heig-vd

Haute Ecole d'Ingénierie et de Gestion
du Canton de Vaud

Chapitre I

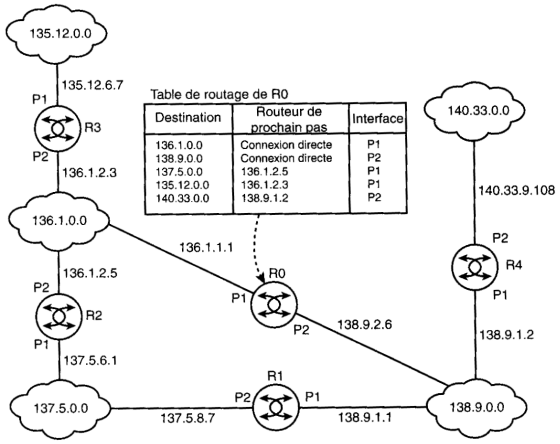
La couche réseau

§ 5.2 Tutorial IBM + support +
routage IP-Luc Saccavini

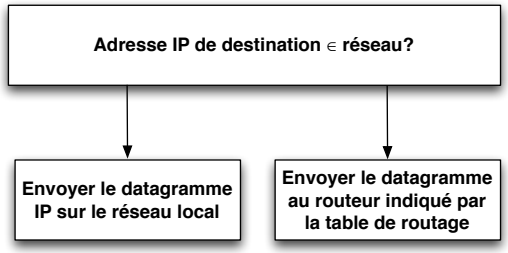
Routage

- La fonction principale de la couche réseau est l'acheminement des datagrammes
  - Les routeurs utilisent des tables de routage pour déterminer le prochain saut
  - L'information dans les tables de routage vient
    - De protocoles de routage chargés de trouver des chemins 'optimaux'
    - D'un configuration statique par l'administrateur
- IP
  - Forwarding de datagramme sur la base de tables de routage
- Protocoles de routage
  - Déterminer les chemins optimaux et créer les tables de routage

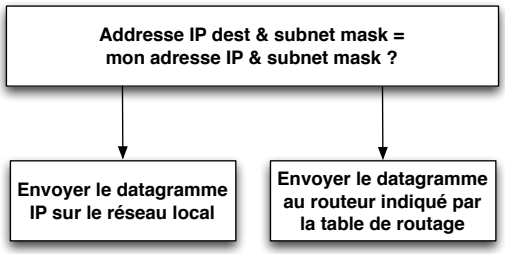
Table de routage



Algorithme de routage (sans sous-réseaux)

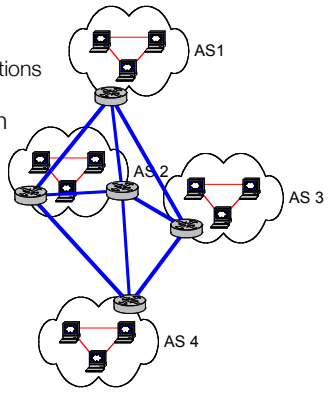


Algorithme de routage (avec sous-réseaux)



Niveaux de routage

- L'Internet a une structure hiérarchique
  - Interconnexion de réseaux
    - Les réseaux sont administrés par des organisations distinctes
  - Un ou plusieurs réseaux regroupés forment un Système Autonome (AS)
- Deux types de routage
  - Routage à l'intérieur d'un AS
    - Interior Gateway Protocol (IGP)
      - Cherche des routes optimales
      - RIP, OSPF
  - Routage entre les AS
    - Exterior Gateway Protocol (EGP)
      - Utilisation de règles qui limitent les routes
      - BGP



Algorithmes de routage

Problème
- Trouver la « meilleur route » vers une destination
- Métriques
  - Nombre de sauts, capacités de liens, trafic, délai

- Inondation (flooding)
  - Similaire au routage par la source dans Token Ring
- Chemin le plus court
  - Statique : topologie et métriques fixes
  - Dynamique : adaptation aux changements de la topologie
    - Vecteur de distance - connaissance locale des métriques
      - RIP, (IGRP)
    - État de liaison - connaissance globale des métriques
      - OSPF, (PNNI)

Routage statique

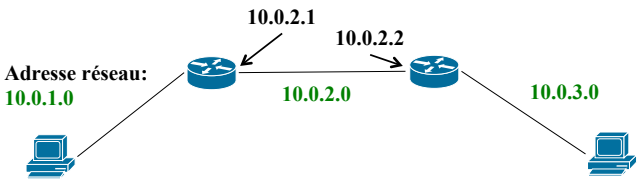
- Configuré manuellement par l'administrateur de réseau.
- Les routeurs vont acheminer les paquets à des ports déterminés à partir des routes.
- Aucune communication entre routeurs
- Désavantages:
  - Coordination nécessaire, fastidieux pour des grands réseaux (>10 routeurs)
  - Aucune adaptation dynamique
  - Ne fonctionne que pour des topologies simples, pas possible de gérer des routes redondantes

Routage statique (2)

- Quand utilise-t-on le routage statique?
  - Définir manuellement une route par défault (poste de travail)
  - Définir une route qui n'est pas donnée par les protocoles de routage
  - Quand nous voulons faire passer du trafic par certains endroits au sein d'une topologie complexe
  - Pour augmenter la sécurité du réseau en autorisant la communication spécifiquement entre certains sous-réseaux
  - Utilisation efficace des ressources (pas besoin d'utiliser de la LB pour les messages de routage)

Routage statique (3)

- Exemple de routage statique



Commandes Cisco:

Routeur R1: >IP route 10.0.3.0 255.255.255.0 10.0.2.2
Routeur R2: >IP route 10.0.1.0 255.255.255.0 10.0.2.1

Routage dynamique

- Caractéristiques
  - Adaptatif (si des liaisons ou des routeurs lâchent)
  - Configuration relativement simple (peu dépendant du nombre de routeurs)
- Objectifs
  - Optimisation: meilleures routes
  - Elimination des boucles de routage
  - Consommation de largeur de bande faible (minimiser les messages échangés)
  - Convergence et reconfiguration rapide
  - Simplicité de configuration

Protocoles de routage dynamiques

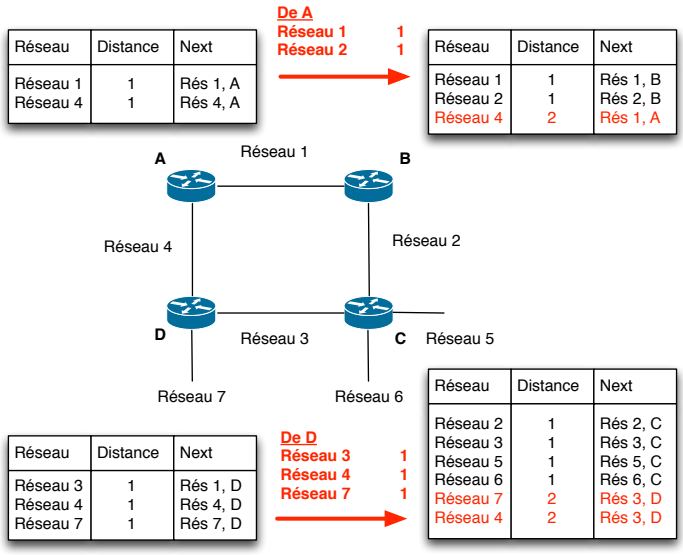
- Protocoles intérieurs (IGP)
  - A vecteur de distance: RIP, IGRP
  - A état de liens: OSPF, IS-IS
  - Une autorité d'administration, taille < 100 routeurs
- Protocoles extérieurs (EGP)
  - EGP, BGP
  - Valables sur tout le réseau Internet
  - Séparation en systèmes autonomes (AS)

Préambule: Routage dynamique à vecteur de distance, Bellman-Ford

- Chaque routeur maintient une table de routage
  - Pour toutes les destinations : Destination, Nœud suivant, Distance
    - Distance: Nombre de sauts, délai, …
    - Distance peut être infinie si aucune route n'est connue
- Le routeur connaît la distance qui le sépare de ses voisins directs
- Les mises à jour (updates) se font directement entre voisins
  - Les voisins échangent les routes connues
  - Périodiquement ou quand la table change (appelé "triggered update")
- Algorithme de Bellman-Ford distribué
  - Le routeur X connaît la distance d(X,Y) vers ses voisins Y
  - Initialement, la distance D(X,n) vers la destination n est
    - D(X,n) = 0, si X est directement connecté au réseau n
    - D(X,n) = ∞ pour toutes les autres destinations
  - Le routeur X reçoit le vecteur des distances {D(Y,n)} du voisin Y vers tous les n
  - Le routeur X calcule la meilleure distance vers la destination n

$$D(X,n) = \min_{\text{voisins Y}} \left(d(X,Y) + D(Y,n) \right)$$

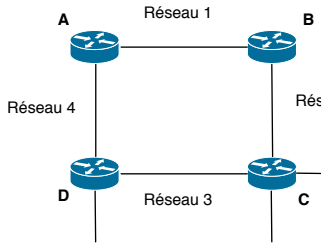
Exemple



Exemple

Réseau	Distance	Next
Réseau 1	1	Rés 1, A
Réseau 4	1	Rés 4, A

Réseau	Distance	Next
Réseau 1	1	Rés 1, B
Réseau 2	1	Rés 2, B
Réseau 3	2	Rés 2, C
Réseau 4	2	Rés 1, A
Réseau 5	2	Rés 2, C
Réseau 6	2	Rés 2, C
Réseau 7	3	Rés 2, C



