

WAN

IP routing

Eric Gaillard - 2015

EPITA - MAJEURES SRS & TCOM

Plan

- **Notions fondamentales**
- **RIP v1/v2**
- **OSPF**
- **BGP4**

Routing IP

- **Deux types de routage**
 - Routage statique
 - Routage dynamique
- **Routage statique**
 - Configuration manuelle des tables de routage
- **Routage dynamique**
 - Apprentissage dynamique des tables de routage
 - Protocole d'échange des tables de routage entre les routeurs
 - Diffusion périodique et/ou sur évènement
 - Interior Gateway Protocol vs exterior Gateway Protocol
 - Les protocoles «Distance vector» ou «Belman-Ford
 - Les protocoles « link state » ou «Short Path First »

Routage IP

- **Les routeurs IP interconnectent des réseaux de nature différente ayant en commun l'usage du protocole IP**
- **Rappels sur les équipements d'interconnexion des réseaux Ethernet**
 - Répéteur Ethernet / Ethernet
 - *1 domaine de collision*
 - *1 domaine de diffusion IP*
 - Pont Ethernet / Ethernet
 - *2 domaines de collision*
 - *1 domaine de diffusion IP*
 - Routeur Ethernet / Ethernet
 - *2 domaines de collision*
 - *2 domaines de diffusion*

Routage IP : les routeurs

- **Interconnexion d 'au moins 2 réseaux IP**
- **Vérification du checksum**
- **Vérification du numéro de version IP**
- **Traitement du TOS**
- **Traitement des options IP**
- **Identification du prochain saut (next hop)**
- **Fragmentation éventuelle du datagramme**
- **Décrément du TTL**

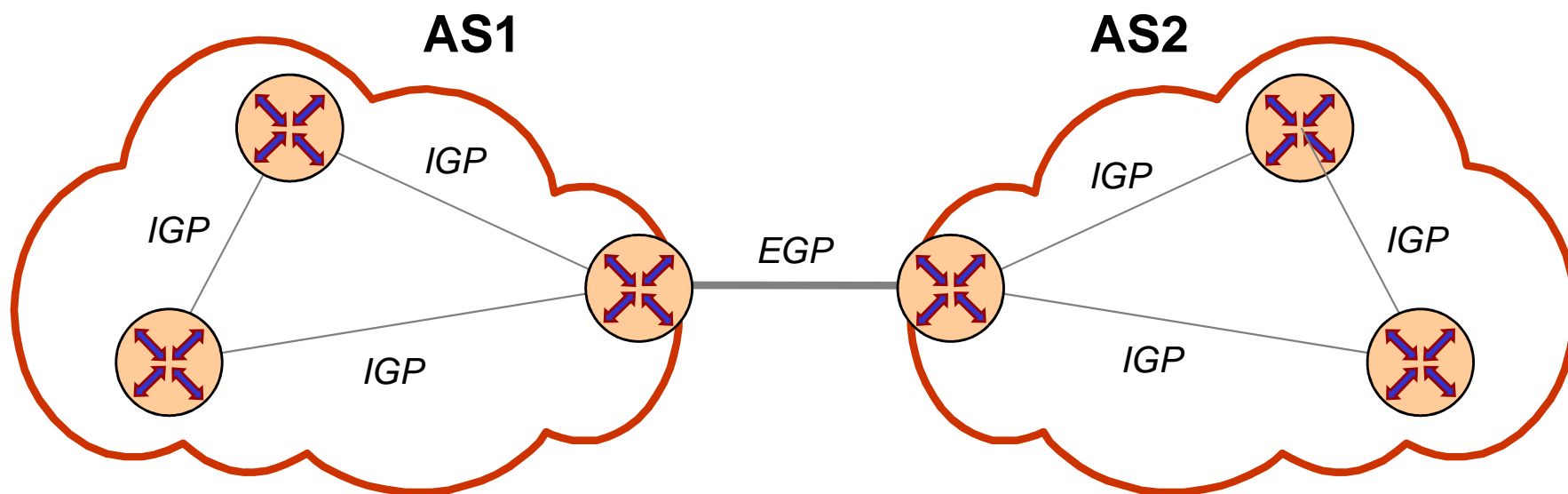
Routage IP : Exemple de table de routage

| Destination | Route Mask | Next Hop | Port | Metrique | Type | Src | Age |
|--------------|---------------|---------------|------|----------|------|-----|-----|
| 195.170.5.0 | 255.255.255.0 | 195.170.5.1 | E3 | 0 | Dir | Loc | 113 |
| 195.170.7.0 | 255.255.255.0 | 195.170.5.2 | E3 | 1 | Rem | RIP | 31 |
| 195.170.9.0 | 255.255.255.0 | 195.170 .11.2 | E4 | 1 | Rem | RIP | 25 |
| 195.170.11.0 | 255.255.255.0 | 195.170.11.1 | E4 | 0 | Dir | Loc | 114 |
| 115.0.0.0 | 255.0.0.0 | 195.170.11.2 | E4 | 7 | Rem | RIP | 27 |

Routage IP : Les systèmes autonomes

- **Autonomous System**
- **Besoin d'une discontinuité de routage dans l'Internet**
 - Au début réseau unifié
 - Ensemble de routeur avec le protocole GGP (Gateway to Gateway Protocol)
 - Administration impossible et tables de routage trop conséquentes
- **Domaine de routage sous la responsabilité d'une administration unique**
- **Architecture de routage indépendante des autres systèmes autonomes**
- **Un Système Autonome peut être constitué d'un ou plusieurs réseaux.**
- **UN routeur au sein d'un AS est appelé « Internal Gateway »**
- **Un AS est identifié par un numéro unique (16 Bits) attribué par le NIC.**
- **Les protocoles EGP et BGP utilisent les numéros de Systèmes Autonomes dans l'échange des informations de routage.**

Routage IP : EGP vs IGP



Routage IP : EGP vs IGP

- **Les protocoles IGP (Interior Gateway Protocol) assurent le dialogue entre deux routeurs situés à l'intérieur d'un même AS**
 - Les protocoles « Distance vector » ou « Belman-Ford »
 - *RIP, IGRP*
 - Les protocoles « link state » ou « Short Path First »
 - *OSPF, ES-IS, IS-IS*
- **Les protocoles de type EGP (Exterior Gateway Protocol) assurent le dialogue entre deux routeurs situés sur des AS différents**
 - *EGP, BGP 2, ... BGP 4*
- **Les routeurs au sein d'un même AS n'échangent que les tables de routage internes à l'AS, mais certains routeurs doivent dialoguer avec les routeurs EGP pour découvrir les réseaux externes à l'AS.**

Routage IP : Distance Vector vs Link State

- **Les protocoles «Distance vector» ou «Belman-Ford »**
 - Diffusion périodique des informations de routage à tous les routeurs voisins
 - Les informations que les routeurs diffusent sont constituées
 - *des adresses réseaux des réseaux IP joignables*
 - *du nombre de routeurs (hop) à traverser pour joindre ces réseaux IP*
 - *de l'adresse IP du prochain routeur à traverser*
 - La mise à jour des tables de routage est automatique à la réception de ces informations

- **Les protocoles « link state » ou «Short Path First »**
 - gestion de l'état des liens
 - Diffusion à tous les routeurs de l'état des liens
 - Chaque routeur construit une vision topologique générale du réseau
 - Choix du chemin le plus court (Dijkstra)

Routage IP : Distance Vector vs Link State

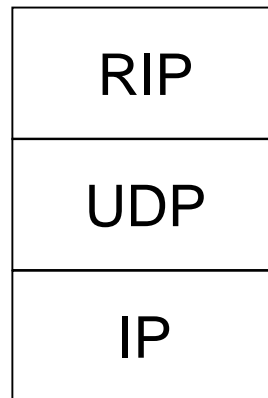
- **Algorithmes « distance vector » ou Belman-Ford**
- **Le choix du meilleur chemin se fait par le routeur en fonction de la métrique.**
- **Chaque routeur diffuse à ses proches voisins les routes dont il a connaissance**
- **Informations diffusées**
 - adresse IP du réseau distant « connu »
 - adresse IP du proche routeur
 - métrique (nombre de saut pour joindre le réseau distant « connu »)
- **Un routeur qui reçoit ces informations, les analyse, et met éventuellement à jour sa table de routage**
 - métrique favorable
 - nouveau réseau

RIP : généralités

- **Routing Information Protocol**
- **RIPv1 : RFC1058 (obsolète) / RIPv2 : RFC1387 et 1388 / RFC 1723**
- **Issu de l'architecture XNS de Xerox**
- **Utilisation première dans Arpanet vers 1969**
- **Implémentation par l'Université de Bekerley (routed)**
- **RIP est un protocole Interior Gateway Protocol de type Bellman-Ford (1957) ou distance vector**
- **RIP est le plus ancien des protocoles de routage dynamique**
- **Les paquets RIP sont diffusés toutes les 30 s avec la liste des réseaux accessibles et leur distance (hop count)**

RIP v1 : principe de fonctionnement

- Utilise les services de UDP (port 520)

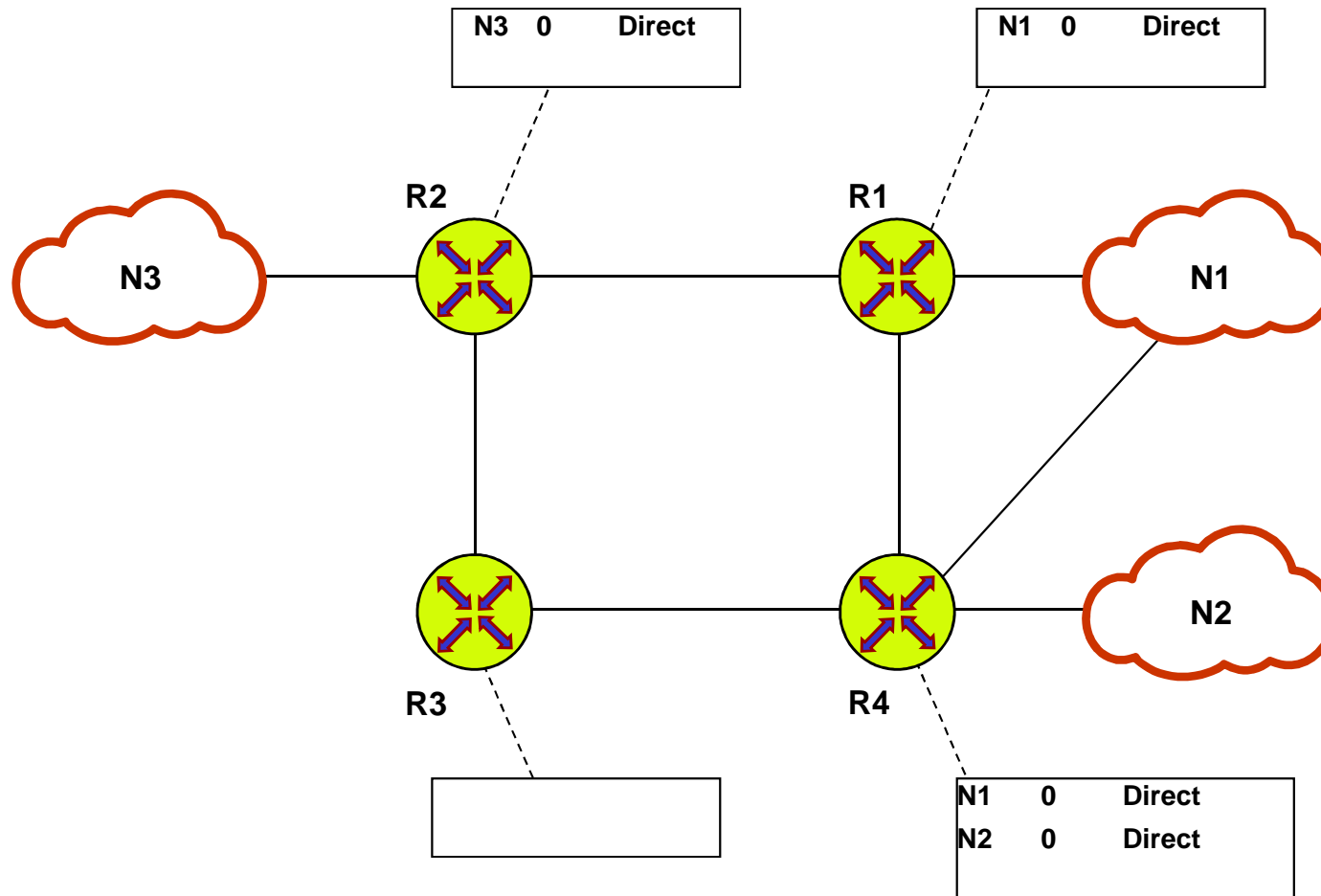


- Le processus de routage dans RIP est basé sur le comptage du nombre de saut. Les routes sont détruites pour un nombre maximal de 16 sauts.
- RIP ne tient pas compte pour le choix d'une route des débits des lignes, de la congestion sur les lignes et de la qualité du service demandé.
- A router running RIP will broadcast its entire routing table out each interface every 30 seconds.
- Adjacent routers will hear this update, add one “hop” and calculate the best path to each destination.

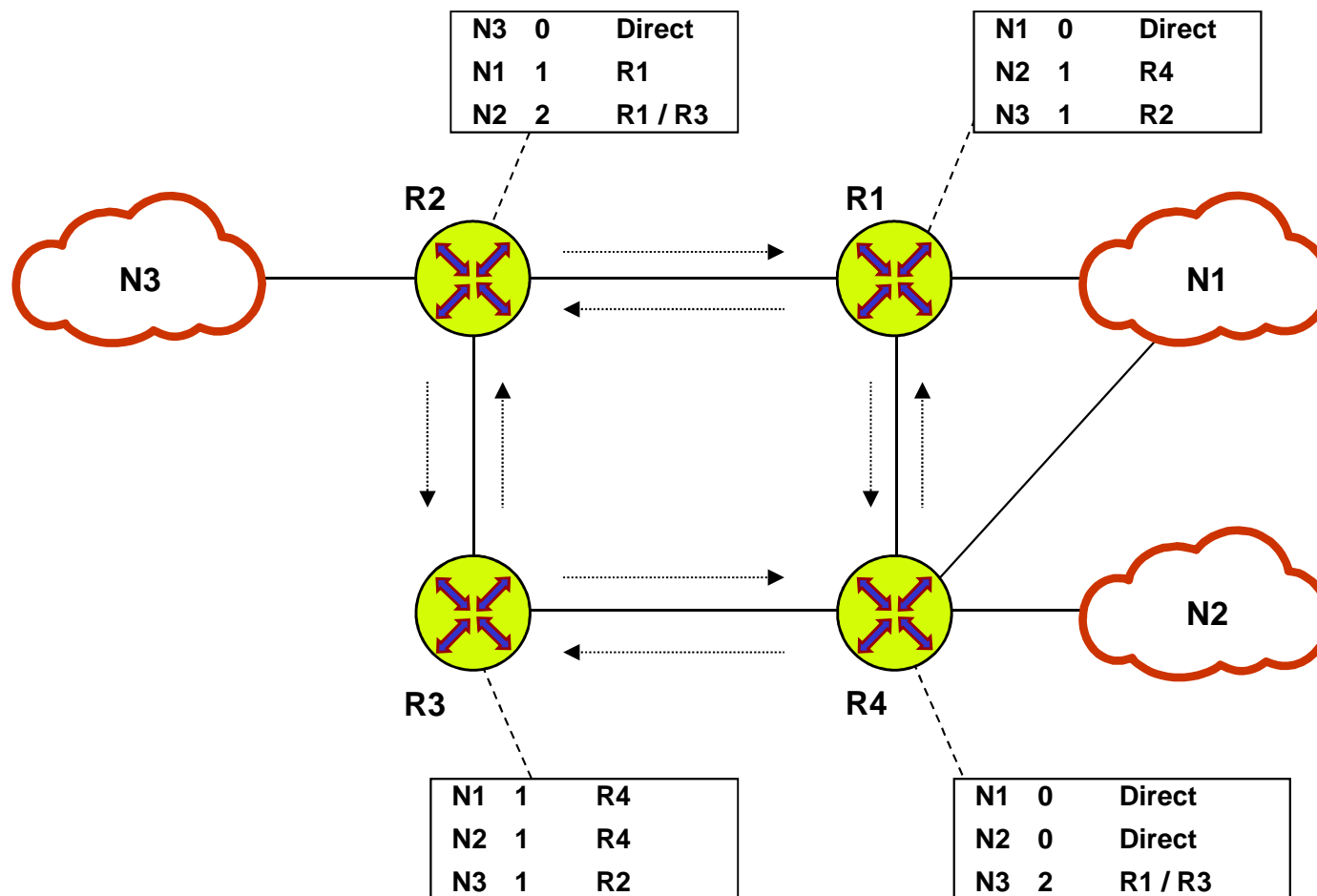
RIP v1 : principe de fonctionnement

- **Temps de convergence : environ 1 minute**
- **Limité à 15 nœuds entre deux extrémités (métrique maximale de 16)**
- **On ne transmet que les changements et pas les tables complètes**
- **Version classique**
 - A expiration du timer
 - Routed améliore l'heuristique en choisissant le timer le moins écoulé
- **Triggered update**
 - Lors d'un changement de topologie
 - Route rompue : sans nouvelle d'une route depuis 3 minutes, elle est considérée comme rompue, sa métrique devient 16
- **Mise à jour : envoi d'information toutes les 30 secondes**
- **Timer aléatoire : les mises à jour sont espacées de 1 à 5 secondes pour éviter une congestion**

