




Architecture des Réseaux - Routage IP

M. Abderrezak RACHEDI


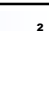
Université Paris-Est Marne-la-Vallée (UPEMLV)
École Supérieure d'Ingénieurs Paris-Est Marne-la-Vallée (ESIPE-MLV)

Filière: 2^{ème} année Electronique et Informatique Systèmes Communicants (EISC)

Année 2013/2014

© 2011 - Développé par M. Abderrezak RACHEDI - droits réservés - reproduction interdite

1






Modalités

- 5 séances de 2 heures de Cours/TD
- 4 séances de 4 heures de TP
 - Un compte rendu pour chaque TP
- Note finale = $0.60 \cdot DS + 0.40 \cdot \text{Moy}(CCi)$

© 2011 - Développé par M. Abderrezak RACHEDI - Droits réservés - reproduction interdite

2

Bibliographie (Pour savoir plus ...)

- Livres :
 - Titre « Les Réseaux », Auteur : Guy Pujolle, Eyrolles, 2008
 - Titre « Réseaux », Auteur: Andrew Tanenbaum, Pearson Education, 2004
- Autres :
 - Cisco (Ressources en line)
 - ...


© 2011 - Développé par M. Abderrezak RACHEDI - Droits réservés - reproduction interdite

3

4

Plan

- Introduction au routage IP
- Routage : principe et table de routage
- Routage statique
- Routage dynamique
- Algorithme de routage
 - Vecteur de distance : « Distance-Vector »
 - Etat de lien : « Link-State »
- Protocole RIP (Routing Information Protocol)
- Protocole OSPF (Open Shortest Path First)
- Protocole EIGRP (Enhanced IGRP)
- Routage externe
- Protocole BGP (Border Gateway Protocol)


 © 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite

5

Introduction au routage IP

Qu'est ce que le routage ?
Processus qui assure l'acheminement d'un élément (courrier, trains, paquets IP, ...) vers la destination


- Le processus du routage doit connaître :
 - La destination,
 - Les itinéraires possibles pour atteindre la destination,
 - Le(s) meilleur(s) itinéraire(s) pour atteindre la destination,
 - Un moyen d'actualiser les itinéraires.

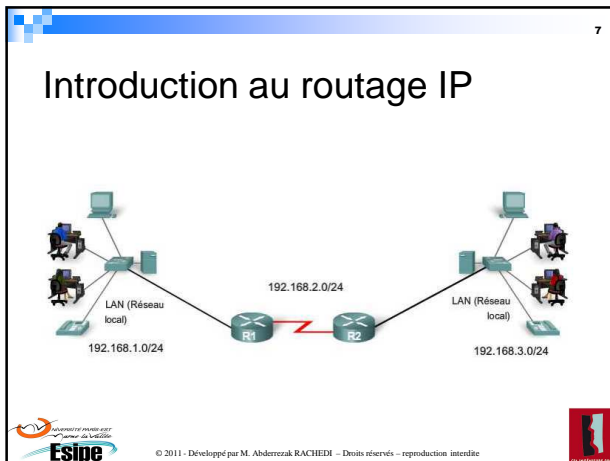
 © 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite

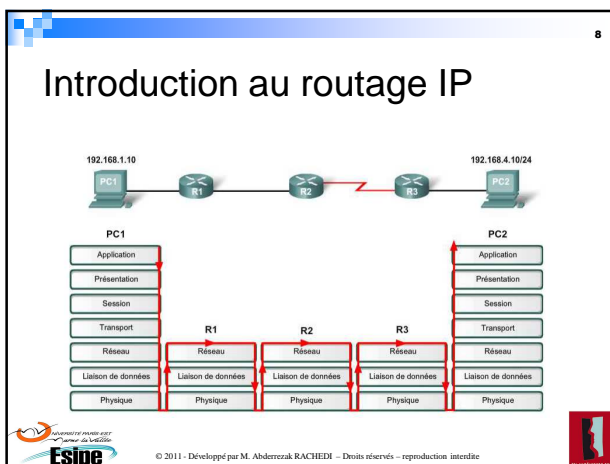
6

Introduction au routage IP

- Un équipement sur un réseau local
 - Peut atteindre directement les machines sur le même segment sans routage (ARP) => routage direct
 - Ne peut pas atteindre les équipements sur un autre réseau (ou sous-réseau) sans un intermédiaire
- Qui doit faire du routage sur un réseau ?
 - Équipement connecté à 2 réseaux ou sous-réseaux au moins => routage indirect
 - Station de travail avec 2 interfaces réseau au moins,
 - Routeur (CISCO, Juniper, ...)

 © 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite





9

Problèmes de routage

- La réception des paquets par le récepteur dans un ordre qui n'est pas forcément celui de l'envoi par l'émetteur
- La longueur de l'adresse, qui doit être suffisamment importante pour pouvoir représenter tous les récepteurs potentiels du réseau (IPv4: 4 octets – IPv6 16 octets)
- La taille de la table de routage
 - Si le réseau a beaucoup de clients, le nombre de lignes de la table de routage peut être très important, et de nombreux paquets de supervision sont nécessaires pour la maintenir à jour

➔ Pour un routage efficace, il faut essayer de limiter le nombre de lignes des tables de routage


Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

10

Principe du routage IP


- Routage IP basé uniquement sur l'adresse du destinataire
- Chaque équipement du réseau sait atteindre un équipement d'un autre réseau, s'il existe au moins un équipement de routage pour acheminer les paquets à l'extérieur du réseau local
- Les informations de routage sont mémorisées dans la table de routage des équipements (routeurs)
- Cette table doit être périodiquement mise à jour
 - Manuellement : routage STATIQUE
 - Automatiquement : routage DYNAMIQUE

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

11

Table de routage


- Accès:
 - Station Unix: « netstat -r[n] » (linux: ip route)
 - Routeur Cisco/zebra: « show ip route »
- Contenu minimal
 - Adresse du réseau auquel l'équipement est directement connecté
 - Route par default
- Mise-à-jour
 - Manuelle : routage statique
 - Unix : commande « route »
 - Cisco/Zebra : « ip route »
 - Automatique : routage dynamique
 - Protocole de routage : échanger d'information de routage
 - Mixte : routage statique et routage dynamique

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

12

Principe du routage IP


- Le routage s'effectue sur deux opérations :
 - La sélection du meilleur chemin (optimal)
 - Niveau 3 du modèle OSI
 - Métrique: nombre de sauts, bande passante, délai, etc
 - La commutation du paquet sur l'interface appropriée

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

13

La politique d'acheminement de paquet


- **Déterministe :**
 - Une seule route est possible par rapport à la destination.
 - Les tables de routage peuvent être fixées à la configuration du réseau
 - Les mises à jour périodiquement par le(s) centre(s) de gestion (gestion centralisée ou décentralisée)
- **Adaptative :**
 - Aucun chemin n'est prédéterminé, le chemin sera fixé au moment du routage en fonction de données sur l'état du réseau (charge, indisponibilité d'un nœud, ...)
- **Mixte :**
 - La politique est adaptative à l'établissement du chemin et déterministe durant le reste de la session
 - Cette technique est utilisée dans les réseaux en mode orienté connexion

 © 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite

14

Classification des algorithmes de routage


- **Routage statique**
 - Mise à jour manuelle de tous les équipements réseau
 - Pour les réseaux les plus stable
 - Complexe et risque d'erreur pour les grand réseaux (> 10 routeurs)
- **Routage dynamique**
 - Adaptation dynamique à l'évolution du réseau :
 - Changement de la topologie réseau
 - Changement des conditions réseau (paramètres de Qualité de Service)
 - Nécessite un protocole de routage

 © 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite

15

Routage IP statique

- Une route statique est basée sur :
 - L'adresse du réseau + Le masque de sous-réseau du réseau distant
 - L'adresse du routeur du tronçon suivant (next-hop) + l'interface de sortie
- **Route par défaut**
 - Facilite la circulation des données sur un réseau de grande taille
 - Pour atteindre une destination inconnue
 - Utilisée si le prochain saut ne figure pas explicitement dans la table de routage

 © 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite

16

Routage IP statique

```

R1#show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP
       O - OSPF, EX - OSPF external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, ia - IS-IS inter area
       * - candidate default, U - per-user static route, o - ODR
       P - periodic downloaded static route

Gateway of last resort is not set

C 192.168.1.0/24 is directly connected, FastEthernet0/0
C 192.168.2.0/24 is directly connected, Serial0/0/0
S 192.168.3.0/24 [1/0] via 192.168.2.2
  
```

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

17

Routage IP statique - Exemple

Table de routage

Réseau	Masque	Moyen de l'atteindre
192.168.2.0	255.255.255.0	eth0
100.0.0.0	255.0.0.0	eth1
101.0.0.0	255.0.0.0	eth2
192.168.1.0	255.255.255.0	100.0.0.1
192.168.3.0	255.255.255.0	101.0.0.2

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

18

Routage IP statique – Exercice

■ Donner la table de routage du routeur R

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

19



Routage IP statique

Problèmes du routage statique

- Mise à jour **manuelle** de tous les équipements du réseau
- Une station ne peut atteindre que les réseaux qu'on lui indique par la **commande route**
- Boucles de routage
- Routages asymétriques

■ **Recommandations générales**

- Stations, Routeurs d'extrémité => Routage statique
- Routeurs => **Routage dynamique**



 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite 

20

Routage IP statique

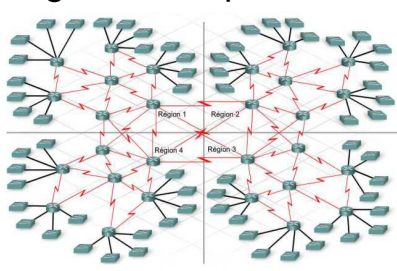
Avantages d'un routage statique

- Sécurité par masquage de certaines parties d'un inter-réseau
- Moins de surcharge par rapport au routage dynamique



 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite 

21

Routage IP statique




■ Comment assurer la maintenance du routage statique avec une telle architecture réseau?

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite 

22

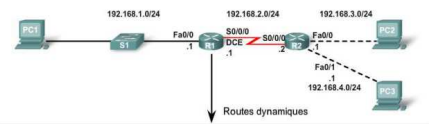
Routage IP dynamique

- **Plusieurs routes possibles** pour rejoindre une destination
=> l'usage d'un protocole de routage dynamique
- Une route statique privilégiée (une seule route) et ignore les autres
- Existence de **plusieurs routes** est une nécessité pour assurer la redondance du service, voire même **l'équilibrage du trafic** sur plusieurs liens

 © 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite


23

Routage IP dynamique



```


R1#show ip route
Codes: C - connected, S - static, I - IGRP, E - EIGRP, M - mobile, B - BGP
        D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2, S - EIGP
        I - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, Ia - IS-IS inter
            area
        * - candidate default, U - per-user static route, o - ODR
        P - periodic downloaded static route
Gateway of last resort is not set
C 192.168.1.0/24 is directly connected, FastEthernet0/0
C 192.168.2.0/24 is directly connected, Serial0/0/0
O 192.168.3.0/24 [1/0] via 192.168.2.2
O 192.168.4.0/24 [120/1] via 192.168.2.2, Serial0/0/0
  
```

 © 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite

24

Routage IP dynamique


- **Objectifs des protocoles de routage :**
 - ☐ Sélectionner les meilleures routes
 - ☐ Éliminer les boucles de routage
 - ☐ Éviter la configuration manuelle
 - ☐ Gérer dynamiquement le changement des routes
 - ☐ Maintenir la cohérence des informations associées aux routes
 - ☐ Limiter la taille des tables de routage pour réduire le temps de traitement
 - ☐ Réduire la consommation de la bande passante et du CPU
 - Hiérarchie des échanges d'informations
 - ☐ Notion du système autonome
 - ☐ Classification des protocoles : IGP et EGP

 © 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite

25

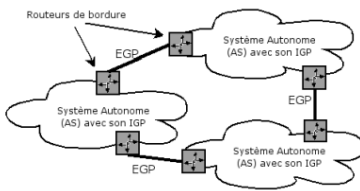
Routage IP dynamique

- **Système autonome (AS) : Domaine de routage**
 - Internet est le résultat de l'interconnection des AS
 - Domaine est géré par une seule autorité
 - Architecture de chaque AS est indépendante des autres domaines
 - Chaque AS possède un numéro unique (2 octets) attribué par :
 - RIPE NCC (Réseaux IP Européens - Network Coordination Centre)
 - ARIN (American Registry for Internet Numbers), etc
 - Ex: 1717 Renater, 1307 Jussieu, 3 MIT, 11 Harvard, 32 Stanford


 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

26

Routage IP dynamique

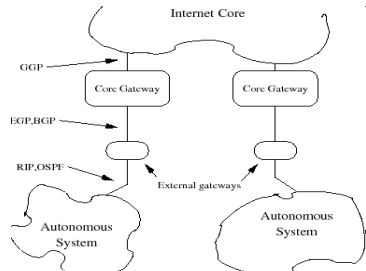


IGP: Interior Gateway Protocol
EGP: Exterior Gateway Protocol

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite


27

Routage IP dynamique



GGP: Gateway to Gateway Protocol
EGP: Exterior Gateway Protocol
BGP: Border Gateway Protocol
RIP: Routing Information Protocol
OSPF: Open Shortest Path First


Un AS, le monde extérieur et intérieur

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

28

Routage IP dynamique


- Tous les routeurs de même AS :
 - sont interconnectés entre eux
 - échangent leur tables de routage
- Deux familles de protocoles de routage
 - Protocoles entre routeurs d'un AS (Intra AS)
 - IGP : Interior Gateway Protocol (RIP, OSPF, etc)
 - Protocoles de routage entre AS (Inter AS)
 - EGP : Exterior Gateway Protocol (EGP, BGP)

 © 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite

29

Routage IP dynamique


- Types de routes
 - Routes statiques
 - configurées manuellement
 - Routes "connectées"
 - créées automatiquement quand une interface réseau est "active"
 - Routes dites "intérieures"
 - routes au sein d'un AS
 - routes apprises par un IGP
 - Routes dites "extérieures"
 - routes n'appartenant pas à l'AS local
 - apprises par un EGP

