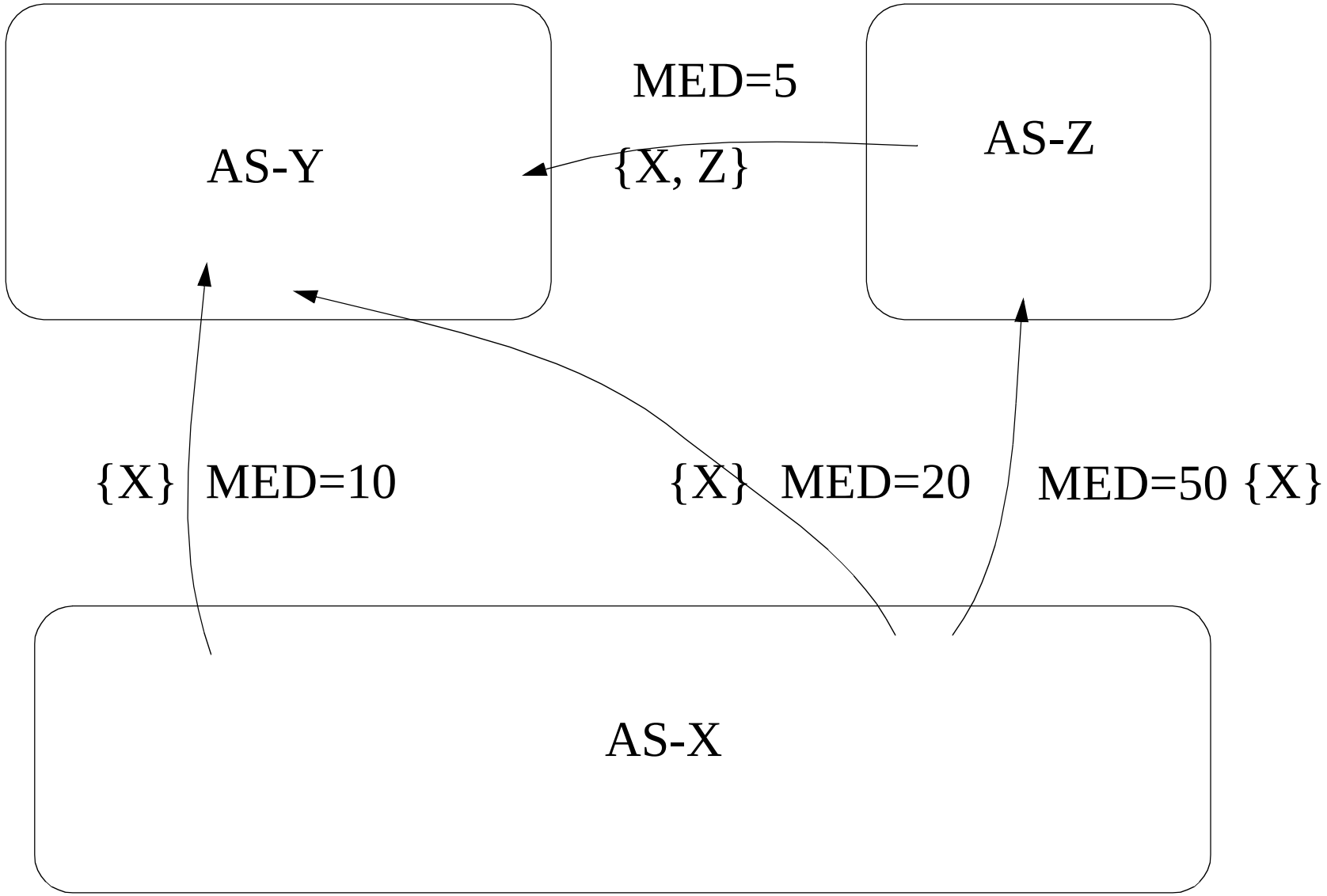
+Certains opérateurs ajoutent plusieurs fois leur numéro d'AS pour défavoriser cette route