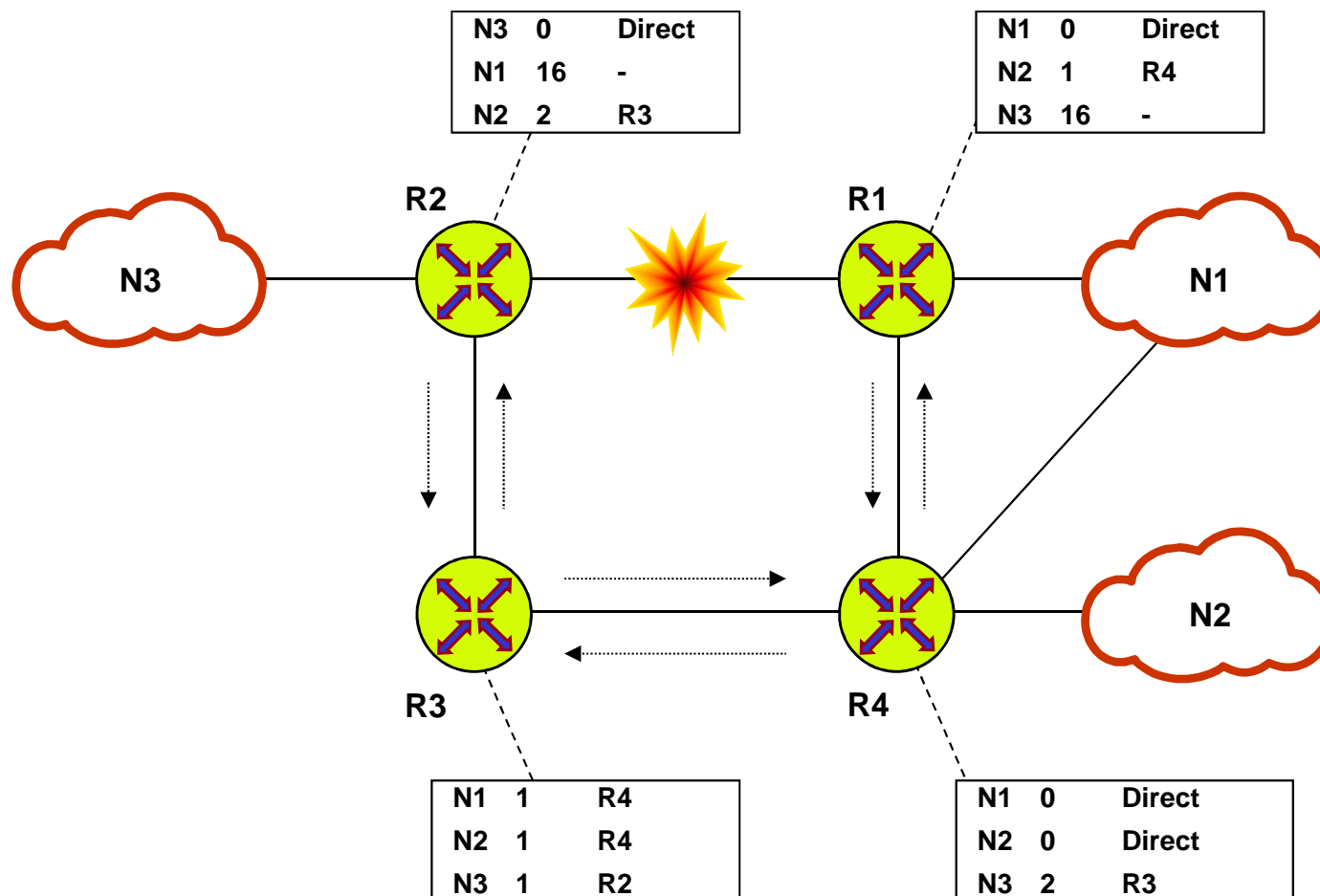
RIP v1 : principe de fonctionnement



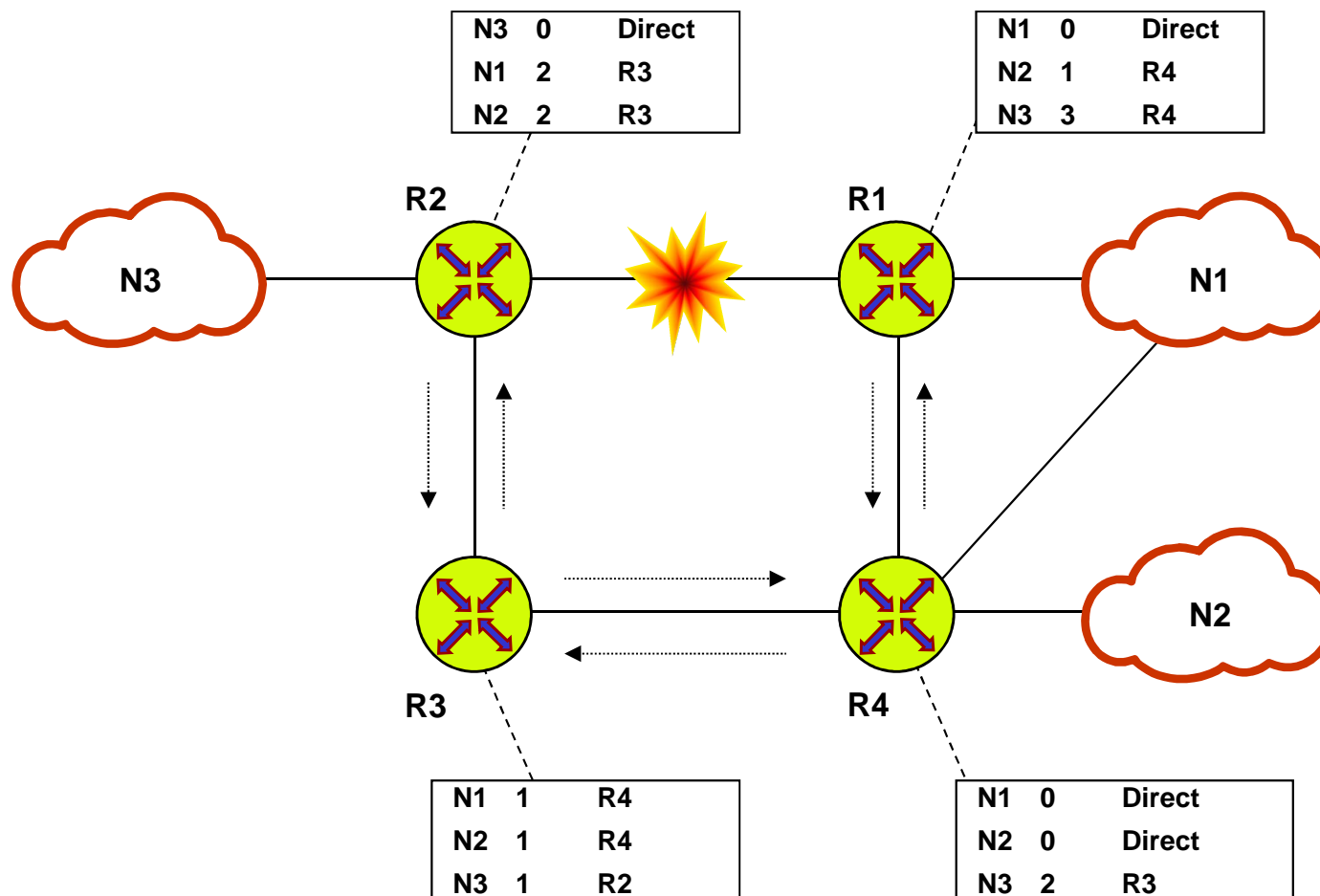
RIP v1 : principe de fonctionnement



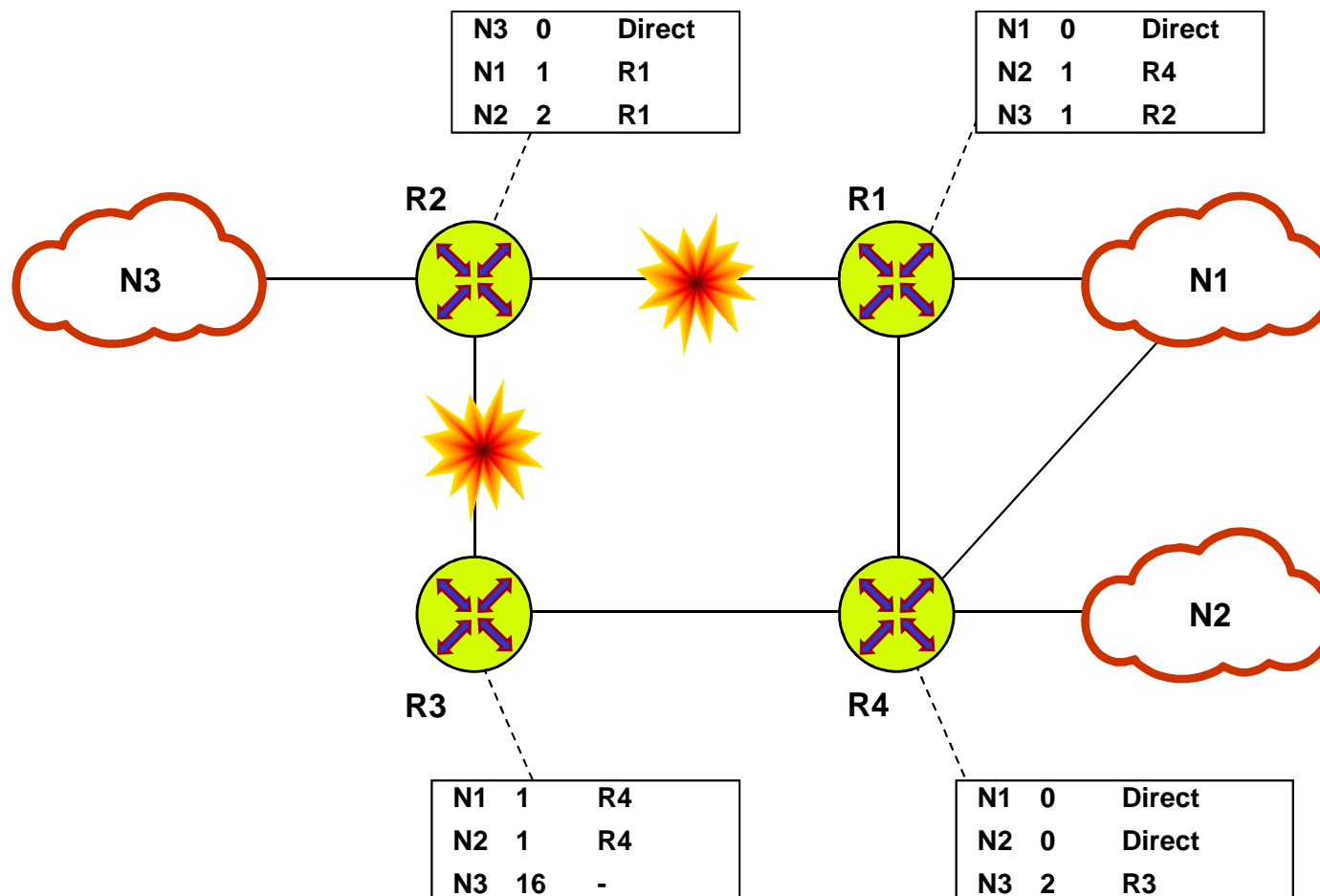
RIP v1 : reroutage sur incident



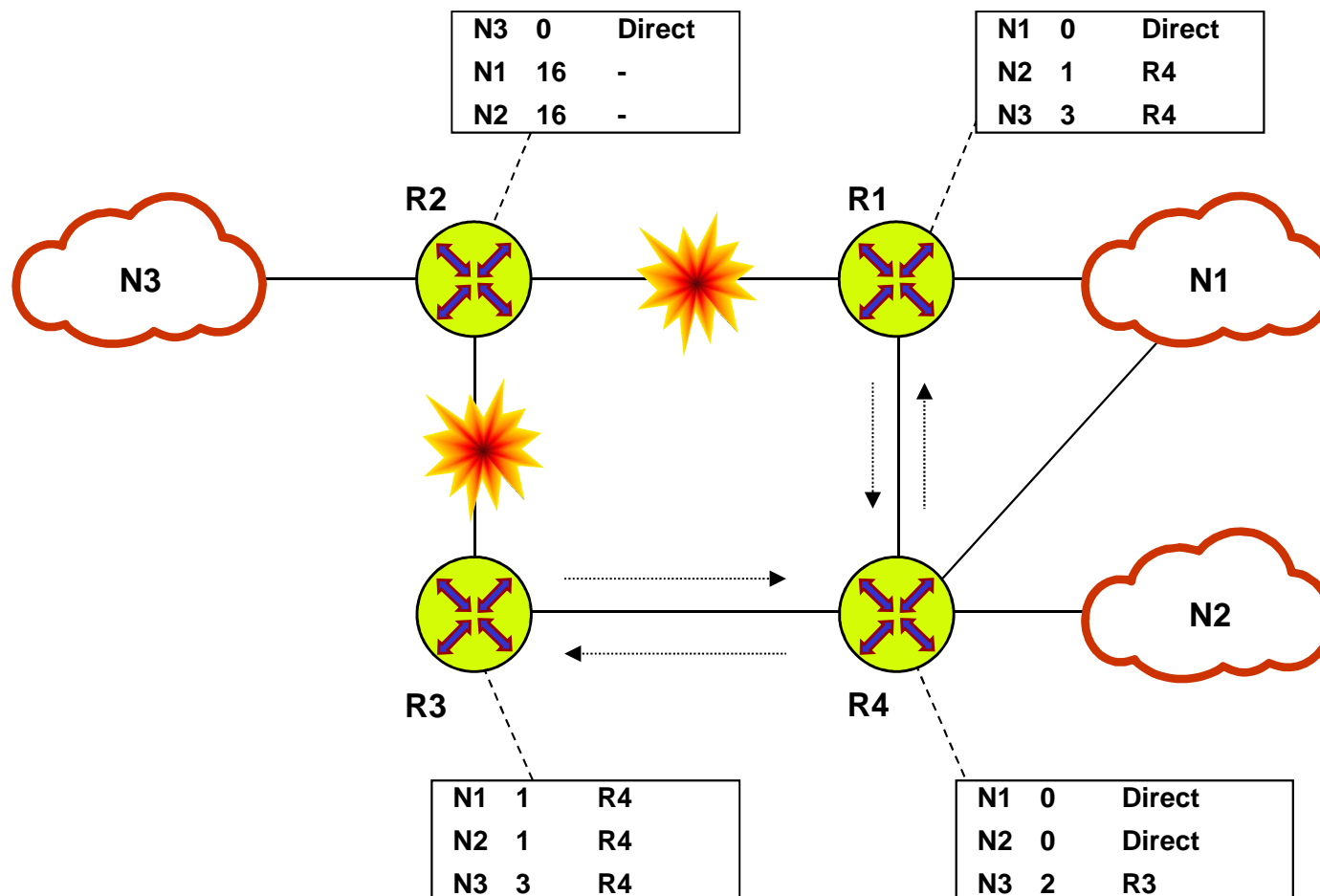
RIP v1 : reroutage sur incident



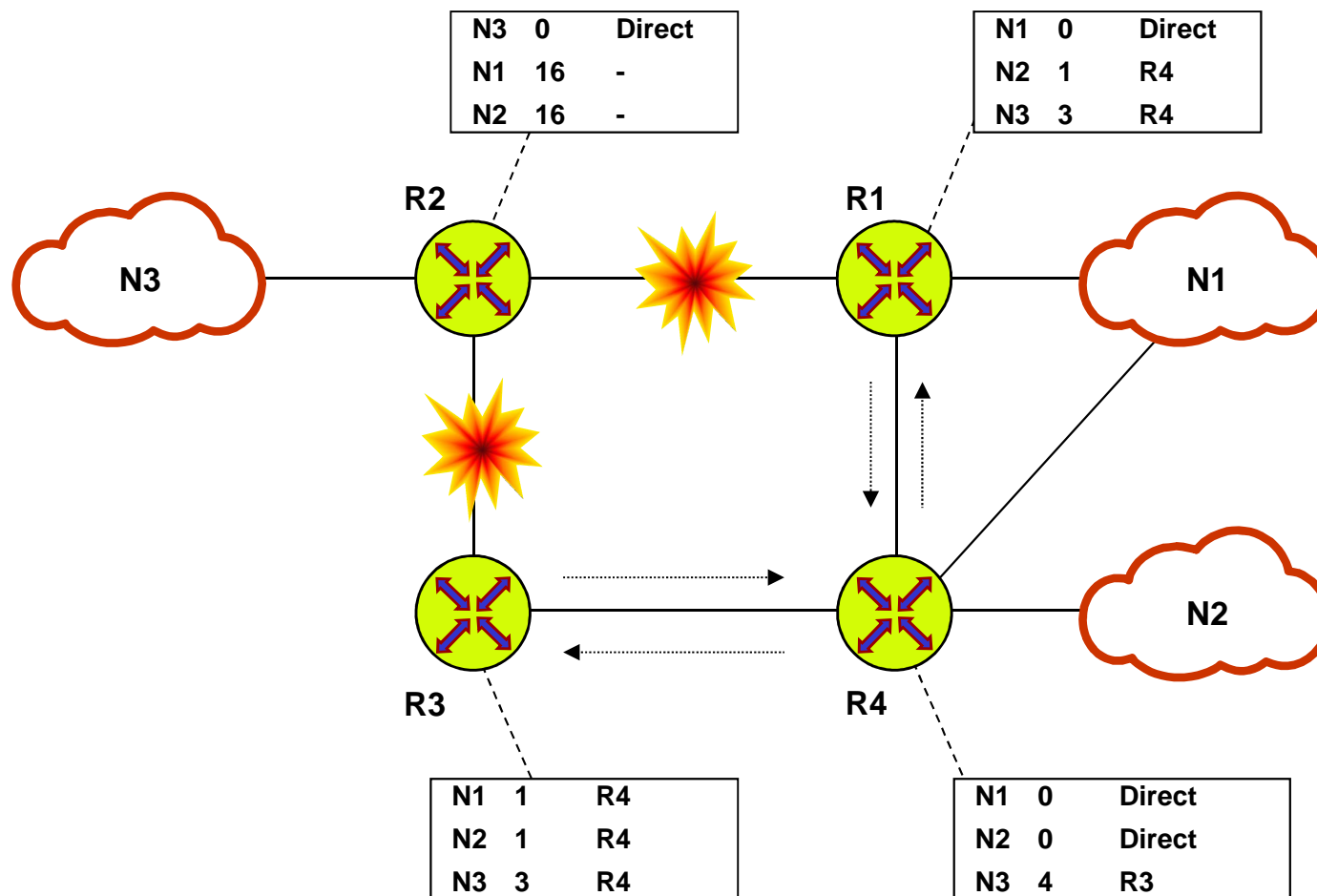
RIP v1 : compte à l'infini



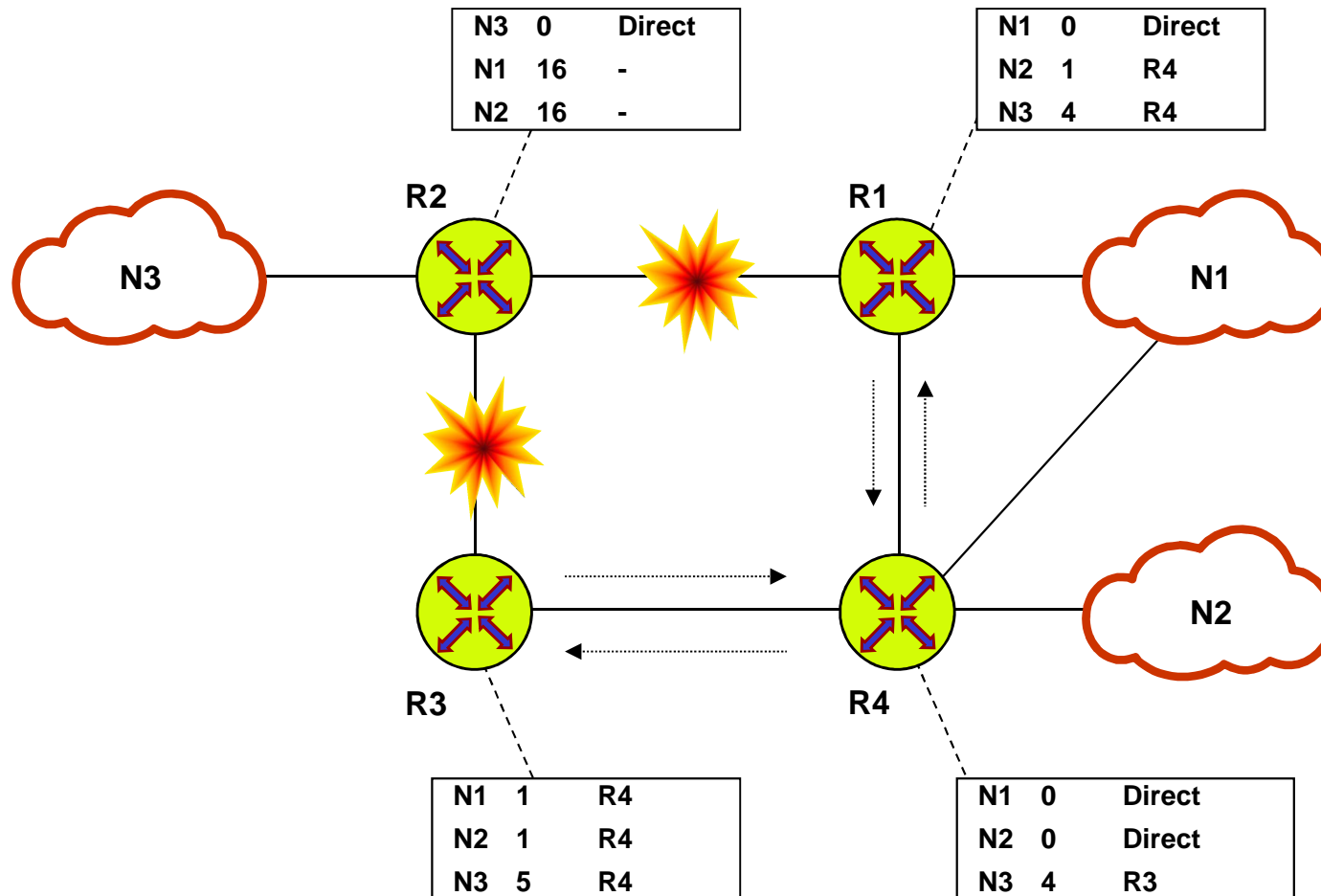
RIP v1 : compte à l'infini



RIP v1 : compte à l'infini

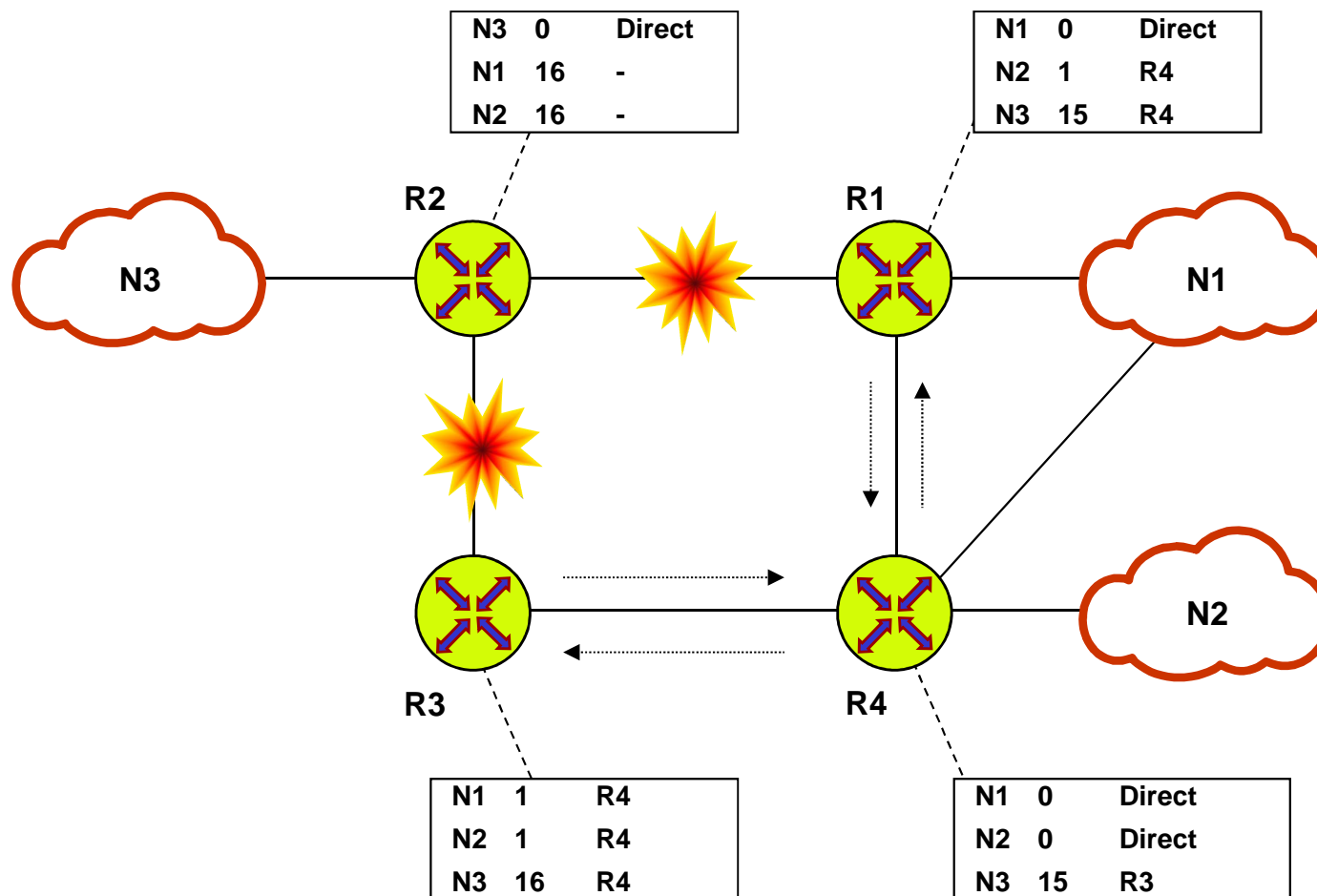


RIP v1 : compte à l'infini



...