 © 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite

30

Routage IP dynamique


- Politique de routage
 - Définition de ce que vous acceptez ou envoyez aux autres
 - connexion économique, partage de charge, etc...
 - Accepter des routes de certains FAI et pas d'autres
 - Envoyer des routes à certains FAI et pas à d'autres
 - Préférer les routes d'un FAI plutôt que d'un autre

 © 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite

31

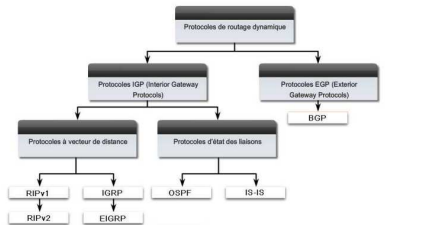
Routage IP dynamique

- Deux grandes classes de protocoles de routage
 - **Distance Vector** (Vecteur de distance)
 - Plus court chemin calculé à l'aide du **nombre de sauts** et de manière distribuée
 - Envoi **périodique** de tables de routage aux voisins
 - **Link State** (état des liaisons)
 - Chaque routeur envoie une liste complète des routeurs voisins et **des coûts des liens** (ex. bande passante) contenue dans sa base de données
 - Construction du **graphe localement**
 - Calcul du **plus court chemin dans le graphe** avec l'algorithme Dijkstra

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

32


Routage IP dynamique



```

graph TD
    A[Protocoles de routage dynamique] --> B[Protocoles IGP (Interior Gateway Protocols)]
    A --> C[Protocoles EGP (Exterior Gateway Protocols)]
    B --> D[Protocoles à vecteur de distance]
    B --> E[Protocoles d'état des liaisons]
    C --> F[BGP]
    D --> G[RIPv1]
    D --> H[RIPv2]
    D --> I[IGRP]
    D --> J[EIGRP]
    E --> K[OSPF]
    E --> L[IS-IS]
  
```


EIGRP: Enhanced Interior Gateway Routing Protocol
IS-IS: Intermediate System to Intermediate System

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

33

Protocoles à vecteur de distance

- Algorithme **Belman-Ford** : calcul distribué des routes
- Routeur **diffuse régulièrement** les routes qu'il connaît pour ses voisins
- Une route est composée de :
 - L'adresse du réseau de destination
 - L'adresse du prochain routeur (next-hop)
 - La **métrique** : nombre de sauts = nombre de routeurs traversés pour atteindre la destination
 - Le routeur compare les routes qu'ils reçoivent avec les siennes
=> MAJ sa propre table de routage si :
 - La route reçue est **nouvelle**
 - La route reçue est **meilleure** (métrique inférieure)

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

34

Protocoles à vecteur de distance

- Exemple de convergence

- Initialement chaque routeur n'a dans sa table de routage qu'une information de routage vers la destination locale
- Le coût de tous les liens = 1

Destination	Direction	Coût
A	local	0

Destination	Direction	Coût
B	local	0

Destination	Direction	Coût
C	local	0

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

35

Protocoles à vecteur de distance

- Exemple de convergence

A envoie son vecteur de distance (A=0) à ses voisins directs B et C

Destination	Direction	Coût
A	local	0

Destination	Direction	Coût
B	local	0

Destination	Direction	Coût
C	local	0

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

36

Protocoles à vecteur de distance

- Exemple de convergence

B reçoit le vecteur (A=0) en provenance de A. Il regarde s'il a déjà une entrée A. Ce n'est pas le cas. Il ajoute une entrée A, le numéro d'interface 1 et 0+1 = 1 pour la distance

Destination	Direction	Coût
A	local	0

Destination	Direction	Coût
B	local	0
A	1	1

Destination	Direction	Coût
C	local	0
A	2	1

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

37

Protocoles à vecteur de distance

- Exemple de convergence

B envoie son vecteur de distance (A=1, B=0) à ses voisins directs A et C

A | local | 0
A | 1 | 0

B | local | 0
A | 1 | 1

C | local | 0
A | 2 | 1

(A=1, B=0)

(A=1, B=0)

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

38

Protocoles à vecteur de distance

- Exemple de convergence

A reçoit le vecteur (A=1, B=0) en provenance de B. B est une nouvelle entrée. Il l'ajoute avec une distance de 0+1. A est locale, il ne modifie pas son entrée.

A | local | 0
B | 1 | 1

B | local | 0
A | 1 | 1

C | local | 0
A | 2 | 1
B | 3 | 1

(A=1, B=0)

(A=1, B=0)

C reçoit le vecteur (A=1, B=0) en provenance de B sur la voie 3. B est une nouvelle entrée. Il l'ajoute avec une distance de 0+1. A est une entrée déjà présente avec une distance de 1, inférieure à la nouvelle (2). Aucune modification n'est apportée à l'entrée.

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

39

Protocoles à vecteur de distance

- Exemple de convergence

A | local | 0
B | 1 | 1

B | local | 0
A | 1 | 1

C | local | 0
A | 2 | 1
B | 3 | 1

(A=1, B=1, C=0)

(A=1, B=1, C=0)

C envoie le vecteur (A=1, B=1, C=0) à ses voisins directs A et B.

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

40

Protocoles à vecteur de distance

- Exemple de convergence

A reçoit sur l'interface 2 le vecteur (A=1, B=1, C=0).
C est une nouvelle entrée qu'elle ajoute avec une distance de 0+1.
A est locale, elle n'est pas modifiée.
La nouvelle distance pour (1+1+2) est supérieure à la distance existante dans la table (1). Aucune modification n'est apportée à B.

A | local | 0
B | 1 | 1
C | 2 | 1
(A=1, B=1, C=0)

B | local | 0
A | 1 | 1
C | 3 | 1

B reçoit sur l'interface 3 le vecteur (A=1, B=1, C=0).
C est une nouvelle entrée qu'il ajoute avec une distance de 0+1.
B est locale, elle n'est pas modifiée.
La nouvelle distance pour A (1+1+2) est supérieure à la distance existante dans la table (1). Aucune modification n'est apportée à A.

C | local | 0
A | 2 | 1
B | 3 | 1

(A=1, B=1, C=0)

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

41

Protocoles à vecteur de distance

- Exemple de convergence

A | local | 0
B | 1 | 1
C | 2 | 1
(A=0, B=1, C=1)

(A=1, B=0, C=1)

(A=1, B=1, C=0)

(A=1, B=0, C=1)

(A=1, B=1, C=0)

B | local | 0
A | 1 | 1
C | 3 | 1
(A=1, B=0, C=1)

C | local | 0
A | 2 | 1
B | 3 | 1
(A=1, B=1, C=0)

Toutes les tables de chaque routeur ont convergé vers un état stable.

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

42

Protocoles à vecteur de distance

- Exemple d'adaptation

A | local | 0
B | 1 | 1
C | 2 | 1
(C=0, A=1, B=inf.)

(C=0, A=1, B=inf.)

C | local | 0
A | 2 | 1
B | 3 | 1

Rupture du lien BC.
Détectée par B et C, voisins directs.
Ici C est le premier à réagir, mais le comportement est le même pour B.
C envoie son vecteur de distance avec B marqué à l'infini.

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

43

Protocoles à vecteur de distance

- Exemple d'adaptation

A reçoit de la voie 2 le vecteur de distance de C.
Toutes les entrées sont déjà présentes dans la table.
Toutes ont un coût identique ou inférieur, il n'y a donc aucune modification (de surcroît pour B à l'infini).

A	local	0
B	1	1
C	2	1

B	local	0
A	1	1
C	3	1

C	local	0
A	2	1
B	3	1

(C=0, A=1, B=inf.)

Esipe

© 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite

44

Protocoles à vecteur de distance

- Exemple d'adaptation

A	local	0
B	1	1
C	2	1

B	local	0
A	1	1
C	3	1

C	local	0
A	2	1
B	2	1

(A=0, B=1, C=1)

A envoie à son tour son vecteur de distance.
C apprend de A que celui-ci joint B en 1 saut.
Il déduit qu'il peut rejoindre à nouveau B en passant par A voie 2 pour un coût de 1+1=2 (distance pour aller à A plus distance de A à B)

Esipe

© 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite

45

Protocoles à vecteur de distance

- Exemple de comptage à l'infini

Soit le réseau stabilisé ayant les liens AB et BC
(A passe par B pour rejoindre C et inversement)

A	local	0
B	1	1
C	1	2

B	local	0
A	1	1
C	3	1

C	local	0
B	3	1
A	3	2

Esipe

© 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite

46

Protocoles à vecteur de distance

- Exemple de comptage à l'infinie

Soudain B ne reçoit plus de paquets de C. Il le marque alors à l'infini et en informe ses voisins directs à savoir A

A | local | 0
B | 1 | 1
C | 1 | 2

(A=1,
B=0,
C=inf)

B | local | 0
A | 1 | 1
C | 3 | inf

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

47

Protocoles à vecteur de distance

- Exemple de comptage à l'infinie

Après réception, A a constaté que toutes les entrées du vecteur reçu étaient déjà présentes dans sa table et ce pour des distances égales ou supérieures (de surcroît pour la distance vers C). Il n'apporte aucune modification à sa table. Et renvoie son vecteur à ses voisins directs en l'occurrence ici B, en lui indiquant qu'il peut rejoindre C en 2 sauts...

A | local | 0
B | 1 | 1
C | 1 | 2

(A=0,
B=1,
C=1)

B | local | 0
A | 1 | 1
C | 3 | inf

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

48

Protocoles à vecteur de distance

- Exemple de comptage à l'infinie

B reçoit le vecteur de A et apprend qu'il peut rejoindre à nouveau C par la voie 1 pour un coût de 2+1=3 (distance de B à A plus coût de A pour rejoindre C).

A | local | 0
B | 1 | 1
C | 1 | 2

(A=0,
B=1,
C=1)

B | local | 0
A | 1 | 1
C | 1 | 3

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

49

Protocoles à vecteur de distance

- Exemple de comptage à l'infinie

B envoie son vecteur de distance à A.
Celui-ci constate que la distance de C en provenance de la voie 1 a augmenté et est passée à 3.
Il met à jour son entrée vers C avec la distance $3+1=4$
(coût de B pour joindre C plus distance pour atteindre B de A).

(A=1, B=0, C=1)

Même raisonnement pour B et ainsi de suite jusqu'à l'infini...

Esipe

© 2011 - Développé par M. Abdelrazak RACHEDI - Droits réservés - reproduction interdite

50

Protocoles à vecteur de distance

«Distance-Vector»

- **Avantages**
 - ☐ Algorithme simple
 - ☐ Totalement décentraliser
 - ☐ Interopérabilité (stations/routeurs)
- **Inconvénients**
 - ☐ Convergence lente pour les grands réseaux
 - ☐ La taille des informations de routage est proportionnel au nombre de réseau
 - ☐ Pas de chemins de multiples
 - ☐ Coût des routes externes est arbitraire
 - ☐ Bouclage à l'infini

Esipe

© 2011 - Développé par M. Abdelrazak RACHEDI - Droits réservés - reproduction interdite

51

Protocoles à vecteur de distance

- **Solutions aux inconvénients**
 - ☐ Fixer une valeur finie pour l'infinie (16)
 - ☐ Le routeur n'envoie pas à un voisin les information qui passe par la voie pour joindre ce voisin
(Split Horizon : L'horizon coupé)


Esipe

© 2011 - Développé par M. Abdelrazak RACHEDI - Droits réservés - reproduction interdite

52

Protocoles d'état des liaisons


- Algorithme à état des liens de **Dijkstra**
- Chaque routeur communique à tous les routeurs l'état des liens avec ses voisins directs
- Métrique: débit, délai, charge, fiabilité, distance
- Les étapes à suivre pour chaque routeur
 - Découvrir les voisins directs
 - Evaluer les coûts pour les atteindre
 - Diffuser ces informations à tous les autres routeurs
 - Construire la **matrice des coûts** (représente la topologie réseau)
 - Calculer le **plus court chemin** vers tous les routeurs

 © 2011 - Développé par M. Abdelrazak RACHEDI - Droits réservés - reproduction interdite

53

Protocoles d'état des liaisons


- Etape 1: Découverte des voisins
 - Envoi du paquet **HELLO** sur toutes les liaisons
 - Chaque routeur reçoit ce paquet répond pour se présenter
- Etape 2 : Mesure le coût du lien (état du lien)
 - Calcul du temps d'aller-retour du **paquet Echo**
 - Avec/sans prise en compte de la charge du réseau
- Etape 3 : Formation du paquet d'état du liens à transmettre
 - Emetteur, liste des routeurs voisins directs et le coût associé

 © 2011 - Développé par M. Abdelrazak RACHEDI - Droits réservés - reproduction interdite

54

Protocoles d'état des liaisons

- Etape 4 : Diffusion de ces informations à tout le réseau
 - Par inondation (flooding)
- Etape 5 : Calcul de la matrice de coûts
 - Constitue une représentation de la topologie réseau
- Une route peut avoir trois états :
 - **Validé** : à partir de la source il n'existe aucun autre chemin pour atteindre la destination
 - **Découverte** : nouvelle route pour joindre le nœud suivant (next-hop) à partir d'un nœud validé
 - **En attente** : nouvelle route dont on ne sait pas si elle peut être validée ou pas

 © 2011 - Développé par M. Abdelrazak RACHEDI - Droits réservés - reproduction interdite

55

Protocoles d'état des liaisons

- Exemple

Formation des paquets pour diffuser des informations sur le voisinage

Calcul de la matrice de coûts

	A	B	C	D	E	F
A	0	7	0	0	0	4
B	7	0	3	0	2	0
C	0	3	0	5	0	0
D	0	0	5	0	7	4
E	0	2	0	7	0	3
F	4	0	0	4	3	0

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

56

Protocoles d'état des liaisons

- Exemple

Construction de la table de routage de A

Table des coûts						
A \ De	A	B	C	D	E	F
A	0	7	0	0	0	4
B	7	0	3	0	2	0
C	0	3	0	5	0	0
D	0	0	5	0	7	4
E	0	2	0	7	0	3
F	4	0	0	4	3	0