+Problème en iBGP car les boucles ne sont pas détectées avec cet attribut

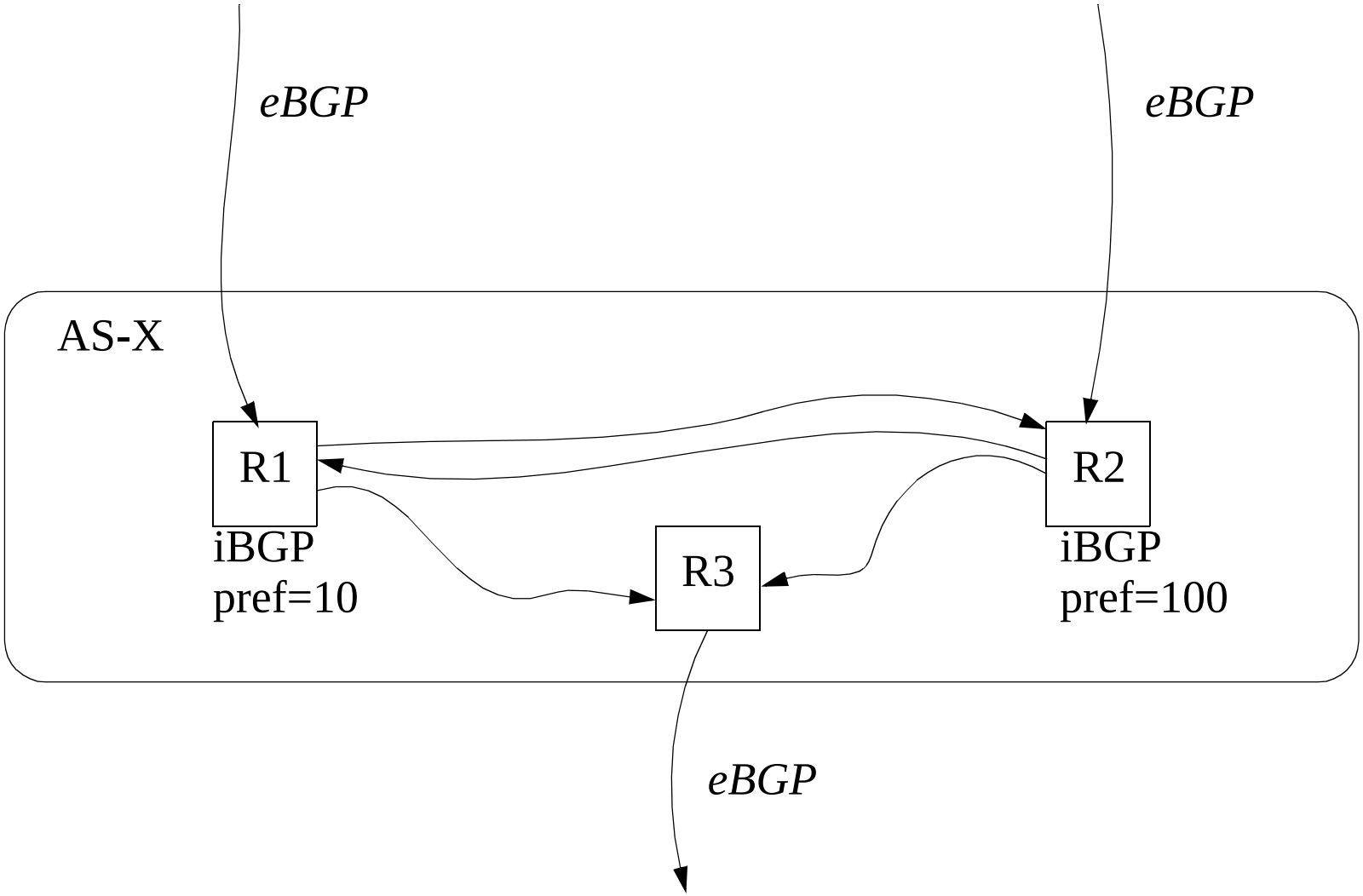
NEXT_HOP



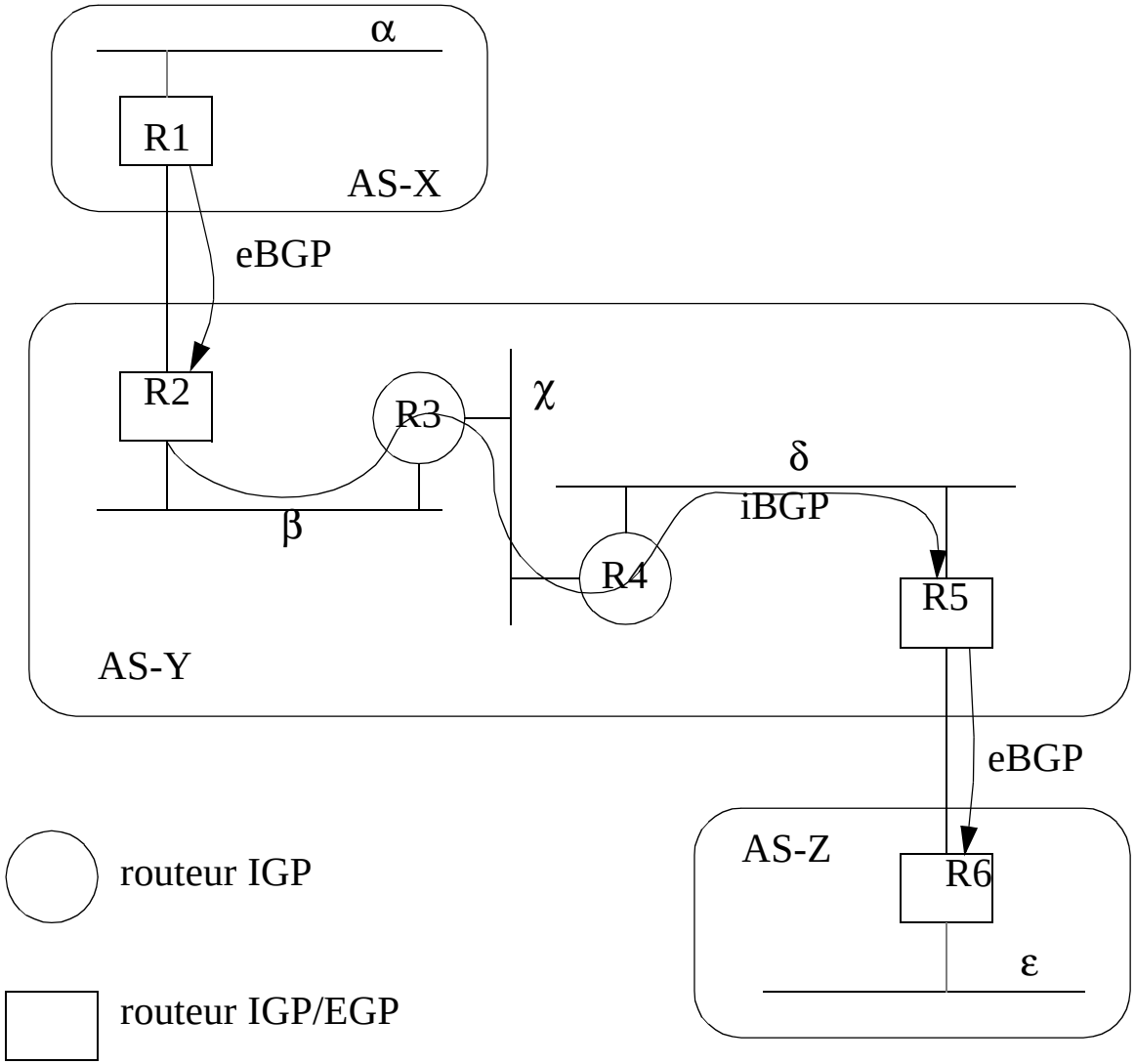
MULI_EXIT_DISC



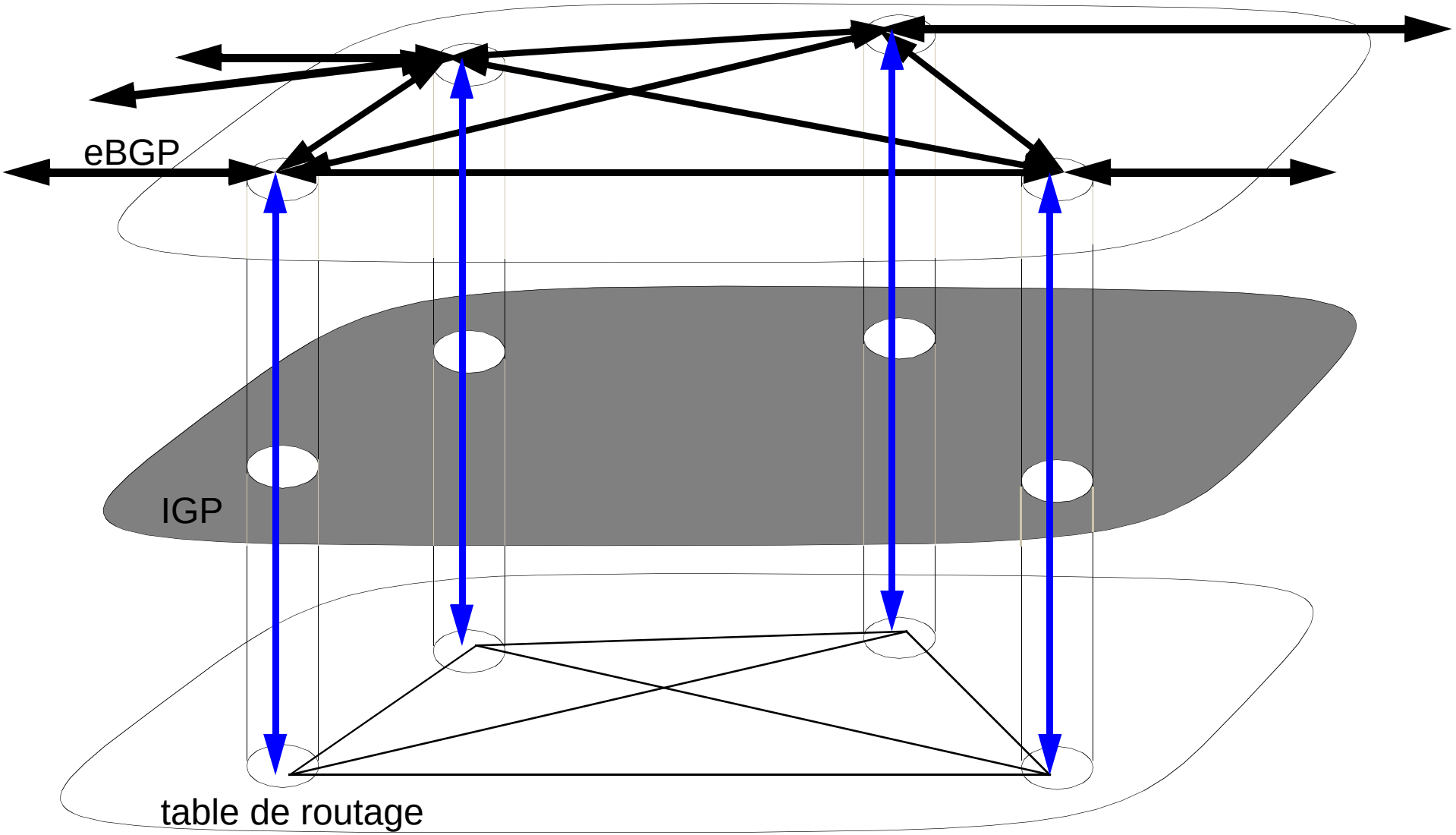
LOCAL_PREF



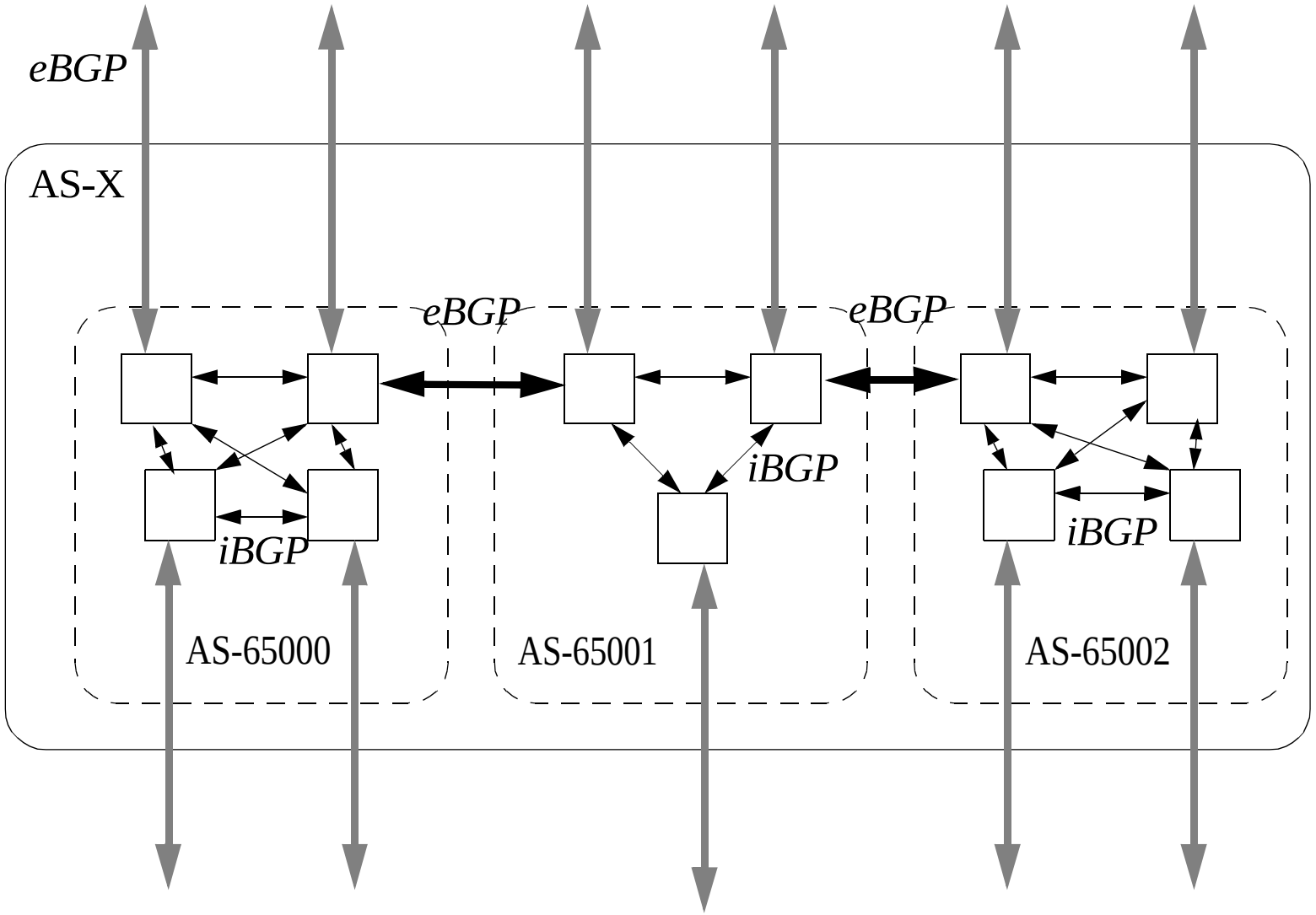
SYNCHRONISATION



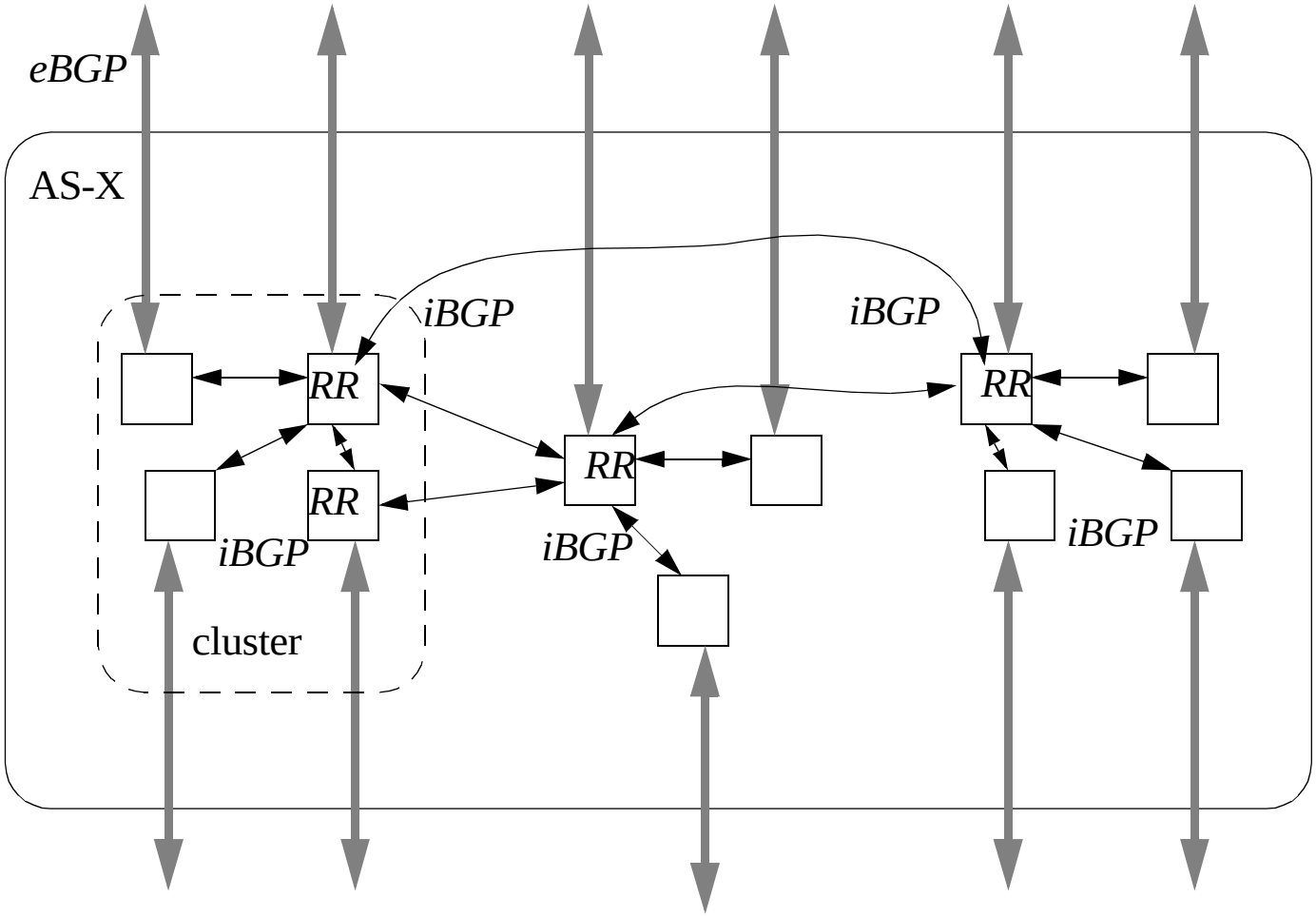
SYNCHRONISATION



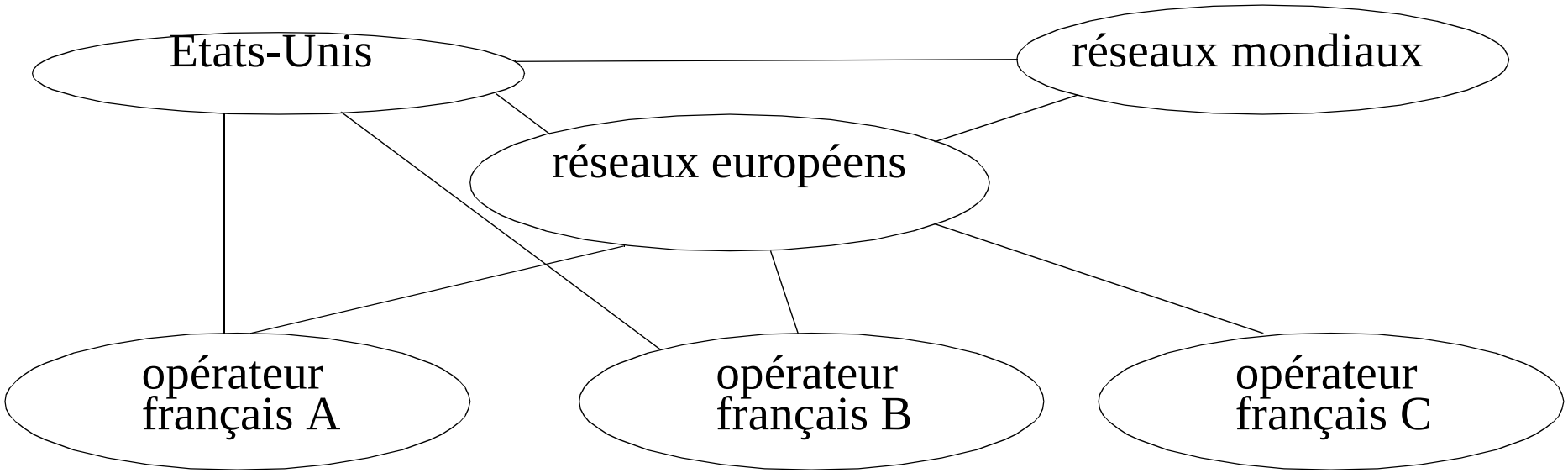
ATTRIBUTS POUR IBGP



ATTRIBUTS POUR IBGP

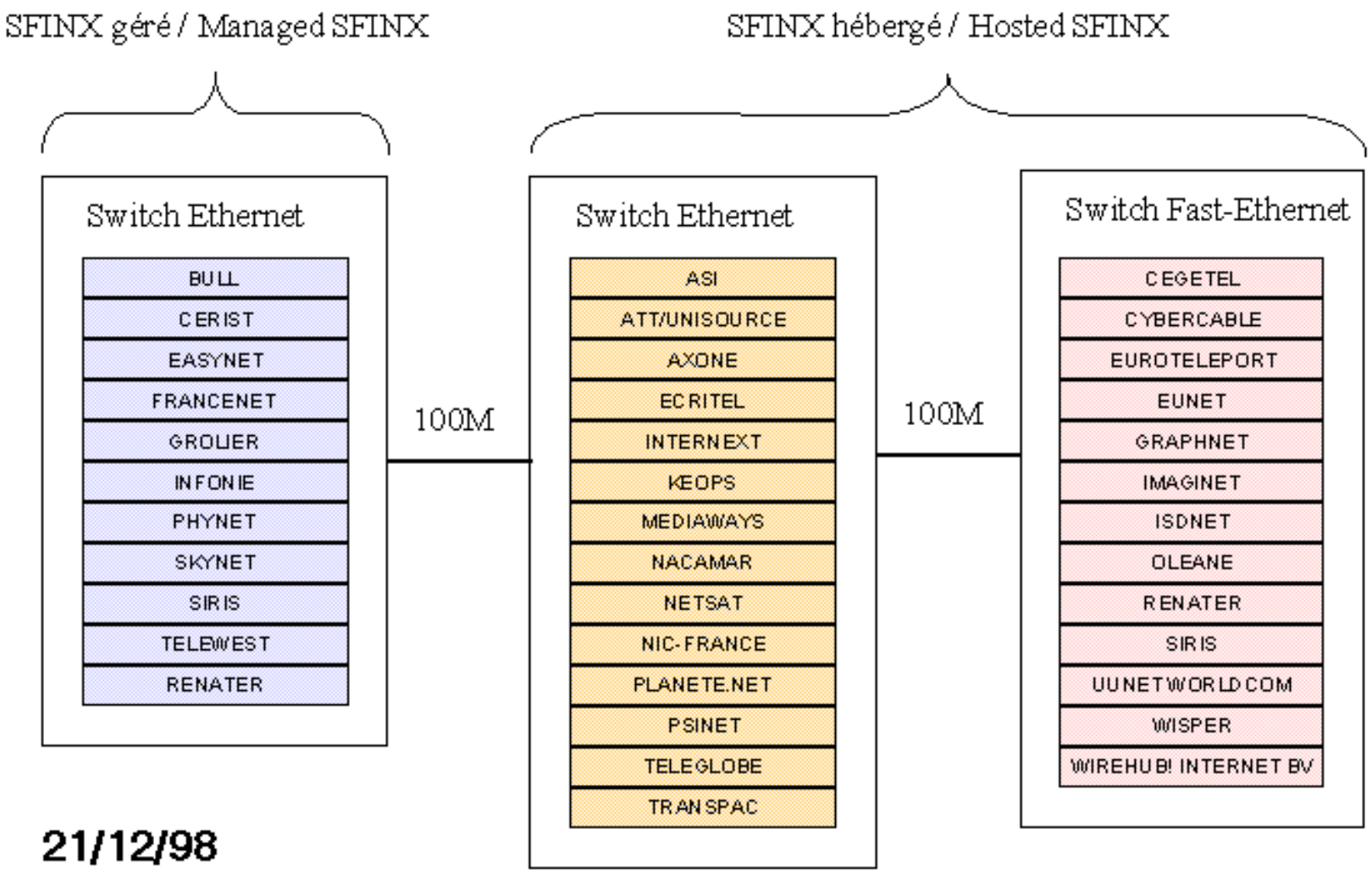


POINTS D'ÉCHANGE



+ Eviter d'utiliser des liaisons internationales pour le trafic national

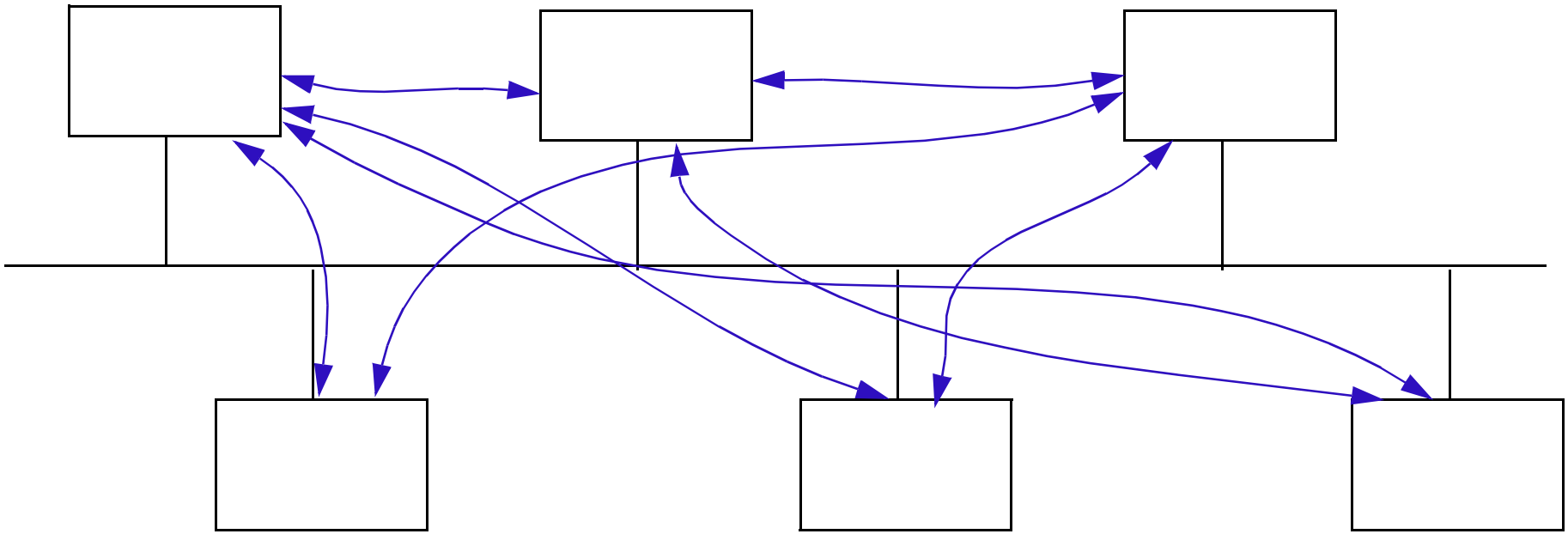
POINT D'INTERCONNEXION : FRANCE SFINX



ECHANGE D'ANNONCES SUR LE SFINX

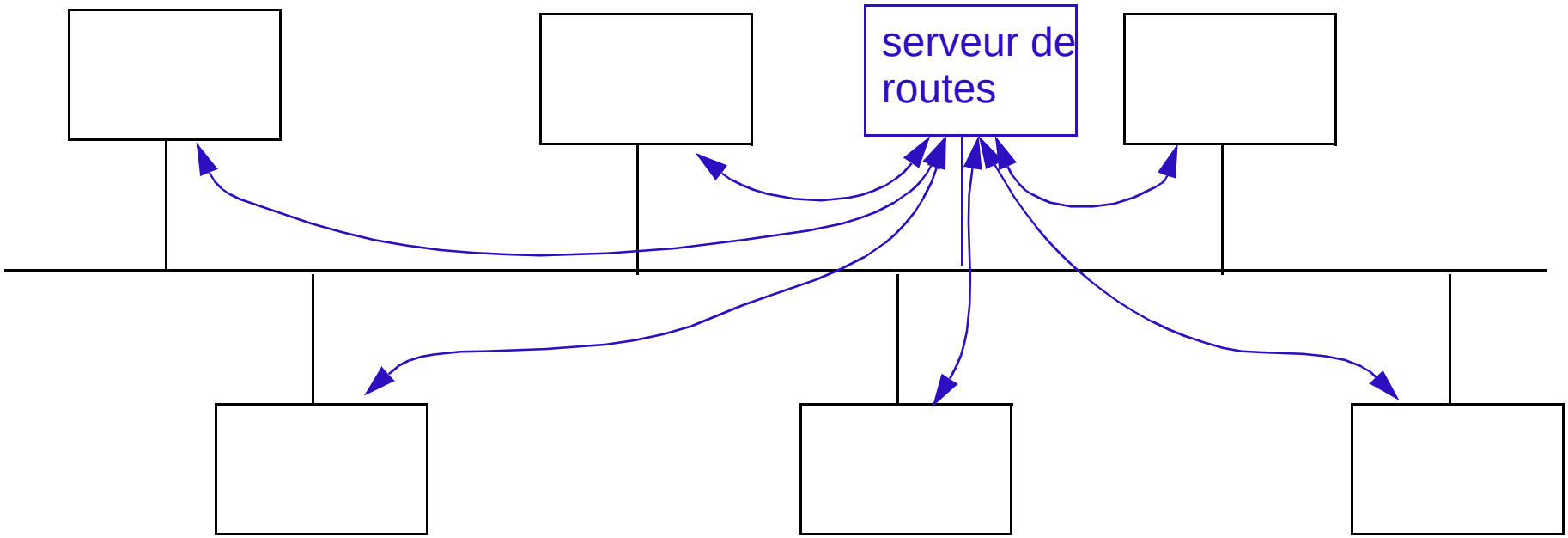
11/01/99			ASI	ATT	AXO	BUL	CER	CYB	CEG	EAS	ECR	EUN	EUR
			AS 6804	AS 3300	AS 2686	AS 6	AS 3208	AS 6678	AS 8228	AS 6727	AS 8304	AS 1899	AS 6771