RIP v1 : compte à l'infini

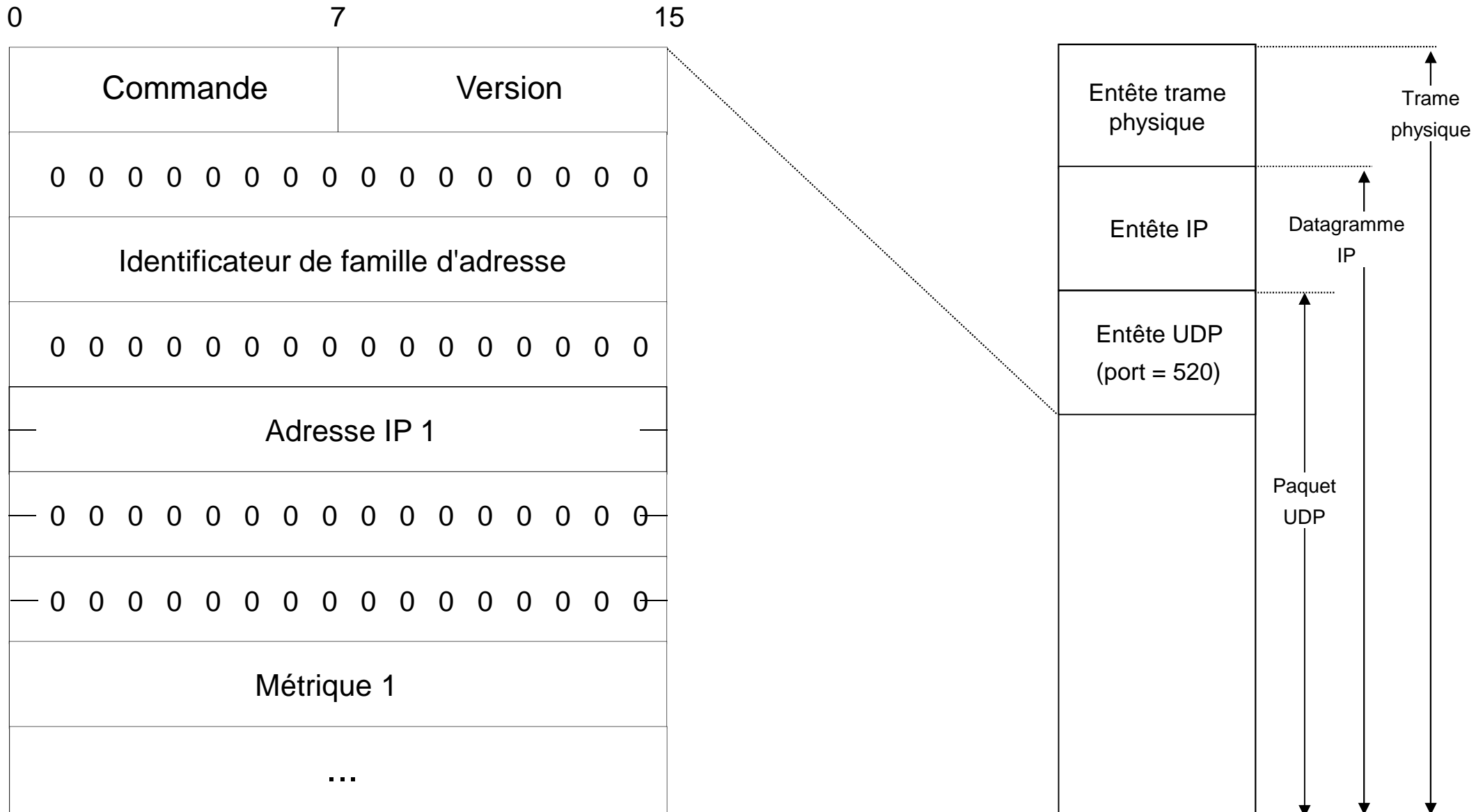


...

RIP v1 : quelques solutions aux problèmes

- **Triggered update** : émission non pas à expiration du timer mais dès la détection d'un problème
- **Split horizon** : rétention d'information :
 - In Split horizon, a route is not advertised out the interface from which it was learned.
 - *A passe par B pour joindre C*
 - *B n'envoie jamais à A une indication de métrique plus faible que celle de A*
 - Simple : mise à jour diffusée sauf vers la source de la mise à jour
 - Avec route empoisonnée : pour l'origine de l'information, on donne une métrique maximale
- **Split horizon with Poisoned Reverse**
 - Another anti-loop mechanism is called Poison Reverse. In poison reverse, routes are advertised with a hop count of 16 out the interface from which they were learned.
- **Hold on** :
 - La rupture est diffusée mais sa propre table n'est modifiée qu'à expiration d'un timer
 - Réduit les boucles mais ralentit l'algorithme

RIP v1 : format des messages



RIP v1 : format des messages

- **COMMANDE** **1 octet**
 - Type du message
 - 1 : Request (pull)
 - 2 : Response (push)
 - 3 : Trace ON
 - 4 : Trace OFF
 - 5 : Reserved
- **VERSION** **1 octet**
 - Indique la version de RIP utilisée
 - 1 : RIP v1
 - 2 : RIP v2
- **Jusqu'à 25 informations (taille de paquet maximale)**
- **Identificateur de la famille d'adresse**
 - 2 pour IP
- **Adresse IP / Métrique (25 entrées possibles)**

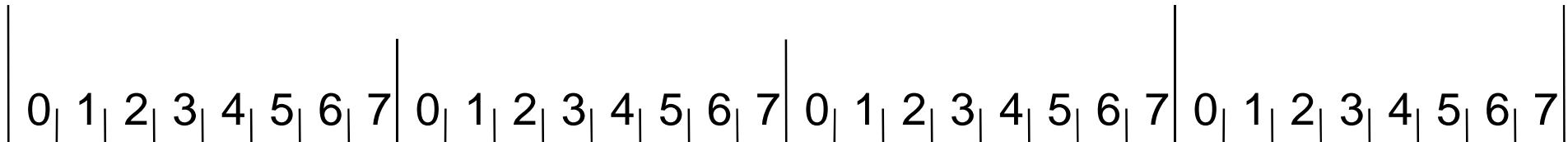
RIP v1 : Envoi et réception de demande

- **Généralement en broadcast sur les interfaces du nœud**
- **Emission**
 - Démarrage : adresse 0.0.0.0 et métrique 16
 - Débogage : demande de route particulière
- **Réception**
 - Vérification des champs (version, 0 pour les champs vides et checksum)
 - Envoi de toute la table ou des routes demandées

RIP v2 : généralités

- **RFC 1387 / 1388**
- **Justifications de RIP 2 ?**
 - Limitations de RIPv1
 - Apparition de IS-IS ou OSPF
- **Simplification des configurations**
- **Overhead limité**
- **Utilisation du multicast pour la diffusion des tables**
 - 224.0.0.9
- **Authentication**
- **RIPv2 est compatible RIPv1**

RIP v2 : format des paquets



| | | |
|-------------------------------------|---------|-------------------------|
| Commande | Version | Domaine de routage |
| Identificateur de famille d'adresse | | Identifiant de la route |
| Adresse IP | | |
| Masque de sous-réseau | | |
| Adresse du prochain relais | | |
| Métrique | | |

RIP v2 : format des paquets

- **COMMANDE** **1 octet**
 - Identique à RIP v1
- **VERSION** **1 octet**
 - Présent dans RIP v1
 - Version 2
- **DOMAINE DE ROUTAGE**
 - Plusieurs sous-réseaux logiques partageant leurs informations de routage ; par défaut 0
- **Identificateur de famille d'adresse**
 - 2 IP
 - 0xFFFF Authentification
- **Identifiant de route**
 - Utile aux EGP mais inutile à l'IGP Exemple : numéro d'AS
 - Ignoré et dupliqué

OSPF : pourquoi un nouveau protocole ?

- **Limites de RIP1**

- 15 « sauts » maximum
- Pas de support de VLSM
- Envoi périodique de l'ensemble de la table de routage
- Pas de sécurisation par authentification
- Les tables de routage convergent lentement
- Décision de routage basée sur le nombre de saut
- Pas de hiérarchie de routage

- **Conséquences**

- Augmentation de la taille des AS
- Augmentation de la taille des tables de routage RIP
- Augmentation de la consommation de bande passante par RIP
- Augmentation du temps de convergence des routes
- Augmentation du nombre de changements des routes

- **RIP2 pallie une partie des « faiblesses » de RIP1**

OSPF : généralités

- **Open Shortest Path First**
- **OPSF Working Group**
- **Protocole de type IGP**
- **Dijkstra algorithm (link state / SPF)**
- **Utilise les services de IP (protocole = 89)**
- **Diffusion des mises à jour en multicast (224.0.0.xxx)**
- **Diffusion des informations de routage en cas d'événement**
- **Link State Advertisements**
- **Hiérarchie de routage à 2 niveaux**
- **Coûts des liaisons / Multi-chemins / répartition de charge**
- **Variable Length Subnet Masks / sous-réseaux répartis**
- **Sécurisation par authentification des paquets OSPF**
- **“OSPF's implementation are at the implementor's own risk”**

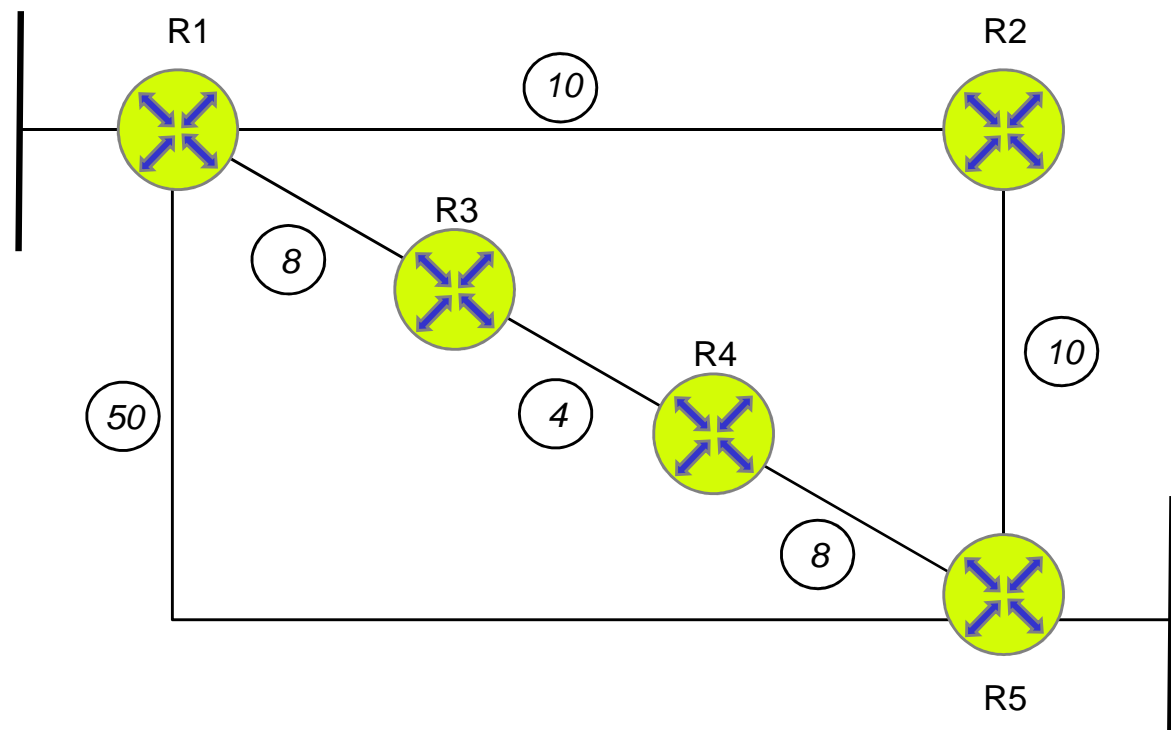
OSPF : principe général

- **Chemin de meilleur coût**

- RFC 1253 : 100 000 000 / débit du lien en Bps
- FDDI C=1, 10BaseT C=10, E1 C=48, T1=64, 56Kbps=1785, ...

- **Multi chemins**

- **Répartition de charge**



OSPF : historique

- **1987 / 1988**
 - Le protocole IS-IS est retenu par l'ANSI pour le routage intra-domaine (environnement CLNP)
 - début des travaux sur OSPF (OSPF Working Group) sur la base d'IS-IS
- **1989**
 - publication des spécifications de OSPFv1 (rfc 1131) / 1ère implémentation OSPF (UMD)
 - 1ère implémentation d'OSPF par Proteon
 - IS-IS est proposé comme standard ISO
- **1990 / 1991**
 - 1er test d'interopérabilité d'OSPF
 - publication des spécifications de OSPFv2 (rfc 1247) / démonstration OSPF à Interop
 - 1ère implémentation d'OSPF par Cisco
- **1992**
 - ajout d'OSPF à « gated » / OSPF recommandé comme IGP
 - OSPF est mis en œuvre sur les réseaux au détriment d'IS-IS
- **1993**
 - Mise à jour des spécifications d'OSPFv2 (rfc 1583)
 - support du multicasting (MOSPF)
- **1994**
 - ajout d'une fonction d'authentification
 - Mise à jour des spécifications d'OSPFv2 (rfc 2178)
- **1996 / 1998**
 - IS-IS et OSPF utilisés par lesgrans ISP
- **1999 / 2001**
 - Extensions à OSPF

OSPF : origine des travaux

- **Travaux de Bolt, Beranek, and Newman sur le Landmark packet-switching développé dans les années 70 sur la base d'un algorithme de type SPF**
- **Travaux de recherche du Dr. Perlman sur le routage par diffusion**
- **Travaux du BBNen 1986 sur les aires de routage**
- **Travaux OSI sur le protocole de routage Intermediate System-to-Intermediate System (IS-IS)**

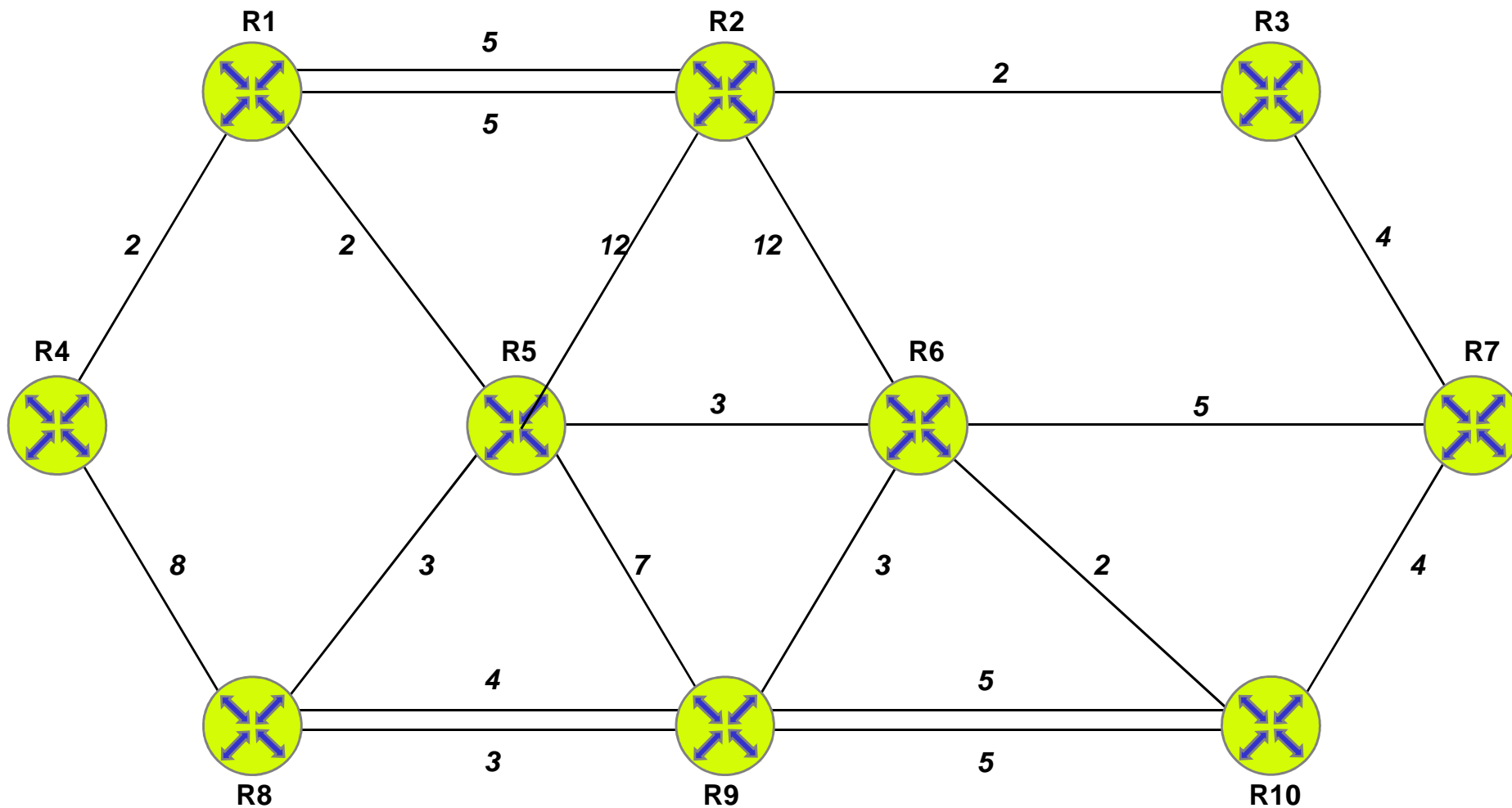
OSPF : Routage de type link-state

- Chaque routeur garde la trace de son propre état
- Chaque routeur envoie et reçoit des paquets Hello pour connaître ses voisins
- Chaque routeur construit des LSP (Link-State Packet) avec ses informations locales
- Tous les routeurs diffusent (flooding) en multicast leur LSA (Link State Advertisement)
- Avec les LSP reçus, chaque routeur construit sa base de données d'information (LSPDB)
- Tous les routeurs ont la même LSPDB qui est indépendante de la table de routage
- Les LSPs sont transmis lorsqu'un changement survient
- Un rafraichissement périodique de la base est réalisé
- Chaque routeur construit sa topologie du réseau avec l'algorithme SPF de Dijkstra
 - la topologie est établie sous la forme dite de Shortest-Path Tree
 - chaque routeur est la racine du SPT
 - boucles impossibles
- A partir de la topologie SPT la table de routage (Routing Information Table) est construite

OSPF : algorithme Shortest Path de Dijkstra

- **Begin.**
- $\forall v \in V,$
 - add v to set U ,
 - initialize $\text{Distance}(v) = \text{cost}(s, v)$
- **Distance(s) = 0; Remove s from U.**
- **while U is not empty do**
 - $v \mapsto$ any member of G with minimum distance.
 - Remove v from U .
 - For each neighbor w of v , do
 - *if member(w, U)*
 - *$\text{distance}(w) = \min(\text{distance}(w), \text{cost}(w, v) + \text{distance}(v));$*
- **Stop.**

OSPF : algorithme Shortest Path First



OSPF : algorithme Shortest Path First

- Construction des LSP
- R1 : R4 - 2 / R5 - 2 / R2 - 5 / R2 - 5
- R2 : R1 - 5 / R1 - 5 / R3 - 2 / R5 - 12 / R6 - 12
- R3 : R2 - 2 / R7 - 4
- R4 : R1 - 2 / R8 - 8
- R5 : R1 - 2 / R2 - 12 / R6 - 3 / R9 - 7 / R5 - 3
- R6 : R2 - 12 / R7 - 5 / R10 - 2 / R9 - 3 / R5 - 3
- R7 : R3 - 4 / R10 - 4 / R6 - 5
- R8 : R4 - 8 / R5 - 3 / R9 - 4 / R9 - 3
- R9 : R8 - 4 / R8 - 3 / R5 - 7 / R6 - 3
- R10 : R9 - 5 / R9 - 5 / R6 - 2 / R7 - 4

OSPF : algorithme Shortest Path First

- **Table des routeurs adjacents pour R9**

| ● | Cost | Neighbor | Interface |
|---|-------------|-----------------|-----------------------------|
| ● | | R8 | serial0 3 |
| ● | | R8 | serial1 4 |
| ● | | R5 | serial2 7 |
| ● | | R6 | serial3 3 |
| ● | | R10 | serial4 5 |
| ● | | R10 | serial5 5 |

OSPF : algorithme Shortest Path First

● Candidats

- R5 C=7 via S2
- R6 C=3 via S3
- R8 C=3 via S0
- R10 C=5 via S4 ou S5

● Chemins calculés

- R9

● Inconnus

- R1 R2 R3 R4 R7

● Candidats

- R5 C=6 via S0
- R6 C=3 via S3
- R10 C=5 via S4 ou S5
- R4 C=11 via S0

● Chemins calculés

- R9
- R8 C=3 via S0

● Inconnus

- R1 R2 R3 R7

R8

● Candidats

- R2 C=15 via S3
- R4 C=11 via S0
- R5 C=6 via S0 ou S3
- R7 C=8 via S3
- R10 C=5 via S3 ou S4 ou S5

● Chemins calculés

- R9
- R6 C=3 via S3
- R8 C=3 via S0

● Inconnus

- R1 R3

R6

OSPF : algorithme Shortest Path First

- **Candidats**

- R2 C=15 via S3
- R4 C=11 via S0
- R5 C=6 via S0 ou S3
- R7 C=8 via S3

- **Chemins calculés**

- R9
- R6 C=3 via S3
- R8 C=3 via S0
- R10 C=5 via S3 ou S4 ou S5

- **Inconnus**

- R1 R3 R10

- **Candidats**

- R1 C=8 via S0 ou S3
- R2 C=15 via S3
- R4 C=11 via S0
- R7 C=8 via S3

- **Chemins calculés**

- R9
- R5 C=6 via S0 ou S3
- R6 C=3 via S3
- R8 C=3 via S0
- R10 C=5 via S3 ou S4 ou S5

- **Inconnus**

- R3 R5

OSPF : algorithme Shortest Path First

● Candidats

- R1 C=8 via S0 ou S3
- R2 C=15 via S3
- R4 C=11 via S0
- R3 C=12 via S3

● Chemins calculés

- R9
- R5 C=6 via S0 ou S3
- R6 C=3 via S3
- R7 C=8 via S3
- R8 C=3 via S0
- R10 C=5 via S3 ou S4 ou S5

● Inconnus R7

● Candidats

- R2 C=13 via s0 ou S3
- R3 C=12 via S3
- R4 C=10 via S0 ou S3

● Chemins calculés

- R9
- R1 C=8 via S0 ou S3
- R5 C=6 via S0 ou S3
- R6 C=3 via S3
- R7 C=8 via S3
- R8 C=3 via S0
- R10 C=5 via S3 ou S4 ou S5

● Inconnus R1

OSPF : algorithme Shortest Path First

● Candidats

- R2 C=13 via s0 ou S3
- R3 C=12 via S3

● Chemins calculés

- R9
- R1 C=8 via S0 ou S3
- R4 C=10 via S0 ou S3
- R5 C=6 via S0 ou S3
- R6 C=3 via S3
- R7 C=8 via S3
- R8 C=3 via S0
- R10 C=5 via S3 ou S4 ou S5

● Inconnus R4

● Candidats

- R2 C=13 via s0 ou S3

● Chemins calculés

- R9
- R1 C=8 via S0 ou S3
- R3 C=12 via S3
- R4 C=10 via S0 ou S3
- R5 C=6 via S0 ou S3
- R6 C=3 via S3
- R7 C=8 via S3
- R8 C=3 via S0
- R10 C=5 via S3 ou S4 ou S5