Routes validées	Routes Découvertes	En attente
A,0	AB,7 en attente AF,4 validée	AB,7

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

57

Protocoles d'état des liaisons

- Exemple

Table des coûts						
A \ De	A	B	C	D	E	F
A	0	7	0	0	0	4
B	7	0	3	0	2	0
C	0	3	0	5	0	0
D	0	0	5	0	7	4
E	0	2	0	7	0	3
F	4	0	0	4	3	0

Routes validées	Routes Découvertes	En attente
A,0	AB,7 en attente AF,4 validée	AB,7
AF,4	FE,7 validée FD,8 en attente	AB,7 FD,8

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

58

Protocoles d'état des liaisons

- Exemple

	A	B	C	D	E	F
A	0	7	0	0	0	4
B	7	0	3	0	2	0
C	0	3	0	5	0	0
D	0	0	5	0	7	4
E	0	2	0	7	0	3
F	4	0	0	4	3	0

Routes validées	Routes Découvertes	En attente
A,0	AB,7 en attente AF,4 validée	AB,7
AF,4	FE,7 validée FD,8 en attente	AB,7 FD,8
FE,7	EB,9 fin et validation AB,7 ED,14 fin et validation FD,8	AB,7 FD,8

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

59

Protocoles d'état des liaisons

- Exemple

	A	B	C	D	E	F
A	0	7	0	0	0	4
B	7	0	3	0	2	0
C	0	3	0	5	0	0
D	0	0	5	0	7	4
E	0	2	0	7	0	3
F	4	0	0	4	3	0

Routes validées	Routes Découvertes	En attente
A,0	AB,7 en attente AF,4 validée	AB,7
AF,4	FE,7 validée FD,8 en attente	AB,7 FD,8
FE,7	EB,9 fin et validation AB,7 ED,14 fin et validation FD,8	AB,7 FD,8
AB,7	BC,10 en attente BE,9 fin on sait aller à E pour 7	BC,10

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

60

Protocoles d'état des liaisons

- Exemple

	A	B	C	D	E	F
A	0	7	0	0	0	4
B	7	0	3	0	2	0
C	0	3	0	5	0	0
D	0	0	5	0	7	4
E	0	2	0	7	0	3
F	4	0	0	4	3	0

Routes validées	Routes Découvertes	En attente
A,0	AB,7 en attente AF,4 validée	AB,7
AF,4	FE,7 validée FD,8 en attente	AB,7 FD,8
FE,7	EB,9 fin et validation AB,7 ED,14 fin et validation FD,8	AB,7 FD,8
AB,7	BC,10 en attente BE,9 fin on sait aller à E pour 7	BC,10
FD,8	DC,13 fin et validation BC,10 DF,16 fin on sait aller à F pour 4	BC,10

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

61

Protocoles d'état des liaisons - Exemple

	A	B	C	D	E	F
A		7	0	0	0	4
B	7		3	0	2	0
C	0	3		5	0	0
D	0	0	5		7	4
E	0	2	0	7		3
F	4	0	0	4	3	

Routes validées	Routes Découvertes	En attente
A,0	AB,7 en attente AF,4 validée	AB,7
AF,4	FE,7 validée FD,8 en attente	AB,7 FD,8
FE,7	EB,9 fin et validation AB,7 ED,14 fin et validation FD,8	AB,7 FD,8
AB,7	BC,10 en attente BE,9 fin on saut aller à E pour 7	BC,10
FD,8	DC,13 fin et validation BC,10 DF,15 fin on saut aller à F pour 4	BC,10
BC,10	CD,15 fin on saut aller à D pour 5	vide

© 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite

62

Protocoles d'état des liaisons

- **Avantage**
 - Convergence **rapide et sans boucle**
 - Possibilité de **chemins multiples**
 - Plusieurs métriques précises couvrant différents besoins
 - Chaque routeur calcule ses routes indépendamment des autres
 - Les algorithmes d'états de liaisons sont mieux adaptés au **facteur d'échelle** que les algorithmes de Vector-Distance
- **Inconvénients**
 - **Complexe** à mettre en œuvre
 - **Consommation de ressources** non négligeable

© 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite

63


RIP (Routing Information Protocol)

© 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite

64

Introduction

- Origine : RFC 1058 (RIPv1 : 6/1988)
- Type d'algorithme : Distance-Vector
- Algorithme de routage : Ford-Fulkerson (1962)
- Premières implémentations : Unix BSD/routed
- Deux versions : RIPv1 et RIPv2 (RIPng pour IPv6)
- Fonctionne au-dessus d'UDP/IP avec le port 520
- Simple d'utilisation et à mettre en œuvre
- Le rafraîchissement des routes est tous les 30 secondes
- La route est considérée indisponible si elle n'est pas rafraîchie durant 3 minutes

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite


65

RIPv1

- La seule métrique utilisée par RIP est **nombre de sauts**
- Les annonces sont effectuées **en diffusion généralisée** (broadcast)
- Les routes validées sont les routes dont le nombre de sauts est inférieure à 16

Message RIPv1 encapsulé

En-tête de trame liaison de données	En-tête de paquet IP	En-tête de segment UDP	Message RIP (512 octets ; jusqu'à 25 routes)
-------------------------------------	----------------------	------------------------	----------------------------------------------


 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

66

RIPv1 : Format de message

En-tête de trame liaison de données		En-tête de paquet IP		En-tête de segment UDP		Message RIP (512 octets ; jusqu'à 25 routes)	
Bit		0 7 8		15 16		23 24 31	
Commande = 1 ou 2		Version = 1		ID de famille d'adresses (2 = IP)		Doit être zéro	
Adresse IP (adresse réseau)		Doit être zéro		Doit être zéro		Doit être zéro	
Mesure (sauts)		Entrées de route multiples, jusqu'à 25					



- **Commande** : 1 pour une demande, 2 pour une réponse
- **Version** : 1 pour Protocole RIP version 1
- **Identificateur de famille d'adresses** : 2 pour le protocole IP sauf si un routeur exige une table de routage complète, auquel cas ce champ doit avoir la valeur zéro
- **Adresse IP** : adresse de la route destination
- **Mesure** : Nombre de saut compris entre 1 et 16. Le routeur émetteur augmente la mesure avant d'envoyer le message

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

67

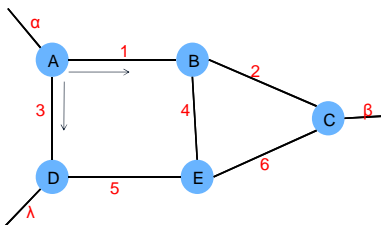
RIPv1 : Processus de mise à jour

- Si entrée n'existe pas et $< \infty$
 - Ajout avec bonne métrique et prochain routeur
 - Initialisation temporisation
- Si entrée présente et nouvelle $< \text{ancienne}$
 - Mise à jour métrique et prochain routeur
 - Réinitialisation temporisation calcule le meilleur
- Si entrée présente et routeur suivant = émetteur
 - Mise à jour de la métrique
 - Réinitialisation temporisation


© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite


68

RIPv1 : Exemple





A		
α	A	0
1	A	0
3	A	0

B		
1	B	0
2	B	0
4	B	0

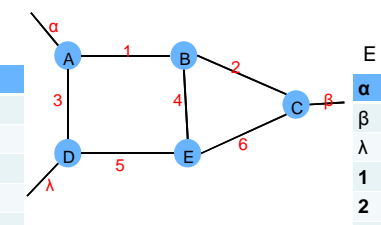
....

E		
4	E	0
5	E	0
6	E	0


© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite


69



RIPv1 : Exemple



A		
α	A	0
β	B	2
λ	D	1
1	A	0
2	B	1
3	A	0
4	B	1
5	D	1
6	D	2

....

E		
α	D	2
β	C	1
λ	D	1
1	B	A
2	B	1
3	D	1
4	E	0
5	E	0
6	E	0


© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite


70

RIPv1 : Les mises à jours

- Le vecteur de distance est envoyé :**
 - Périodiquement (chaque) **30 secondes**
 - Par déclenché : dès qu'une entrée est modifiée. Uniquement les entrées modifiées sont transmises
- Chaque message contient une entrée et une mesure**
 - Valeur 16 correspond à l'infini
 - Mesure infinie non transmise (sauf en déclenché)
 - Maximum de 25 entrées par message
- Minuteur de temporisation**
 - Si aucune mise à jour n'a été reçue pour actualiser une route existante dans les **180 secondes** (par défaut), la route est marquée comme **non valide** (valeur 16 attribuée à la mesure)
 - La route est conservée dans la table de routage jusqu'à l'expiration du minuteur d'annulation
- Minuteur d'annulation**
 - = 240 secondes**
 - Lorsque le délai du minuteur d'annulation expire, la route est supprimée de la table de routage

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

71

RIPv1 : Un réseau inaccessible

Table A:

α	A	16
β	B	2
λ	D	1
1	A	0
2	B	1
3	A	0
4	B	1
5	D	1
6	D	2

Table E:

α	D	2
β	C	1
λ	D	1
1	B	A
2	B	1
3	D	1
4	E	0
5	E	0
6	E	0

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

72

RIPv1 : Un réseau inaccessible

Table A:

α	A	16
β	B	2
λ	D	1
1	A	0
2	B	1
3	A	0
4	B	1
5	D	1
6	D	2

Table E:

α	D	16
β	C	1
λ	D	1
1	B	A
2	B	1
3	D	1
4	E	0
5	E	0
6	E	0

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

73

RIPv1 : Une boucle est possible

- Le sous réseau 2 n'est plus accessible
- A envoie sa table de routage à B avant que ce dernier envoie sa table par déclenché
- B va apprendre par A qu'il peut joindre le sous réseau 2 en passant par A en deux sauts, et A apprend que le sous réseau 2 n'est plus accessible
- B et A envoient par déclenché leurs table de routage
- A va apprendre par B qu'il peut joindre le sous réseau 2 en passant par B en trois sauts, et B apprend que le sous réseau 2 n'est plus accessible
-
- L'opération se répète jusqu'à ce que le nombre de saut atteigne 16, et dans ce cas là, A et B se rendent compte que le sous réseau 2 n'est plus accessible

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

74

RIPv1 : Une boucle est possible

- Les sous réseaux 1 et 5 ne sont plus accessibles
- B sélectionne C pour joindre le sous réseaux 5 en 2 sauts
- E sélectionne B pour joindre le sous réseaux 5 en 3 sauts
- C sélectionne E pour joindre le sous réseaux 5 en 4 sauts
- B sélectionne C pour joindre le sous réseaux 5 en 5 sauts
-
- L'opération se répète jusqu'à ce que le nombre de saut atteigne 16, et dans ce cas là, B, C et E se rendent compte que le sous réseau 5 n'est plus accessible

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

75

RIPv1 : Une boucle est possible

- Le problème est appelé : **comptage à l'infini**
- Pour réduire le problème, la distance 16 borne le comptage
- La solution consiste à utiliser l'**horizon partagé**

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

76

RIPv1 : Horizon partagé

Une route ne doit pas être annoncée sur la liaison où elle a été apprise

Réseau	Interface	Saut	Réseau	Interface	Saut	Réseau	Interface	Saut
10.1.0.0	Fa0/0	0	10.2.0.0	S0/0/0	0	10.3.0.0	S0/0/1	0
10.2.0.0	S0/0/0	0	10.3.0.0	S0/0/1	0	10.4.0.0	Fa0/0	0
10.3.0.0	S0/0/0	1	10.1.0.0	S0/0/0	1	10.2.0.0	S0/0/1	1
10.4.0.0	S0/0/0	2	10.4.0.0	S0/0/1	1	10.1.0.0	S0/0/1	2

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

77

RIPv1 : Horizon partagé

- Une route ne doit pas être annoncée sur la liaison où elle a été apprise

R3 n'annonce que le réseau 10.4.0.0 à R2.

Réseau	Interface	Saut	Réseau	Interface	Saut	Réseau	Interface	Saut
10.1.0.0	Fa0/0	0	10.2.0.0	S0/0/0	0	10.3.0.0	S0/0/1	0
10.2.0.0	S0/0/0	0	10.3.0.0	S0/0/1	0	10.4.0.0	Fa0/0	0
10.3.0.0	S0/0/0	1	10.1.0.0	S0/0/0	1	10.2.0.0	S0/0/1	1
10.4.0.0	S0/0/0	2	10.4.0.0	S0/0/1	1	10.1.0.0	S0/0/1	2

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

78

RIPv1 : Horizon partagé

- Une route ne doit pas être annoncée sur la liaison où elle a été apprise

Réseau	Interface	Saut	Réseau	Interface	Saut	Réseau	Interface	Saut
10.1.0.0	Fa0/0	0	10.2.0.0	S0/0/0	0	10.3.0.0	S0/0/1	0
10.2.0.0	S0/0/0	0	10.3.0.0	S0/0/1	0	10.4.0.0	Fa0/0	0
10.3.0.0	S0/0/0	1	10.1.0.0	S0/0/0	1	10.2.0.0	S0/0/1	1
10.4.0.0	S0/0/0	2	10.4.0.0	S0/0/1	1	10.1.0.0	S0/0/1	2

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

79

RIPv1 : Horizon partagé

- Une route ne doit pas être annoncée sur la liaison où elle a été apprise

R2 n'annonce que les réseaux 10.3.0.0 et 10.4.0.0 à R1.
R2 n'annonce que les réseaux 10.2.0.0 et 10.1.0.0 à R3.

Réseau	Interface	Saut	Réseau	Interface	Saut	Réseau	Interface	Saut
10.1.0.0	Fa0/0	0	10.2.0.0	S0/0/0	0	10.3.0.0	S0/0/1	0
10.2.0.0	S0/0/0	0	10.3.0.0	S0/0/1	0	10.4.0.0	Fa0/0	0
10.3.0.0	S0/0/0	1	10.1.0.0	S0/0/0	1	10.2.0.0	S0/0/1	1
10.4.0.0	S0/0/0	2	10.4.0.0	S0/0/1	1	10.1.0.0	S0/0/1	2

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

80

RIPv1 : Horizon partagé

- Une route ne doit pas être annoncée sur la liaison où elle a été apprise

R1 n'annonce que le réseau 10.1.0.0 à R2.