De C

Réseau 2	**1**
Réseau 3	**1**
Réseau 5	**1**
Réseau 6	**1**
Réseau 4	**2**
Réseau 7	**2**

Réseau	Distance	Next
Réseau 3	1	Rés 1, D
Réseau 4	1	Rés 4, D
Réseau 7	1	Rés 7, D

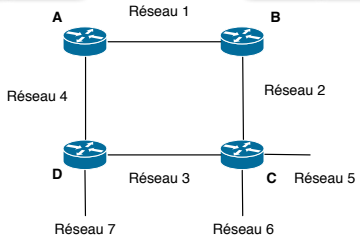
Réseau	Distance	Next
Réseau 2	1	Rés 2, C
Réseau 3	1	Rés 3, C
Réseau 5	1	Rés 5, C
Réseau 6	1	Rés 6, C
Réseau 7	2	Rés 3, D
Réseau 4	2	Rés 3, D

Exemple

après convergence

Réseau	Distance	Next
Réseau 1	1	Rés 1, A
Réseau 2	1	Rés 1, B
Réseau 3	2	Rés 4, D
Réseau 4	1	Rés 4, A
Réseau 5	3	Rés 4, D
Réseau 6	3	Rés 4, D
Réseau 7	2	Rés 4, D

Réseau	Distance	Next
Réseau 1	1	Rés 1, B
Réseau 2	1	Rés 2, B
Réseau 3	2	Rés 2, C
Réseau 4	2	Rés 1, A
Réseau 5	2	Rés 2, C
Réseau 6	2	Rés 2, C
Réseau 7	3	Rés 2, C



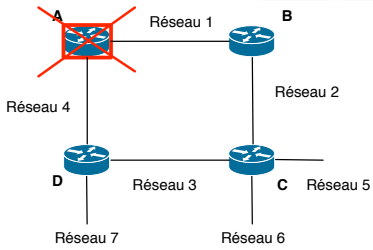
Réseau	Distance	Next
Réseau 1	2	Rés 4, A
Réseau 2	2	Rés 3, C
Réseau 3	1	Rés 3, D
Réseau 4	1	Rés 4, D
Réseau 5	2	Rés 3, C
Réseau 6	2	Rés 3, C
Réseau 7	1	Rés 7, D

Réseau	Distance	Next
Réseau 1	2	Rés 2, B
Réseau 2	1	Rés 2, C
Réseau 3	1	Rés 3, C
Réseau 4	2	Rés 3, C
Réseau 5	1	Rés 5, C
Réseau 6	1	Rés 6, C
Réseau 7	2	Rés 3, D

Exemple 1

Défaillance du routeur A

Réseau	Distance	Next
Réseau 1	1	Rés 1, B
Réseau 2	1	Rés 2, B
Réseau 3	2	Rés 2, C
Réseau 4	2	Rés 1, A
Réseau 5	2	Rés 2, C
Réseau 6	2	Rés 2, C
Réseau 7	3	Rés 2, C



Réseau	Distance	Next
Réseau 1	2	Rés 4, A
Réseau 2	2	Rés 3, C
Réseau 3	1	Rés 3, D
Réseau 4	1	Rés 4, D
Réseau 5	2	Rés 3, C
Réseau 6	2	Rés 3, C
Réseau 7	1	Rés 7, D

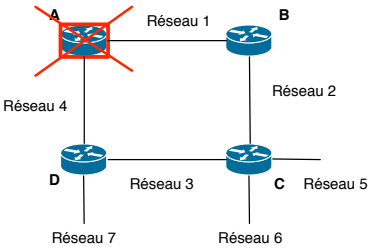
Réseau	Distance	Next
Réseau 1	2	Rés 2, B
Réseau 2	1	Rés 2, C
Réseau 3	1	Rés 3, C
Réseau 4	2	Rés 3, C
Réseau 5	1	Rés 5, C
Réseau 6	1	Rés 6, C
Réseau 7	2	Rés 3, D

Exemple 1

Défaillance du routeur A

Timeout

Réseau	Distance	Next
Réseau 1	1	Rés 1, B
Réseau 2	1	Rés 2, B
Réseau 3	2	Rés 2, C
Réseau 4	2	Rés 1, A
Réseau 5	2	Rés 2, C
Réseau 6	2	Rés 2, C
Réseau 7	3	Rés 2, C



Réseau	Distance	Next
Réseau 1	2	Rés 4, A
Réseau 2	2	Rés 3, C
Réseau 3	1	Rés 3, D
Réseau 4	1	Rés 4, D
Réseau 5	2	Rés 3, C
Réseau 6	2	Rés 3, C
Réseau 7	1	Rés 7, D

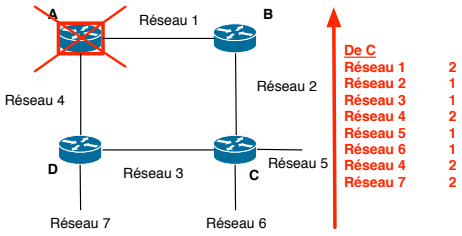
Timeout

Réseau	Distance	Next
Réseau 1	2	Rés 2, B
Réseau 2	1	Rés 2, C
Réseau 3	1	Rés 3, C
Réseau 4	2	Rés 3, C
Réseau 5	1	Rés 5, C
Réseau 6	1	Rés 6, C
Réseau 7	2	Rés 3, D

Exemple 1

Après la défaillance

Réseau	Distance	Next
Réseau 1	1	Rés 1, B
Réseau 2	1	Rés 2, B
Réseau 3	2	Rés 2, C
Réseau 4	3	Rés 2, A
Réseau 5	2	Rés 2, C
Réseau 6	2	Rés 2, C
Réseau 7	3	Rés 2, C



De C

Réseau 1	2
Réseau 2	1
Réseau 3	1
Réseau 4	2
Réseau 5	1
Réseau 6	1
Réseau 4	2
Réseau 7	2

Réseau	Distance	Next
Réseau 1	3	Rés 3, C
Réseau 2	2	Rés 3, C
Réseau 3	1	Rés 3, D
Réseau 4	1	Rés 4, D
Réseau 5	2	Rés 3, C
Réseau 6	2	Rés 3, C
Réseau 7	1	Rés 7, D

De C

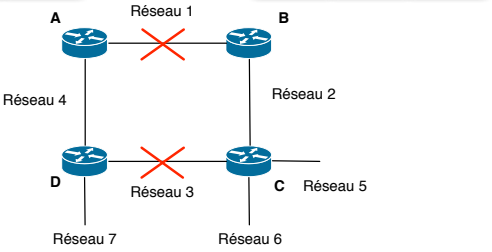
Réseau 1	2
Réseau 2	1
Réseau 3	1
Réseau 4	2
Réseau 5	1
Réseau 6	1
Réseau 4	2
Réseau 7	2

Réseau	Distance	Next
Réseau 1	2	Rés 2, B
Réseau 2	1	Rés 2, C
Réseau 3	1	Rés 3, C
Réseau 4	2	Rés 3, C
Réseau 5	1	Rés 5, C
Réseau 6	1	Rés 6, C
Réseau 7	2	Rés 3, D

Exemple 2 (l'algorithme a convergé) Défaillance de deux liens simultanément

Réseau	Distance	Next
Réseau 1	1	Rés 1, A
Réseau 2	2	Rés 1, B
Réseau 3	2	Rés 4, D
Réseau 4	1	Rés 4, A
Réseau 5	3	Rés 4, D
Réseau 6	3	Rés 4, D
Réseau 7	2	Rés 4, D

Réseau	Distance	Next
Réseau 1	1	Rés 1, B
Réseau 2	1	Rés 2, B
Réseau 3	2	Rés 2, C
Réseau 4	2	Rés 1, A
Réseau 5	2	Rés 2, C
Réseau 6	2	Rés 2, C
Réseau 7	3	Rés 2, C