● Inconnus R3

OSPF : algorithme Shortest Path First

- **Candidats**

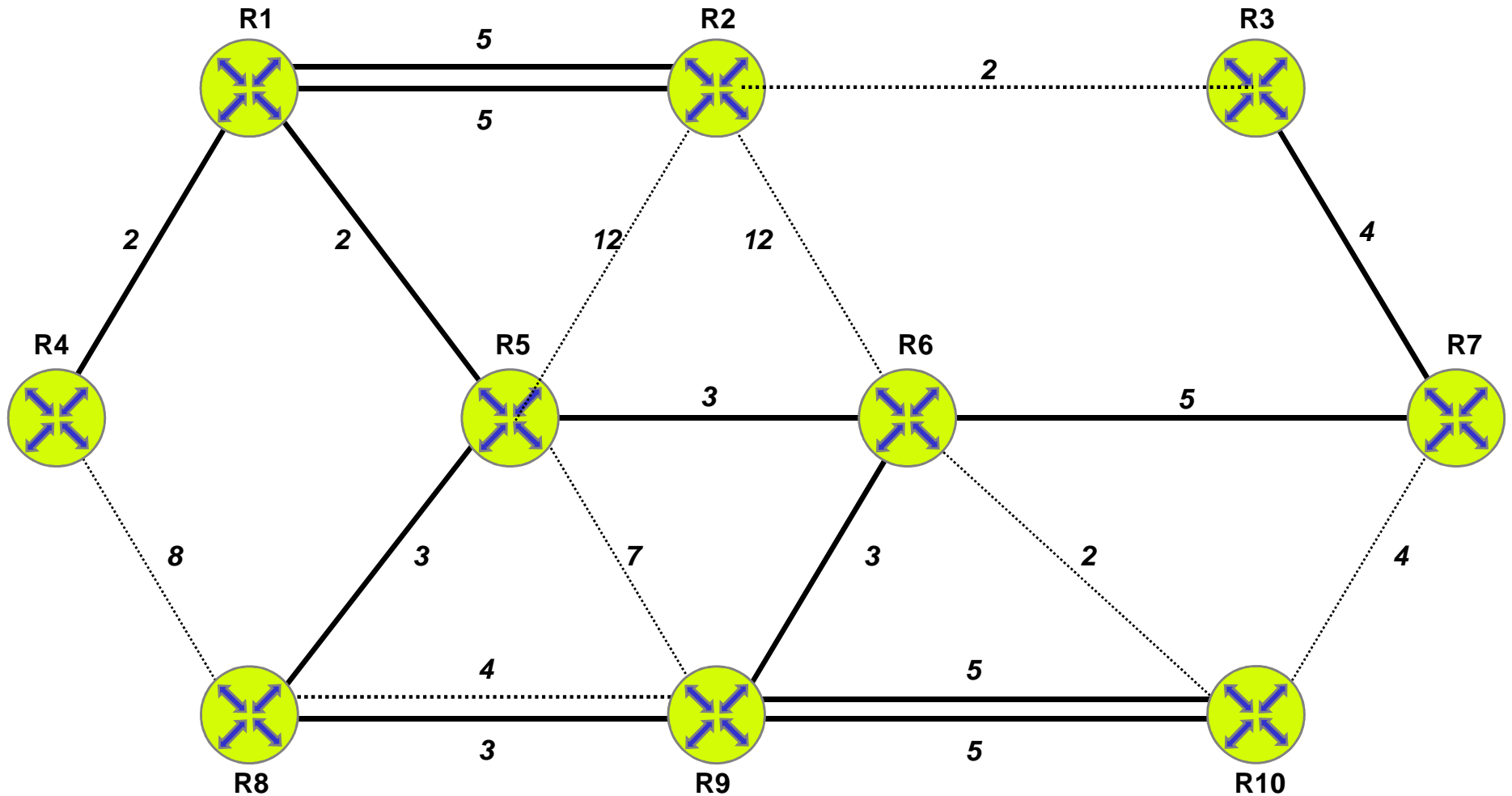
- **Chemins calculés**

- R9
- R1 C=8 via S0 ou S3
- R2 C=13 via s0 ou S3
- R3 C=12 via S3
- R4 C=10 via S0 ou S3
- R5 C=6 via S0 ou S3
- R6 C=3 via S3
- R7 C=8 via S3
- R8 C=3 via S0
- R10 C=5 via S3 ou S4 ou S5

- **Inconnus**

R2

OSPF : algorithme Shortest Path First



OSPF : aires de routage

- Une aire de routage OSPF est constituée par un groupe contigue de machines et de réseaux
- La topologie à l'intérieur d'une aire est invisible de l'extérieur
- Diminution du trafic de routage
- **Backbone area**
 - obligatoire
 - aire 0
 - toutes les autres aires doivent avoir une connexion physique ou logique avec l'aire backbone
 - chaque aire envoie ses informations de routage à l'aire backbone qui rediffusent l'information
- **Transit area**
- **xxx.xxx.xxx.xxx**
- **Stub area**
- **Not-So-stubby-area**

OSPF : aires de routage

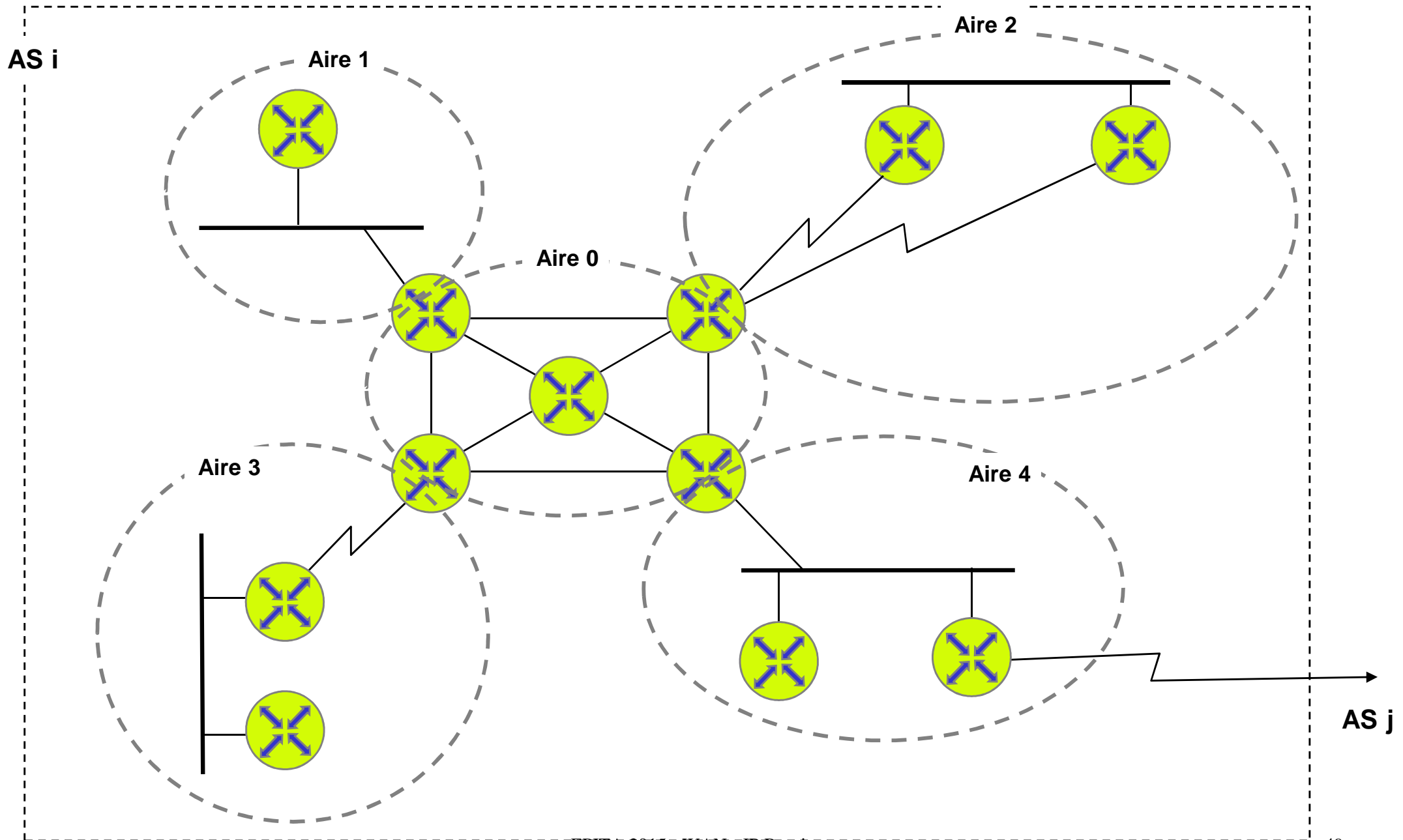
● Stub Areas

- Pas de diffusion des annonces externes à partir des Stub Areas et vers les Stub Areas
- Création de Stub Areas quand la majeure partie de la base de données d'information est constituée d'annonces de routes externes
- réduction de la taille des base de données
- utilisation de route par défaut à la place des routes externes
- Pas de lien virtuel au sein d'une Stub Area
- Aucun ASBR dans une Stub Area

● Not-So-Stubby Areas

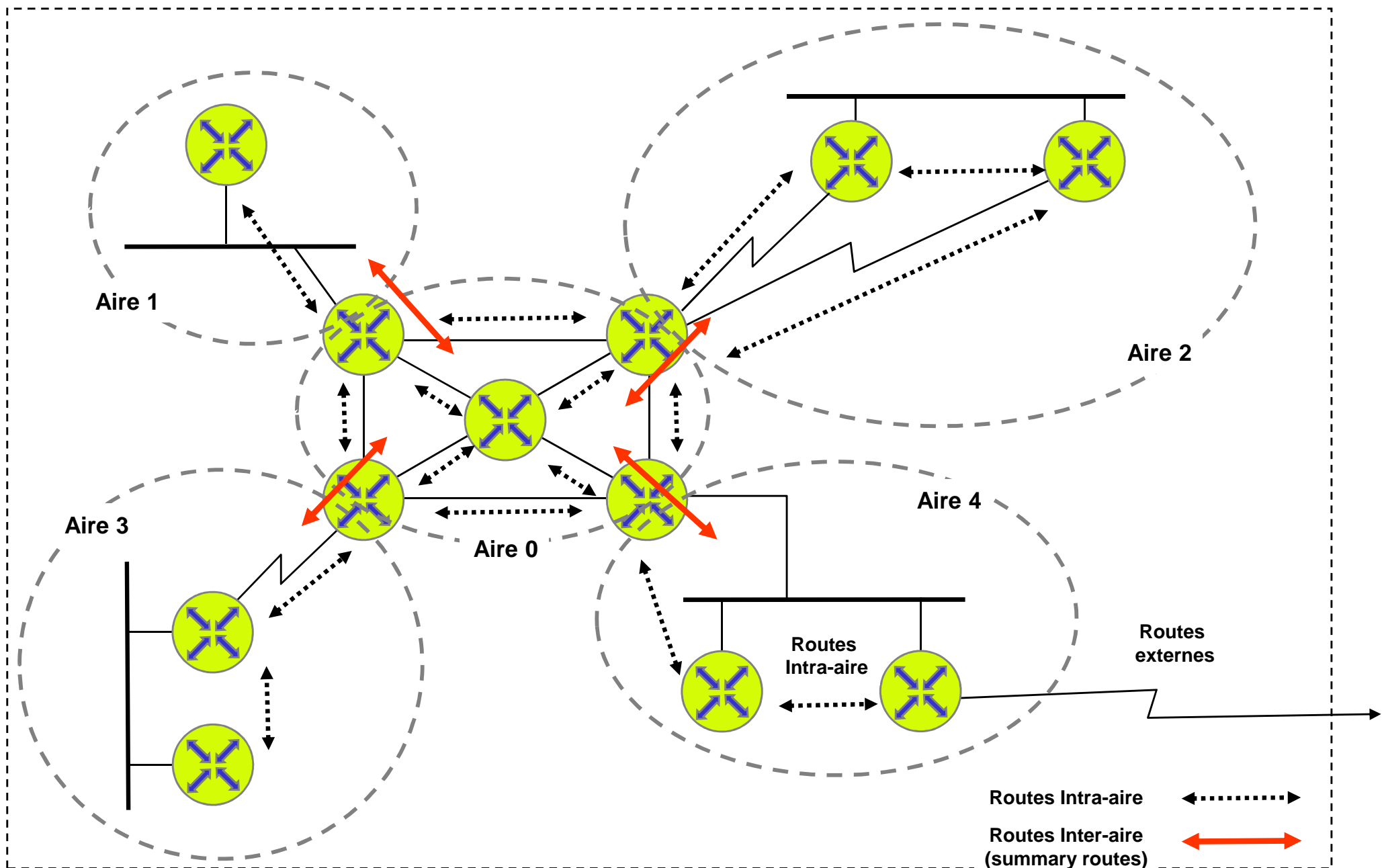
- Diffusion des routes externes à l'intérieur de l'aire
- Les routes externes provenant d'autres aires ne rentrent pas dans les NSSA
- ASBR possible dans une NSSA

OSPF : aires de routage

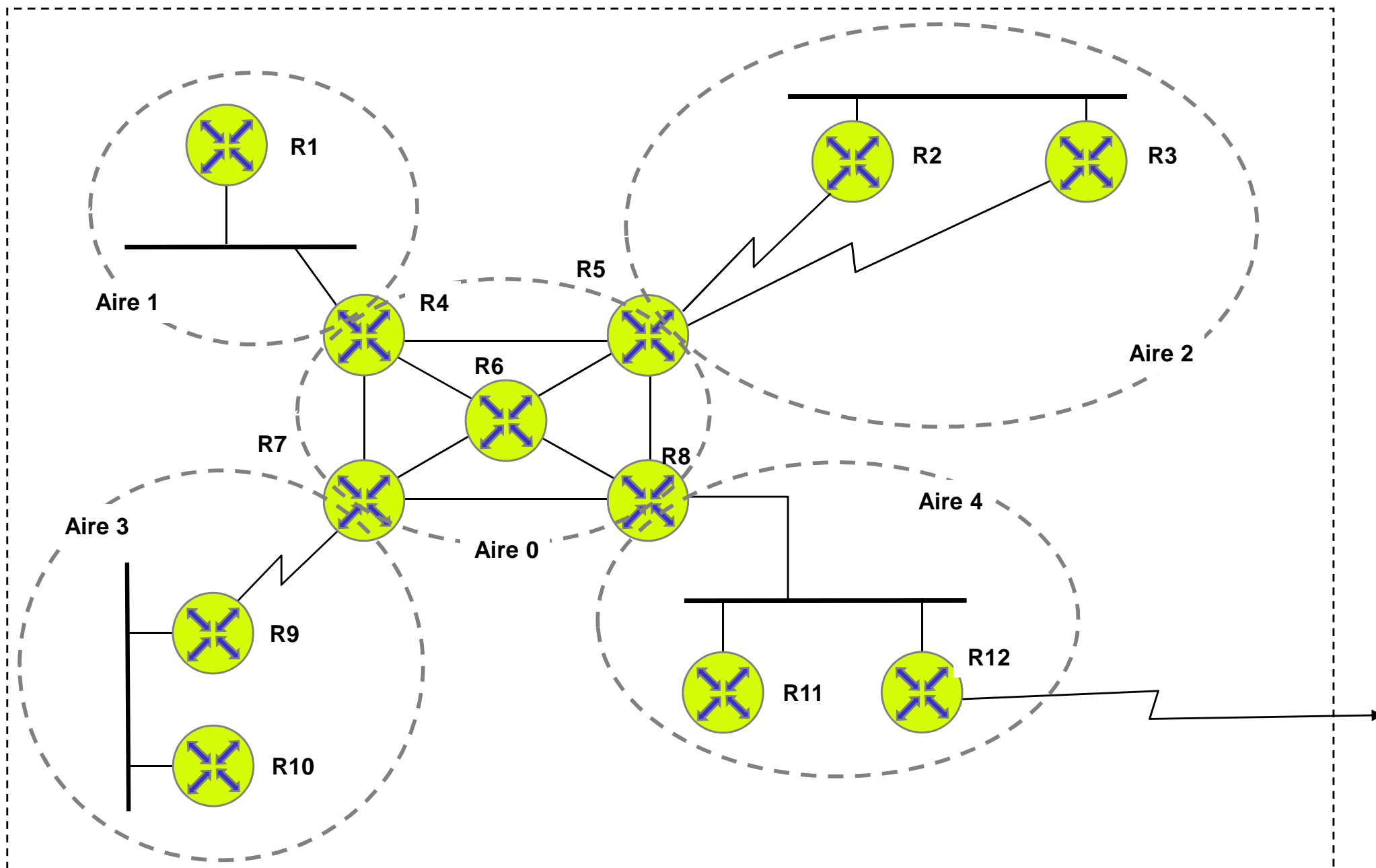




OSPF : typologie des routes



OSPF : routeurs voisins



OSPF : routeurs voisins

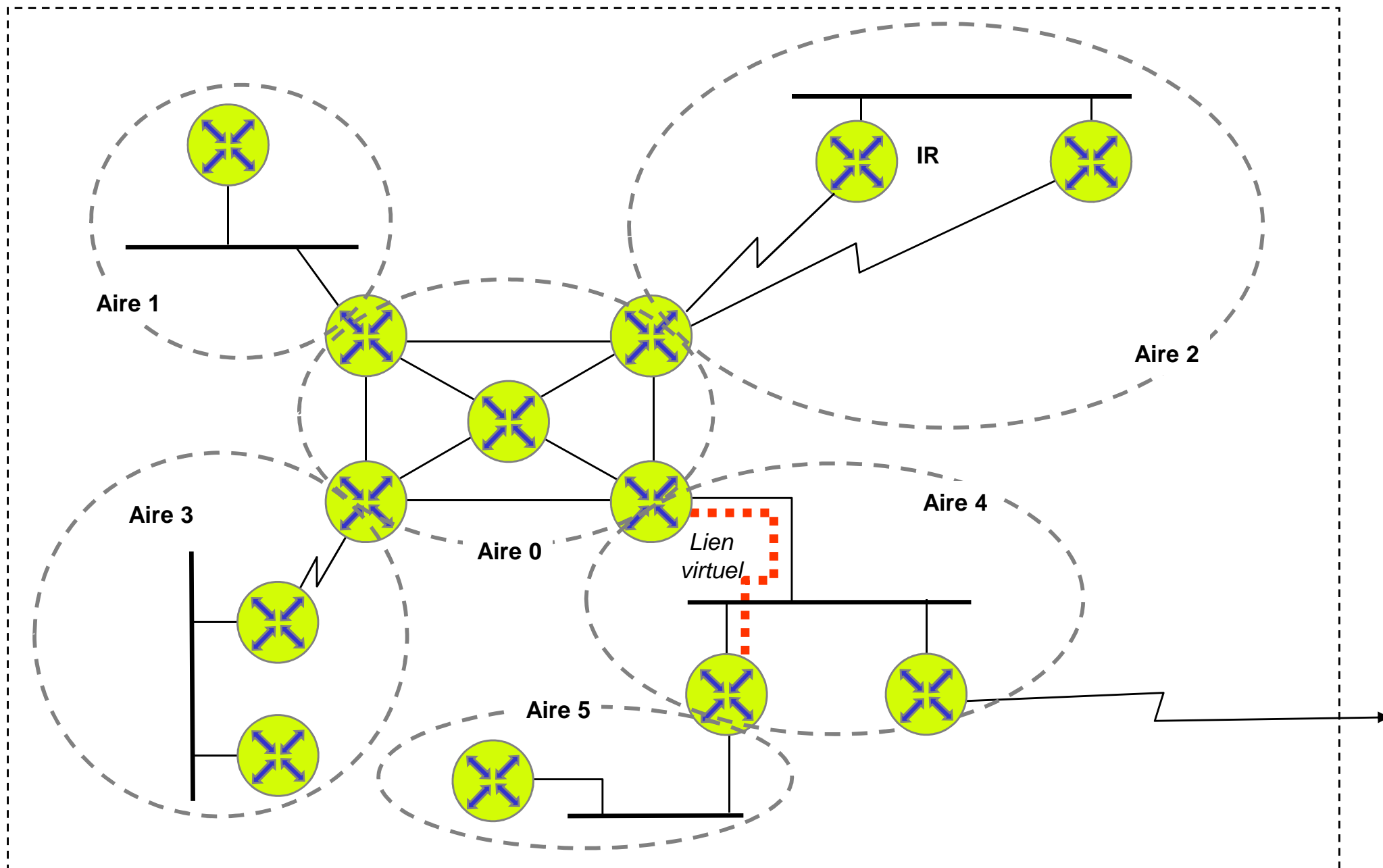
- **Autant d 'ensemble de routeurs voisins que de liens physiques ayans des adresses IP différentes**
- **Les interfaces des routeurs voisins appartiennent au même réseau ou sous-réseau IP**
- **Scénarios possibles**
 - R1 R4 / R2 R3 / R2 R5 / R3 R5 / R4 R5 / R5 R6 / R5 R8 / R4 R6 / R4 R7 / R6 R7 / R6 R8 / R7 R8 / R7 R9 / R10 R9 / R8 R11 R12
 - ou
 - R1 R4 / R2 R3 R5 / R4 R5 R6 R7 R8 / R7 R9 / R10 R9 / R8 R11 R12
 - ou
 - ...
- **Authentication**
- **Protocole Hello**