Réseau	Interface	Saut	Réseau	Interface	Saut	Réseau	Interface	Saut
10.1.0.0	Fa0/0	0	10.2.0.0	S0/0/0	0	10.3.0.0	S0/0/1	0
10.2.0.0	S0/0/0	0	10.3.0.0	S0/0/1	0	10.4.0.0	Fa0/0	0
10.3.0.0	S0/0/0	1	10.1.0.0	S0/0/0	1	10.2.0.0	S0/0/1	1
10.4.0.0	S0/0/0	2	10.4.0.0	S0/0/1	1	10.1.0.0	S0/0/1	2

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

81

RIPv1 : Horizon partagé

- Annonce d'un distance infini sur la liaison où la route a été apprise

Le réseau 10.4.0.0 tombe en panne.

Réseau	Interface	Saut	Réseau	Interface	Saut	Réseau	Interface	Saut
10.1.0.0	Fa0/0	0	10.2.0.0	S0/0/0	0	10.3.0.0	S0/0/1	0
10.2.0.0	S0/0/0	0	10.3.0.0	S0/0/1	0	10.4.0.0	Fa0/0	0
10.3.0.0	S0/0/0	1	10.1.0.0	S0/0/0	1	10.2.0.0	S0/0/1	1
10.4.0.0	S0/0/0	2	10.4.0.0	S0/0/1	1	10.1.0.0	S0/0/1	2

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

82

RIPv1 : Horizon partagé

- Annonce d'un distance infini sur la liaison où la route a été apprise

R3 « empoisonne » la route avec une mesure « infinie ».

Réseau	Interface	Saut	Réseau	Interface	Saut	Réseau	Interface	Saut
10.1.0.0	Fa0/0	0	10.2.0.0	S0/0/0	0	10.3.0.0	S0/0/1	0
10.2.0.0	S0/0/0	0	10.3.0.0	S0/0/1	0	10.4.0.0	Fa0/0	16
10.3.0.0	S0/0/0	1	10.1.0.0	S0/0/1	1	10.2.0.0	S0/0/1	1
10.4.0.0	S0/0/0	2	10.4.0.0	S0/0/1	1	10.1.0.0	S0/0/1	2

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

83

RIPv1 : Horizon partagé

- Annonce d'un distance infini sur la liaison où la route a été apprise

R3 envoie une mise à jour de poison déclenchée à R2.

Réseau	Interface	Saut	Réseau	Interface	Saut	Réseau	Interface	Saut
10.1.0.0	Fa0/0	0	10.2.0.0	S0/0/0	0	10.3.0.0	S0/0/1	0
10.2.0.0	S0/0/0	0	10.3.0.0	S0/0/1	0	10.4.0.0	Fa0/0	16
10.3.0.0	S0/0/0	1	10.1.0.0	S0/0/0	1	10.2.0.0	S0/0/1	1
10.4.0.0	S0/0/0	2	10.4.0.0	S0/0/1	1	10.1.0.0	S0/0/1	2

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

84

RIPv1 : Horizon partagé

- Annonce d'un distance infini sur la liaison où la route a été apprise

R2 « empoisonne » la route avec une mesure « infinie ».

Réseau	Interface	Saut	Réseau	Interface	Saut	Réseau	Interface	Saut
10.1.0.0	Fa0/0	0	10.2.0.0	S0/0/0	0	10.3.0.0	S0/0/1	0
10.2.0.0	S0/0/0	0	10.3.0.0	S0/0/1	0	10.4.0.0	Fa0/0	16
10.3.0.0	S0/0/0	1	10.1.0.0	S0/0/0	1	10.2.0.0	S0/0/1	1
10.4.0.0	S0/0/0	2	10.4.0.0	S0/0/1	16	10.1.0.0	S0/0/1	2

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

85

RIPv1 : Horizon partagé

- Annonce d'un distance infini sur la liaison où la route a été apprise

R1 « empoisonne » la route avec une mesure « infinie ».

Réseau	Interface	Saut	Réseau	Interface	Saut	Réseau	Interface	Saut
10.1.0.0	Fa0/0	0	10.2.0.0	S0/0/0	0	10.3.0.0	S0/0/1	0
10.2.0.0	S0/0/0	0	10.3.0.0	S0/0/1	0	10.4.0.0	Fa0/0	16
10.3.0.0	S0/0/1	1	10.1.0.0	S0/0/0	1	10.2.0.0	S0/0/1	1
10.4.0.0	S0/0/0	16	10.4.0.0	S0/0/1	16	10.1.0.0	S0/0/1	2

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

86

RIPv1 : Problèmes liés aux mises à jour synchronisées

- Lorsque plusieurs routeurs transmettent **simultanément des mises à jour** de routage sur des segments LAN à accès multiples
- Les paquets de mise à jour peuvent **entrer en collision** et causer des délais ou **consommer trop** de bande passante.
- La solution : l'intervalle de mise à jour varie **aléatoirement de 25 à 30 secondes** pour l'intervalle par défaut de 30 secondes

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

87

RIPv1 : Les Limites

- La principale limite du protocole RIPv1 est qu'il s'agit d'un **protocole de routage par classe**
- Diamètre de réseau < à 15
- Métrique statique ne doit pas changer avec le temps
- RIPv1 **n'inclut pas le masque de sous-réseau** avec l'adresse réseau dans les mises à jour de routage
- Problèmes avec les sous-réseaux discontinus ou les réseaux qui utilisent le masquage de sous-réseau de longueur variable (**VLSM : Variable Length Subnet Mask**)

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

88

RIPv2

- RFC 1387, 1388, 1723 (RIPng : RFC 2453)
- Diffusion multicast (224.0.0.9)
- Protocole avec possibilité d'authentification :
 - Mot de passe
 - MD5
- Classless Inter Domain Routing (CIDR) ce qui permet de faire un meilleur regroupement de route (agrégation)
- Supporte le VLSM
 - Véhicule le masque de réseau

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

89

RIPv2 : calcul le regroupement de route

Étape 1 : Énumérez les réseaux en format binaire.

172.20.0.0	10101100 . 00010100 . 00000000 . 00000000
172.21.0.0	10101100 . 00010101 . 00000000 . 00000000
172.22.0.0	10101100 . 00010110 . 00000000 . 00000000
172.23.0.0	10101100 . 00010111 . 00000000 . 00000000

Étape 2 : Comptez le nombre de bits en correspondance les plus à gauche pour déterminer le masque.
14 bits en correspondance, /14 ou 255.252.0.0

Étape 3 : Copiez les bits en correspondance et complétez avec des bits à 0 pour terminer l'adresse réseau.

172.20.0.0	10101100 . 00010100 . 00000000 . 00000000
------------	-------------------------------------------

Copier Ajouter des bits 0

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

90

RIPv2

- RIPv2 comme RIPv1 comporte les fonctions :
 - Mise hors service et autres minuteurs pour tenter d'éviter les boucles de routage
 - Découpage d'horizon, avec ou sans empoisonnement inverse, dans le même but
 - Mises à jour déclenchées en cas de modification de la topologie pour une convergence plus rapide
 - Nombre de sauts maximum limité à 15, un nombre de sauts de 16 indique un réseau inaccessible

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

91

Comparaison entre RIPv1 et RIPv2

RIP v1	RIP v2
Facile à configurer	Facile à configurer
Prend en charge uniquement un protocole de routage par classes (classful).	Prend en charge l'utilisation du routage CIDR (Classless).
La mise à jour de routage ne contient aucune information de sous-réseau.	Envoie des informations sur les masques de sous-réseau avec les mises à jour des routes.
Ne supporte pas le routage CIDR ce qui oblige tous les équipements d'un même réseau à utiliser le même masque de sous-réseau	Supporte le routage CIDR ce qui permet à des équipements d'un même réseau d'utiliser différents masques de sous-réseau
Aucune authentification dans les mises à jour	Permet l'authentification dans ses mises à jour de routage
Envoie les broadcasts sur 255.255.255.255.	Envoie les mises à jour de routage en multicast sur 224.0.0.9 ce qui est plus efficace.

Esipe
© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

92

OSPF (Open Shortest Path First)

- Pourquoi OSPF?
 - Pour combler les limites du protocole RIP
 - Possibilité de gérer des domaines de diamètre **> 16**
 - Amélioration du **temps de convergence**
 - Métrique **plus adaptée** (prise en compte des débits)
 - Meilleure possibilité **d'agrégation des routes**
 - Segmentation possible du domaine en aires
 - Mais : OSPF est plus **complexe** et configuration pas simple

Esipe
© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

93

OSPF (Open Shortest Path First)

- Protocole de routage à état de liaison (Link-state)
 - Ouvert (**Open** SPF) : spécification dans la RFC1247(v1) et RFC2328(v2)
 - Créer pour remplacer RIP et les autres protocoles internes (IGP)
- Utilise IP, le multicast via **224.0.0.5** (pour tous les routeurs OSPF de AS) et **224.0.0.6** (pour l'élection du DR -Designated Router)
- Calcul des **coûts** en guise de métrique
- Utilise les types de service (champ ToS) d'IP
 - Permet la gestion de **plusieurs routes** pour la même destination
 - Permet **l'équilibrage de charge** parmi les différentes routes vers le même réseau
- Supporte VLSM
- OSPF v2 version récente publiée 1998 dans le RFC2328
- OSPF V3 (extension à IPv6) publié en 1999 dans le RFC2740

Esipe
© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

94



OSPF (Open Shortest Path First)

OSPF assure la détection de tout changement topologique et ensuite il le signale à **tous les routeurs** du réseau.

Notamment les modifications :

- Perte de liaison
- Routeur hors service



Les routeurs sont connectés à leurs voisins par des lignes dont **ils testent leur état régulièrement** par l'envoi d'un message périodique.

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite 

95

OSPF : Routage hiérarchique



- Avec un réseau de **grande taille** :
 - **Surcharge de trafic** dû au transit des informations de routage
 - Des calculs de **route longues**
- OSPF utilise un **routage hiérarchique** basé sur le découpage du réseau en aires (areas)
 - Réseaux contigus reliés par une dorsale (Backbone)
 - Chaque aire constitue un ensemble indépendants de réseaux
 - Table des liaisons ne contient que des liaisons de l'aire
 - Protocole d'inondation s'arrête au frontière de l'aire
 - Les routeurs ne calculent que les routes internes de l'aire
 - Il existe des **routeurs de bordure** reliés à plusieurs aires et ils transmettent des informations récapitulatives sur les aires

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite 

96

OSPF : Topologie de réseaux

- Aires OSPF
- Liens virtuels
- Types de Routeurs
- Types de routes OSPF
- Différents types de LSA (Link State Advertisements)
- Plusieurs chemins de couts identiques

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite 

97

OSPF : les aires

- Dans un grand domaine chaque changement provoque une diffusion de la table d'état des liens de tous les routeurs :
 - Consommation de bande passante importante
 - Charge CPU important sur les routeurs
 - Alors que la portée d'une modification reste localiser
- D'où l'idée de découper le domaine en aires
 - Chaque aire est plus simple et plus stable
- Pour garder une cohérence globale, une aire principale (backbone):
 - Relie toute les aires entre elles
 - Connaît toutes les informations du routage, mais ne diffuse que des condensés (en agrégeant les routes)

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

98

OSPF : les aires

- 3 classes d'aires
 - Aire Backbone (Area 0) unique: chemin obligatoire pour passer d'une aire à une autre
 - Aires secondaires :
 - Tous les routeurs ont une vue complète de la carte du réseau
 - Routeurs calculent localement la meilleure route entre la source et la destination
 - Aires terminales (sub areas) :
 - Pas de vue complète de la carte du réseau,
 - Il s'agit d'une zone qui ne contient qu'un seul routeur frontière. La zone ne peut pas être traversée.

ABR (Area Border Router)
ASBR (Autonomous System Border Router)
Backbone Router - Routeur situé dans l'Aire 0 (Backbone)

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

99

OSPF : Aires OSPF

- L'utilisation et la configuration de plusieurs zones permet de réduire la taille des bases de données d'états des liaisons

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

100

OSPF : Aires OSPF

- Lorsqu'il y a une modification de topologie, **seuls les routeurs de la zone concernée reçoivent le LSA et exécutent l'algorithme SPF**
- Chaque aire définit une base de données **invisible hors de l'aire**
- Toutes les aires **doivent être connectées** au backbone
- Quand doit-on subdiviser le réseau en aires ?
 - ☐ Lorsque le backbone **a plus de 10 à 15 routeurs**
 - ☐ Lorsque la topologie du backbone **devient complexe**

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

101

OSPF : Liens virtuels

- **Pas recommandés**
- A quoi servent-ils ?
 - ☐ Lorsque le backbone **a plus de 10 à 15 routeurs**
 - ☐ Utiliser dans **un scénario de backup**
 - ☐ Permet d'assurer la connectivité **d'une aire à une aire autre que le backbone**
 - ☐ Permet d'assurer **la connectivité d'une aire déconnectée**

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

102

OSPF : Liens virtuels

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

103

OSPF : Liens virtuels

Lien virtuel Entre area 1 et area 2

Area 1

Area 2

Area 0

Area 3

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

104

OSPF : Les sous-protocoles

- **AS Border Router** : routeurs **frontières de l'AS** qui apprend les routes extérieures à l'AS par EGP (ex: BGP)
- **Area Border Router** : routeur appartenant à plusieurs aires (interface multiples => aires multiples)
 - Maintient de BD de topologie différentes
- **Designated router** : routeur principale sélectionné dans chaque réseau pour limiter les transactions entre les routeurs

Area 0 (backbone area)

Area 1

Area 2

Area 3

ASBR

RIP-enabled network

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

105

OSPF : Les sous-protocoles

- OSPF : protocole complexe
 - Nécessite plusieurs sous-protocoles pour fonctionner et chaque routeur gère plusieurs BD
- 1) Protocole Hello
 - Identification de voisinage et maintenance des liens
- 2) Protocole d'échange /description
- 3) Flooding (inondation)
 - Mise à jour incrémentale des tables

Area 0 (backbone area)

Area 1

Area 2

Area 3

ASBR

RIP-enabled network

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

106

OSPF : L'algorithme OSPF

OSPF utilise l'algorithme SPF de Dijkstra

Base de données à état de liens

Algorithme SPF

Arbre SPF

Table de routage

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

107

OSPF : Encapsulation de message

- La partie données d'un message OSPF est encapsulée dans un paquet. Cette zone de données peut inclure un des 5 types de paquets LSP OSPF.