Réseau	Distance	Next
Réseau 1	2	Rés 4, A
Réseau 2	2	Rés 3, C
Réseau 3	1	Rés 3, D
Réseau 4	1	Rés 4, D
Réseau 5	2	Rés 3, C
Réseau 6	2	Rés 3, C
Réseau 7	1	Rés 7, D

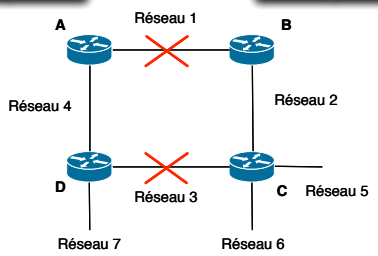
Réseau	Distance	Next
Réseau 1	2	Rés 2, B
Réseau 2	1	Rés 2, C
Réseau 3	1	Rés 3, C
Réseau 4	2	Rés 3, C
Réseau 5	1	Rés 5, C
Réseau 6	1	Rés 6, C
Réseau 7	2	Rés 3, D

Exemple 2

D détecte la défaillance

Réseau	Distance	Next
Réseau 1	1	Rés 1, A
Réseau 2	2	Rés 1, B
Réseau 3	2	Rés 4, D
Réseau 4	1	Rés 4, A
Réseau 5	3	Rés 4, D
Réseau 6	3	Rés 4, D
Réseau 7	2	Rés 4, D

Réseau	Distance	Next
Réseau 1	1	Rés 1, B
Réseau 2	1	Rés 2, B
Réseau 3	2	Rés 2, C
Réseau 4	2	Rés 1, A
Réseau 5	2	Rés 2, C
Réseau 6	2	Rés 2, C
Réseau 7	3	Rés 2, C



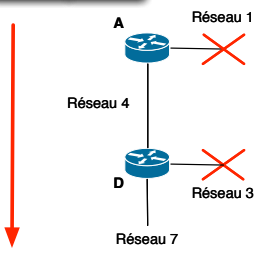
Réseau	Distance	Next
Réseau 1	2	Rés 4, A
Réseau 2	∞	NR
Réseau 3	∞	NR
Réseau 4	1	Rés 4, D
Réseau 5	∞	NR
Réseau 6	∞	NR
Réseau 7	1	Rés 7, D

Réseau	Distance	Next
Réseau 1	2	Rés 2, B
Réseau 2	1	Rés 2, C
Réseau 3	1	Rés 3, C
Réseau 4	2	Rés 3, C
Réseau 5	1	Rés 5, C
Réseau 6	1	Rés 6, C
Réseau 7	2	Rés 3, D

Exemple 2
Focus sur A-D

Réseau	Distance	Next
Réseau 1	1	Rés 1, A
Réseau 2	2	Rés 1, B
Réseau 3	2	Rés 4, D
Réseau 4	1	Rés 4, A
Réseau 5	3	Rés 4, D
Réseau 6	3	Rés 4, D
Réseau 7	2	Rés 4, D

Réseau	Distance
Réseau 1	1
Réseau 2	2
Réseau 3	2
Réseau 4	1
Réseau 5	3
Réseau 6	3
Réseau 7	2

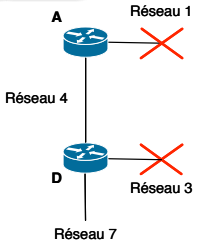


Réseau	Distance	Next
Réseau 1	2	Rés 4, A
Réseau 2	3	Rés 4, A
Réseau 3	3	Rés 4, A
Réseau 4	1	Rés 4, D
Réseau 5	4	Rés 4, A
Réseau 6	4	Rés 4, A
Réseau 7	1	Rés 7, D

Exemple 2

A détecte la défaillance

Réseau	Distance	Next
Réseau 1	∞	NC
Réseau 2	∞	NC
Réseau 3	2	Rés 4, D
Réseau 4	1	Rés 4, A
Réseau 5	3	Rés 4, D
Réseau 6	3	Rés 4, D
Réseau 7	2	Rés 4, D

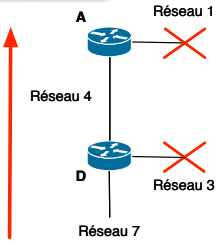


Réseau	Distance	Next
Réseau 1	2	Rés 4, A
Réseau 2	3	Rés 4, A
Réseau 3	3	Rés 4, A
Réseau 4	1	Rés 4, D
Réseau 5	4	Rés 4, A
Réseau 6	4	Rés 4, A
Réseau 7	1	Rés 7, D

Exemple 2

Réseau	Distance	Next
Réseau 1	3	Rés 4, D
Réseau 2	4	Rés 4, D
Réseau 3	4	Rés 4, D
Réseau 4	1	Rés 4, A
Réseau 5	5	Rés 4, D
Réseau 6	5	Rés 4, D
Réseau 7	2	Rés 4, D

Réseau	Distance
Réseau 1	2
Réseau 2	3
Réseau 3	3
Réseau 4	1
Réseau 5	4
Réseau 6	4
Réseau 7	1



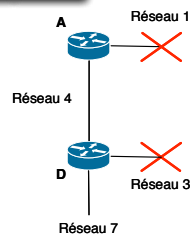
Réseau	Distance	Next
Réseau 1	2	Rés 4, A
Réseau 2	3	Rés 4, A
Réseau 3	3	Rés 4, A
Réseau 4	1	Rés 4, D
Réseau 5	4	Rés 4, A
Réseau 6	4	Rés 4, A
Réseau 7	1	Rés 7, D

Exemple 2

Comptage à l'infini

Réseau	Distance	Next
Réseau 1	3	Rés 4, D
Réseau 2	4	Rés 4, D
Réseau 3	4	Rés 4, D
Réseau 4	1	Rés 4, A
Réseau 5	5	Rés 4, D
Réseau 6	5	Rés 4, D
Réseau 7	2	Rés 4, D

Réseau	Distance
Réseau 1	3
Réseau 2	4
Réseau 3	4
Réseau 4	1
Réseau 5	5
Réseau 6	5
Réseau 7	2

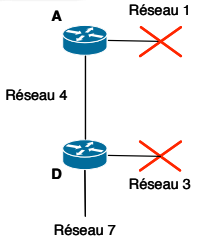


Réseau	Distance	Next
Réseau 1	4	Rés 4, A
Réseau 2	5	Rés 4, A
Réseau 3	5	Rés 4, A
Réseau 4	1	Rés 4, D
Réseau 5	6	Rés 4, A
Réseau 6	6	Rés 4, A
Réseau 7	1	Rés 7, D

Exemple 2

Réseau	Distance	Next
Réseau 1	5	Rés 4, D
Réseau 2	6	Rés 4, D
Réseau 3	6	Rés 4, D
Réseau 4	1	Rés 4, A
Réseau 5	7	Rés 4, D
Réseau 6	7	Rés 4, D
Réseau 7	2	Rés 4, D

Réseau	Distance
Réseau 1	4
Réseau 2	5
Réseau 3	5
Réseau 4	1
Réseau 5	6
Réseau 6	6
Réseau 7	1

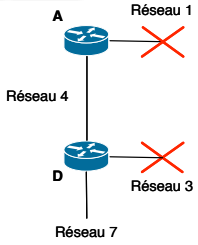


Réseau	Distance	Next
Réseau 1	4	Rés 4, A
Réseau 2	5	Rés 4, A
Réseau 3	5	Rés 4, A
Réseau 4	1	Rés 4, D
Réseau 5	6	Rés 4, A
Réseau 6	6	Rés 4, A
Réseau 7	1	Rés 7, D

Exemple 2

Réseau	Distance	Next
Réseau 1	5	Rés 4, D
Réseau 2	6	Rés 4, D
Réseau 3	6	Rés 4, D
Réseau 4	1	Rés 4, A
Réseau 5	7	Rés 4, D
Réseau 6	7	Rés 4, D
Réseau 7	2	Rés 4, D

Réseau	Distance
Réseau 1	5
Réseau 2	6
Réseau 3	6
Réseau 4	1
Réseau 5	7
Réseau 6	7
Réseau 7	2



Réseau	Distance	Next
Réseau 1	6	Rés 4, A
Réseau 2	7	Rés 4, A
Réseau 3	7	Rés 4, A
Réseau 4	1	Rés 4, D
Réseau 5	8	Rés 4, A
Réseau 6	8	Rés 4, A
Réseau 7	1	Rés 7, D