OSPF : lien virtuel

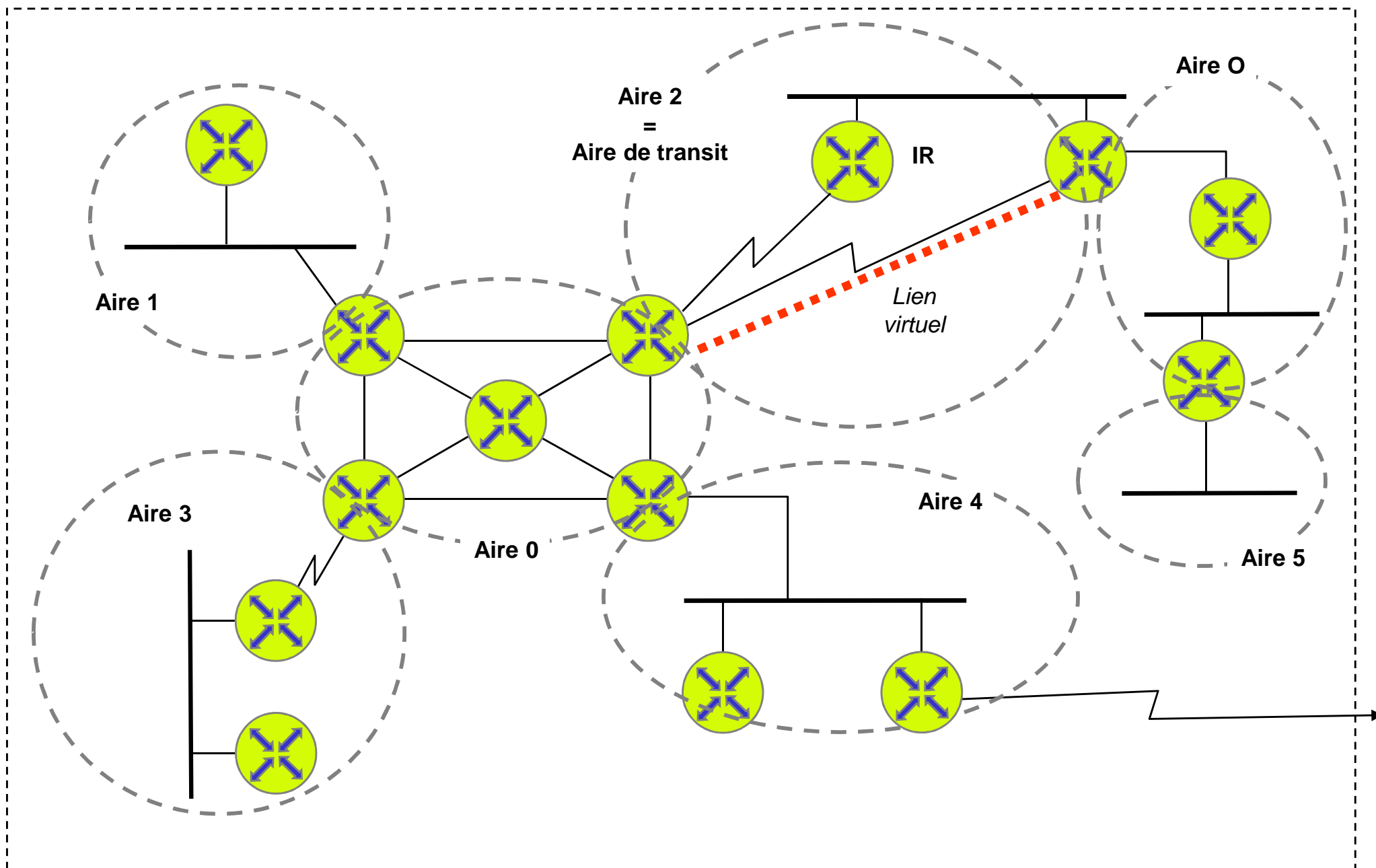
- **Raccordement d'une aire qui n'a pas d'attachement physique avec l'aire backbone**
 - création d'un lien logique avec l'aire backbone
 - entre deux ABR attachés à une même aire et dont l'un des deux est attachés l'aire backbone

- **Assurer la continuité de l'aire backbone**
 - aire backbone répartie
 - lien logique entre deux parties de l'aire backbone
 - utilisation d'une aire de transit

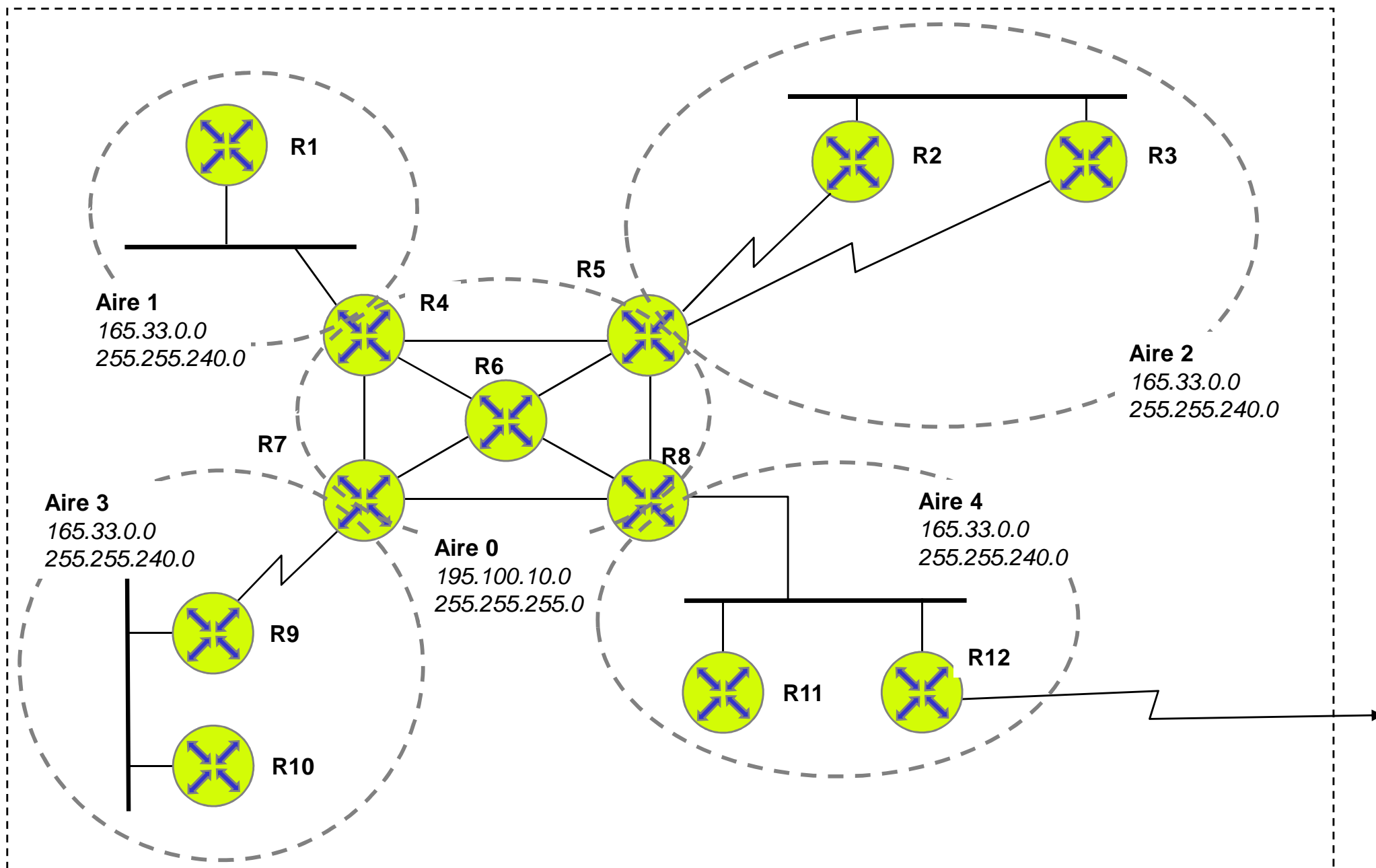
OSPF : lien virtuel



OSPF : lien virtuel



OSPF : sous-réseaux répartis



OSPF : typologie des messages

- **HELLO: Type 1**

- Identification des voisins
- Election du DR dans le cas des réseaux multiaccès
- Recherche du DR
- “I’m alive”

- **DATABASE DESCRIPTION: Type 2**

- Envoi d’information permettant au routeur d’identifier les informations manquantes ou obsolètes

- **LINK STATE REQUEST: Type 3**

- Demande d’information pour mise à jour de la base de données d’information

- **LINK STATE UPDATE: Type 4**

- Réponse à un message de type Link State Request

- **LINK STATE ACK: Type 5**

- Acquittement transmis après réception d’un message Link State Update

OSPF : format de l'entête des messages

| | | |
|-------------------------------|---------------------|----------------|
| VERSION | TYPE | MESSAGE LENGTH |
| SOURCE ROUTER IP ADDRESS | | |
| AREA ID | | |
| CHECKSUM | AUTHENTICATION TYPE | |
| AUTHENICATION DATA 8 BYTES | | |

OSPF : format des paquets

- **VERSION**

- Version du protocole

- **Type**

- 1 Hello
 - 2 Database description (topologie)
 - 3 Link Status Request
 - 4 Link Status Update
 - 5 Link Status Acknowledgement
-

- **SOURCE IP ADDRESS**

- Adresse IP de l'émetteur

- **AREA ID**

- Identification de l'aire

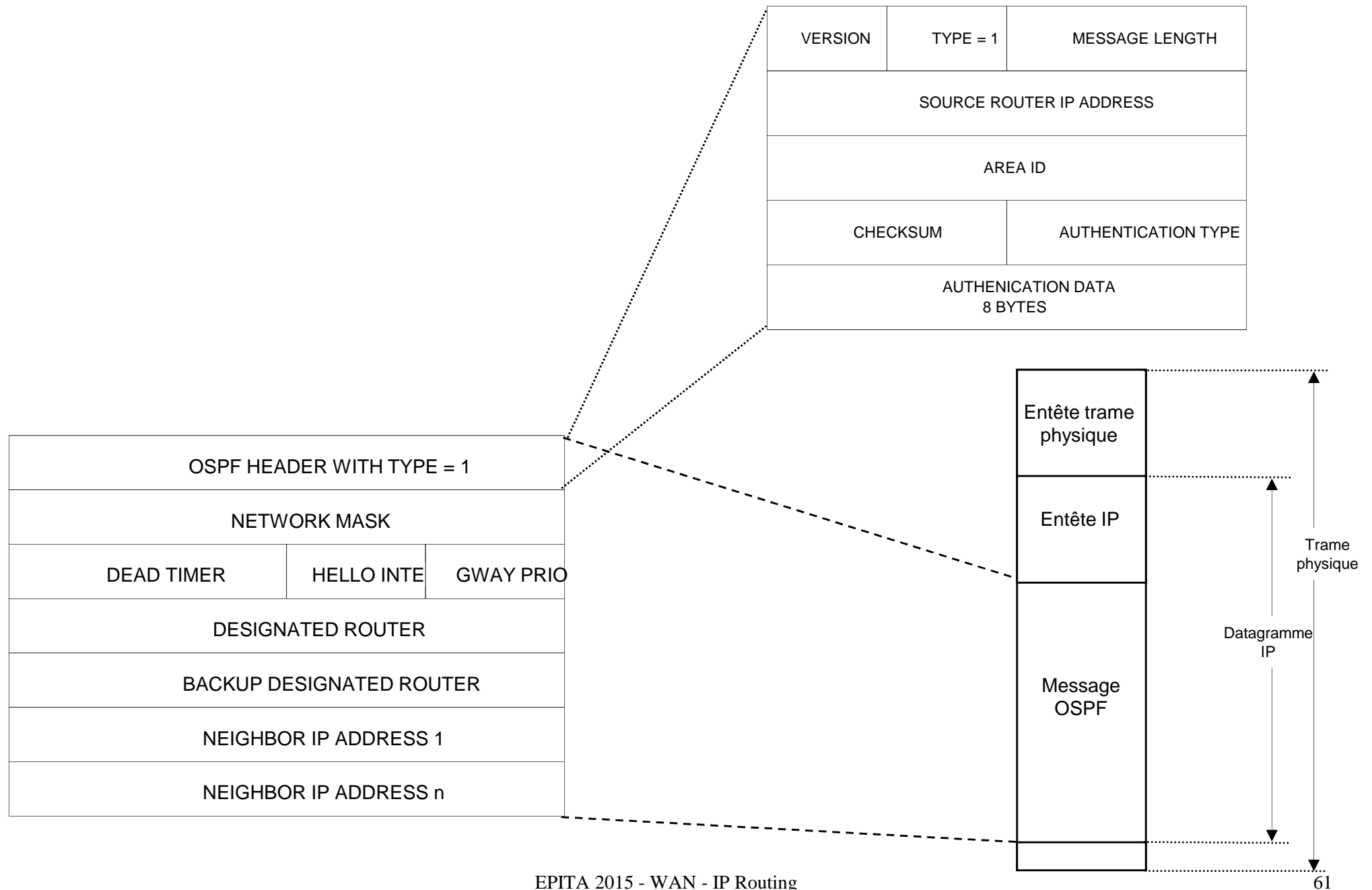
- **AUTHENTICATION TYPE**

- 0 pas d'authentification
- 1 Authentification existante

- **AUTHENTICATION**

Mot de passe

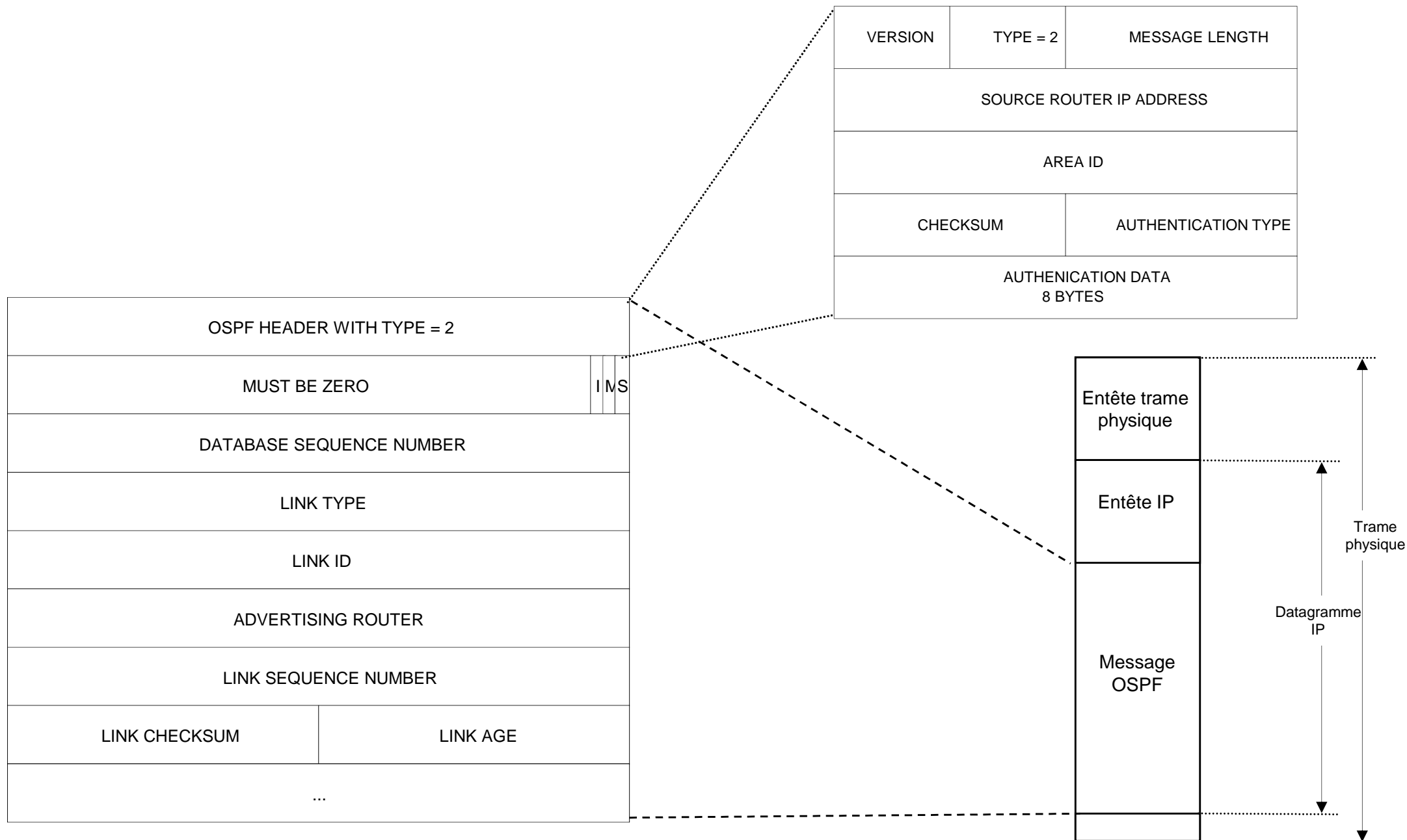
OSPF : format des messages de type 1



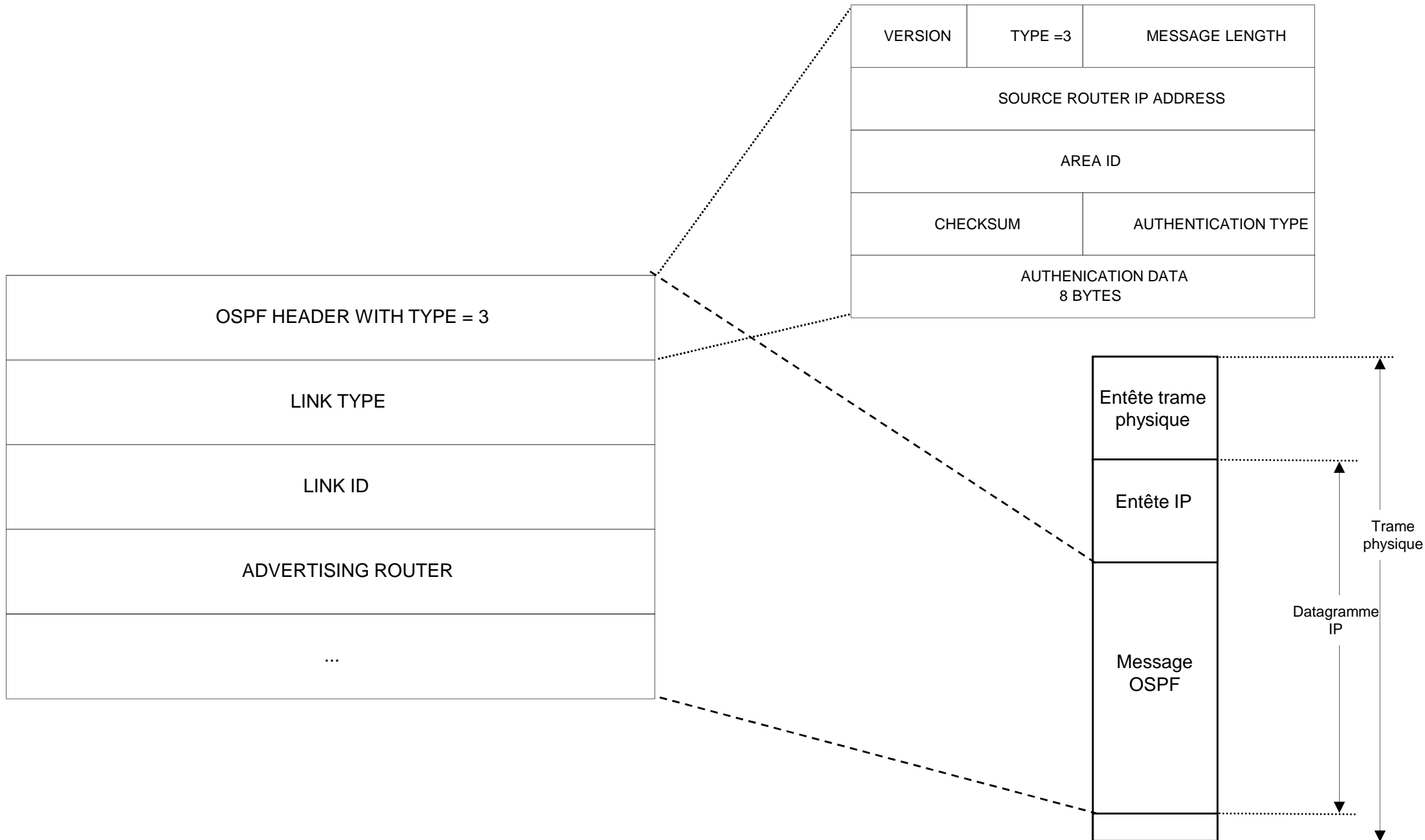
OSPF : format des messages de type 1

- **Network Mask**
 - Masque de sous-réseau du réseau concerné
- **DEAD TIMER**
 - après expiration de ce délai, le voisin est considéré comme inopérant
- **HELLO INTER**
 - Intervalle de temps entre deux messages Hello
- **GWAY PRIO**
 - Numéro interne fixant la priorité du routeur
 - Utilisé pour l'élection du BDR
- **DR et BDR**
 - Identification du DR et du BDR selon l'émetteur
- **NEIGHBOR IP ADDRESS**
 - Adresses IP de tous les voisins qui ont récemment émis des messages Hello