En-tête de trame de liaison de données	En-tête de paquet IP	En-tête de paquet OSPF	Données spécifiques de type de paquet OSPF
<p>Trame de liaison de données (champs Ethernet affichés ici)</p> <p>Adresse MAC source = adresse de l'interface d'envoi</p> <p>Adresse MAC de destination = multibroadcast : 01-00-5E-00-00-05 ou 01-00-5E-00-00-06</p> <p>Paquet IP</p> <p>Adresse IP source = adresse de l'interface d'envoi</p> <p>Adresse IP de destination = multibroadcast : 224.0.0.5 ou 224.0.0.6</p> <p>Champ de protocole = 89 pour OSPF</p>			
		<p>En-tête de paquet OSPF</p> <p>Code du type de paquet OSPF</p> <p>ID du routeur et ID de la zone</p>	<p>Type de paquet OSPF</p> <p>0x01 Hello</p> <p>0x02 Description de base de données (DD)</p> <p>0x03 Requête d'état de liens</p> <p>0x04 Mise à jour d'état de liens</p> <p>0x05 Accusé de réception d'état de liens</p>

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

108

OSPF : Type de paquet LSP


- Hello** - les paquets Hello servent à vérifier si les liaisons sont opérationnels - Permettre l'élection du DR ainsi que BDR
- DBD** - (DataBase Description) : le paquet de description de base de données contient une liste abrégée de la base de données d'état des liaisons du routeur expéditeur et est utilisé par les routeurs de destination pour contrôler la base de données d'état des liaisons locale.
- LSR** - (Link-State Request) : les routeurs de destination peuvent alors demander plus d'informations sur n'importe quelle entrée de la DBD, en envoyant une requête LSR - (Requête d'état des liaisons).

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

109

OSPF : Type de paquet LSP


- **LSU** (Link-State Update - Mise à jour d'état de liaisons) –
 - Ces paquets sont utilisés pour répondre aux LSR, ainsi que pour annoncer de nouvelles informations.
 - Les LSU contiennent **sept types** différents de LSA (Link-State Advertisements – Annonces d'état des liaisons).
- **LSAck** (Link-State Acknowledgement - Accusé de réception d'état des liaisons)
 - Lors de la réception d'une LSU, le routeur envoie un LSAck pour confirmer la bonne réception de cette LSU.

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

110

OSPF : Le protocole HELLO

- Les paquets Hello sont utilisés pour :
 - découvrir des voisins OSPF et établir des contiguïtés
 - annoncer les paramètres sur lesquels les deux routeurs doivent s'accorder pour devenir voisins
 - définir le routeur désigné (DR) et le routeur désigné de secours sur les réseaux à accès multiple, de type Ethernet et Frame Relay
- Les paquets Hello sont envoyés toutes les 10 secondes

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

111

OSPF : Le protocole HELLO


En-tête de trame de liaison de données En-tête de paquet IP En-tête de paquet OSPF Données spécifiques de type de paquet OSPF: Paquet Hello.

Bit(s) : 0 7 8 15 16 23 24 31

En-tête(s) de paquet OSPF:

- Version
- Type=1
- ID du routeur
- ID de la zone
- Somme de contrôle
- Authentification
- Authentification
- Masque de réseau
- Intervalle Hello
- Option
- Priorité du routeur
- Intervalle d'arrêt du routeur
- Routeur désigné (DR)
- Routeur désigné de secours (BDR)
- Liste des voisins


Paquet(s) Hello OSPF:

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

112

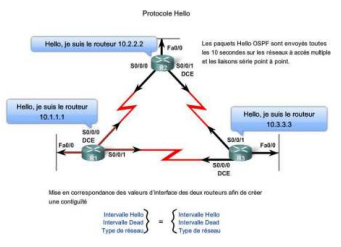
OSPF : Le protocole HELLO

- **Type** : type de paquet OSPF : Hello (1), DD (2), LS Request (3) (requête d'état des liaisons), LS Update (4) (mise à jour d'état des liaisons), LS ACK (5) (accusé de réception d'état des liaisons)
- **ID du routeur** : ID du routeur d'origine
- **ID de la zone** : zone d'origine du paquet
- **Masque de réseau** : masque de sous-réseau associé à l'interface expéditrice
- **Intervalle Hello** : nombre de secondes qui s'écoulent entre deux envois de paquets Hello
- **Priorité du routeur** : utilisé dans la sélection du routeur désigné ou du routeur désigné de secours (étudié par la suite)
- **Routeur désigné (DR)** : ID du routeur désigné, le cas échéant
- **Routeur désigné de secours (BDR)** : ID du routeur désigné de secours, le cas échéant
- **Liste des voisins** : indique l'ID de routeur OSPF du ou des routeurs voisins

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite


113

OSPF : Le protocole HELLO



Mise en correspondance des valeurs d'interface des deux routeurs afin de créer une adjacence


Paramètre	R1	R2	R3
Intervalle Hello	10	10	10
Intervalle Dead	40	40	40
Type de réseau	Point à point	Point à point	Point à point

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

114

Mise à jour d'état des liaisons

- Les paquets de mise à jour d'états des liaisons (LSU) sont les paquets utilisés pour la mise à jour du routage OSPF
- Toutes les LSU contiennent une ou plusieurs LSA (Link-State Advertisements) et les deux termes peuvent s'utiliser pour désigner les informations d'état des liaisons propagées par les routeurs OSPF
- Les LSU-LSA sont envoyés chaque 30 minutes, dans le cas contraire, les mis à jour à chaque changement d'état du réseau

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

115

Mise à jour d'état des liaisons

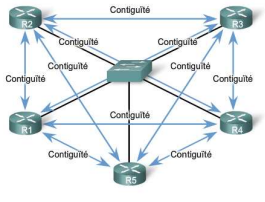
- Les LSA sur les réseaux à accès multiple peuvent présenter deux difficultés pour OSPF :
 - La création de contiguïtés multiples, une pour chaque paire de routeurs.
 - Une diffusion massive de LSA

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

116

OSPF : Echange d'information



Routeurs	Contiguïtés
0	$\frac{n(n-1)}{2}$
5	10
10	45
20	190
100	4950

Nombre de contiguïtés = $\frac{n(n-1)}{2}$
 n = nombre de routeurs
 Exemple : 5 routeurs $(5 - 1) / 2 = 10$ contiguïtés

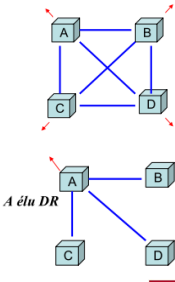
Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

117

OSPF : Le routeur désigné (DR)

- Sur chaque réseau n routeurs
 - Chaque routeur a $(n-1)$ liaisons vers les autres routeurs
 - Au total : $\frac{n(n-1)}{2}$ échanges
- Avec un routeur désigné (DR)
 - Les autres routeurs ont un lien avec le DR uniquement
 - Seul DR annonce vers l'extérieur => réduire à n échanges
- Election : fonction de priorité
 - Election d'un routeur de secours (Backup Designated Router)
 - C'est le mécanisme Hello qui se charge de surveiller le bon fonctionnement des DR (primaire et secondaire)



Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

118

OSPF : Le routeur désigné (DR)

- La solution pour gérer le nombre de contiguïtés et la diffusion des LSA sur un réseau à accès multiple est le routeur désigné (DR)
- OSPF sélectionne un routeur désigné (Designated Router - DR) comme point de collecte et de distribution des LSA envoyées et reçues
- Un routeur désigné de secours (Backup Designated Router - BDR) est également choisi en cas de défaillance du routeur désigné

Esipe © 2011 - Développé par M. Abdelrazak RACHEDI - Droits réservés - reproduction interdite

119

OSPF : Le routeur désigné (DR)

- Un seul DR par réseau d'accès multiple
 - Génère des paquets LSA sur le réseau
 - Le BDR écoute mais ne génère aucun paquet
 - Accélère la synchronisation des bases de données
 - Réduit le trafic sur le réseau d'accès

Esipe © 2011 - Développé par M. Abdelrazak RACHEDI - Droits réservés - reproduction interdite

120

OSPF : Le routeur désigné (DR)

L'ID du routeur est déterminé dans l'ordre suivant :

- 1) Utiliser l'adresse IP configurée à l'aide de la commande OSPF **router-id**
- 2) Si l'ID de routeur n'est pas configuré, celui-ci choisit l'**adresse IP la plus élevée de l'une de ses interfaces de bouclage**
- 3) Si aucune interface de bouclage n'est configurée, le routeur choisit l'**adresse IP active la plus élevée de l'une de ses interfaces physiques**;

Topologie avec interfaces de bouclage

Esipe

121

OSPF : Métrique

- La métrique OSPF s'appelle **coût**
- Plus le **coût est faible** plus l'interface est sollicitée pour acheminer le trafic des données
- Pour calculer le coût avec les routeurs cisco : il faut **cumuler les bandes passantes** des interfaces de sortie depuis le routeur vers le réseau de destination
- Coût = 10^8 / bande passante de l'interface**

Valeurs de coût OSPF Cisco

Type d'interface	10^8 bits/s = Coût
Fast Ethernet et plus rapide	$10^8 / 100\,000\,000$ bits/s = 1
Ethernet	$10^8 / 10\,000\,000$ bits/s = 10
E1	$10^8 / 2\,048\,000$ bits/s = 48
T1	$10^8 / 1\,544\,000$ bits/s = 64
128 Kbits/s	$10^8 / 128\,000$ bits/s = 781
64 Kbits/s	$10^8 / 64\,000$ bits/s = 1562
56 Kbits/s	$10^8 / 56\,000$ bits/s = 1785

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

T1 (ligne série 1544kbps)
 E1 (ligne série 2048kbps)

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

122

OSPF : Chemins multiples à coûts identiques

- Lorsque **n** chemins vers une même destination ont **des coûts égaux**, OSPF installe **n** entrées dans la table de transmission
 - On a une **répartition de la charge** du réseau vers les **n** chemins
 - Solution idéale pour étendre les capacités des liens sur les backbones du FAI (ISP)

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

123

OSPF : Authentification

- L'authentification **ne chiffre pas la table de routage**
- L'authentification permet aux routeurs d'accepter les informations provenant **uniquement des routeurs authentifiés**

Authentification

Esipe

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

124

RIP vs OSPF

	RIP	OSPF
Caractéristiques		
Type d'algorithme	Vecteur de distance	Etat des liens
Métrique	Nombre de sauts	Coût de la liaison
Origine	IETF	IETF
Architecture du réseau		
Type	à plat	hiérarchique
Nombre de routeurs	15	illimité
Performance		
Charge du réseau	élevée	faible
Périodicité des mises à jour	30 sec	3 minutes
Temps de convergence	plusieurs minutes	quelques secondes
Mode de mise à jour	information du voisin	multicast aux routeurs connus

Esipe © 2011 - Développé par M. Abdelrazak RACHEDI - Droits réservés - reproduction interdite

125

EIGRP

- EIGRP (Enhanced IGRP) est un protocole propriétaire développé par Cisco qui s'exécute uniquement sur les routeurs Cisco
- Les caractéristiques du protocole EIGRP sont les suivantes
 - Mises à jour déclenchées (EIGRP n'a pas de mises à jour régulières).
 - Utilisation d'une table topologique pour maintenir toutes les routes reçues des voisins (pas seulement les meilleurs chemins).
 - Établissement de contiguïtés avec des routeurs voisins par le biais du protocole Hello EIGRP.
 - Prise en charge des masques de sous-réseau de longueur variable
 - Bien que les routes soient propagées selon un vecteur de distance, la mesure est plutôt basée sur la bande passante minimale et le délai global du chemin que sur le nombre de sauts.

Esipe © 2011 - Développé par M. Abdelrazak RACHEDI - Droits réservés - reproduction interdite

126

EIGRP : Exemple

Diagram showing network topology with nodes A, B, C, D, E and their connections (bandwidths and link IDs):

- A to B: 8Mbps (Link 1)
- A to D: 2Mbps (Link 3)
- B to C: 2Mbps (Link 2)
- B to E: 16Mbps (Link 4)
- C to E: 16Mbps (Link 6)
- D to E: 4Mbps (Link 5)

Topology Tables:

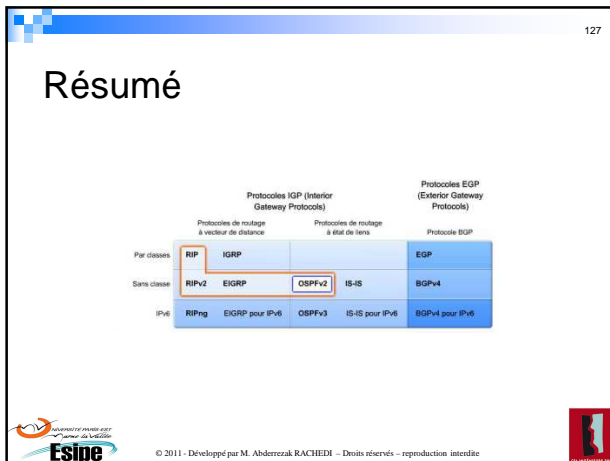
α	A	10
1	A	8
3	A	2

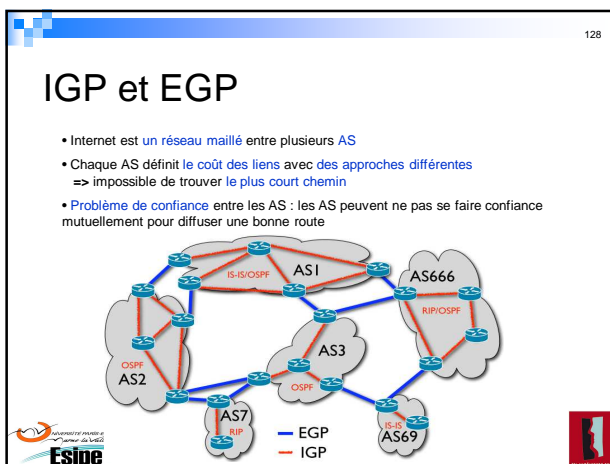
1	B	8
2	B	2
4	B	16

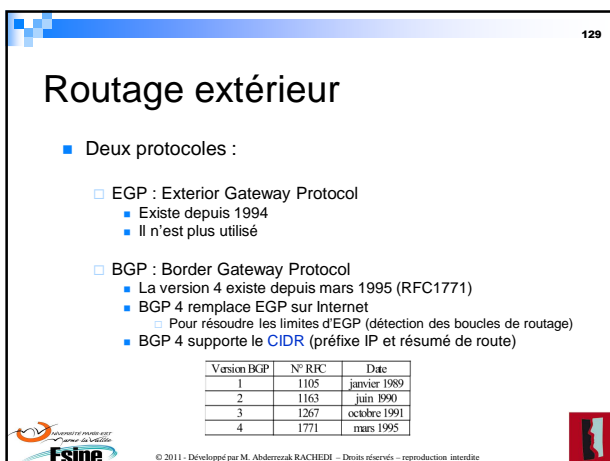
.....