ASI	ASI	AS 6804								X			
AT&T UNISOURCE	ATT	AS 3300											
AXONE IBM	AXO	AS 2686				X		X		X	X	X	X
BULL IBT	BUL	AS 6			X			X	X	X	X	X	X
CERIST	CER	AS 3208											
CYBERCABLE	CYB	AS 6678			X	X				X			X
CEGETEL	CEG	AS 8228				X				X	X		X
EASYNET	EAS	AS 6727	X		X	X		X	X		X		X
ECRITEL	ECR	AS 8304			X	X			X	X			X
EUNET France	EUN	AS 1899			X	X							
EUROTELEPORT	EUR	AS 6771			X	X		X	X	X	X		
FRANCENET	FRA	AS 5436	X		X					X	X		X
GRAPHNET	GRA	AS 7148								X	X		
GROLIER INTERACTIVE	GRO	AS 5410			X	X		X	X	X	X	X	X
IMAGINET	IMA	AS 6675						X	X	X	X	X	
INFONIE	INF	AS 8332	X		X			X			X		X
INTERNEXT	INE	AS 6761				X		X		X	X		X
ISDNET	ISD	AS 5594			X	X	X	X	X	X	X		X
KEOPS	KEO	AS 8740								X	X		

NÉGOCIATION ENTRE OPÉRATEUR



- le peering BGP est de la responsabilité de chaque opérateur