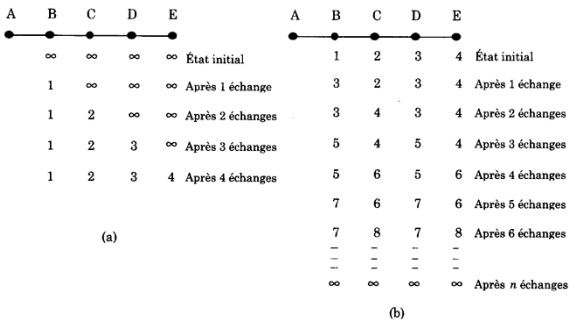
Propagation des bonnes nouvelles sur un réseau linéaire

- Une meilleure route se propage rapidement
- Exemple simple :
  - Réseau linéaire (distance à des nœuds et pas des réseaux, pour simplifier)
  - Distance: nombre de sauts
  - Nœud A vient de démarrer

∞	∞	∞	∞	État initial
1	∞	∞	∞	Après 1 échange
1	2	∞	∞	Après 2 échanges
1	2	3	∞	Après 3 échanges
1	2	3	4	Après 4 échanges

Propagation de mauvaises nouvelles

- Après une panne, le routage converge très lentement
- Exemple :
  - Lien entre A et B tombe en panne



A	B	C	D	E	
∞	∞	∞	∞	∞	État initial
1	∞	∞	∞	∞	Après 1 échange
1	2	∞	∞	∞	Après 2 échanges
1	2	3	∞	∞	Après 3 échanges
1	2	3	4	∞	Après 4 échanges

(a)

A	B	C	D	E	
1	2	3	4	4	État initial
3	2	3	4	4	Après 1 échange
3	4	3	4	4	Après 2 échanges
5	4	5	4	6	Après 3 échanges
5	6	5	6	6	Après 4 échanges
7	6	7	6	8	Après 5 échanges
7	8	7	8	8	Après 6 échanges
∞	∞	∞	∞	∞	Après *n* échanges

(b)

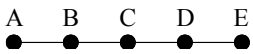
  - ➢ Problème de la valeur infinie

Leçons à tirer du deuxième exemple

- Le coût des réseaux 1,2,3,5,6 augmente sans limites (comptage à l'infini)
  - La vraie valeur des coûts est infinie
  - Ceci est un comportement attendu de l'algorithme (Bellman-Ford)
- Convergence vers un état stable si nous fixons
  - ∞= grand nombre (RIP: ∞= 16)
- La convergence reste très lente. Il faut prendre des mesures…

Route poisoning et Split horizon

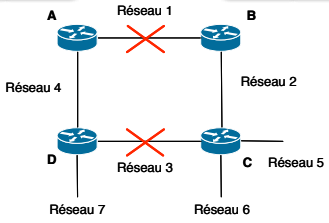
- Pratiquement
  - Il faut essayer d'éviter de compter jusqu'à l'infini
  - Il faut essayer d'éviter de faire du ping-pong dans des boucles
- Retour empoisoné (Route poisoning)
  - Si A détecte une route inaccessible vers X, il va envoyer « distance=∞ » à tous ses voisins
  - Lorsqu'une distance ∞ est reçue pour X, chaque mise à jour à propos de X est ignorée pendant un certain temps
- Horizon éclaté (Split horizon)
  - La distance vers une destination n'est pas annoncée au nœud suivant dans cette direction
    - Exemple
      - C annonce à D une distance D(C,A) = 2
      - C annonce à B une distance D(C,A) = ∞



Exemple 2 avec retour empoisoné

Réseau	Distance	Next
Réseau 1	1	Rés 1, A
Réseau 2	2	Rés 1, B
Réseau 3	2	Rés 4, D
Réseau 4	1	Rés 4, A
Réseau 5	3	Rés 4, D
Réseau 6	3	Rés 4, D
Réseau 7	2	Rés 4, D

Réseau	Distance	Next
Réseau 1	1	Rés 1, B
Réseau 2	1	Rés 2, B
Réseau 3	2	Rés 2, C
Réseau 4	2	Rés 1, A
Réseau 5	2	Rés 2, C
Réseau 6	2	Rés 2, C
Réseau 7	3	Rés 2, C



Réseau	Distance	Next
Réseau 1	2	Rés 4, A
Réseau 2	∞	NR
Réseau 3	∞	NR
Réseau 4	1	Rés 4, D
Réseau 5	∞	NR
Réseau 6	∞	NR
Réseau 7	1	Rés 7, D

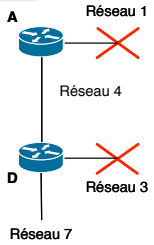
Réseau	Distance	Next
Réseau 1	2	Rés 2, B
Réseau 2	1	Rés 2, C
Réseau 3	1	Rés 3, C
Réseau 4	2	Rés 3, C
Réseau 5	1	Rés 5, C
Réseau 6	1	Rés 6, C
Réseau 7	2	Rés 3, D

Exemple 2

D détecte la défaillance

Réseau	Distance	Next
Réseau 1	1	Rés 1, A
Réseau 2	2	Rés 1, B
Réseau 3	2	Rés 4, D
Réseau 4	1	Rés 4, A
Réseau 5	3	Rés 4, D
Réseau 6	3	Rés 4, D
Réseau 7	2	Rés 4, D

Réseau	Distance
Réseau 1	2
Réseau 2	∞
Réseau 3	∞
Réseau 4	1
Réseau 5	∞
Réseau 6	∞
Réseau 7	1



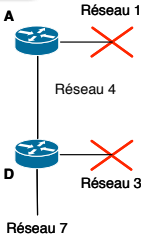
Réseau	Distance	Next
Réseau 1	2	Rés 4, A
Réseau 2	∞	NR
Réseau 3	∞	NR
Réseau 4	1	Rés 4, D
Réseau 5	∞	NR
Réseau 6	∞	NR
Réseau 7	1	Rés 7, D

Exemple 2

A détecte la défaillance

Réseau	Distance	Next
Réseau 1	∞	NR
Réseau 2	∞	NR
Réseau 3	∞	NR
Réseau 4	1	Rés 4, A
Réseau 5	∞	NR
Réseau 6	∞	NR
Réseau 7	2	Rés 4, D

Réseau	Distance
Réseau 1	∞
Réseau 2	∞
Réseau 3	∞
Réseau 4	1
Réseau 5	∞
Réseau 6	∞
Réseau 7	2

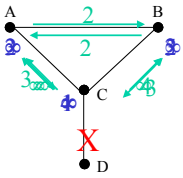


Réseau	Distance	Next
Réseau 1	∞	NR
Réseau 2	∞	NR
Réseau 3	∞	NR
Réseau 4	1	Rés 4, D
Réseau 5	∞	NR
Réseau 6	∞	NR
Réseau 7	1	Rés 7, D

Après un certain temps (timeout) les destinations = ∞ sont éliminées

Solutions pour accélérer la convergence

- Ne fonctionne pas toujours
  - Exemple:
    - Distance vers D
    - Lien C – D tombe en panne

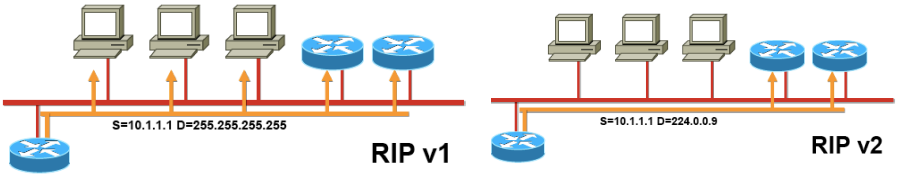


Protocole RIP

- RIP : Routing Information Protocol
  - Protocole de routage par vecteur de distance
  - Métrique: nombre de sauts (routeurs à traverser)
    - Valeur 'infinie' : 16 sauts
    - Utilise l'horizon éclaté (métrique max: 15 pour les routes des voisins) ou le retour empoisonné (distance infinie pour les routes des voisins)
  - Temps de convergence: quelques minutes
  - Basé sur l'algorithme de Bellman-Ford pour sélectionner les meilleurs routes
- Encore utilisé dans de petits réseaux (<15 routeurs)
  - Facile à configurer
- Version améliorée: RIP2

Historique et standardisation de RIP

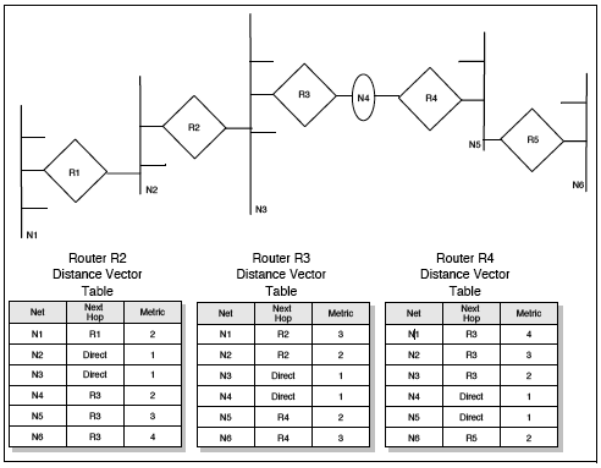
- RIPv1: RFC 1058 (1988)
- RIPv2: RFC 1387, RFC 1388, RFC 1723 (1994)
  - Routage CIDR
  - VLSM and route summarization
  - Authentification des routeurs
  - Diffusion multicast plutôt que broadcast