4	E	16
5	E	4
6	E	16

Esipe Déroulement de l'exercice : un volontaire ! © 2011 - Développé par M. Abdelrazak RACHEDI - Droits réservés - reproduction interdite








130

BGP : Objectifs

- Echange des routes entre AS
 - Opérateurs, gros sites mono ou multi-connectés
- Implémentation de la politique de routage de chaque AS
- Indépendance des IGP utilisés dans les AS
- Supporte le passage à l'échelle (scalability)
- Minimise le trafic induit sur les liens
- Offre une bonne stabilité au routage
- Le meilleur chemin : de plus court chemin en nombre d'AS




© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

131

BGP

- Protocole à «[Path-Vector](#)»
- Supporte le Classless Inter-Domain Routing (CIDR)
- Fonctionne au-dessus de TCP :
 - Port : 179
 - Fiabilité de transmission
- Principe
 - Les routeurs BGP échangent des l'informations ([chemin: liste des numéro d'AS](#)) concernant l'accessibilité de certains préfixes IP
 - Les sessions BGP sont établies entre les routeurs de bords d'AS
 - Protocole symétrique
 - Apprentissage des chemins
 - Sélectionne le meilleur chemin et configure la table de routage avec une variante de protocole « Vector-Distance »
 - La politique d'admission au réseau




© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

132

BGP : Différence avec protocole de Distance-Vector

- Pas de transmission périodique des meilleures routes mais uniquement [les modifications](#)
- BGP mémorise [toutes les routes](#) vers [toutes les destinations](#)
- BGP construit des routes [sans boucles](#)
 - Le chemin suivi est décrit explicitement à l'aide de la liste des AS traversés
 - Les boucles sont facilement détectées




© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

133

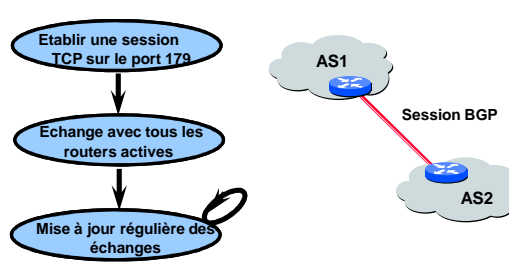
BGP

- BGP est utilisé lorsque :
 - Le réseau est « **multihomed** » (possède son propre @ IP public et son numéro d'AS)
 - Fournir un routage **complet ou partial** à un client en aval
 - Lorsqu'une information sur le chemin vers un AS est nécessaire
- BGP n'est pas nécessaire :
 - Si l'AS est connecté **par un seul point**
 - Et pas transmission d'informations sur le routage en aval
 - Une route est utilisée **par défaut**

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

134

BGP : Opérations simplifiées




```

graph TD
    A(Etablir une session TCP sur le port 179) --> B(Echange avec tous les routeurs actifs)
    B --> C(Mise à jour régulière des échanges)
    C --> A
  
```

AS1

Session BGP

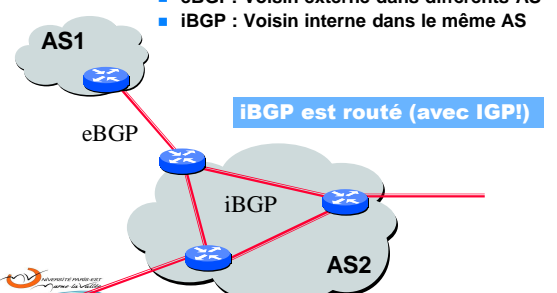
AS2

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

135

BGP : Deux types de relations de voisinage

- eBGP: Voisin externe dans différents AS
- iBGP : Voisin interne dans le même AS




AS1

eBGP

iBGP

AS2


iBGP est routé (avec IGP!)

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

136

BGP : Les règles pour les AS multi-connectés

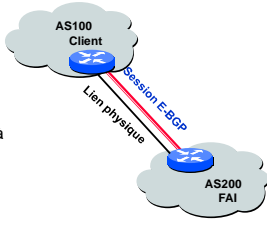
- Les routeurs de bord de **même AS** échangent leurs informations en i-BGP
- Les connexions en i-BGP forment un **maillage complet** sur les routeurs de bord d'un AS
- Les IGP internes à l'AS qui **assurent et maintiennent la connectivité** entre les routeurs de bord
- Le numéro d'AS est un numéro officiel (si connexions entre 2 différents AS)
- Si un routeur de bord **n'arrive pas à atteindre** une route de son AS, **il ne la propage pas** à ses voisins BGP


 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

137

BGP : Exemple 1

- Client connecté à un **seul** Fournisseur d'Accès Internet (FAI).
- Les routeurs qui échangent leurs informations en BGP doivent être **directement connectés** (liaison point-à-point ou LAN partagé).
- L'utilisation des numéros d'AS privés est à éviter pour les AS terminaux (clients).
- Les AS officiels (enregistrés) : de **1 à 64511**
- Les AS privés (non-enregistrés) : de **64512 à 65535**

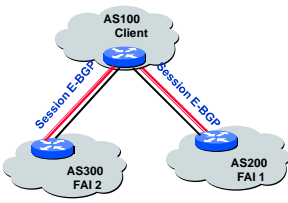



 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

138

BGP : Exemple 2

- Client connecté à deux FAI :
 - Faire passer tout son trafic par FAI1 et garder sa liaison avec FAI2 en cas de panne
 - Equilibrer son trafic entre FAI1 et FAI2
- C'est un cas typique qui illustre l'utilisation du protocole de routage BGP pour réagir de manière dynamique en cas de défaillance d'un lien



 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

139

BGP : Exemple 3

- Client connecté à deux FAI par deux routeurs :
 - Protection contre la défaillance de l'un d'entre eux ou de l'un de ses routeurs
- Connexion BGP (i-BGP) entre les routeurs de bord de l'AS100
 - Maintenir une cohérence entre les deux routeurs qui doivent posséder les mêmes informations de routage

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

140

BGP : Exemple 4

- Client connecté à **trois** FAI avec redondance sur l'un
- Maillage complet de sessions i-BGP
- Pour les autres AS, les routeurs de bord de l'AS100 (client) sont vue de point vue fonctionnel comme un seul routeur avec 4 interfaces

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

141

BGP : Différents types de domaines

- Les domaines de transit
 - Un domaine de transit **autorise l'utilisation de son infrastructure** par certains domaines pour transmettre des paquets vers d'autres domaines
 - Exemple : Renater, GEANT, UUNet, etc
- Les domaines de souche
 - Un domaine de souche **n'autorise pas d'autres domaines** à utiliser son infrastructure pour transmettre des paquets
 - Un « stub-domain » est connecté à au moins un domaine de transit

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

142

BGP : Différents types de domaines

- Single-homed stub domain (ex. S1)
- Dual-homed stub domain (ex. S2)
- Transit domain (ex. T1, T2 et T3)

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

143

BGP: Exemple de domaine de transit GEANT

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

144

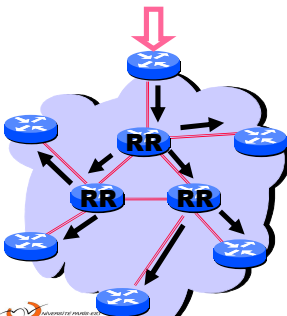
Le Maillage iBGP ne permet pas le passage à l'échelle

- N routers de bordure => $N(N-1)/2$ sessions peering
- Chaque routers doit avoir N-1 iBGP sessions
- L'ajout d'un seul iBGP nécessite le changement de configuration de tous l'autres iBGP
- Chaque router doit écouter ses voisins pour mettre à jour sa table de routage
- Il existe 4 solutions :
 - (0) Acheter des grands routeurs !
 - (1) Diviser l'AS à plusieurs petits ASes
 - (2) BGP Route reflectors
 - (3) BGP confederations

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

146

RR : Route Reflectors

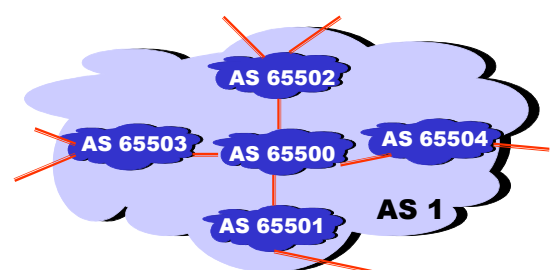


- Les réflecteurs de route "Route reflectors" peuvent transmettre les mises à jour d'iBGP aux clients
- Chaque RR ne transmet que les meilleurs chemins
- Les attributs **ORIGINATOR_ID** et **CLUSTER_LIST** sont nécessaires pour éviter les boucles

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

146

BGP : Confédérations



De l'extérieur ressemble à AS 1

eBGP confédération (entre les membres d'ASes)

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

147

BGP : Peering

- Il existe 2 types de peering BGP :
 - Customer-Provider peering
 - Relation asymétrique dans laquelle un client (un domaine de routage) achète une connectivité à Internet auprès d'un fournisseur d'accès (un autre domaine de routage)
 - Le fournisseur annonce à sons client toutes les routes qu'il connaît
 - Shared-Cost peering
 - Relation symétrique où deux domaines de routage acceptent d'échanger gratuitement leurs paquets à travers un point d'interconnexions
 - Chaque « peer » envoie à l'autre ses propres routes et celles de ses clients

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

148

BGP : Les messages

- 4 types de messages
 - ☐ OPEN
 - ☐ KEEPALIVE
 - ☐ NOTIFICATION
 - ☐ UPDATE
- Taille de message de 19 à 4096 octets
- Eventuellement l'intégrité est assuré par MD5

Esipe © 2011 - Développé par M. Abdelrazak RACHEDI - Droits réservés - reproduction interdite

149

BGP : Le message OPEN

- 1er message envoyer **après l'ouverture de la session** TCP et informe ses voisins de :
 - ☐ Sa version de BGP
 - ☐ Son numéro d'AS
 - ☐ L'identifiant du processus BGP

0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
Version
Mon système autonome
Hold Time
Identifiant BGP
Opt Perm Len
Paramètres optionnels

- Propose une **valeur de temps** pour maintenir la session
 - Valeur par défaut : **90 secondes**
 - Si 0 maintenir la sessions sans limite de durée
- Déclenche le processus d'attente d'un **KEEPALIVE**

Esipe © 2011 - Développé par M. Abdelrazak RACHEDI - Droits réservés - reproduction interdite

150

BGP : Le message KEEPALIVE

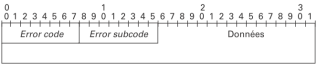
- Confirme l'ouverture de session (OPEN)
- Ajuste le « timer » pour contrôler le temps de **maintient de la session**
- Si le temps de maintient est différent de 0
 - ☐ Le message est renvoyer chaque **30 secondes**
- Message de **taille** minimum de **19 octets**
- En cas d'absence de modification de leur table de routage, les routeurs n'échangent plus que des messages **KEEPALIVE** toute les 30 secondes, ce qui génère un trafic limité au niveau de BGP

Esipe © 2011 - Développé par M. Abdelrazak RACHEDI - Droits réservés - reproduction interdite

151

BGP : Le message NOTIFICATION

- Ferme la session BGP
- Ferme aussi la session TCP
- Fourni un code pour renseigner sur l'erreur
- Annule toutes les routes apprises par BGP
 - Peut provoquer des instabilités de routage injustifiées
- Emis sur incident :
 - Pas de KEEPALIVE pendant 90 secondes
 - Message incorrect
 - Problème lié au processus BGP



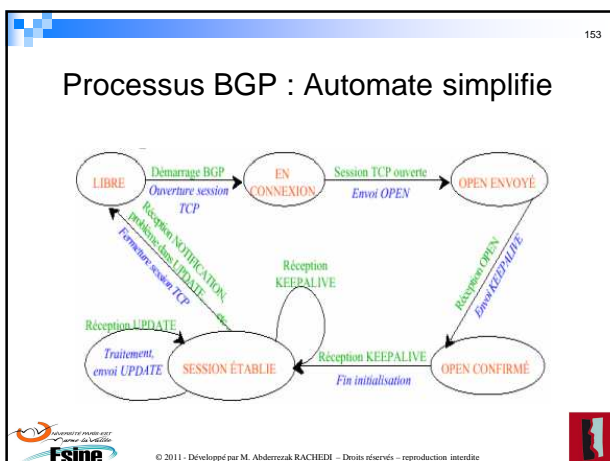
© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