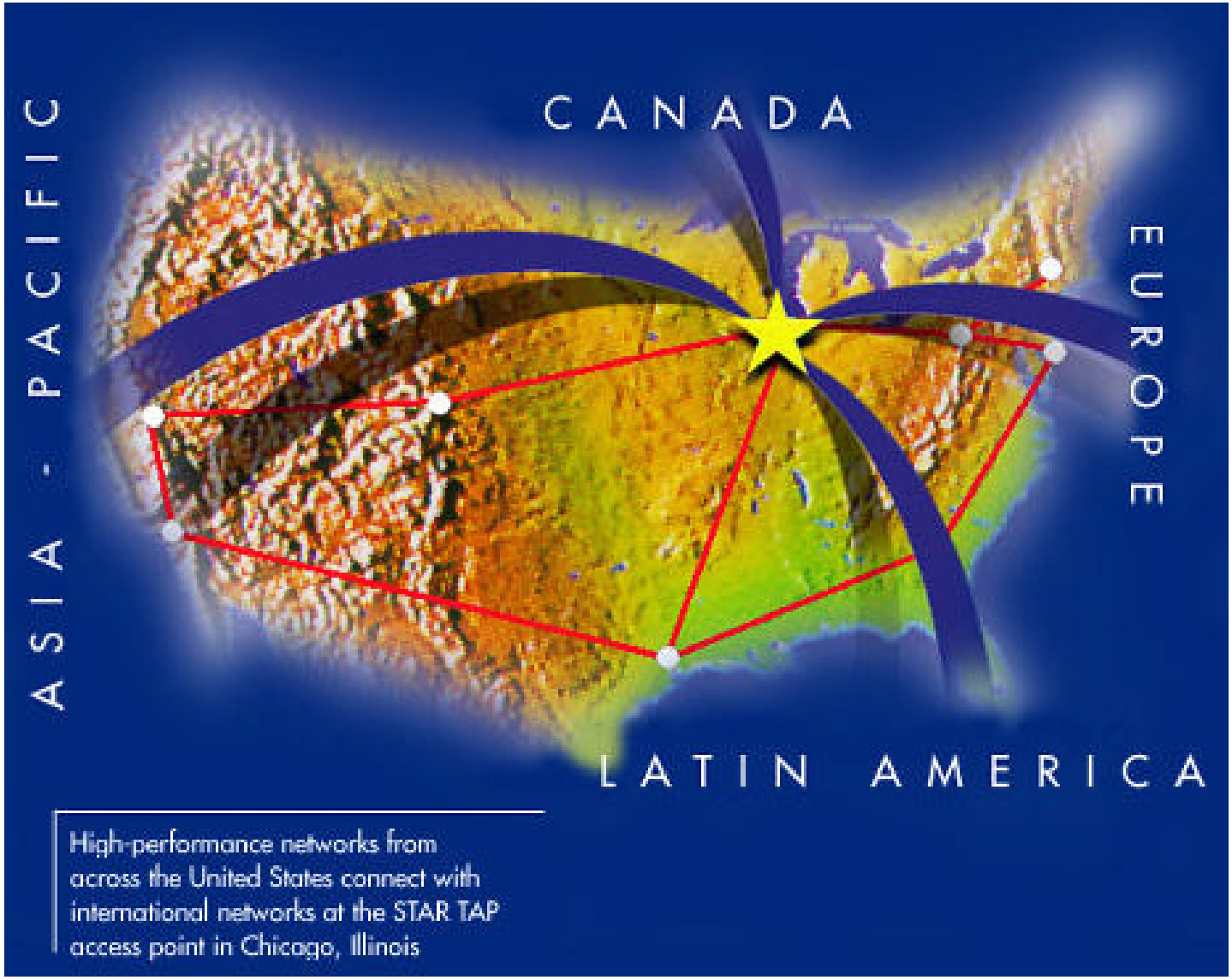
SERVEUR DE ROUTES



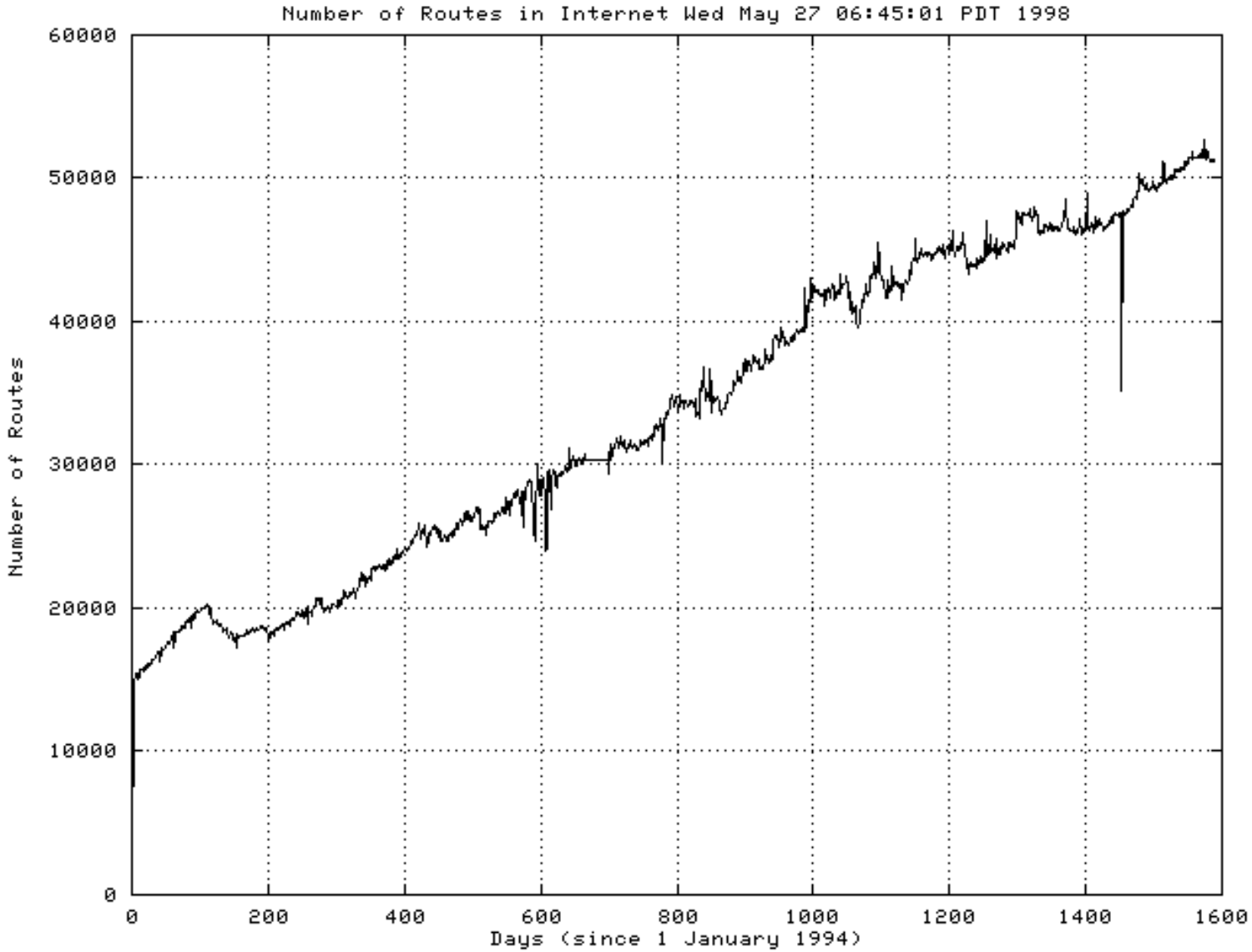
- Le serveur de route connaît la politique de routage des opérateurs
  - dérivée des bases de données comme RIPE

+ Est ce que les paquets de données passent par le serveur de route ?

LE CENTRE DU MONDE UNIVERSITAIRE



AGRÉGATION


Number of Routes in Internet Wed May 27 06:45:01 PDT 1998
Number of Routes
Days (since 1 January 1994)
0
10000
20000
30000
40000
50000
60000
0
200
400
600
800
1000
1200
1400
1600

AGRÉGATION TRÈS CONTRÔLÉE

1) Gains by aggregating at the origin AS level

--- 12Feb99 ---

ASnum	NetsNow	NetsCIDR	NetGain	% Gain	Description
AS271	385	139	246	63.9%	BCnet Backbone
AS2493	275	157	118	42.9%	iSTAR Internet, Inc.
AS174	656	540	116	17.7%	Performance Systems International
AS11305	116	8	108	93.1%	UNKNOWN

...