- RIPng (pour IPv6): RFC 2080, RFC 2453 (1998)

Exemple avec RIP

- Les routeurs envoient leurs tables de routage à leurs voisins
- Note: Next hop = R1 signifie « adresse IP de R1 »



Router R2 Distance Vector Table

Net	Next Hop	Metric
N1	R1	2
N2	Direct	1
N3	Direct	1
N4	R3	2
N5	R3	3
N6	R3	4

Router R3 Distance Vector Table

Net	Next Hop	Metric
N1	R2	3
N2	R2	2
N3	Direct	1
N4	Direct	1
N5	R4	2
N6	R4	3

Router R4 Distance Vector Table

Net	Next Hop	Metric
N1	R3	4
N2	R3	3
N3	R3	2
N4	Direct	1
N5	Direct	1
N6	R5	2

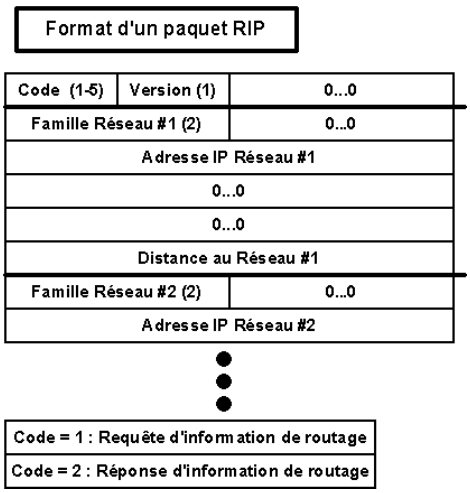
Source: tutorial IBM

Paquets RIP

- Types de paquets envoyés
  - Requêtes RIP: paquets envoyés aux voisins pour leur demander leur table de routage (vecteur de distance)
  - Réponses RIP: paquets envoyés par un routeur pour annoncer les informations qu'il connaît (tout ou une partie de la table de routage)
    - Automatiquement envoyée toutes les 30 secondes
    - Ou envoyée suite à une requête d'un voisin

Format des paquets RIPv1

- Paquets UDP (port 520 pour IPv4 et 521 pour IPv6)
- Version: version de RIP (2 car 1 n'est plus utilisée)
- Famille Réseau: X'0002' pour IP
- Taille maximale: 512 octets (25 entrées). Sinon plusieurs paquets
- Distance au réseau: métrique, de 1 à 15.
- Route par défaut: 0.0.0.0

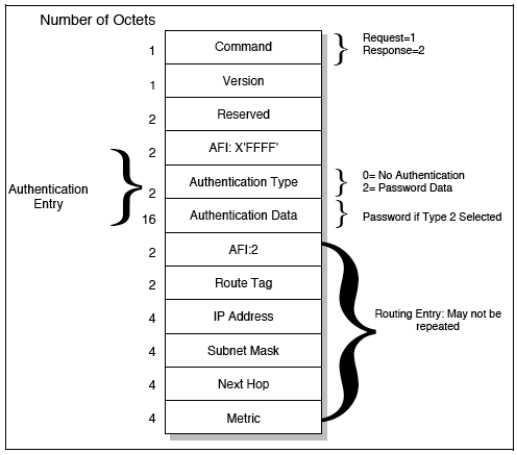


Minuteurs de RIP

- Routing-update
  - 30 secondes +/- 0 à 5 secondes
  - Période maximale entre deux annonces
- Route-timeout
  - 180 secondes
  - Durée de vie associée à chacune des routes apprises par RIP. Après ce temps la route est invalide. Elle sera réellement effacée après le temps “route-flush”
- Route-flush
  - 120 secondes
  - Périodicité du nettoyage de la tables de routage de RIP. Les routes invalides sont effacées.

Sécurité de RIPv2

- On veut se prémunir
  - Des routeurs intrus
  - De fausses informations (erreurs de paramétrage, mauvaises routes par défaut)
- Authentification MD5
  - Cisco (secret partagé: key-string: class):
    - Ip rip authentication mode md5
    - Ip rip authentication key-chain name-of-chain



Route tag: distinction entre route interne (même AS) et externe
Subnet mask: Supporte VLSM!

RIPng

- RIP est une adaptation de RIPv2 pour IPv6. Quelques modifications:
  - Usage d'un autre port UDP: 521 au lieu de 520
  - L'authentification se base sur IPSec et plus MD5
  - La limitation du nombre de routes n'est plus 25 mais limitée par le MTU
  - Possibilité d'annoncer sépcifiquement un "next hop"

§ 5.7 Tutorial IBM
Cours TCP/IP, Le Boudec, EPFL

EIGRP

Enhanced interior Gateway Routing Protocol

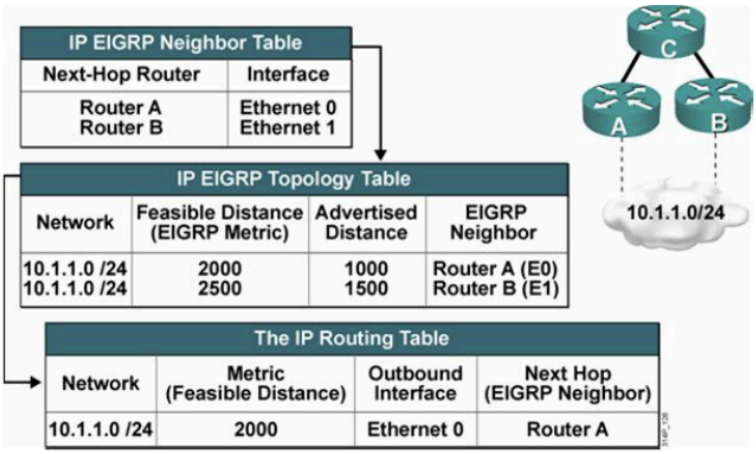
- Protocole propriétaire (Cisco), pas de standard IETF
- Utilise une métrique pour estimer le délai global
- Maintient une liste de routes alternatives
  - Peuvent être utilisées en cas de défaillance d'un lien
  - Peut partager de la largeur de bande (load sharing)
- Nouveau routeur voisin: échange des tables de routage mais propagation des changements seulement
- Supporte VLSM et CIDR

EIGRP (2)

- Broadcast toutes les 90 secondes
- Pas de limite à 15
  - # routeurs inclu dans les messages
- Peut faire un résumé des routes
- Maintient une table pour connaître l'état de ses voisins adjascents (neighbor table)
- Maintient une table de topologie pour créer la table de routage IP
- Peut garantir l'ordre d'arrivée des paquets à son voisin
  - sans garantir la fiabilité de l'arrivée des paquets
- Typiquement un processus pour IPv4 et un autre pour IPv6

Exemple EIGRP (de Cisco)

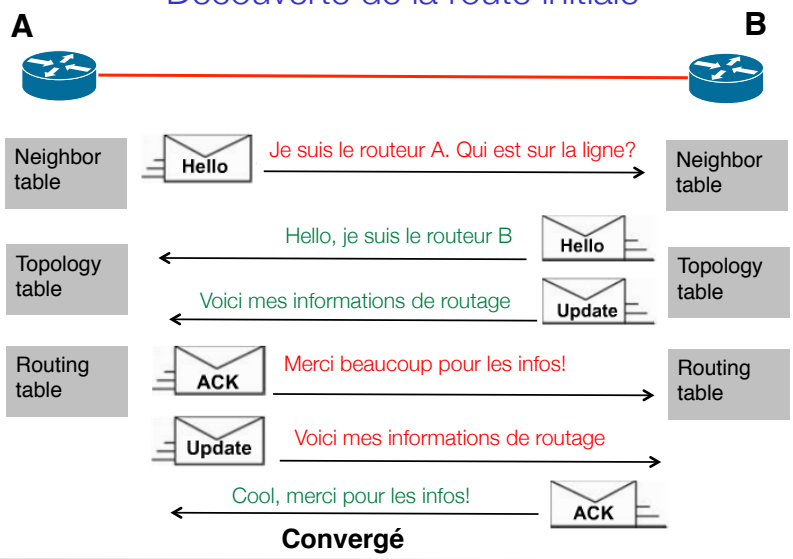
Router C's tables:



Paquets EIGRP

- Hello: établi une relation avec ses voisins
- Update: envoie des mises à jour de routage
- Query: interroge les voisins sur des informations de routage
- Reply: répond à une interrogation de routage
- ACK: aquitte un paquet

Découverte de la route initiale



EIGRP

- Métriques
  - Largeur de bande
  - Délai
  - Fiabilité
  - Charge
  - MTU
- EIGRP est la métrique IGRP multipliée par 256

Calcul de la métrique EIGRP

- Par défaut la métrique EIGRP est

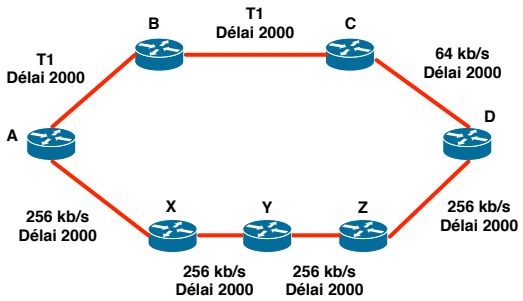
  Métrique=Largeur de bande (ligne la plus lente)+délai (somme des délais)
- Délai=somme des délais le long du chemin, en dizaines de microsecondes, multiplié par 256.
- Largeur de bande (BW: Bandwidth)

  $[10^7$/largeur de bande minimum le long du chemin en kb/s]*256
- Métrique
  - K de défault: K1=1, K2=0, K3=1, K4=0, K5=0

  Métrique = [K1*BW+(K2*BW)/(256-load)+K3*délai]
- Si K n'est pas =0 (reliability=fiabilité)

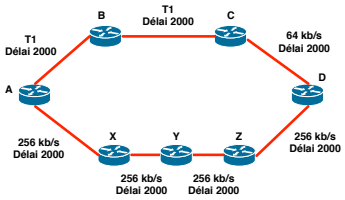
  Métrique = Métrique*[K5/(reliability+K4)]

Calcul de la métrique EIGRP



A -> B -> C -> D BW la plus faible: 64 kb/s Délai total: 6000
A -> X -> Y -> Z -> D BW la plus faible: 256 kb/s Délai total: 8000

Calcul de la métrique EIGRP



A -> X -> Y -> Z ->D

BW=(10^7/least BW in kb/s)*256
BW=(10^7/256)*256
BW=10'000'000

Délai=(4*2000)*256
Délai=2'048'000

Métrique=BW+Délai=12'048'000

A -> B -> C -> D

BW=(10^7/least BW in kb/s)*256
BW=(10^7/64)*256
BW=40'000'000

Délai=(2000+2000+2000)*256
Délai=6000*256
Délai=1'536'000

Métrique=BW+Délai=41'536'000

Choix de A:
chemin le meilleur marché, par X (entrée dans sa table de routage)

Métriques dynamiques

- Certains proposent des métriques dynamiques pour améliorer le chemin le plus court
- Plus de charge sur le réseau -> coûts hauts -> utilisation moindre de la ligne
  - Utilisé par EIGRP
- Il peut y avoir un certain nombre de questions: paradoxe de Braess

Routage avec un délai minimum et équilibre de Wardrop

- Hypothèse: Tous les flots choisissent un chemin au délai minimum
- Des chemins alternatifs existent et les flots peuvent les utiliser
- Il peut y avoir un équilibre (équilibre de Wardrop) tel que les délais de tous les chemins sont égaux.

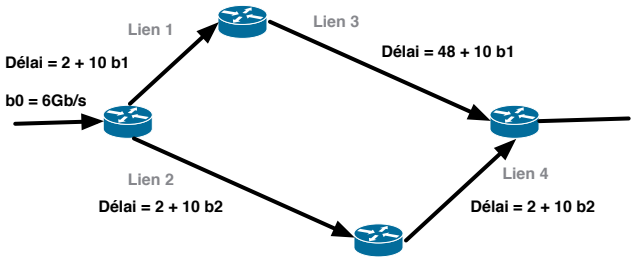
Premier principe de Wardrop

The journey times in all routes actually used are equal and less than those which would be experienced by a single vehicle on any unused route
(wikipédia)

Deuxième principe de Wardrop

At equilibrium the average journey time is minimum
(wikipédia)

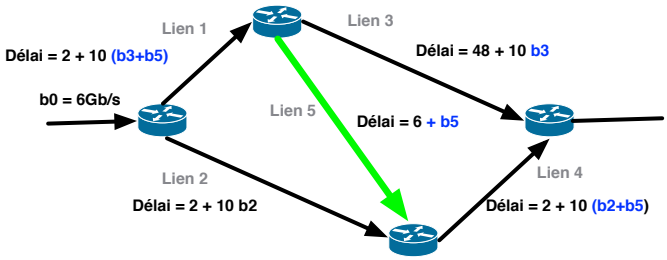
Equilibre de Wardrop



- Pour quelles valeurs de b1, b2 avons-nous un équilibre de Wardrop?

b1=b3=3

Ajout d'un lien (5)

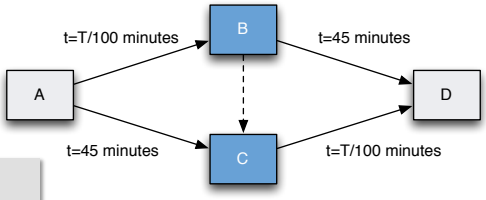


Equations pour le délai:
50 + 11.b3 +10.b5 =
50 + 11.b2 + 10.b5 =
10 + 10.b3 + 10.b2 + 21.b5

Flot total:
b5 + b2 + b3 = b0 = 6 Gb/s

Solution:
b5 = b2 = b3 = 2 Gb/s
Délai = 92 ms pour tous

Paradoxe de Braess



1. Liaison entre B-C n'existe pas
   - 4000 conducteurs
   - Répartition égale entre A-B-D et A-C-D
   - Temps = 2000/100 + 45 = 65 minutes
2. Liaison entre B-C: t=0
   - Trajet préférentiel: A-B-C-D
   - Temps A-B-C-D: 4000/100 + 4000/100 = 80 minutes
   - Temps A-C-D: 45+ 4000/100 = 85 minutes
   - Aucune incitation au changement!!!!

Conclusion:
Si chaque conducteur était d'accord de ne pas emprunter B-C, chaque conducteur en profiterait (15 minutes) mais chaque conducteur y gagnerait en empruntant B-C-> instable

Paradoxe de Braess

- Calcul du délai
  - Lorsque la ligne 5 est désactivée ……
  - Lorsque la ligne 5 est activée ……
- On a à faire avec le paradoxe de Braess
- Conclusion:
  - Le routage basé sur le délai n'est pas optimal

Routage optimal

- A la place de calculer le trajet ayant le délai minimum nous pourrions essayer de résoudre un problème d'optimisation en maximisant une sorte de « fonction d'utilité (utility function) » comme par exemple:
  - Minimisation du délai total soumis aux contraintes des flots (la solution optimale dépend de tous les flots)
  - Référence pour un exemple d'implémentation d'algorithme distribué pour le contrôle de congestion TCP: Data Networks, Bertsekas et Gallager.

Conclusion sur le routage à vecteur de distance

- Le routage à vecteur de distance est bien pensé
  - Complétement distribué
- Déployé à grande échelle
- Simple
- Par contre: convergence lente
  - Pas adapté pour les grands réseaux comlexes
  - Les protocoles à état de liens devraient être utilisés à la place

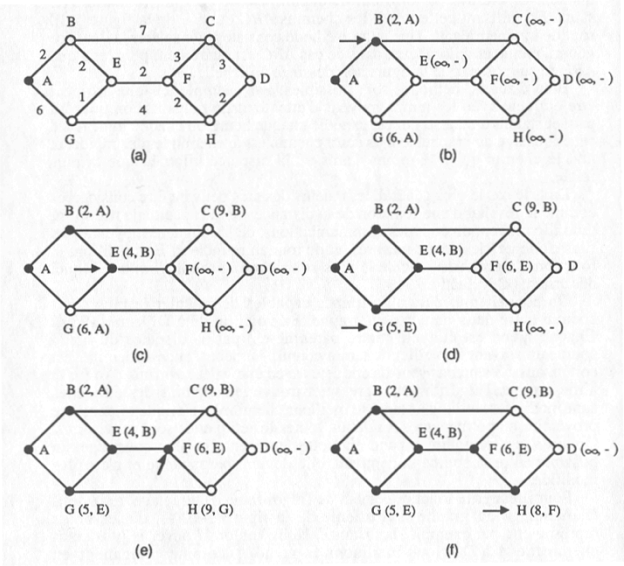
Routage par état de liaison

- Le routage par vecteur de distance fut utilisé dans l'ARPANET jusqu'en 1979
  - Métrique originale : longueur des files d'attente
    - Idée: chemin avec un délai de transfert minimal
    - Applicable si toutes les lignes ont le même débit
  - Problèmes :
    - Évolution vers des interconnexions hétérogènes
    - Convergence trop lente dans un réseau important
- ➢ Introduction du routage par état de lien (link state routing) dans ARPANET