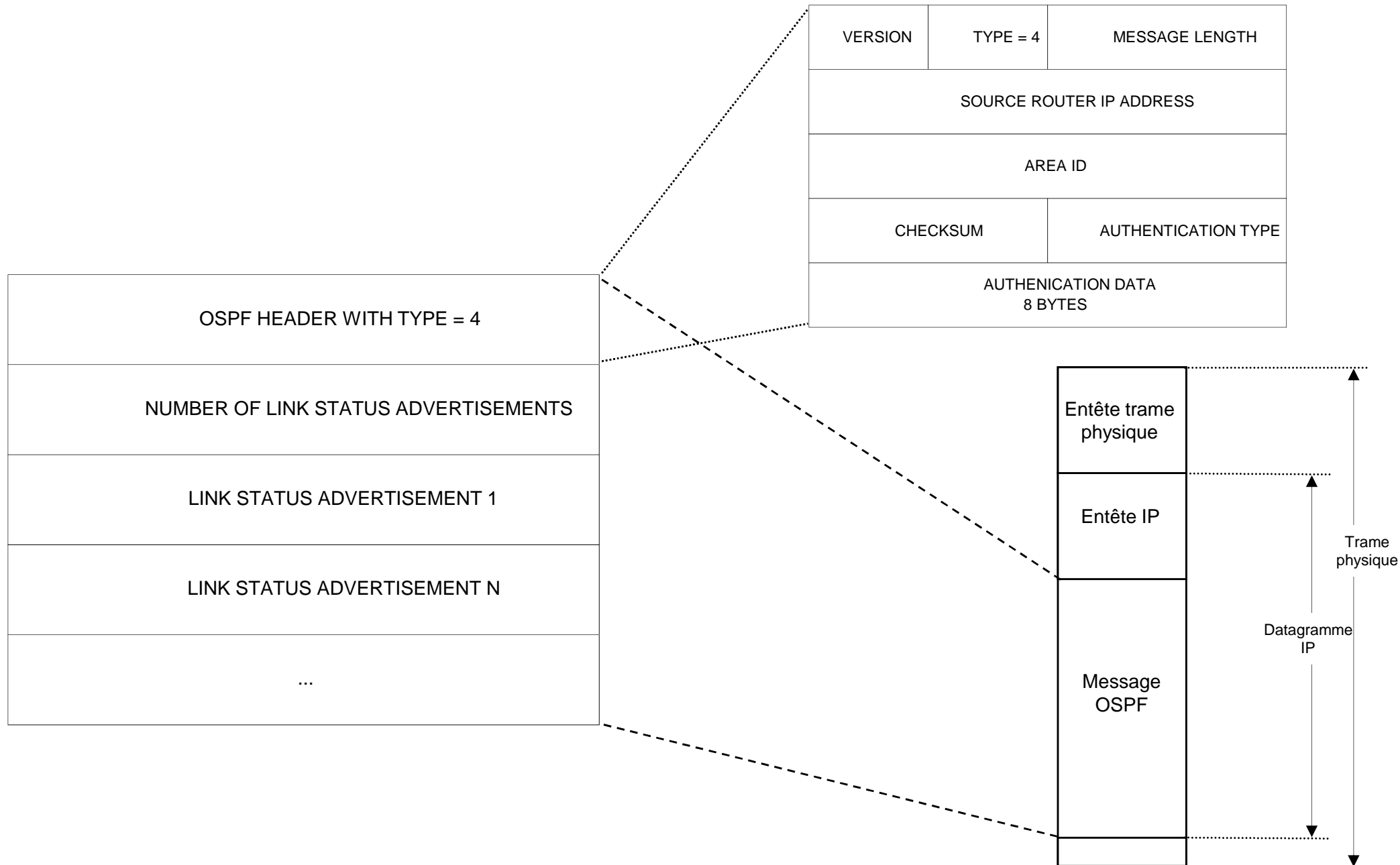
OSPF : format des messages de type 2



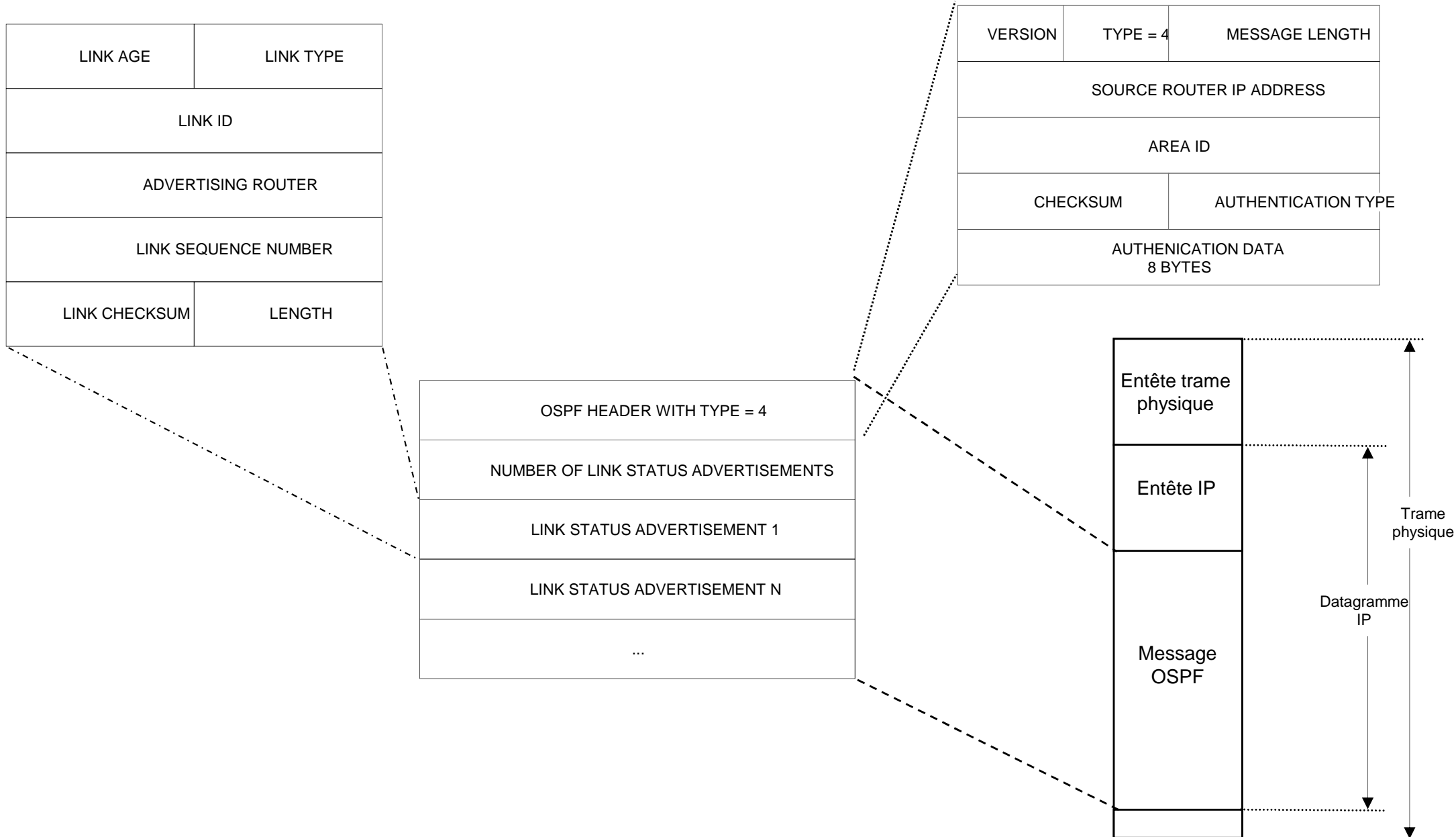
OSPF : format des messages de type 3



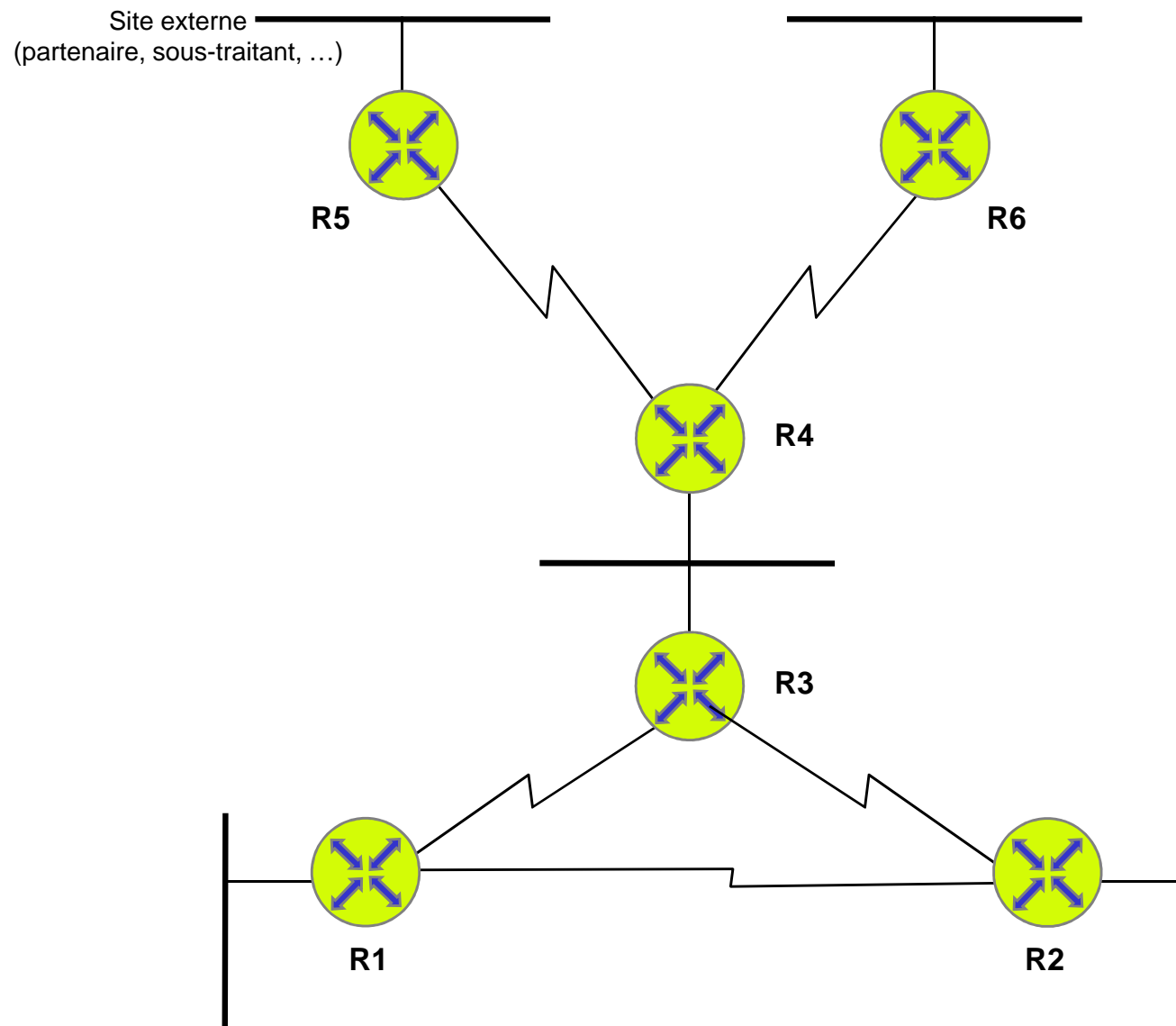
OSPF : format des messages de type 4



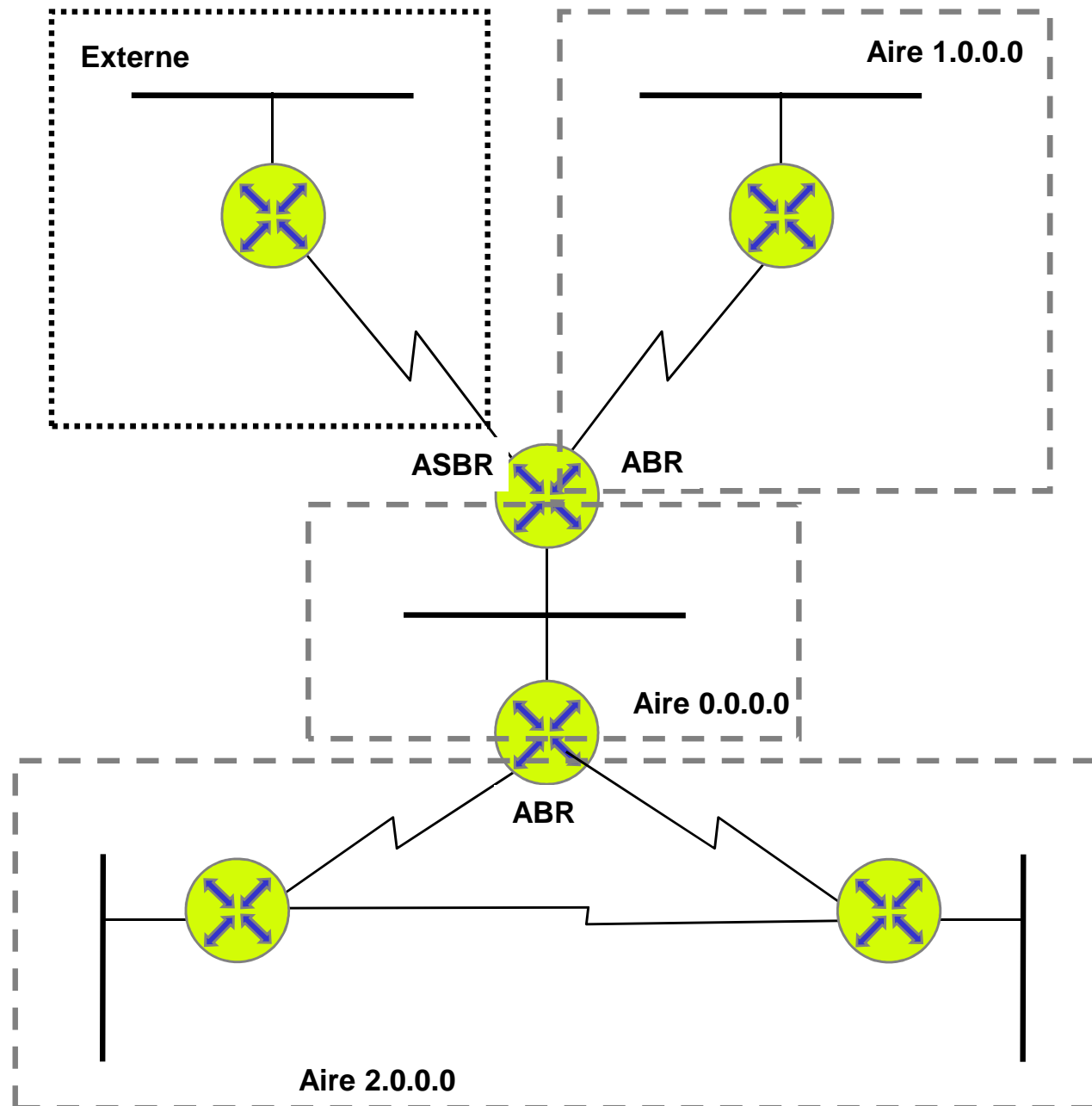
OSPF : format des LSAs



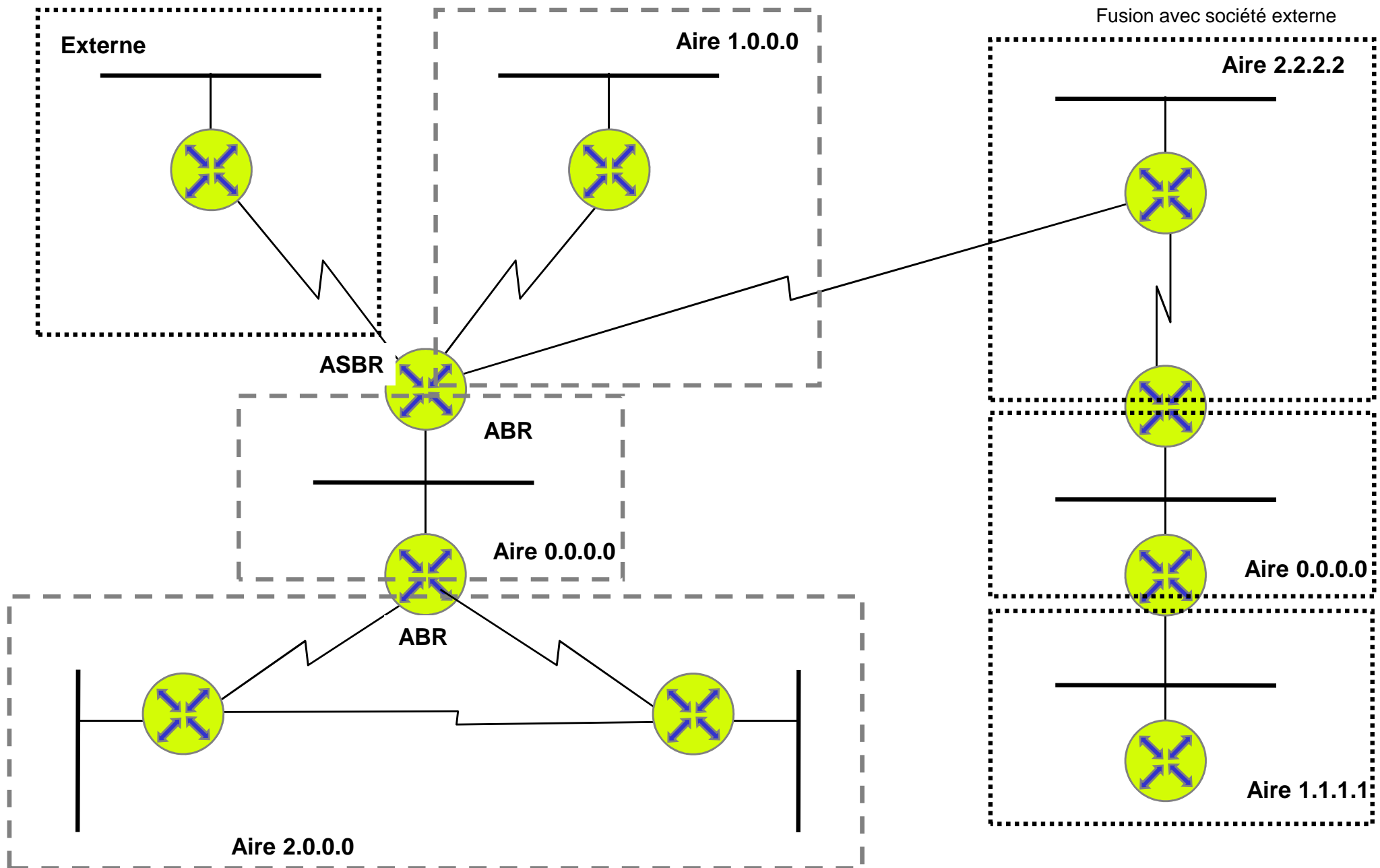
OSPF : étude de cas



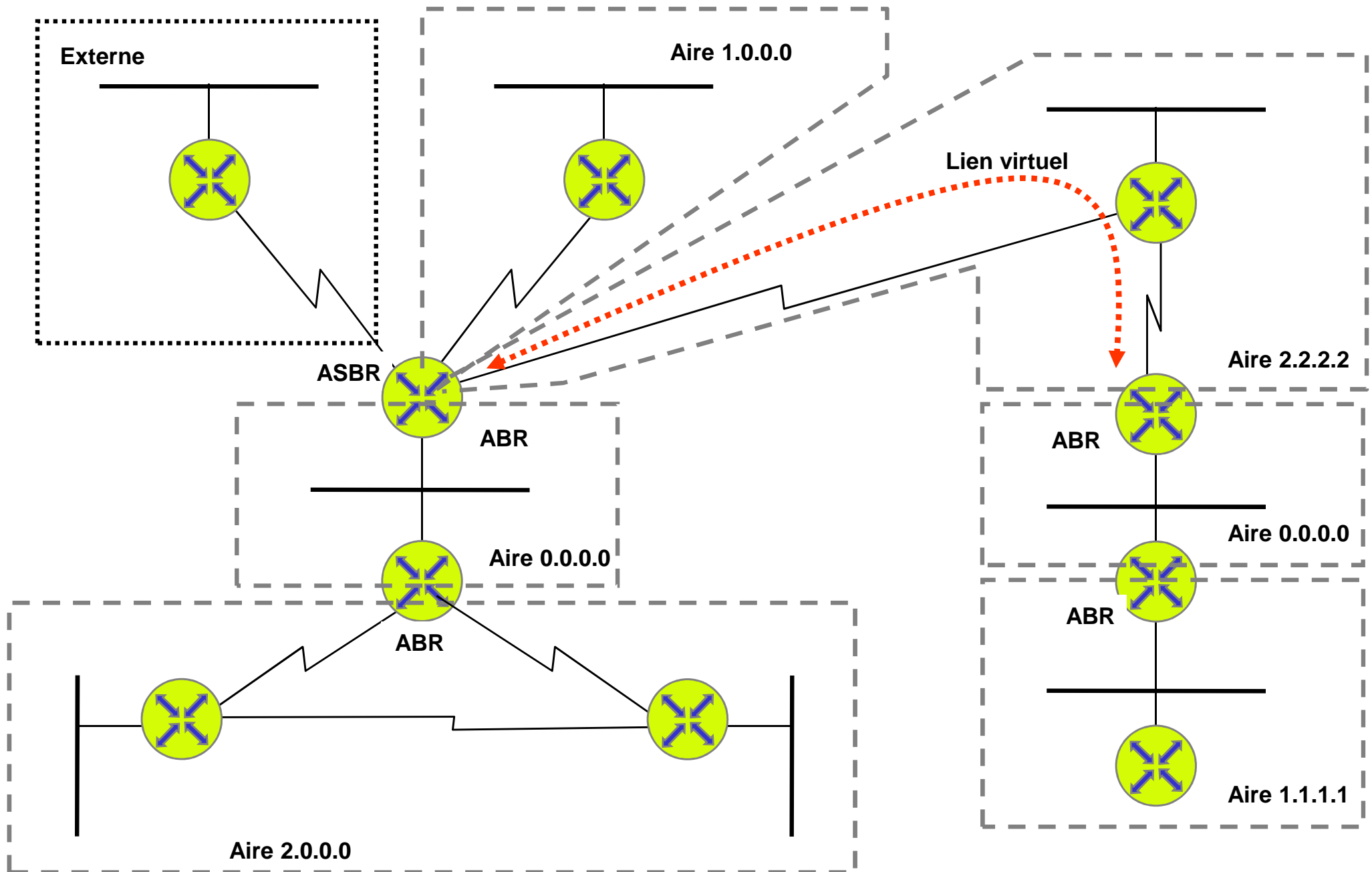
OSPF : étude de cas



OSPF : étude de cas



OSPF : étude de cas



OSPF : RFC de référence (1/2)

- RFC 1131 **OSPF specification**
- RFC 1245 **OSPF Protocol Analysis**
- RFC 1246 **Experience with the OSPF Protocol**
- RFC 1247 **OSPF Version 2**
- RFC 1248 **OSPF Version 2 Management Information Base**
- RFC 1252 **OSPF Version 2 Management Information Base**
- RFC 1253 **OSPF Version 2 Management Information Base**
- RFC 1364 **BGP OSPF Interaction**
- RFC 1370 **Applicability Statement for OSPF**
- RFC 1403 **BGP OSPF Interaction**
- RFC 1583 **OSPF Version 2**
- RFC 1584 **Multicast extensions to OPSF**
- RFC 1585 **MOSPF: analysis and experience**
- RFC 1586 **Guidelines for Running OSPF Over Frame Relay Networks**

OSPF : RFC de référence (2/2)

- RFC 1587 The OSPF NSSA Option
- RFC 1745 BGP4/IDRP for IP - OSPF interaction
- RFC 1765 OSPF Database Overflow
- RFC 1793 Extending OSPF to Support Demand Circuits
- RFC 1850 OSPF Version 2 Management Information Base
- RFC 2154 OSPF with digital signatures
- RFC 2178 OSPF Version 2
- RFC 2328 OSPF Version 2
- RFC 2329 OSPF Standardization Report
- RFC 2370 The OSPF Opaque LSA Option
- RFC 2740 OSPF for IPv6
- RFC 2676 QoS Routing Mechanisms and OSPF Extensions
- RFC 2844 OSPF over ATM and Proxy PAR
- RFC 3137 OSPF Stub Router Advertisement