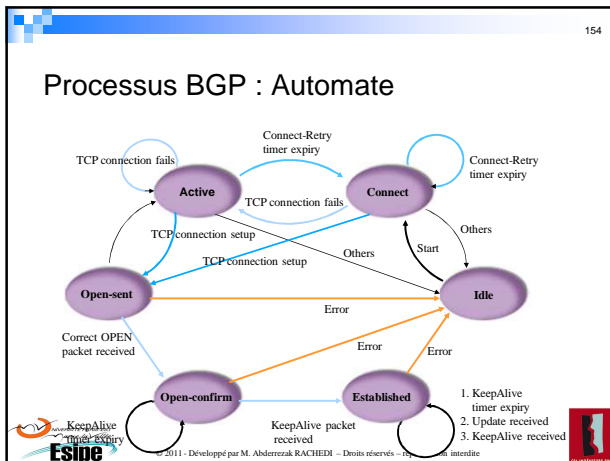
152

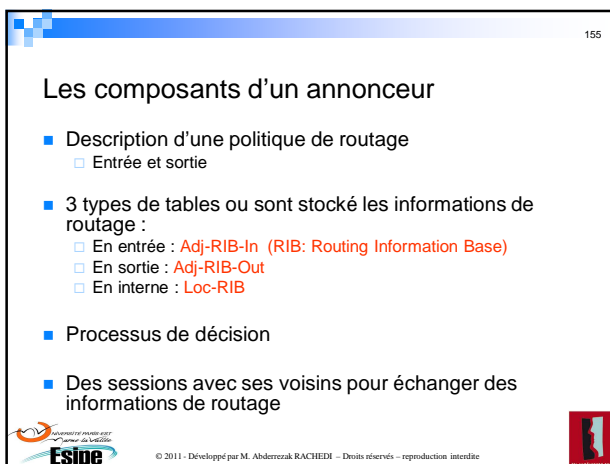
BGP : Le message UPDATE

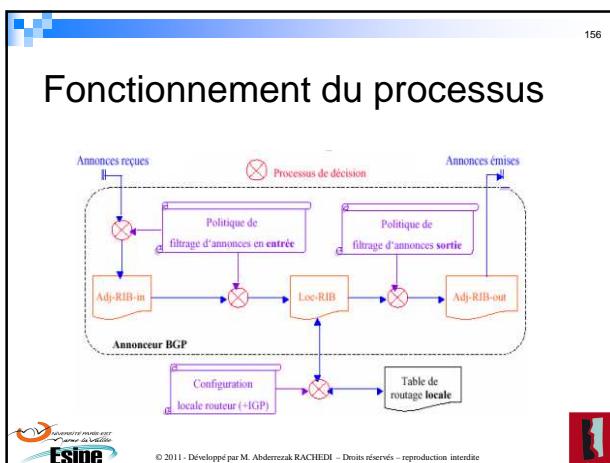
- UPDATE sert à échanger des informations de routage
 - Route à éliminer
 - Ensemble d'attributs de la route
 - Ensemble des réseaux accessibles (NLRI)
 - Chaque réseau est défini par (préfixe et longueur)
- Envoyer uniquement si il a un changement
- Active le processus BGP
 - Modification des RIB (Update, politique de routage, conf.)
 - Envoyer un message UPDATE vers les autres voisins

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite









157

Les attributs BGP

- Ils décrivent les **caractéristiques** associés à un préfixe particulier :
 - Leur transmission est soit **transitives** soit **non-transitive**
 - Certains sont **obligatoires**
- Ils servent à **sélectionner la meilleure route**
- BGP appelle l'**ensemble des préfixes partagent les même attributs** : Network Layer Reachability Information (NLRI)

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

158

Les attributs BGP

Value	Code	Reference
1	ORIGIN	[RFC1771]
2	AS_PATH	[RFC1771]
3	NEXT_HOP	[RFC1771]
4	MULTI_EXIT_DISC	[RFC1771]
5	LOCAL_PREF	[RFC1771]
6	ATOMIC_AGGREGATE	[RFC1771]
7	AGGREGATOR	[RFC1771]
8	COMMUNITY	[RFC1997]
9	ORIGINATOR_ID	[RFC2796]
10	CLUSTER_LIST	[RFC2796]
11	DPA	[Chen]
12	ADVERTISER	[RFC1863]
13	RTD_PATH / CLUSTER_ID	[RFC1863]
14	MP_REACH_NLRI	[RFC2283]
15	MP_UNREACH_NLRI	[RFC2283]
16	EXTENDED_COMMUNITIES	[Rosen]
...
255	Réservé pour le développement	

Les attributs les plus importants

Source IANA: <http://www.iana.org/assignments/bgp-parameters>

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

159

L'attribut « ORIGIN »

- Attribut Obligatoire**
- Indique l'**origine** de l'apprentissage de la route
- Trois valeurs
 - IGP – Route apprise de l'intérieur de l'AS
 - Préfixe obtenu avec une clause "network" (ex. : *network 35.0.0.0*)
 - EGP – Route apprise par un autre AS
 - Redistribué par un EGP
 - Incomplete – Route dont l'apprentissage n'est pas défini
 - Redistribué par un IGP (exemple : *redistribute ospf*)


IGP > EGP > INCOMPLETE

Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

160

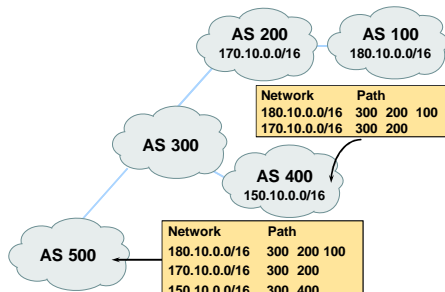
L'attribut « ASPATH »

- **Attribut Obligatoire**
- Utiliser uniquement par eBGP
- Donne la route sous forme d'une liste de *segments AS*
- Chaque routeur **ajoute** son numéro d'AS aux AS_Path des routes qu'il a apprises *avant de les ré-annoncées*
- L'AS_Path le **plus court** qu'il est choisit
- Détection de boucles
- Mise en œuvre de politiques


 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

161

L'attribut « ASPATH »

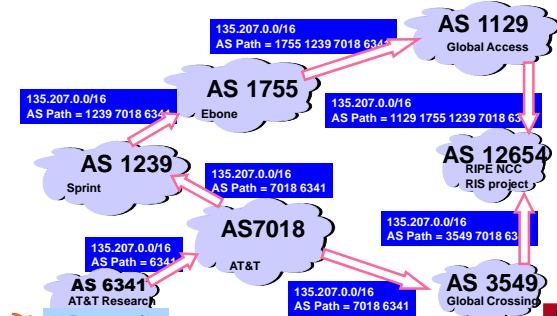



Network	Path
180.10.0.0/16	300 200 100
170.10.0.0/16	300 200
150.10.0.0/16	300 400

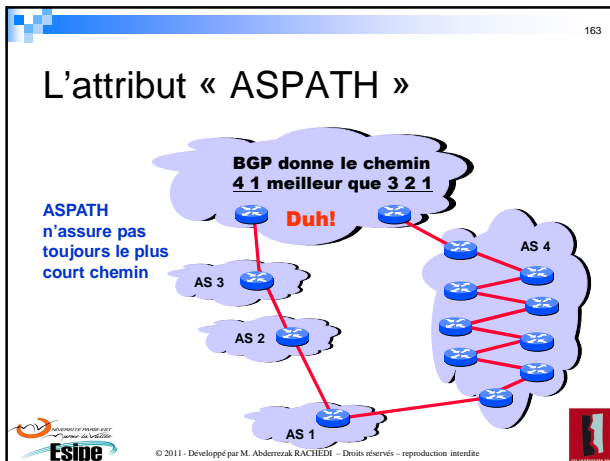
 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

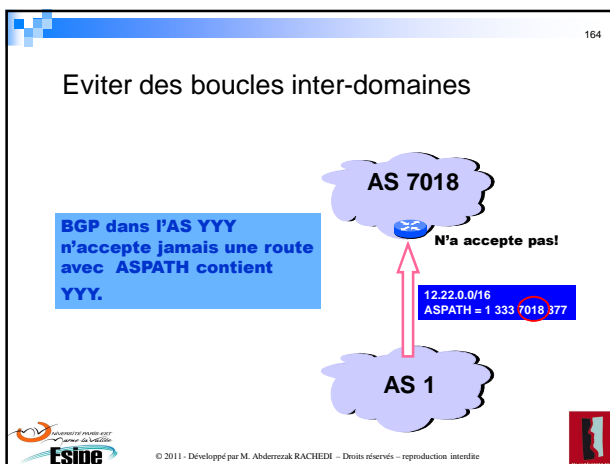
162

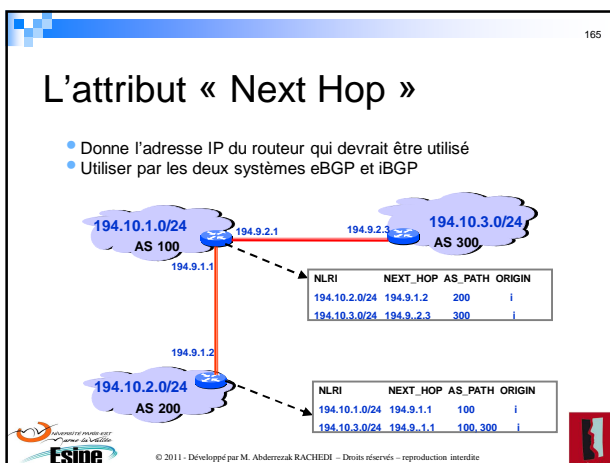
L'attribut « ASPATH »



 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite







166

L'attribut « Next Hop »

Network	Next-Hop	Path
160.10.0.0/16	192.20.2.1	100

- Prochain routeur pour joindre un réseau
- Dans une session eBGP c'est en général une adresse locale

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

167

L'attribut « Next Hop »

Network	Next-Hop	Path
150.10.0.0/16	192.10.1.1	200
160.10.0.0/16	192.10.1.1	200 100

- Prochain routeur pour joindre un réseau
- Dans une session eBGP c'est en général une adresse locale
- Le "next-hop" est mis à jour dans les sessions eBGP

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

168

L'attribut « Next Hop »

Network	Next-Hop	Path
150.10.0.0/16	192.10.1.1	200
160.10.0.0/16	192.10.1.1	200 100


- Le "next-hop" n'est pas modifié dans les sessions iBGP

© 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

169

L'attribut « Next Hop »


- Les adresses des "next-hops" doivent circuler dans l'IGP
- Recherche **réursive** des routes
- Permet de concevoir la topologie BGP **indépendamment** de la topologie physique du réseau
- En interne les bonnes décisions de routage sont faites par l'IGP

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

170

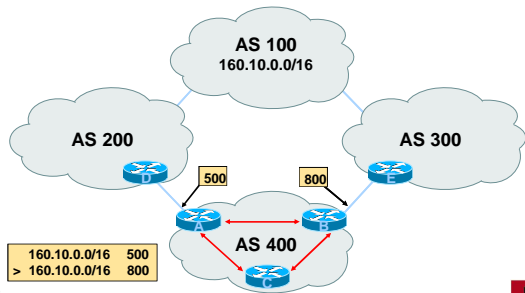
L'attribut «LOCAL PREFERENCE »


- **Attribut optionnel** et non transitif
- Utiliser pour **influencer le processus de sélection** du meilleur chemin
- Interprétation local à l'AS (**iBGP**)
⇒ Jamais annoncer en eBGP
- Valeur par défaut chez Cisco : **100**
- Prise en compte **avant** la longueur de l'AS_Path
- Le chemin avec la plus grande valeur de « Local_Pref » est sélectionné

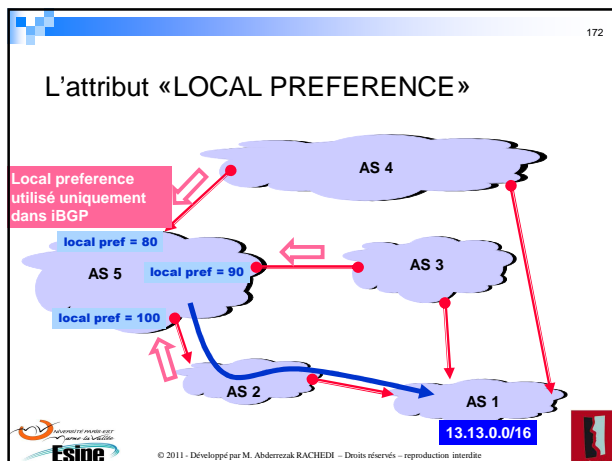
 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

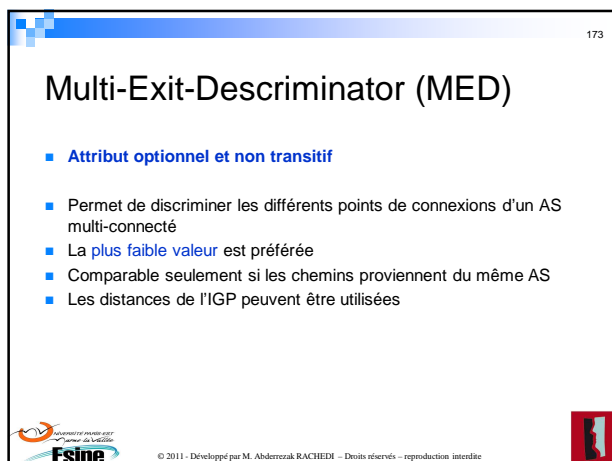
171

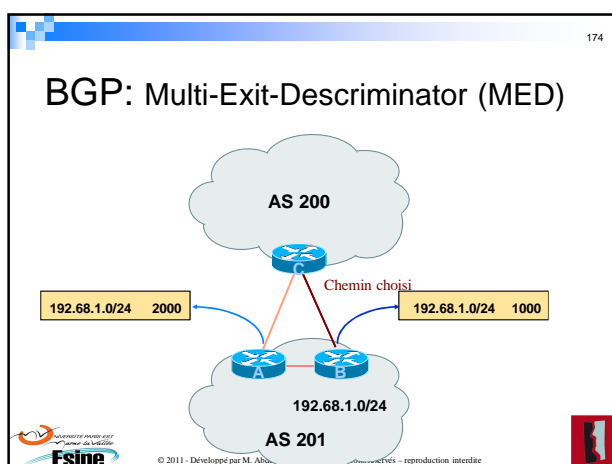
L'attribut «LOCAL PREFERENCE »