Préambule: le plus court chemin

- Algorithme de Dijkstra
  - Représenter le réseau par un graphe
  - Pondérer chaque arête k par un coût p_k
  0. Marquer chaque nœud par un doublet (C_i,N_x)
     - C_i : Distance totale de la source
     - N_x : Nœud précédent (pour reconstruire le chemin)
  1. Doublet de chaque nœud initialisé à (¥,-) à l 'exception du nœud d 'origine initialisé à (0,-)
  2. Choisir le nœud N_i avec le coût C_i le plus bas et qui n'est pas marqué et le marquer comme 'permanent'
  3. Calculer les coûts des chemins de tous les voisins N_j du nœud N_i : $C_j=C_i+p_k$
  4. Si la nouvelle valeur C_j est plus petite que l'ancienne, --> actualiser le doublet de N_j : (C_j,N_i)
  5. Répéter à partir de 2 jusqu'à ce que la destination soit marquée 'permanent'

Exemple



Routage par état de lien : principe

- Chaque routeur doit périodiquement effectuer les opérations suivantes:
  1. Découvrir ses voisins et apprendre leur adresse respective
  2. Déterminer la distance vers chacun des voisins
  3. Construire un paquet contentant l'information apprise
  4. Envoyer ce paquet spécial à tous les autres routeurs du sous-réseau
  5. Calculer le plus court chemin vers tous les autres routeurs
- Un routeur apprend alors la topologie complète du réseau
- Calculer le plus court chemin

Découvrir ses voisins

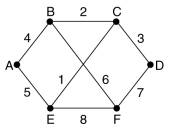
- Un routeur envoie périodiquement des messages Hello sur toutes les lignes de sortie
  - Un routeur voisin répond avec
    - son nom,
    - son adresse IP, …
  - Ainsi, un routeur détecte rapidement l'état des liens de sortie (up, down)

Déterminer la métrique des liens

- Un protocole d'état de liens peut se baser sur plusieurs métrique pour le calcul du plus court chemin
  - Exemples :
    - délai,
    - throughput,
    - fiabilité de transmission
- Les métriques peuvent être mesurées à l'aide de paquets de test

Diffusion de l'information

- Chaque routeur construit des paquets contentant l'information sur l'état des liens locaux (LSP: link state packet)



Link State Packets

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

- Les LSP d'un routeur sont diffusés dans le réseau entier
  - Inondation fiable
    - Un routeur transmet un LSP reçu sur tous les ports sauf le port de réception
    - Un numéro de séquence unique permet d'éliminer des LSP dupliqués
      - Un routeur incrémente le no. de séquence pour chaque LSP émis
    - Les autres routeurs enregistrent le LSP le plus récent de chaque nœud
    - Éliminer les LSP quand sa durée de vie est terminée
    - La réception d'un LSP est confirmé par un accusé de réception

Calcul du plus court chemin

- Un routeur apprend l'état des liens du réseau entier
- Le calcul du plus court chemin peut être effectué en local à chaque nœud
  - Aucune dépendance du calcul d'autres routeurs
  - Convergence rapide et garantie
  - La synchronisation des bases de données garanti l'absence de boucles persistantes
  - Chaque nœud calcule un arbre de chemins minimaux à partir de lui même comme racine
- Méthodes de calcul :algorithme de Dijkstra
  - C'est le meilleur algorithme connu pour ceci

Comparaison
Vecteur de distance – État de lien

- Vecteur de distance
  - Transmission des vecteurs de distance entre voisins
    - Information globale: distances vers toutes les destinations
    - Peuvent devenir très longs dans des réseaux importants
  - Calcul distribué
    - Convergence peut être lente
    - Problème du comptage à infini
      - La distance maximale doit être limitée, p.ex. à 15
- État de lien
  - Diffusion de l'information topologique par inondation
    - Information sur la topologie locale vue d'un routeur
    - Nécessite la limitation de la taille d'un réseau
  - Calcul local --> convergence rapide et fiable

Exercices 25, 26, 27, 31, 33, 34, 35, 36, 37, 49

OSPF
Open Shortest Path First

- Caractéristiques du protocole
  - Protocole 'ouvert' : standard Internet non-propriétaire
  - Protocole d'état de lien
  - Permet l'utilisation de plusieurs métriques
    - Routage peut dépendre du type de trafic
  - Permet l'équilibrage de la charge sur plusieurs chemins à coût égal (load balancing)
  - Domaines > 16 routeurs
  - Protocole IGP (à l'intérieur d'un Système Autonome)
    - Introduit une hiérarchie supplémentaire dans l'AS : les zones
- Protocole de routage le plus utilisé actuellement

OSPF standarisation

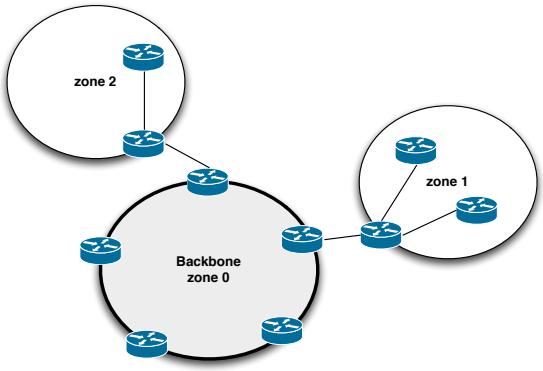
- OSPFv1: RFC1131, RFC 1247
- OSPFv2: RFC 2328, RFC 3630
- OSPFv3: RFC 2740 (adaptation pour IPv6), RFC 3101 (aires NSSA), RFC 4552 (confidentialité et authentification des échanges)

Diviser les grands réseaux

- Pourquoi faut-il diviser les grands réseaux (nombre de routeurs) ?
- Coût de la mise à jour des tables de routage
  - Si la topologie change les tables sont mises à jour
  - Algorithme de Dijstra: n routeurs, k chemins -> complexité O(n*k)
  - Taille de la base de données (augmente avec la taille du réseau)
- Utilisation d'un routage hiérarchique pour limiter les mises à jour et le temps de calcul
  - Plusieurs zones (areas)
  - Calculs des chemins indépendamment dans chaque zone
  - Aggréger les informations de routage et les injecter dans les autres zones

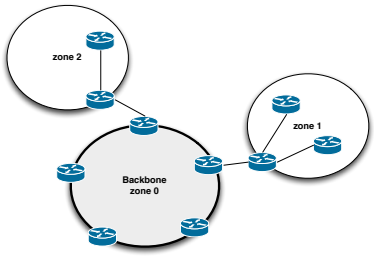
Principe du routage hiérarchique

- Un système autonome (AS) est divisé en zones (areas)
  - Zone 0 : réseau backbone
  - Interconnecte les autres zones

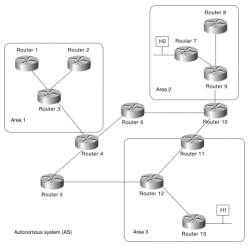


Principe du routage hiérarchique

- Tout le trafic des zones va passer par la zone 0
  - Hiérarchie stricte
- A l'intérieur d'une zone: routage à état de lien « classique »
  - Une base de données par zone
- Routage à plus haut niveau: vecteur de distance (pas de problèmes de boucles avec un backbone, zone 0)



Routeurs



- OSPF comprend 4 types de routeurs
  - Routeur intra-zone
    - Entièrement à l'intérieur d'une zone
  - Routeur inter-zones (Area Border Router ABR)
    - Connecté à plus d'une zone
  - Routeur fédérateur (Backbone Router)
    - Connecté à l'épine dorsale (zone 0)
  - Routeur inter-systèmes autonomes (AS Boundary Routers)
    - Connecté aux routeurs d'autres Systèmes Autonomes

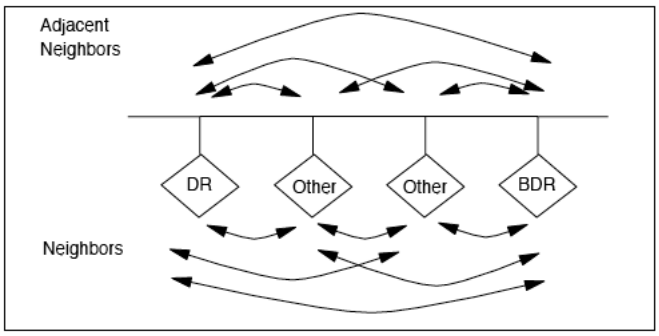
OSPF: Échange d'information de topologie