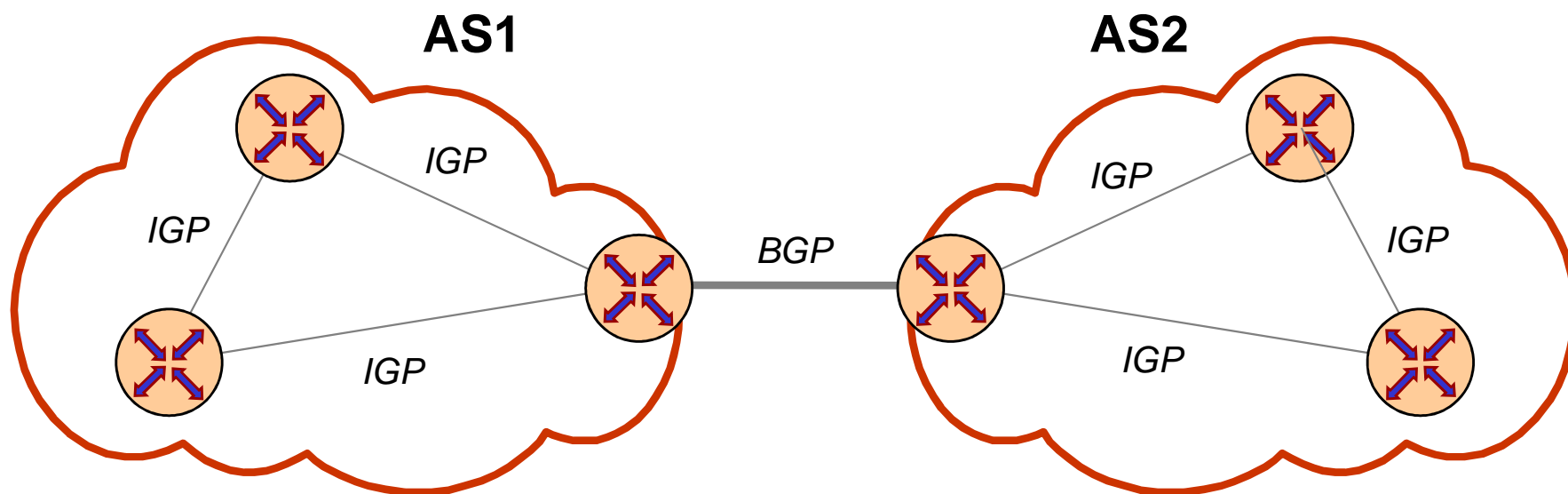
OSPF : comparaison RIPv1, RIPv2 et OSPF

| | RIPv1 | RIPv2 | OSPF |
|---------------------------|-----------------|-----------------|------------|
| Algorithme | Distance Vector | Distance Vector | Link State |
| Mise à jour | Broadcast | Multicast | Multicast |
| Authentification | Non | Oui | Oui |
| Multi Chemins | Non | Non | Oui |
| Convergence | Lente | Rapide | Rapide |
| Support des sous-réseaux | Non | Oui | Oui |
| Hiérarchisation | Non | Non | Oui |
| Facilité d'administration | Facile | Facile | Complexe |

BGP 4 : généralités

- **Border Gateway Protocol**
- **RFC 1771**
- **Utilise les services de TCP (port 179)**
- **Protocole de routage de type Path Vector**
- **Successeur de EGP**
- **Les routeurs frontières à l'intérieur d'un AS doivent s'échanger les informations qu'ils ont apprises**
- **iBGP**
 - BGP interne
 - Echange dans un AS
 - Maillage logique entre les routeurs
- **eBGP**
 - BGP externe
 - échange entre AS
- **Un message iBGP reçu par un routeur frontière n'est transmis qu'à un routeur d'un autre AS**
- **BGP standard de facto des protocoles EGP utilisés sur Internet**
- **Gère le CIDR**

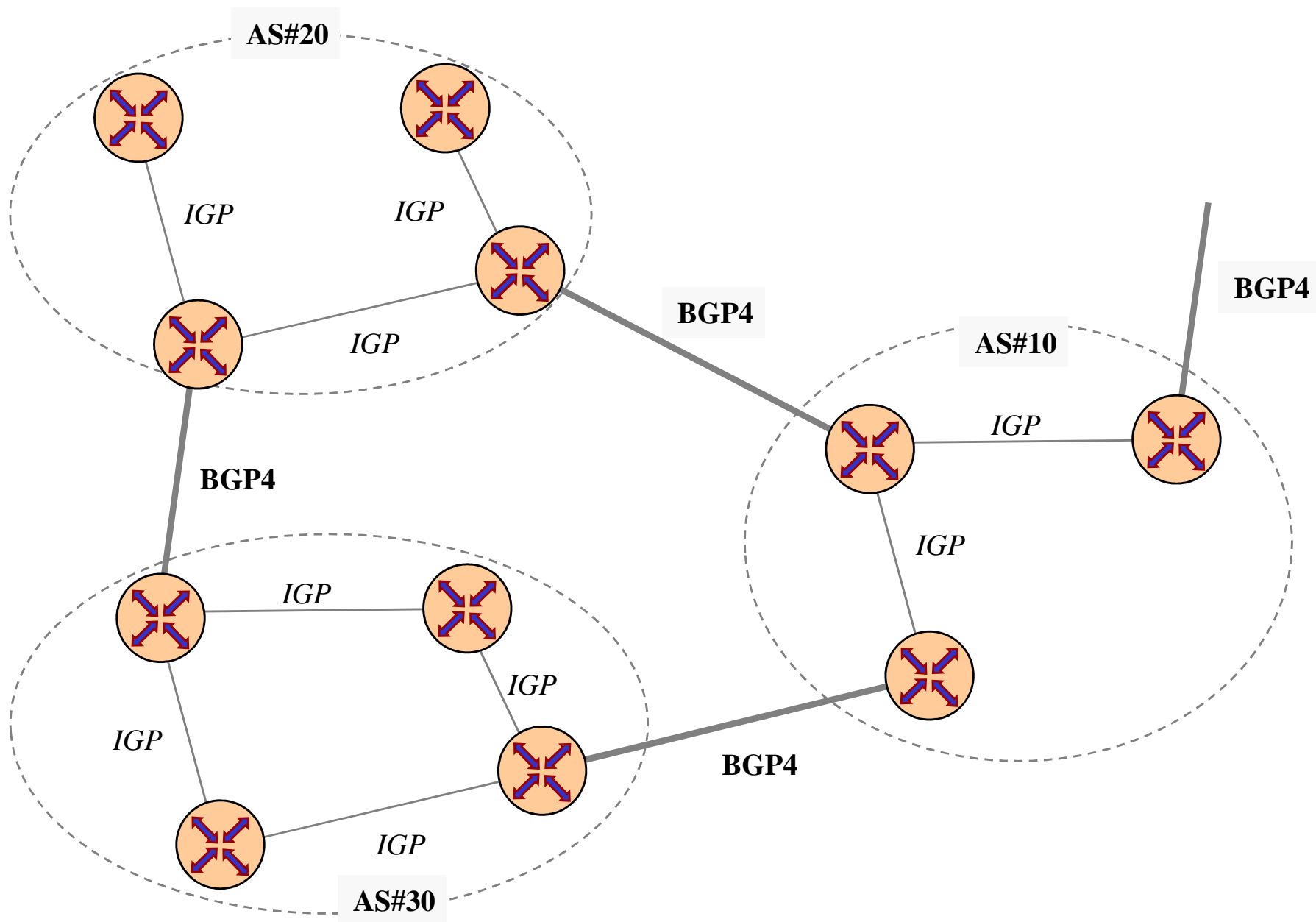
BGP 4 : généralités



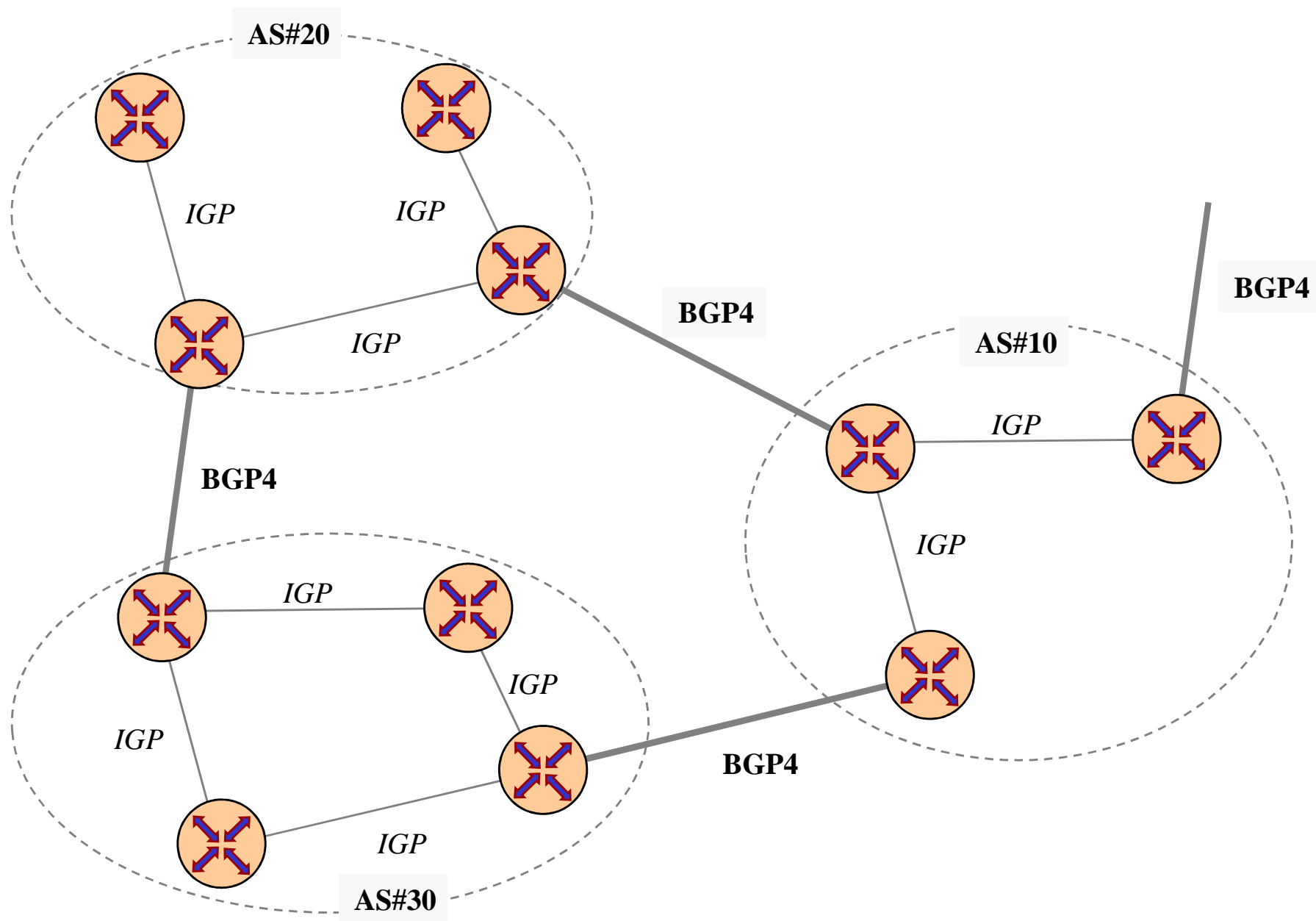
BGP 4 : généralités

- **Mode connecté**
 - Utilise les service de TCP
 - Analogue à des échanges point à point
- **Avantages**
 - Fiable
 - Simplifie le protocole
 - Permet une mise à jour incrémentale
- **Inconvénients**
 - Sensible aux congestions
 - Nécessite un message sonde
 - Moins précis que UDP pour l'état d'un lien

BGP4 : architecture de routage



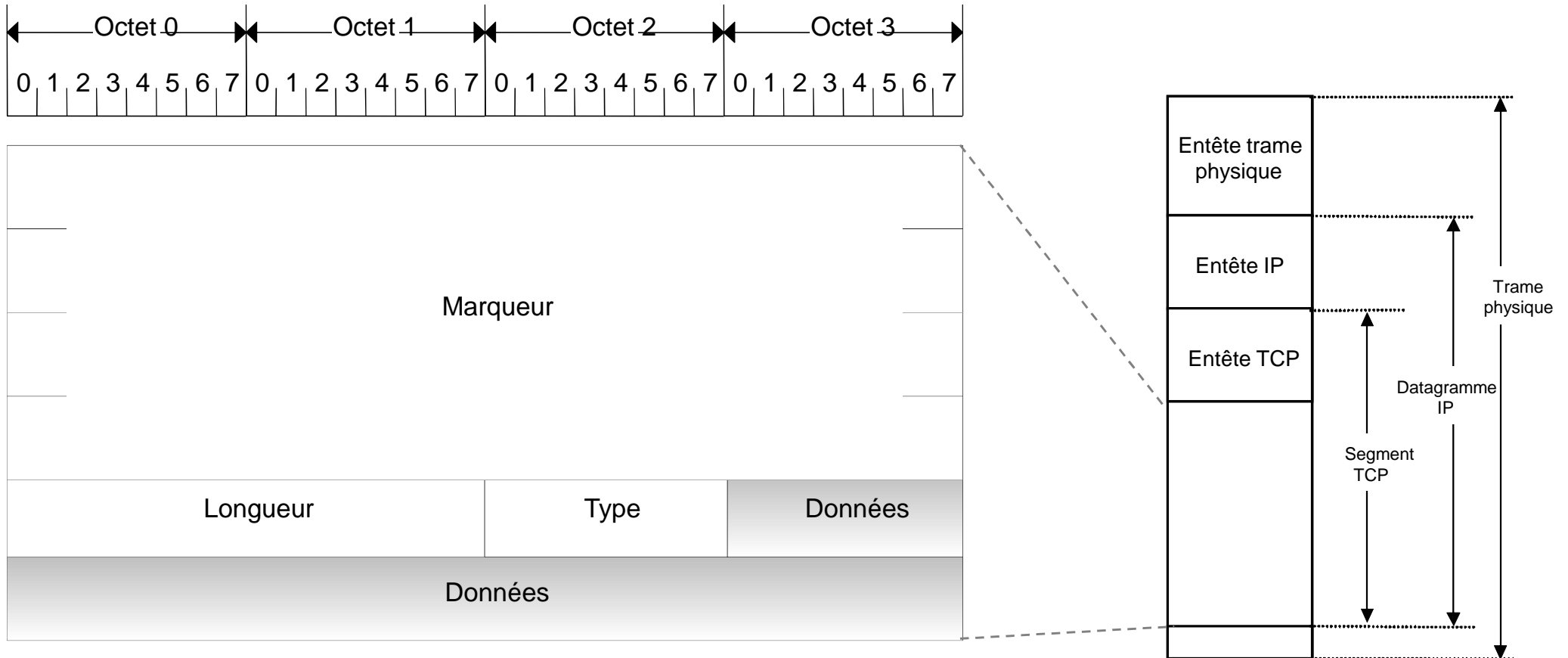
BGP4 : architecture de routage



BGP 4 : format des messages

- **BGP 4 utilisent principalement 4 types de messages**
 - OPEN
 - UPDATE
 - NOTIFICATION
 - KEEPALIVE
- **La taille minimum des messages est de 19 octets**
- **La taille maximum des messages est de 4096 octets**

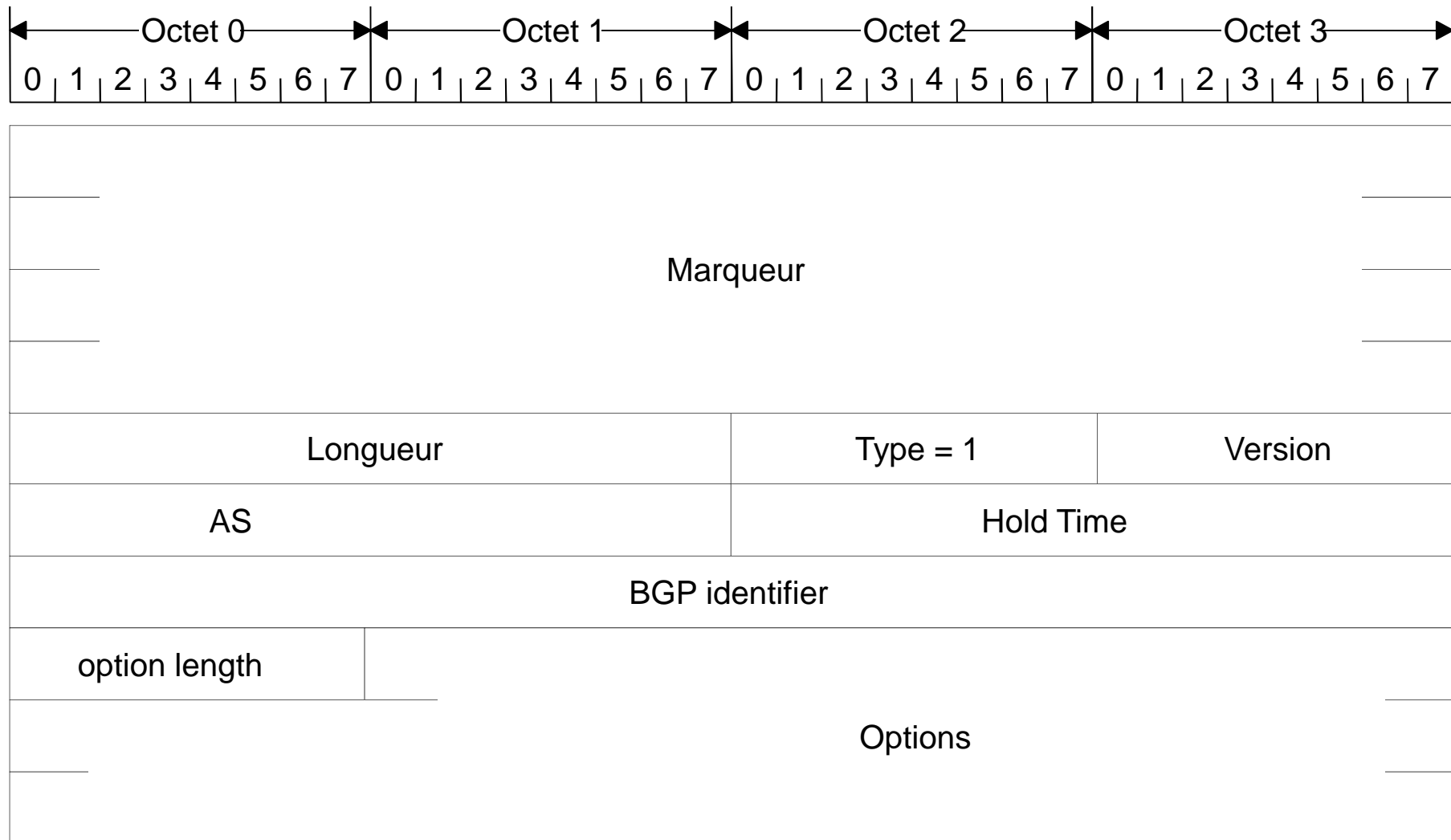
BGP 4 : format des messages



BGP 4 : format des messages

- **Système de délimitation pour distinguer les messages**
- **Champ marqueur pour resynchroniser, authentifier**
- **Champ longueur entre 19 (entête seul) et 4096 octets**
- **Champ type :**
 - 1 pour une ouverture
 - 2 pour une mise à jour
 - 3 pour une notification
 - 4 pour un sondage

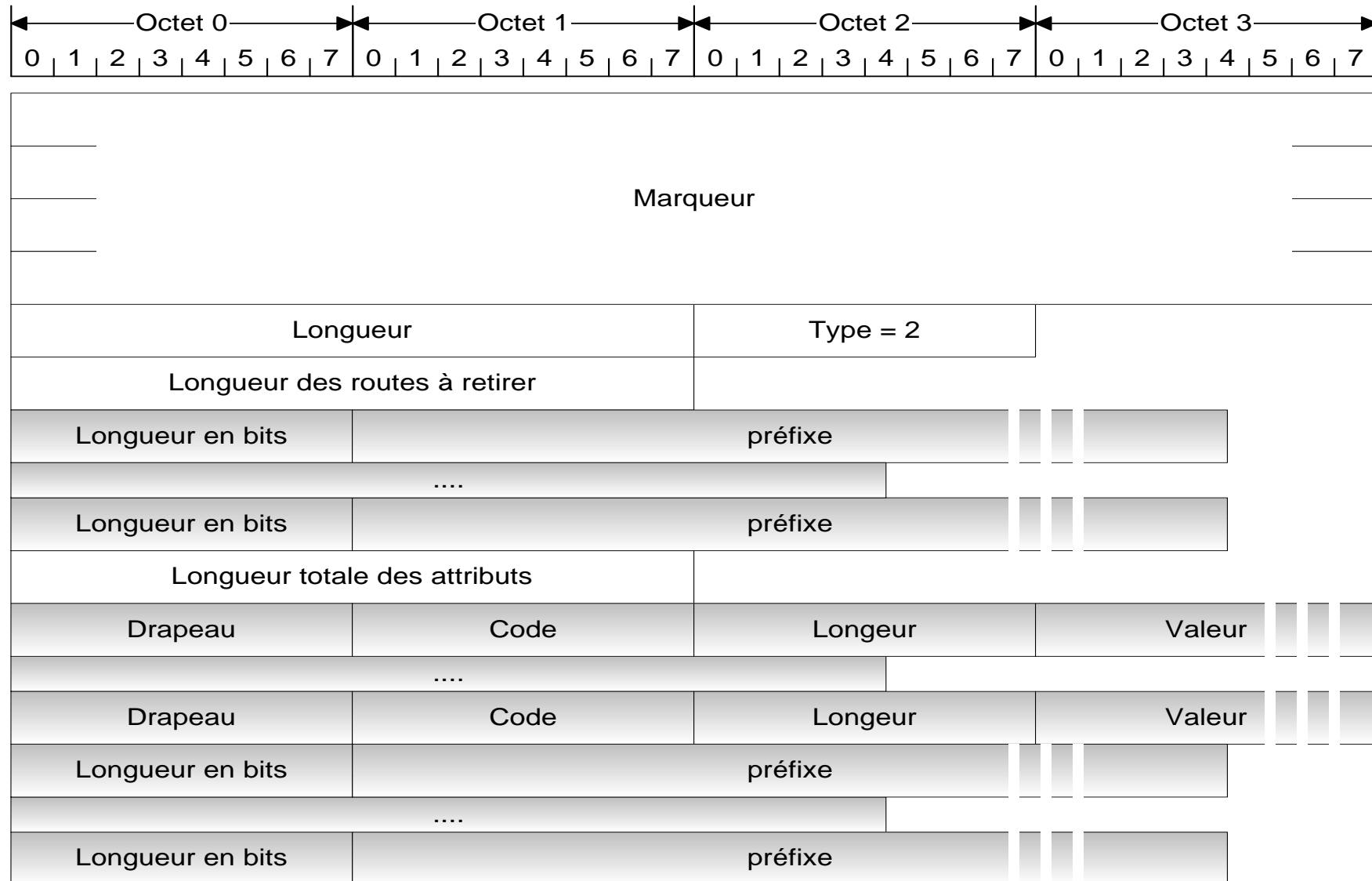
BGP 4 : les message OPEN (1/2)