 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite








175

L'attribut « Poids (WEIGHT) »


- Attribut optionnel et non-transitif
- Spécifique Cisco utilisé lorsqu'il y a plus d'une route vers la même destination
- Attribut local à un routeur (non propagé ailleurs)
- Lorsqu'il y a plusieurs choix, on préfère la route dont le poids est le plus élevé
- Le poids est une valeur dans l'intervalle [0 – 65535]
- Valeur par défaut "32768" pour les chemins dont l'origine est le routeur et 0 pour les autres

 © 2011 - Développé par M. Abdelrazak RACHEDI - Droits réservés - reproduction interdite

176

L'attribut « Communautés »

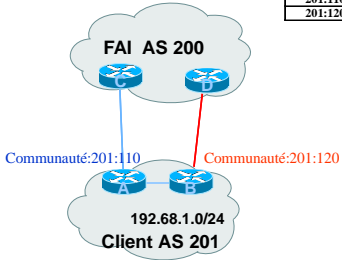
- Attribut optionnel et transitif
- Permettent de créer des **groupes de destinations**
- Regroupement par destination et identifier par un numéro (de 0 à 4×10^9)
- Il y a 3 attributs de communautés prédéfinis :
 - **No-Export** : La route n'est pas annoncée à des voisins eBGP
 - **No-Advertise** : La route n'est pas annoncée à n'importe quel réseau
 - **Internet** : Route est annoncée à toute la communauté Internet
- Une destination peut être membre de plusieurs communauté
- Attribut très flexible, car il permet de faire des choix avec des critères inter ou intra-AS


 © 2011 - Développé par M. Abdelrazak RACHEDI - Droits réservés - reproduction interdite

177

L'attribut « Communautés »

Communauté	Local Preference
201:110	110
201:120	120




 © 2011 - Développé par M. Abdelrazak RACHEDI - Droits réservés - reproduction interdite

178


Les attributs de route optionnels

- **ATOMIC_AGGREGATE :**
 - Indicateur d'agrégation
 - Quand des routes plus précises ne sont pas annoncées
- **AGGREGATOR :**
 - Donne l'AS qui a formé la route agrégée
 - L'adresse IP du routeur qui fait l'agrégation

 © 2011 - Développé par M. Abdelrazak RACHEDI - Droits réservés - reproduction interdite

179


BGP : Sélection des routes



Local Preference le plus grand Enforce relationships

ASPATH le plus court
MED le plus petit
i-BGP < e-BGP : route externe est préférée à une route interne Traffic engineering
Le plus proche voisin local (au sens IGP)


Router ID le plus petit

 © 2011 - Développé par M. Abdelrazak RACHEDI - Droits réservés - reproduction interdite

180

Sélection des routes


- La route doit être **synchronisée**
C'est à dire être dans la table de routage
- Le "Next-hop" doit être **joignable**
Il se trouve dans la table de routage
- Prendre la valeur la plus élevée pour **le poids** (weight)
Critère spécifique Cisco et local au routeur
- Choisir la préférence locale la plus élevée
Appliqué pour l'ensemble des routeurs de l'AS
- La route est d'origine locale
Via une commande BGP "redistribute" ou "network"

 © 2011 - Développé par M. Abdelrazak RACHEDI - Droits réservés - reproduction interdite

181

Sélection des routes


- Choisir le plus court chemin d'AS
en comptant le nombre d'AS dans l'attribut AS-Path
- Prendre l'origine de valeur la plus faible
IGP < EGP < INCOMPLETE
- Choisir le plus petit MED
pour des chemins en provenance d'un même AS
- Préférer une route Externe sur une route Interne
prendre la sortie la plus proche
- Choisir le "next-hop" le plus proche
Plus faible métrique IGP, donc plus proche de la sortie de l'AS
- Plus petit "Router-ID"
- Adresse IP du voisin la plus petite

 © 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite

182

Politique de routage


- Pourquoi ?
 - ☐ Pour envoyer le trafic vers des routes choisies
 - ☐ Filtrage de préfixes en entrée et sortie
 - ☐ Pour forcer le respect des accords Client-ISP
- Comment ?
 - ☐ Filtrage basé sur les AS – « filter list »
 - ☐ Filtrage basé sur les préfixes – « distribute list »
 - ☐ Modification d'attributs BGP – « route maps »

 © 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite

183

Politique de routage


- Elle peut influencer :
 - ☐ Le traitement des routes reçues
 - ☐ Le traitement des routes annoncées
 - ☐ L'interaction avec les IGP de l'AS
- En pratique :
 - ☐ Du filtrage de réseaux
 - ☐ Du filtrage de routes (AS_PATH)
 - ☐ Du manipulation d'attributs de routes

 © 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite

184

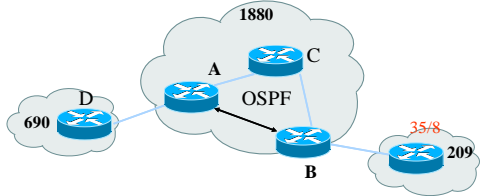
Synchronisation

- Spécifique IOS Cisco :
BGP n'annoncera pas une route **avant que l'ensemble des routeurs de l'AS ne l'ait apprise par un IGP**
- Désactiver la synchronisation si :
 - ☐ L'AS ne sert pas d'AS de **transit**,
ou
 - ☐ Tous les routeurs de transit exécutent BGP,
ou
 - ☐ iBGP est utilisé sur le cœur de réseau (backbone)


 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

185

Synchronisation




- C ne exécute pas BGP (**non-pervasive BGP**)
- A n'annoncera pas 35/8 à D tant qu'il ne l'aura pas appris par l'IGP
- Il faut désactiver la synchronisation pour éviter ce problème

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

186

Les difficultés


- **L'équilibrage de charge :**
 - ☐ BGP n'est pas conçu pour assurer l'équilibrage de charge
=> Il choisit et installe la meilleure route
 - ☐ Multi liens et Multi chemins
- **Le routage asymétrique :**
 - ☐ Impossible de contrôler le suivi du chemin
- **L'inter relation IGP et BGP**
 - ☐ Le protocole IGP distribue les informations de routage internes au domaine
 - ☐ Les routes ne sont jamais distribuées entre IGP et BGP

 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

187

Les difficultés


- Préfixe avec origines multiples
 - BGP recommande qu'un préfixe ait toujours pour **origine le même AS**
 - Les **BGP Multiple AS Origin (MOAS)** sont le résultat d'une erreur de configuration et peuvent créer des incidents de type **déni de service**
 - Si un routeur annonce un préfixe pour lequel il **n'assure pas réellement le transit**, ce dernier peut devenir inaccessible depuis tout ou une partie d'Internet (Incident YouTube de 2008 causé par Pakistan Telecom)
- Solution : Filtrage
 - Ces filtres sont alors mis à jour manuellement si les voisins annoncent de nouvelles routes => **la complexité de la gestion**
 - D'autres approches sont **S-BGP (Secure BGP)** et **soBGP (secure origin BGP)**
 - Le groupe de travail **Secure Inter-Domain Routing (SIDR)** de l'IETF travaille sur un système de validation de la source d'un préfixe appelé **Resource Public Key Infrastructure (RPKI)**

 © 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite

188

Les difficultés

- BGP est **sensible à l'oscillation rapide des routes (Route flap)**
 - Une route apparaît et disparaît, Change d'attributs
 - Consomme des ressources (bande passante, processeur), influence le routage sur Internet
- Les annonces des routes inaccessibles => recalcule des tables de routage des routeurs
- Une fonctionnalité nommée **damping** (ou **dampening**, **BGP Route Flap Damping**) vise à réduire les effets de l'oscillation de routes

 © 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite


189

Les difficultés

La fonction Route flap dempering :

- À chaque oscillation d'une route, le **damping** va accroître une **pénalité numérique** associée à cette route

- Supprime les routes qui **oscillent**
- Les routes au-dessus de « **suppress-limit** » ne sont pas annoncées
- Les routes au-dessous de « **reuse-limit** » sont annoncées

 © 2011 - Développé par M. Abdelrazzak RACHEDI - Droits réservés - reproduction interdite

190

Les difficultés

- **Vitesse de convergence**
 - BGP est un protocole de routage **lent** par rapport aux IGP
 - La vitesse de convergence de BGP dépend de plusieurs paramètres :
 - Le **hold time** qui détermine après combien de temps une session inactive (sans UPDATE ni KEEPALIVE) est fermée. (3 sec au minimum, 90 sec par défaut)
 - Le **nombre de préfixes** qui devront chacun faire l'objet de messages UPDATE
 - Les **performances du CPU** pour le traitement des messages UPDATE
 - Le **Minimum Route Advertisement Interval (MRAI)** qui définit l'intervalle de temps minimum entre deux annonces BGP pour le même ensemble de préfixes et vers le même voisin
La RFC propose un MRAI de 30 s pour eBGP et 5 s pour iBGP
 - L'utilisation de **BGP Protocol Independent Convergence (PIC)**, qui permet un reroutage rapide et indépendant du nombre de préfixe à rerouter

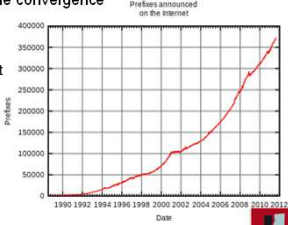
Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

191

Les difficultés

- **Croissance de la table de routage**
 - Avec le temps, les routeurs plus anciens n'ont plus les ressources mémoire ou CPU nécessaires au maintien d'une table complète
 - La taille de la table nuit à la vitesse de convergence

- Des efforts ont été entrepris pour réduire les préfixes en les agrégeant
- En 2011, le nombre de préfixes diffusés sur Internet est d'environ **360 000**

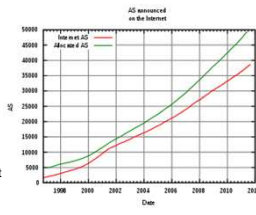


Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

192

Numéro d'AS vers 32 bits

- Les numéros d'AS sont codés sur **16 bits**, (64510 AS publics et le reste pour l'usage privé)
- En 2011, il restait environ **15000** AS libres
- Un épuisement complet des AS disponibles est estimé vers **septembre 2013**
- => Evolution du codage des numéros d'AS sur **32 bits** (RFC4893).
 - Le numéro d'AS est représenté par la notation x.y (les AS 1.x et 65535.65535 sont réservés)
 - Le nombre d'AS disponibles est à plus de **quatre milliards**.
 - L'assignation des AS sur 32 bits a débuté en **2007**




Esipe © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

193

Implémentation libre de BGP

- **GNU Zebra**
- **Quagga Software**, fork de GNU Zebra pour les systèmes dérivés d'Unix
- **OpenBGPD**, implémentation BGP, sous licence ISC, par l'équipe d'**OpenBSD**
- **BIRD**
- **XORP**, pour eXtensible Open Router Platform, sous licence BSD.

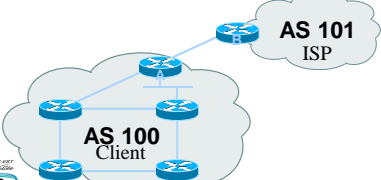
 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite


194

BGP et architecture de réseaux

AS "feuille" (stub AS)

- Situation ne nécessitant pas de BGP
- Route par défaut chez le FAI
- Le FAI annonce vos réseaux dans son AS
- La politique de routage de votre FAI est également la vôtre




 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite


195

BGP et architecture de réseaux

AS multi-raccordé (multi-homed)

- Les routeurs d'extrémité font du BGP
- Sessions IBGP entre ces routeurs
- Il faut redistribuer les routes apprises avec prudence dans l'IGP, ou bien utiliser une route par défaut



 © 2011 - Développé par M. Abdelrazek RACHEDI - Droits réservés - reproduction interdite

196

BGP et architecture de réseaux

Réseau d'un FAI

- iBGP permet de transporter les routes extérieures à l'AS
- Un IGP permet de gérer la topologie du réseau
- Un maillage complet iBGP est requis

© 2011 - Développé par M. Abdelhak RACHEDI - Droits réservés - reproduction interdite

197

Exemple

router bgp 200
 neighbor 172.16.0.4 remote-as 500
 neighbor 10.20.0.0 mask 255.255.0.0

router bgp 100
 neighbor 172.16.0.2 remote-as 500

router bgp 200
 neighbor 172.16.0.6 remote-as 500

router bgp 500
 neighbor 172.16.0.1 remote-as 100
 neighbor 10.50.0.0 mask 255.255.0.0

router bgp 500
 neighbor 172.16.0.5 remote-as 200
 neighbor 10.50.0.1 remote-as 500

router ospf 1
 network 10.50.0.0 0.0.255.255 area 0
 network 10.51.0.0 0.0.255.255 area 0
 redistribute bgp 500

Numéro de l'AS
 neighbor 172.16.0.1 remote-as 100
 neighbor 10.50.0.21 remote-as 500
 network 10.50.0.0 mask 255.255.0.0
 et leur annonce les routes de son AS.

Toutes les routes apprises de BGP (dont les routes des autres AS) sont injectées dans OSPF afin qu'elles soient diffusées dans l'AS 500.

référé - reproduction interdite

198

Exercices

- BGP est un protocole de routage :
 - Par vecteur de distance
 - Par état de liaison
 - Les deux (1 et 2)
 - Aucun des deux (ni 1 ni 2)
- BGP, OSPF, IGRP, EIGRP et RIP sont des protocoles de la couche
 - 4, 3, 5, 6, - Aucune des couches précédentes
 - BGP (5), OSPF, IGRP, EIGRP et RIPv (3)
- Combien de type peering BGP existe-il ? Expliquer la différence de chaque type
 - Peering : Customer-Provider peering et Shared-Cost peering
- S'il existe plusieurs routes entre deux AS, comment les routeurs BGP de chaque AS doivent calculer la meilleure route vers l'autre AS ?

A l'aide des attributs : Multi_Exit_Disc et Local_Pref

© 2011 - Développé par M. Abdelhak RACHEDI - Droits réservés - reproduction interdite