List of possibly interesting aggregates

Aggregate	Origin	AS Description	Netname
166.49.146.0/23	AS9335	Concert Internet Plus	*
166.49.148.0/24	AS9335	Concert Internet Plus	
168.108.2.0/24	AS7741	OPCO-AS	*
168.108.3.0/24	AS7741	OPCO-AS	*
168.108.4.0/24	AS7741	OPCO-AS	*
168.108.5.0/24	AS7741	OPCO-AS	*
168.108.6.0/24	AS7741	OPCO-AS	*
168.108.7.0/24	AS7741	OPCO-AS	*

CONFIGURATION DE BGP SUR UN CISCO

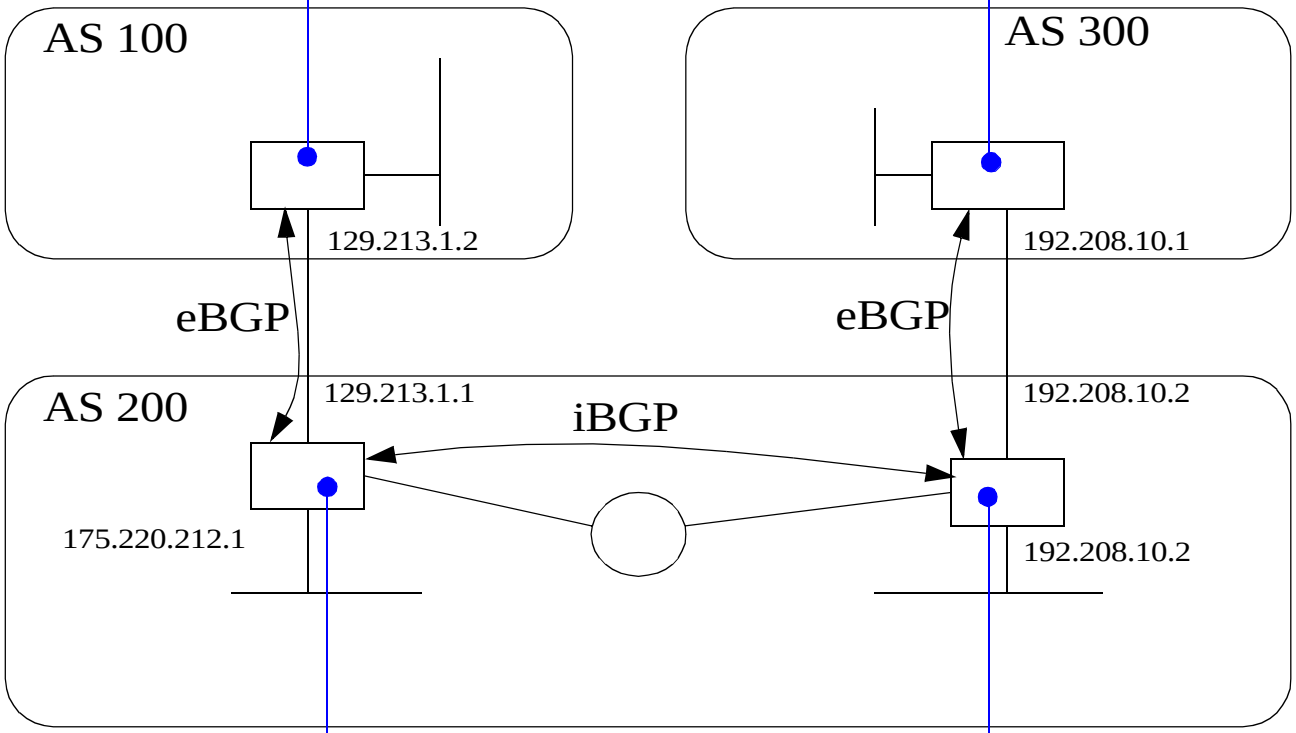
+ Contrairement aux IGP, les EGP doivent être configurés :

- définir l'adresse du routeur (et le numéro de système autonome) vers lequel les NLRI seront envoyés
- accepter les NLRI d'un autre routeur (adresse IP et système autonome)
- définir des filtres pour n'émettre que les NLRI ayant certaines caractéristiques
  - utilisation d'expressions régulières
- définir des filtres pour n'accepter que les NRLI ayant certaines caractéristiques

EXEMPLE SIMPLE

router bgp 100
neighbor 129.213.1.1 remote-as 200

router bgp 300
neighbor 192.208.10.2 remote-as 200

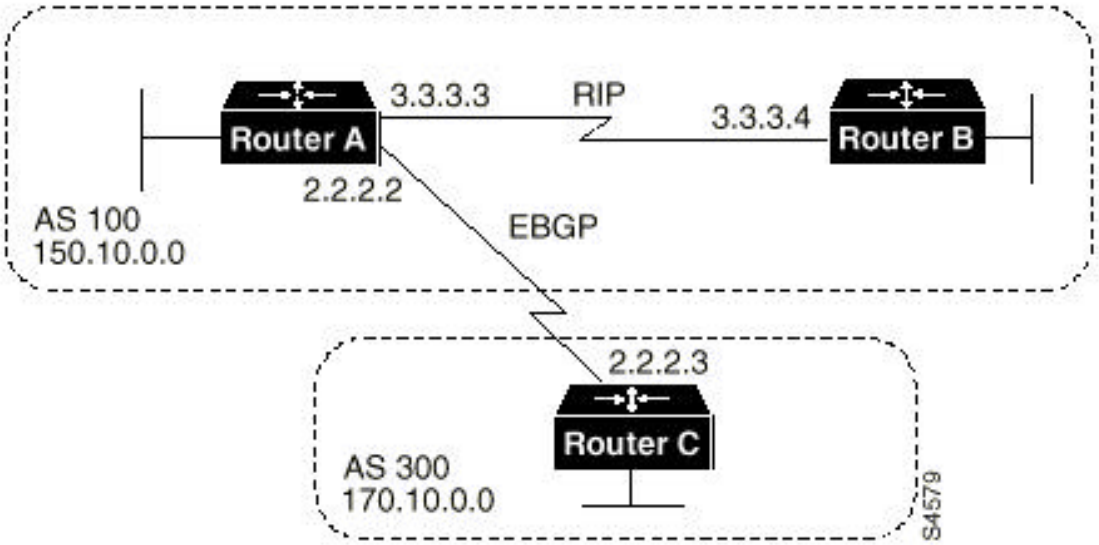


router bgp 200
neighbor 129.213.1.2 remote-as 100
neighbor 175.220.1.2 remote-as 200

router bgp 200
neighbor 175.220.212.1 remote-as 200
neighbor 192.208.10.1 remote-as 300

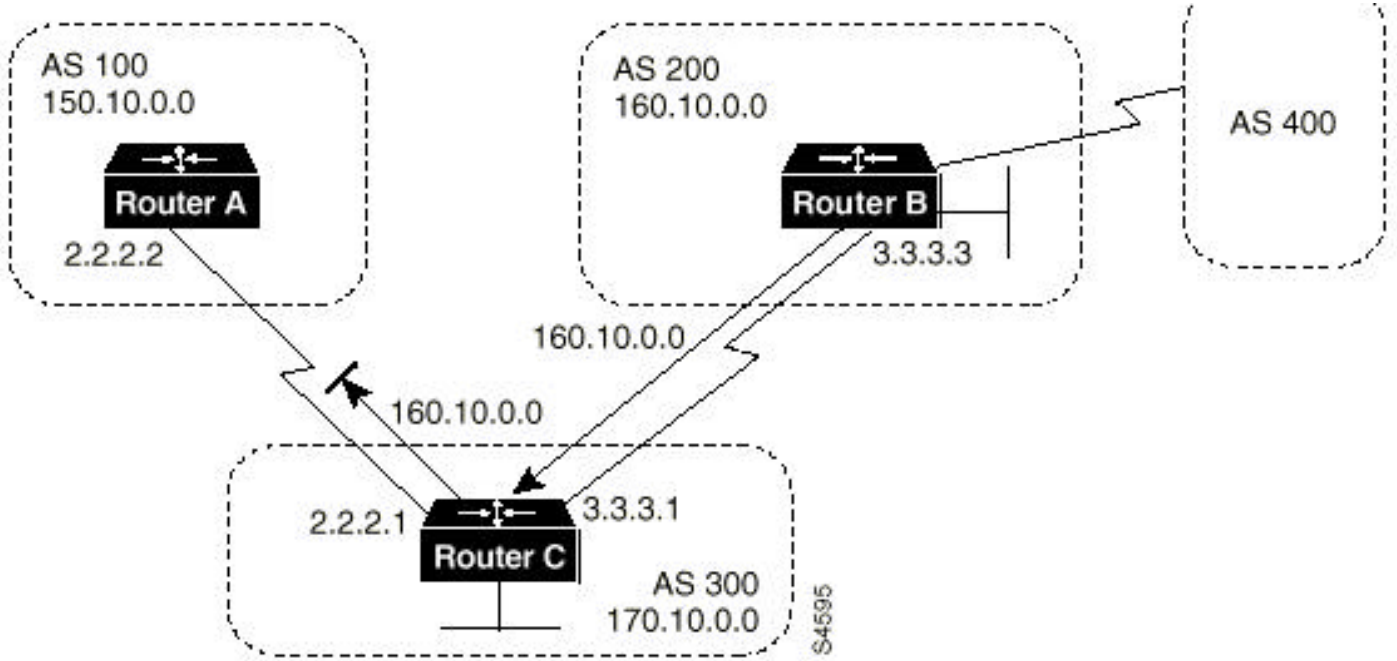
TRAITEMENT DES ATTRIBUTS

!Router A
router rip
network 3.0.0.0
network 2.0.0.0
network 150.10.0.0
passive-interface serial 0
redistribute bgp 100 route-map SETMETRIC
!
router bgp 100
neighbor 2.2.2.3 remote-as 300
network 150.10.0.0
!
route-map SETMETRIC permit 10
match ip-address 1
set metric 2
!
route-map SETMETRIC permit 20
set metric 5
!
access-list 1 permit 170.10.0.0 0.0.255.255