BGP 4 : les message OPEN (2/2)

- **Envoyé après connexion TCP**
- **Négociation des paramètres de l'association**
- **Mon Système Autonome : champ indiquant le système autonome émetteur**
- **Temps de garde en seconde**
- **Identifiant BGP : adresse IP de l'émetteur (toujours la même)**
- **Champ Option pouvant contenir des paramètres d'authentification**

BGP 4 : Le message UPDATE (1/3)



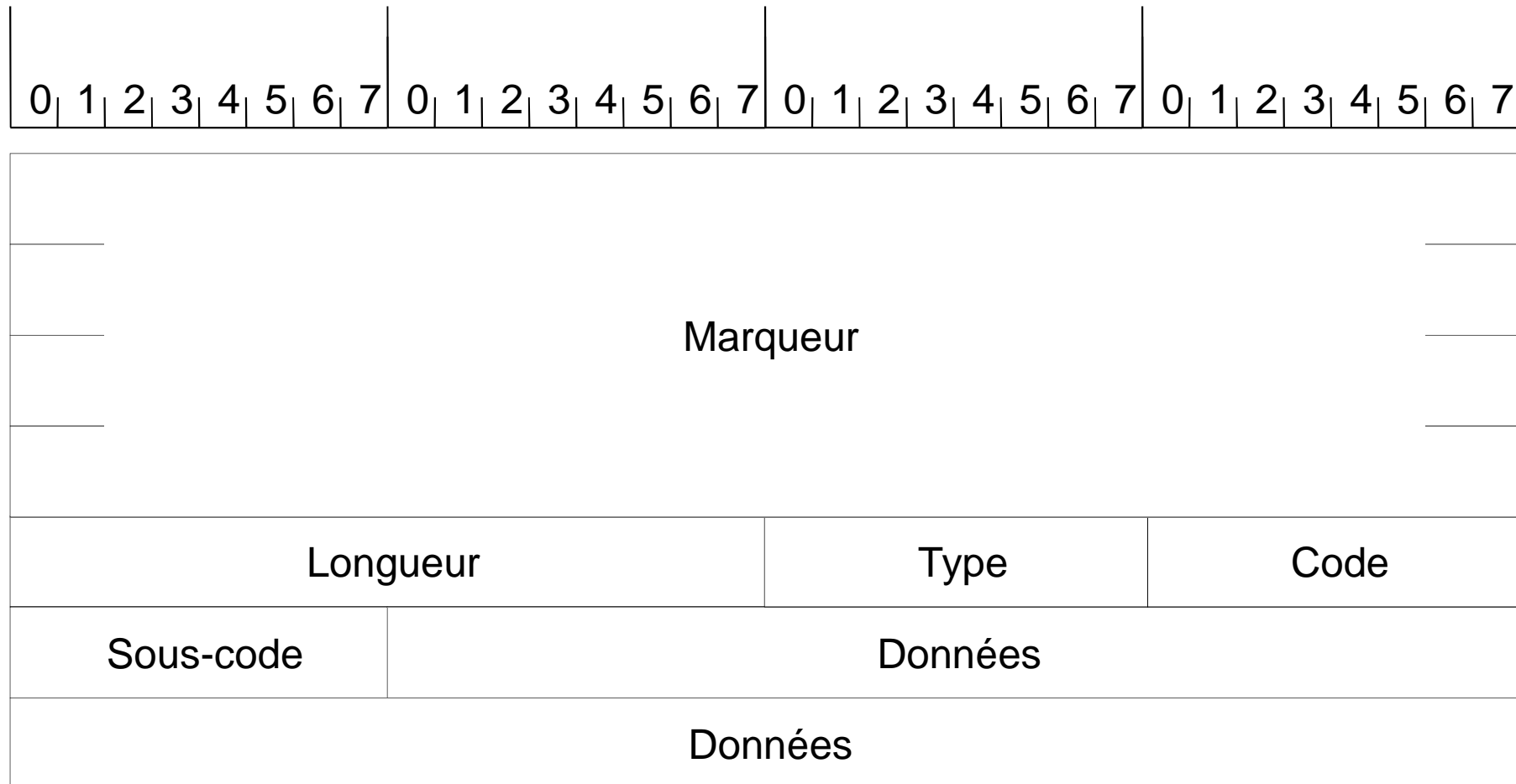
BGP 4 : Le message UPDATE (2/3)

- **Network Layer Reachability Information (NLRI)**
- **Le message est composé de trois parties :**
 - les routes à retirer
 - les attributs
 - les nouvelles routes
- **Drapeau pour classer les attributs suivant plusieurs critères :**
 - le bit O indique si l'attribut est optionnel
 - le bit T indique si l'attribut est transitif
 - le bit P indique si l'attribut est partiel
 - le bit EL indique si le champ longueur sera sur 2 octets au lieu d'un seul

BGP 4 : Le message UPDATE (3/3)

| Drapeau | Code | Attribut | RFC |
|----------------------------|------|----------------------|--------------|
| Bien connu Obligatoire | 1 | ORIGIN | 1771 |
| Bien connu Obligatoire | 2 | AS_PATH | |
| Bien connu Obligatoire | 3 | NEXT_HOP | |
| Optionnel Non transitif | 4 | MULTI_EXT_DISCR | |
| Bien connu Facultatif | 5 | LOCAL_PREF | |
| Bien connu Facultatif | 6 | ATOMIC_AGGREGATE | |
| Optionnel Transitif | 7 | AGGREGATOR | |
| Optionnel Transitif | 8 | COMMUNITY | 1997 |
| Optionnel Non transitif | 9 | ORIGINATOR_ID | 1966 |
| Optionnel Non transitif | 10 | CLUSTER_LIST | 1966 |
| | 11 | DPA | expérimental |
| Optionnel Non transitif | 12 | ADVERTISER | 1863 |
| Optionnel Non transitif | 13 | RCID_PATH/CLUSTER_ID | 1863 |

BGP 4 : les messages NOTIFICATION (1/3)



BGP 4 : les messages NOTIFICATION (2/3)

- **Envoyé lors de messages erronés ou lorsque le temps de garde expire**
- **Message de notification est envoyé au partenaire**
- **Connexion fermée après son émission**
- **Code et sous code afin de définir la raison de l'envoi du message de notification**

BGP 4 : les messages NOTIFICATION (3/3)

| Code | | Sous-code | |
|------|------------------------|-----------|---------------------------------|
| 1 | Erreur dans l'en-tête | 1 | Connexion non synchronisée |
| | | 2 | Longueur de message erronée |
| | | 3 | Type de message erroné |
| 2 | Erreur d'ouverture | 1 | Numéro de version non supporté |
| | | 2 | Mauvais numéro d'AS |
| | | 3 | Identifiant BGP incorrect |
| | | 4 | Paramètre option non supporté |
| | | 5 | Echec d'authentification |
| | | 6 | Temps de garde inacceptable |
| 3 | Erreur de mise à jour | 1 | Liste d'attribut incorrect |
| | | 2 | Attribut bien connu non reconnu |
| | | 3 | Attribut bien connu manquant |
| | | 4 | Erreur dans le drapeau |
| | | 5 | Longueur attribut erroné |
| | | 6 | Attribut ORIGIN invalide |
| | | 7 | Boucle dans le chemin d'AS |
| | | 8 | Attribut NEXT_HOP invalide |
| | | 9 | Erreur dans attribut optionnel |
| | | 10 | Champ réseau invalide |
| | | 11 | AS_PATH mal formé |
| 4 | Temps de garde écoulé | | |
| 5 | Erreur dans l'automate | | |
| 6 | Arrêt | | |

BGP 4 : les messages KEEPALIVE

- Envoyé pour que le temps de garde n'expire pas
- Si les mises à jour ne sont pas assez fréquentes alors on envoie un message sonde
- En général 1/3 du temps de garde
- En-tête avec champ type à 4

| | | | | | |
|--------|----|----|------|----|----|
| 0 | 15 | 16 | 23 | 24 | 31 |
| Marker | | | | | |
| Length | | | Type | | |

BGP 4 : Le regroupement en confédération

- **RFC 1965**
- **Réduction du maillage**
- **Une confédération**
 - ensemble de sous AS privés dans un AS
- **Dans un sous AS privé, les routeurs dialoguent en iBGP**
- **Les sous AS dialoguent entre eux à l'aide d'eBGP**
- **Lorsqu'un message sort de la confédération, les numéros d'AS privés sont retirés de la liste**

BGP 4 : Les réflecteurs de route

- **RFC 1966**
- **L'astuce est de lever la restriction de transfert d'iBGP à certains routeurs et d'ajouter quelques contrôles**
- **Les réflecteurs de routes sont chargés de transférer l'information aux routeurs clients**
- **Le réflecteur et ses clients forment un cluster**
- **CLUSTER_ID identifie le cluster**
- **ORIGINATOR_ID identifie l'initiateur du message**
- **Ces attributs permettent d'éviter les boucles**

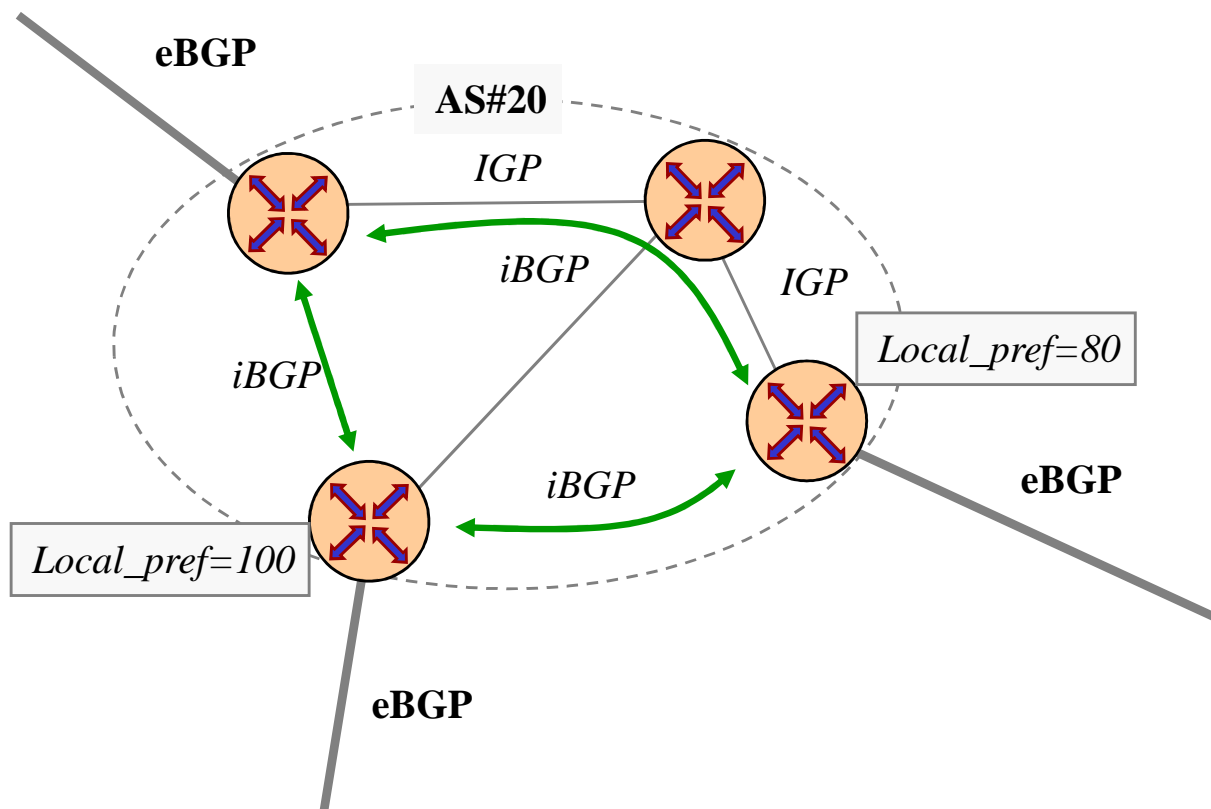
BGP 4 : les attributs

- Transitif ou non transitif
- 13 attributs
- Obligatoires ou optionnels
 - 1 ORIGIN
 - 2 AS_PATH
 - 3 NEXT_HOP
 - 4 MULTI_EXT_DISCR
 - 5 LOCAL_PREF
 - 6 ATOMIC_AGGREGATE
 - 7 AGGREGATOR
 - 8 COMMUNITY
 - 9 ORIGINATOR_ID
 - 10 CLUSTER_LIST
 - 11 DPA
 - 12 ADVERTISER
 - 13 RCID_PATH/CLUSTER_ID

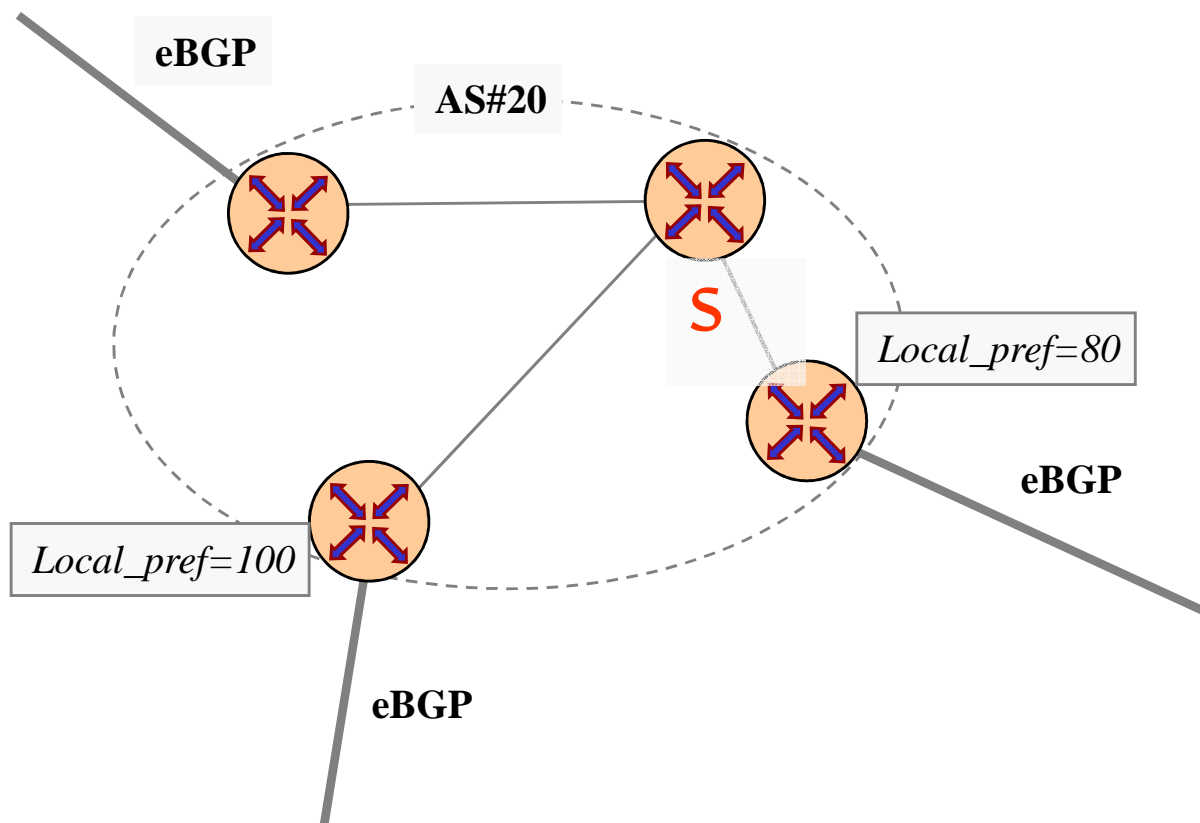
BGP 4 : attribut LOCAL_PREF

- **Permet de choisir entre plusieurs liens externes le point de sortie de l'AS**
- **Transmis dans les messages UPDATE**
 - Code = 5
 - Entier sur 4 octets
- **Attribut non transmis en dehors de l'AS**
 - Transmission de l'attribut en iBGP seulement
 - Le point de sortie correspond à la valeur de LOCAL_PREF la plus élevée
- **Activation optionnelle**
- **Application**
 - Choix d'une liaison plus performante
 - Choix d'un contrat opérateur plus intéressant selon le type de trafic
 - Choix d'un lien de secours

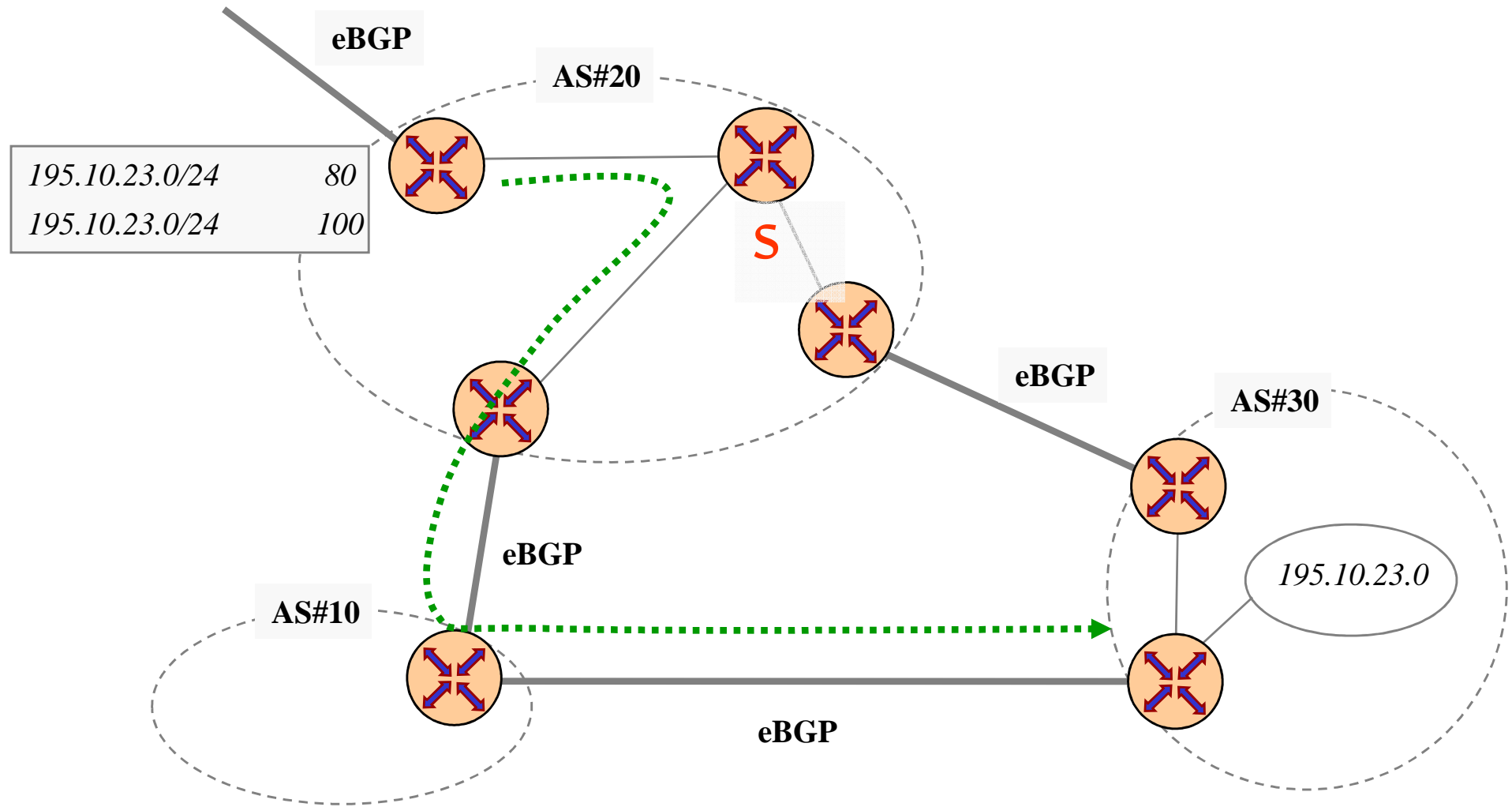
BGP4 : attribut LOCAL_PREF



BGP4 : attribut LOCAL_PREF



BGP4 : attribut LOCAL_PREF



BGP4 : Les règles de sélection des routes

- **Poids mis sur une route par l'administrateur selon provenance ou AS traversé**
- **Le plus grand degré de préférence LOCAL_PREF**
- **Le chemin d'AS le plus court**
- **ORIGIN : l'annonce d'un IGP mieux que d'un EGP mieux que d'un inconnu**
- **MED avec la métrique la plus faible quand dans le même AS ou métrique IGP vers le routeur du NEXT_HOP**
- **Annonce Internet mieux qu'eBGP mieux qu'iBGP**
- **Le routeur ayant le plus petit identifiant**

BGP 4 : Inconvénients

- **Annonces de routes élevées**
- **Demande beaucoup de mémoire et de ressource CPU**
- **Constat d'instabilité des routeurs**
 - Phénomènes d'oscillation (annonce et retrait de routes)
- **Remèdes :**
 - attente avant mise à jour
 - système de pénalité

BGP 4 : Les successeurs de BGP-4

- **Multiprotocol Border Gateway Protocol (MBGP) (RFC 2283)**
 - Extension à d'autres protocoles (IPX, IPv6...)
 - 14 : MP_REACH_NLRI, 15 : MP_UNREACH_NLRI
 - Permet également de faire du multipoint
- **InterDomain Routing Protocol (IDRP)**
 - Pour l'environnement IPv6
 - Vecteur distance mais chemin de bout en bout
 - Définit une politique de routage entre AS
 - Paquets IP spécifiques avec numéro 45 dans l'entête