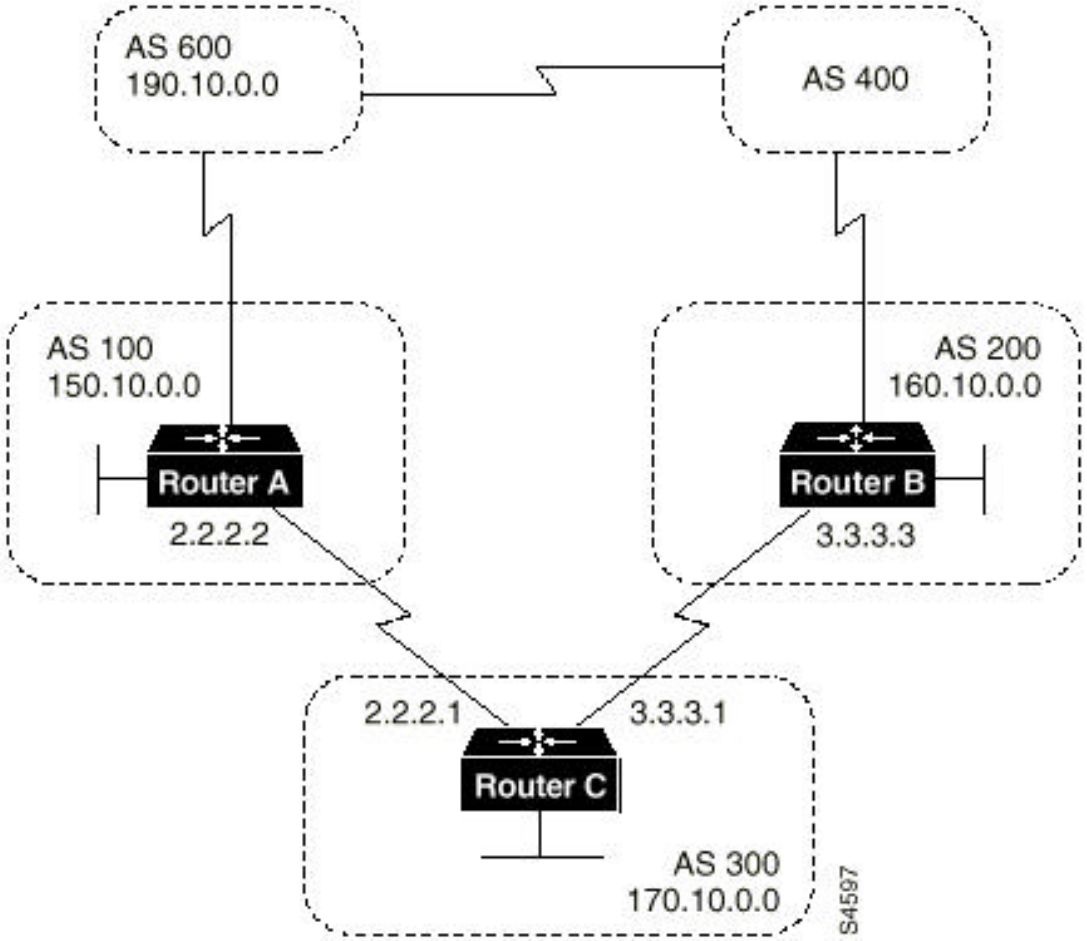
!Router C
router bgp 300
network 170.10.0.0
neighbor 2.2.2.2 remote-as 100
neighbor 2.2.2.2 route-map SETCOMMUNITY out
!
route-map SETCOMMUNITY permit 10
match ip address 1
set community 300
!
access-list 1 permit 0.0.0.0 255.255.255.255

FILTRAGE DES NRLI



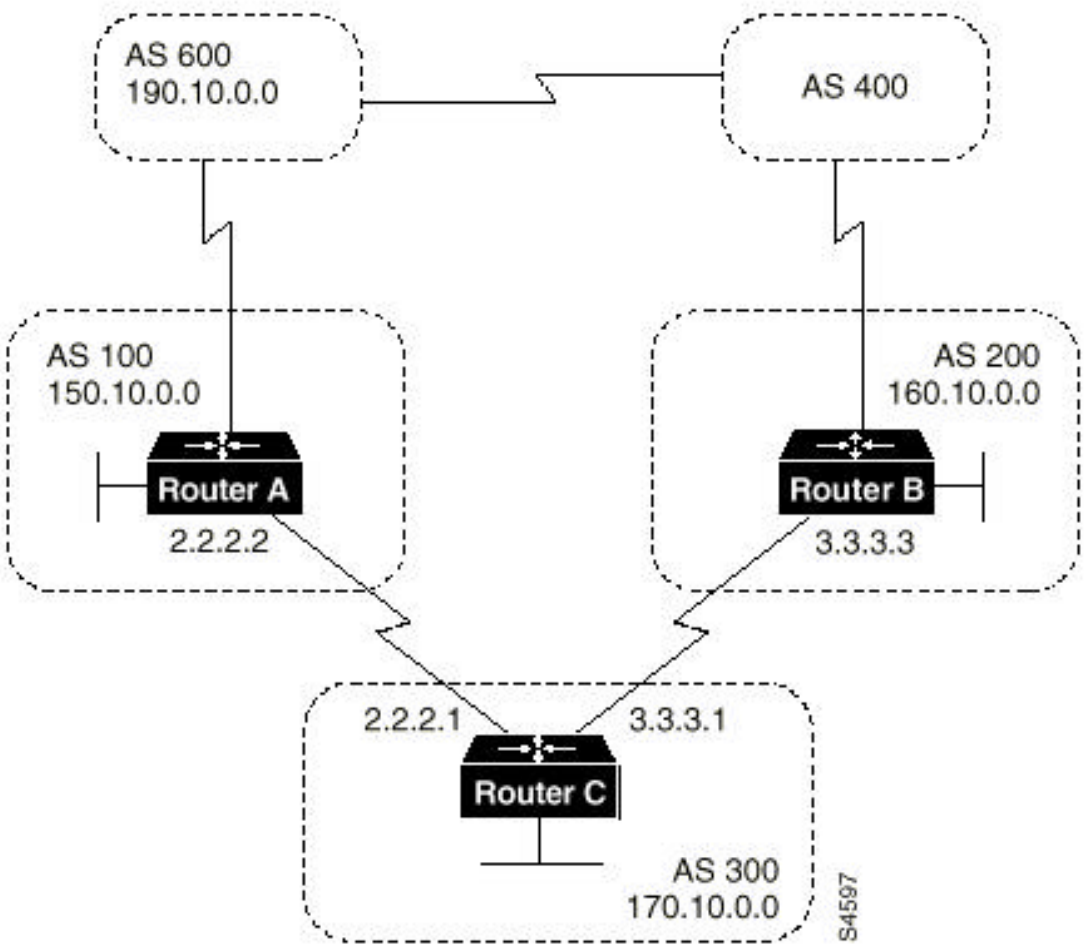

```
!Router C
neighbor 3.3.3.3 remote-as 200
neighbor 2.2.2.2 remote-as 100
neighbor 2.2.2.2 filter-list 1 out
!
ip as-path access-list 1 deny ^200$
ip as-path access-list 1 permit .*
```

FILTRAGE DES NRLI



!Router C
router bgp 300
network 170.10.0.0
neighbor 3.3.3.3 remote-as 200
neighbor 3.3.3.3 route-map STAMP in
!
route-map STAMP permit 10
match as-path 1
set weight 20
!
ip as-path access-list 1 permit ^200$

FILTRAGE DES NRLI




```
!Router C
router bgp 300
network 170.10.0.0
neighbor 3.3.3.3 remote-as 200
neighbor 3.3.3.3 route-map STAMP in
!
route-map STAMP permit 10
match as-path 1
set weight 20
!
route-map STAMP permit 20
match as-path 2
!
route-map STAMP permit 30
set weight 10
!
ip as-path access-list 1 permit ^200$
ip as-path access-list 2 deny _400_
```

+ Que fait cette configuration ?

LES EXPRESSIONS RÉGULIÈRES

symbole		actions
point	.	correspond à n'importe quel caractère y compris un espace.
astérisque	*	correspond à 0 ou plus séquence du modif.
plus	+	correspond à 1 ou plus séquence du motif.
point d'interrogation	?	correspond à 0 ou 1 occurence du motif.
circonflexe	^	correspond au début de la chaine de caractères.
dollar	$	correspond à la fin de la chaine de caractères.
souligné	_	correspond aux caractères , { } ( ) ^ $ espace
crochet	[]	définit un ensemble d'un seul caractère
moins	-	sépare un interval de caractères
parenthèses	()	mémorise la séquence pour un traitement #1, #2

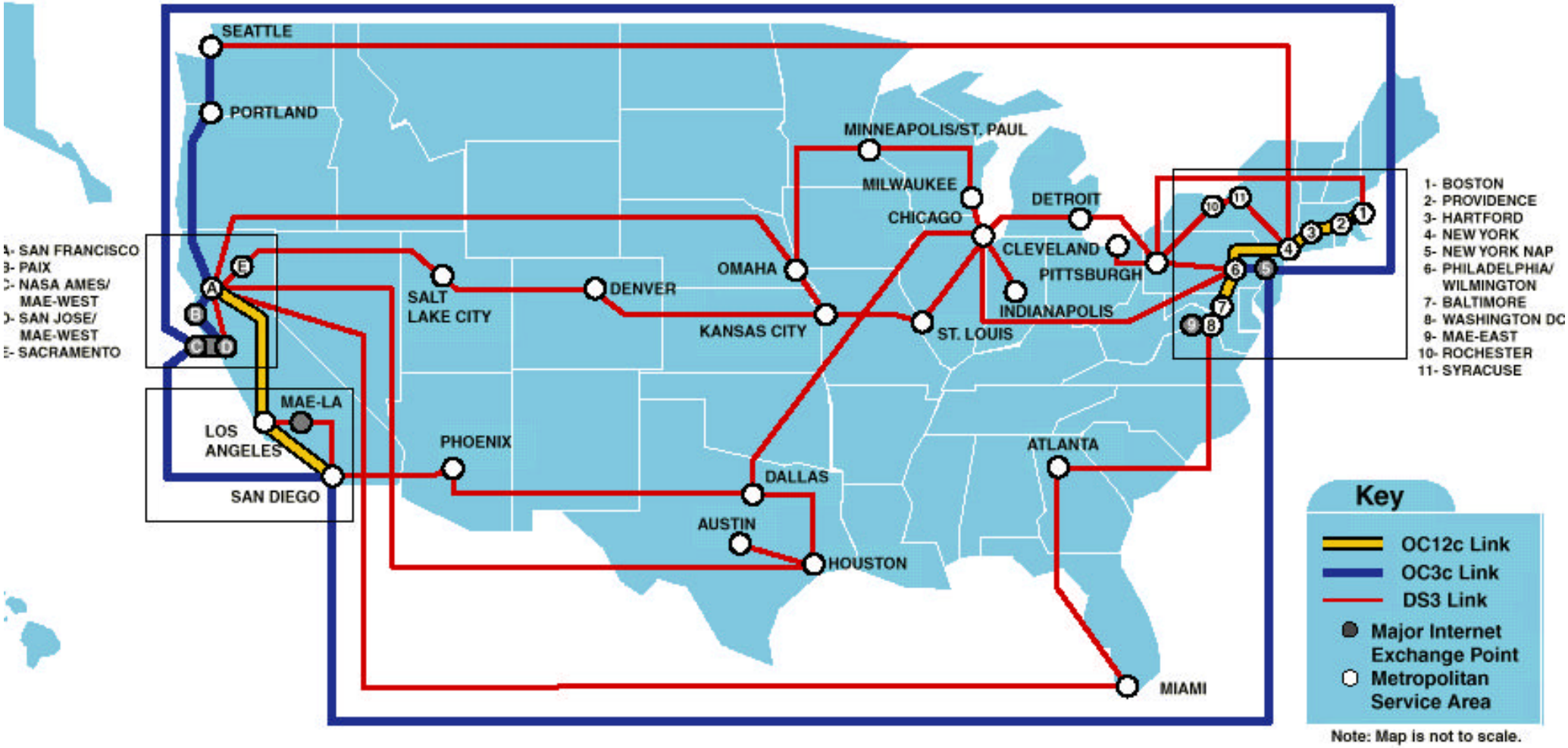
- _100_(via AS100)
- ^100$ (peering direct de l'AS100)
- ^.* 100 (venant de l'AS100)

LES OUTILS DE CONTRÔLE ET DE MESURE

+ Les Looking Glass :

- Possibilité de visualiser la configuration des routeurs (ou des serveurs de routes)
- Accessible :
  - telnet
  - page web.
- Par exemple :
  - http://www.merit.edu/ipma/tools/lookingglass.html
  - telnet://route-server.cerf.net
  - telnet://route-server.ip.att.net

EXEMPLE SUR ROUTE-SERVER.CERF.NET



EXEMPLE SUR ROUTE-SERVER.CERF.NET

################################## ROUTE-SERVER.CERF.NET ################

CERFnet BGP Route Monitor

Reflects CERFnet's Peering Arrangements at the New York Sprint NAP, PAIX,
MAE-West, Private Peers, and MAE-LA. This router should have the global
BGP4 routing table.

The ip forwarding table used by this router does not always reflect the
routing policy of CERFnet AS 1740. This router is in AS 1838 and maintains
ebgp multihop peers with key CERFnet border routers.

192.157.69.5 == New York Sprint NAP 198.32.136.5 == MAE-West (NASA Side)
198.32.176.25 == PAIX 198.32.136.55 == MAE-West (MFS Side)
198.32.146.20 == MAE-LA

The corresponding topology is at:
http://www.cerf.net/cerfnet/about/Bbone-map.html

This route-server is off on a private ethernet off our San Diego nodes.
Questions? email pushp@cerf.net

################################## ROUTE-SERVER.CERF.NET ################

EXEMPLE SUR ROUTE-SERVER.CERF.NET

route-server.cerf.net>**sh ip bgp summary**
BGP table version is 8855242, main routing table version 8855242
53576 network entries (267670/267876 paths) using 14534372 bytes of memory
8467 BGP path attribute entries using 1161132 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
Dampening enabled. 397 history paths, 164 dampened paths

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State
192.157.69.5	4	1740	2569023	104319	8855231	0	0	4d00h	
198.32.136.5	4	1740	2658338	104318	8855238	0	0	4d00h	
198.32.136.55	4	1740	2658774	104264	8855238	0	0	4d00h	
198.32.146.20	4	1740	2551705	104330	8855242	0	0	4d00h	
198.32.176.25	4	1740	2512480	104318	8855231	0	0	4d00h	

route-server.cerf.net>

EXEMPLE SUR ROUTE-SERVER.CERF.NET

```
route-server.cerf.net>sh ip bgp neighbors
BGP neighbor is 192.157.69.5,  remote AS 1740, external link
 Index 1, Offset 0, Mask 0x2
  BGP version 4, remote router ID 134.24.127.201
  BGP state = Established, table version = 8855381, up for 4d00h
  Last read 00:00:02, hold time is 180, keepalive interval is 60 seconds
  Minimum time between advertisement runs is 30 seconds
  Received 2569071 messages, 0 notifications, 0 in queue
  Sent 104322 messages, 0 notifications, 0 in queue
  Outgoing update network filter list is 10
  Connections established 4; dropped 3
  External BGP neighbor may be up to 255 hops away.
Connection state is ESTAB, I/O status: 1, unread input bytes: 0
Local host: 192.215.254.5, Local port: 179
Foreign host: 192.157.69.5, Foreign port: 14070

Enqueued packets for retransmit: 0, input: 0, saved: 0

Event Timers (current time is 0x175660BA8):
Timer          Starts    Wakeups            Next
Retrans          5806         11             0x0
TimeWait            0          0             0x0
AckHold         36490      25431             0x0
SendWnd             0          0             0x0
KeepAlive           0          0             0x0
GiveUp              0          0             0x0
```

PmtuAger 0 0 0x0

iss: 2275911518 snduna: 2276021615 sndnxt: 2276021615 sndwnd: 15928
irs: 2275848384 rcvnxt: 2286517440 rcvwnd: 16341 delrcvwnd: 43

SRTT: 382 ms, RTTO: 1061 ms, RTV: 148 ms, KRTT: 0 ms
minRTT: 60 ms, maxRTT: 860 ms, ACK hold: 300 ms
Flags: passive open, nagle, gen tcbs

Datagrams (max data segment is 536 bytes):
Rcvd: 52815 (out of order: 43), with data: 47388, total data bytes: 10669055
Sent: 48132 (retransmit: 11), with data: 5794, total data bytes: 110096

+ Quel est le débit moyen de la connexion BGP ?

EXEMPLE SUR ROUTE-SERVER.CERF.NET

route-server.cerf.net>**sh ip bgp**
BGP table version is 8854815, local router ID is 39.7.46.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

	Network	Next Hop	Metric	LocPrf	Weight	Path
*	3.0.0.0	198.32.136.55			0	1740 701 80 i
*		198.32.136.5			0	1740 701 80 i
*		192.157.69.5			0	1740 701 80 i
*>		198.32.146.20			0	1740 701 80 i
*		198.32.176.25			0	1740 701 80 i
*	4.0.0.0	198.32.136.55			0	1740 1 i
*		198.32.136.5			0	1740 1 i
*		192.157.69.5			0	1740 1 i
*>		198.32.146.20			0	1740 1 i
*	6.0.0.0	198.32.136.55			0	1740 568 721 1455 i
*		198.32.136.5			0	1740 568 721 1455 i
*		192.157.69.5			0	1740 568 721 1455 i
*>		198.32.146.20			0	1740 568 721 1455 i
*		198.32.176.25			0	1740 568 721 1455 i
*	9.2.0.0/16	198.32.136.55			0	1740 1673 1675 i
*		198.32.136.5			0	1740 1673 1675 i
*		192.157.69.5			0	1740 1673 1675 i
*>		198.32.146.20			0	1740 1673 1675 i

EXEMPLE SUR ROUTE-SERVER.CERF.NET

route-server.cerf.net>**sh ip bgp regexp _1717_**
BGP table version is 8774930, local router ID is 39.7.46.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

	Network	Next Hop	Metric	LocPrf	Weight	Path
*	128.93.0.0	198.32.136.55			0	1740 1239 1800 5511 1717 1309 ?
*		198.32.136.5			0	1740 1239 1800 5511 1717 1309 ?
*		192.157.69.5			0	1740 1239 1800 5511 1717 1309 ?
*>		198.32.146.20			0	1740 1239 1800 5511 1717 1309 ?
*		198.32.176.25			0	1740 1239 1800 5511 1717 1309 ?
*	129.20.0.0	198.32.136.55			0	1740 1239 5511 1717 1938 ?
*		198.32.136.5			0	1740 1239 5511 1717 1938 ?
*		192.157.69.5			0	1740 1239 5511 1717 1938 ?
*>		198.32.146.20			0	1740 1239 5511 1717 1938 ?
*		198.32.176.25			0	1740 1239 5511 1717 1938 ?
*	129.88.0.0	198.32.136.55			0	1740 1239 5511 1717 2477 ?
*		198.32.136.5			0	1740 1239 5511 1717 2477 ?
*		192.157.69.5			0	1740 1239 5511 1717 2477 ?
*>		198.32.146.20			0	1740 1239 5511 1717 2477 ?
*		198.32.176.25			0	1740 1239 5511 1717 2477 ?

EXEMPLE SUR ROUTE-SERVER.CERF.NET

route-server.cerf.net>sh ip bgp regexp ^.* **1717$**
BGP table version is 8853425, local router ID is 39.7.46.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

	Network	Next Hop	Metric	LocPrf	Weight	Path
*	132.166.0.0/15	198.32.136.55			0	1740 1239 5511 1717 i
*		198.32.136.5			0	1740 1239 5511 1717 i
*		192.157.69.5			0	1740 1239 5511 1717 i
*>		198.32.146.20			0	1740 1239 5511 1717 i
*		198.32.176.25			0	1740 1239 5511 1717 i
*	132.168.0.0/15	198.32.136.55			0	1740 1239 5511 1717 i
*		198.32.136.5			0	1740 1239 5511 1717 i
*		192.157.69.5			0	1740 1239 5511 1717 i
*>		198.32.146.20			0	1740 1239 5511 1717 i
*		198.32.176.25			0	1740 1239 5511 1717 i
*	161.104.0.0/15	198.32.136.55			0	1740 1239 5511 1717 i
*		198.32.136.5			0	1740 1239 5511 1717 i
*		192.157.69.5			0	1740 1239 5511 1717 i
*>		198.32.146.20			0	1740 1239 5511 1717 i
*		198.32.176.25			0	1740 1239 5511 1717 i
*	164.81.0.0	198.32.136.55			0	1740 1239 5511 1717 ?
*		198.32.136.5			0	1740 1239 5511 1717 ?
*		192.157.69.5			0	1740 1239 5511 1717 ?

EXEMPLE SUR ROUTE-SERVER.CERF.NET

- Afficher les annonces de routes passant par OPEN TRANSIT (As 5511) et originaire de Renater (AS1717)

route-server.cerf.net>**sh ip bgp regexp ^.* 5511 .*1717$**
BGP table version is 8853957, local router ID is 39.7.46.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

	Network	Next Hop	Metric	LocPrf	Weight	Path
*	132.166.0.0/15	198.32.136.55			0	1740 1239 5511 1717 i
*		198.32.136.5			0	1740 1239 5511 1717 i
*		192.157.69.5			0	1740 1239 5511 1717 i
*>		198.32.146.20			0	1740 1239 5511 1717 i
*		198.32.176.25			0	1740 1239 5511 1717 i
*	132.168.0.0/15	198.32.136.55			0	1740 1239 5511 1717 i
*		198.32.136.5			0	1740 1239 5511 1717 i
*		192.157.69.5			0	1740 1239 5511 1717 i
*>		198.32.146.20			0	1740 1239 5511 1717 i
*		198.32.176.25			0	1740 1239 5511 1717 i
*	161.104.0.0/15	198.32.136.55			0	1740 1239 5511 1717 i
*		198.32.136.5			0	1740 1239 5511 1717 i
*		192.157.69.5			0	1740 1239 5511 1717 i
*>		198.32.146.20			0	1740 1239 5511 1717 i
*		198.32.176.25			0	1740 1239 5511 1717 i

EXEMPLE SUR ROUTE-SERVER.CERF.NET

- Pour un NLRI :

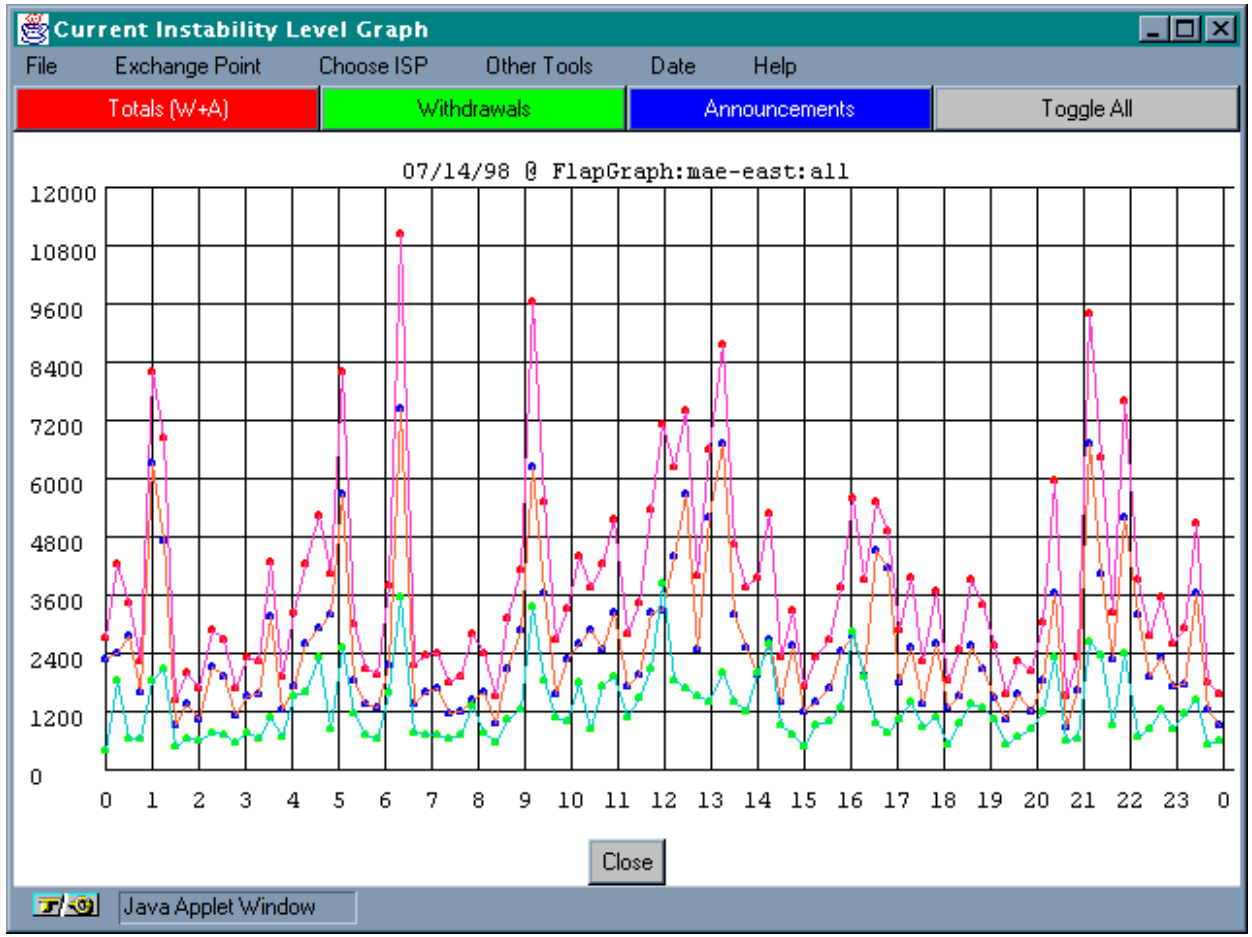
route-server.cerf.net>**sh ip bgp 192.108.119.0**
BGP routing table entry for 192.108.119.0/24, version 8842821
Paths: (5 available, best #4)
  1740 1239 5511 1717 1938
    198.32.136.5 from 198.32.136.5 (134.24.127.29)
      Origin incomplete, valid, external
  1740 1239 5511 1717 1938
    198.32.136.55 from 198.32.136.55 (134.24.127.27)
      Origin incomplete, valid, external
  1740 1239 5511 1717 1938
    192.157.69.5 from 192.157.69.5 (134.24.127.201)
      Origin incomplete, valid, external
  1740 1239 5511 1717 1938
    198.32.146.20 from 198.32.146.20 (134.24.127.23)
      Origin incomplete, valid, external, best
  1740 1239 5511 1717 1938
    198.32.176.25 from 198.32.176.25 (134.24.127.35)
      Origin incomplete, valid, external
route-server.cerf.net>

EXEMPLE SUR ROUTE-SERVER.CERF.NET

- Pour un NLRI :

route-server.cerf.net>**sh ip bgp 24.48.0.0**
BGP routing table entry for 24.48.0.0/18, version 8543708
Paths: (5 available, best #4)
1740 1673 {1324,1329,1327,1325}, (aggregated by 1673 140.223.220.62)
198.32.136.55 from 198.32.136.55 (134.24.127.27)
Origin incomplete, valid, external, atomic-aggregate
1740 1673 {1324,1329,1327,1325}, (aggregated by 1673 140.223.12.62)
198.32.136.5 from 198.32.136.5 (134.24.127.29)
Origin incomplete, valid, external, atomic-aggregate
1740 1673 {1324,1329,1327,1325}, (aggregated by 1673 140.223.217.62)
192.157.69.5 from 192.157.69.5 (134.24.127.201)
Origin incomplete, valid, external, atomic-aggregate
1740 1673 {1324,1329,1327,1325}, (aggregated by 1673 140.223.220.62)
198.32.146.20 from 198.32.146.20 (134.24.127.23)
Origin incomplete, valid, external, atomic-aggregate, best
1740 1673 {1324,1329,1327,1325}, (aggregated by 1673 140.223.220.62)
198.32.176.25 from 198.32.176.25 (134.24.127.35)
Origin incomplete, valid, external, atomic-aggregate

STATBILITÉ DES ROUTES



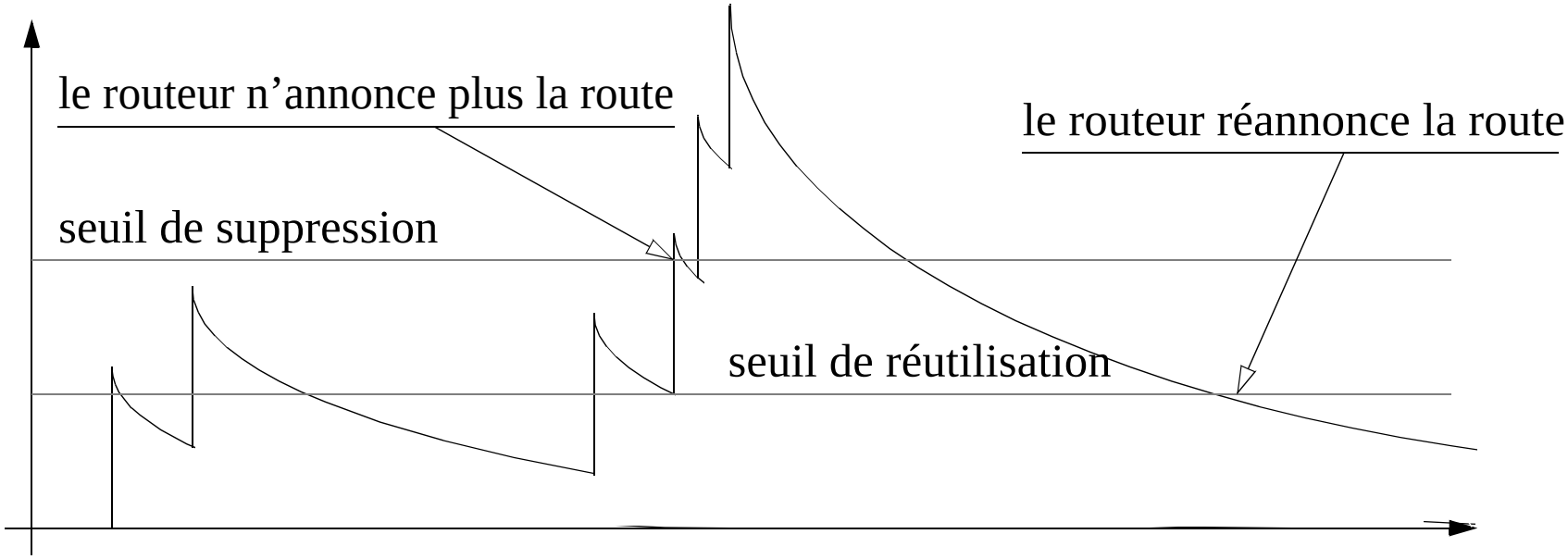
Pour 42 000 routes, entre 3 et 6 millions d'annonces et de retraits par jour.

Comportement oscillatoire mal compris

Réductions des oscillations par rejet des routes instables

DAMPERING

- Décroissance exponentielle



EXEMPLE SUR ROUTE-SERVER.CERF.NET

route-server.cerf.net>sh ip bgp dampened-paths
BGP table version is 8859804, local router ID is 39.7.46.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

Network	From	Reuse	Path
*d 203.55.55.0	198.32.136.55	00:00:1	1740 2828 3840 4433 4740 i
*d 203.55.55.0	192.157.69.5	00:00:1	1740 2828 3840 4433 4740 i
*d 203.55.55.0	198.32.146.20	00:00:2	1740 2828 3840 4433 4740 i
*d 203.55.54.0	198.32.176.25	00:00:4	1740 2828 3840 4433 4740 i
*d 203.26.142.0	198.32.176.25	00:01:0	1740 2828 3840 4433 4740 i
*d 203.93.64.0/19	198.32.136.55	00:01:3	1740 1 5727 4799 4799 4799 4799 4799 4799 i
*d 203.55.54.0	198.32.136.5	00:01:4	1740 2828 3840 4433 4740 i
*d 203.26.142.0	192.157.69.5	00:02:1	1740 2828 3840 4433 4740 i
*d 203.26.142.0	198.32.136.5	00:02:3	1740 2828 3840 4433 4740 i
*d 202.54.73.0	192.157.69.5	00:02:4	1740 701 6453 4755 i
*d 202.54.73.0	198.32.146.20	00:02:5	1740 701 6453 4755 i
*d 194.110.137.0	192.157.69.5	00:02:5	1740 701 6453 5515 i
*d 194.110.137.0	198.32.146.20	00:03:0	1740 701 6453 5515 i
*d 202.54.73.0	198.32.136.55	00:02:5	1740 701 6453 4755 i
*d 194.110.137.0	198.32.136.55	00:03:0	1740 701 6453 5515 i
*d 194.110.137.0	198.32.176.25	00:03:0	1740 701 6453 5515 i
*d 194.110.137.0	198.32.136.5	00:03:1	1740 701 6453 5515 i
*d 194.42.64.0/19	192.157.69.5	00:03:3	1740 174 8218 i

EXEMPLE SUR ROUTE-SERVER.CERF.NET

route-server.cerf.net>**sh ip bgp flap-statistics**
BGP table version is 8860218, local router ID is 39.7.46.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

	Network	From	Flaps	Duration	Reuse	Path
*d	24.232.0.0/19	198.32.136.55	5	00:06:18	00:21:40	1740 7018 3561 6813 4926 4926 4926 4926 7908 10318
*d		198.32.136.5	5	00:06:07	00:21:40	1740 7018 3561 6813 4926 4926 4926 4926 7908 10318
*d		192.157.69.5	5	00:06:06	00:21:40	1740 7018 3561 6813 4926 4926 4926 4926 7908 10318
*d		198.32.146.20	6	00:06:01	00:29:00	1740 7018 3561 6813 4926 4926 4926 4926 7908 10318
*d		198.32.176.25	5	00:06:23	00:21:40	1740 7018 3561 6813 4926 4926 4926 4926 7908 10318
h	53.122.48.0/20	198.32.176.25	1	00:13:13		1740 1800 1755 1273 6805 6878
h		198.32.136.5	1	00:13:16		1740 1800 1755 1273 6805 6878
h		192.157.69.5	2	00:13:19		1740 1239 1800 1755 1273 6805 6878
h		198.32.146.20	2	00:13:33		1740 1239 1800 1755 1273 6805 6878
h		198.32.136.55	2	00:13:20		1740 1239 1800 1755 1273 6805 6878
*	62.172.0.0/16	198.32.136.5	2	00:21:46		1740 2548 2856
*>		198.32.146.20	2	00:22:04		1740 2548 2856
*	128.8.0.0	198.32.136.55	2	00:21:54		1740 2548 27
*		198.32.136.5	2	00:21:46		1740 2548 27
*		192.157.69.5	2	00:21:51		1740 2548 27
*>		198.32.146.20	2	00:22:04		1740 2548 27
*		198.32.176.25	2	00:21:49		1740 2548 27
*	128.164.0.0	198.32.136.55	3	00:20:57		1740 3951 6350 11039