- Principe
  - La topologie d'une zone est invisible aux routeurs d'autres zones
  - Les routeurs intra-zone ne connaissent pas la topologie du réseau backbone
- Fonctionnement
  - Un routeur intra-zone diffuse des LSP à tous les routeurs de sa zone
    - Construction des plus courts chemins à l'intérieur de chaque zone, y compris la zone 0
  - Les routeurs inter-zones injectent des résumés d'état de liens inter-zones dans les zones locales
    - Permet aux routeurs intra-zones de trouver la meilleure sortie vers une autre zone

Routeurs voisins et "adjacences"

- Les routeurs se trouvent dans une même zone OSPF (même mot de passe, les timers pour les paquets Hello sont les mêmes, même stub area)
- Si deux routeurs sont voisins alors ils peuvent établir une relation d'adjacence.
- Deux routeurs sont considérés adjacents lorsqu'ils ont synchronisés leurs bases de données contenant la topologie.
- Multicast
  - 224.0.0.5 – tous les routeurs sur la ligne
  - 224.0.0.6 – tous les routeurs désignés (DR) et de backup (BDR) sur la ligne

Relation entre voisins et adjacences



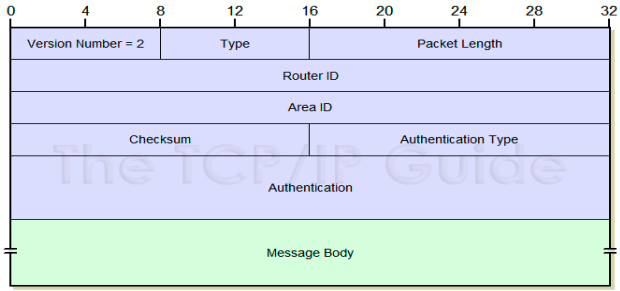
Source: tutorial IBM

Il n'y a pas d'adjacences entre routeurs qui ne pas sélectionnés pour être un DR ou un BDR

Routeur désigné et routeur de backup

- S'il y a plusieurs routeurs dans un réseau, un routeur principal est désigné (Designated Router, DR) ainsi qu'un routeur de backup (Backup Designated Router, BDR)
- Le DR et le BDR ont les mêmes fonctionalités. Le BDR prend le relai si le DR "lâche"
- Le DR existe pour réduire le trafic d'innondation, en réduisant le nombre d'adjacences
- Chaque routeur acquiert la base de données du routeur principal
- Chaque routeur diffuse à ses voisins (LSA)
  - Liste de ses voisins immédiats
  - Coût de la liaison vers chaque voisin
- Chaque routeur met à jour sa base de données –> vision globale du réseau
- Chaque routeur calcule ses meilleurs routes -> tables de routage

Format des paquets OSPF



- Type: 1: Hello, 2: Database description, 3: Link State Request, 4: Link State Update, 5: Link State ACK
- Packet length: Longueur du message
- Router ID: Identificateur du routeurqui a généré le message (en général l'adresse IP d'un de ses interfaces)
- Area ID: Area de laquelle vient le message
- Authentication type: 0: sans authentification, 1: mot de passe, 2: authentification cryptée.

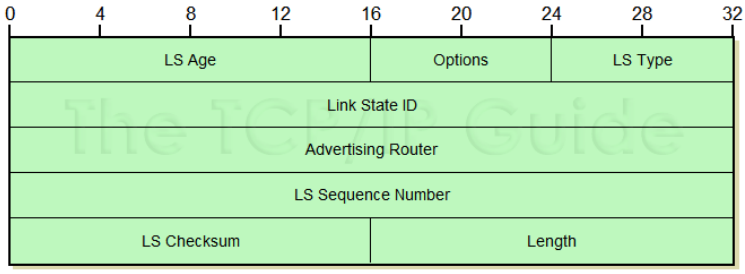
LSAs (Link State Advertisements) et innondation

- Les LSAs sont échangés entre des routeurs adjacents pour synchroniser les bases de données
- Quand un routeur génère ou modifie un LSA il le transmet aux routeurs adjacents qui transmettent le message à leurs voisins
- Chaque LSA est acquitté
- Quand un lien casse, un nouveau LSP (Link State Packet) est envoyé et tous les routeurs recalculent leurs tables de routage

Types d'information des LSAs

- Router LSAs
  - Décrit l'état des interfaces des routeurs. Généré par chaque routeur OSPF
- Network LSAs
  - Liste des routeurs connectés à un réseau multi-accès. Généré par le DR. Innondation dans l'aire
- Summary LSAs
  - Générés par l'ABR (Area Border Router)
    - Type 3: les LSAs décrivent les routes aux destinations dans les autres aires au sein du réseau OSPF
    - Type 4: décrit les routes aux ASBR (AS Border Router, en contact avec d'autres environnements de routage)
- AS external LSAs
  - Décrivent les routes externes au réseau OSPF. Générées par un ASBR. Innodation dans toutes les aires

Format de l'entête des LSAs



Source: The TCP/IP Guide

- LS Age: Nombre de secondes depuis que le LSA a été créé
- LS Type: 1: Router LSA, 2: Network LSA, 3: Summary LSA (IP Net), 4: Summary LSA (ASBR), 5: AS External
- Link State ID: En général adresse IP du routeur ou de la ligne
- Advertising Router: ID du routeur qui est à l'origine du LSA
- LS Checksum: Checksum du LSA, protection des données
- Length: Longueur du LSA, y compris les 20 octets d'entête

Communication entre voisins

1. Découverte des voisins (HELLO)
2. Election d'un routeur désigné
3. Etablissement des adjascences

Paquets HELLO

Maintient la relation entre les routeurs voisins. Envoyés périodiquement par les interfaces des routeurs. Contient le Router ID, priorité, DR et BDR Id.

Election du routeur désigné

Le DR et le BDR sont sélectionnés sur la base des paquets HELLO. Le routeur avec la plus haute priorité OSPF (1-255) devient le DR sur un segment. La même méthode est appliquée pour le BDR. En cas d'égalité c'est le routeur avec le plus haut RID qui gagne. Si un routeur a une priorité 0, il ne sera jamais DR ou BDR

Communication entre voisins (2)

Etablissement des adjascences et synchronisation des bases de données

1. Processus d'échange entre les bases de données

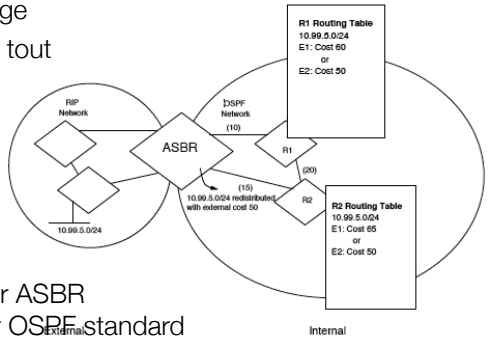
Lorsque deux voisins essaient d'établir une adjascence: échange de paquets de description (liste de LSAs). Mémorisé dans la base de données locale. Etablissement d'une relation "maître-esclave".

2. Chargement de la base de données

Quand le processus d'échange est terminé chaque routeur demande l'information la plus récente à son voisin avec lequel il a une relation d'adjascence (avec un paquet de requête).

OSPF: routes redistribution

- Introduction de routes externes dans le réseau OSPF
  - Peuvent être statiques, apprises d'un autre protocole de routage
  - ASBR publie ces routes dans tout le réseau OSPF (innondation)

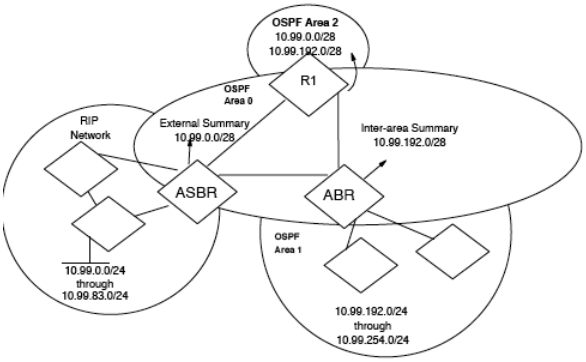


- Routes (de bout-en-bout)
  Portion externe: coût attribué par ASBR
  Portion interne: coût attribué par OSPF standard

OSPF: stub areas

- Aire dans laquelle on n'a pas d'innondation de LSAs
- But: réduction de la base de données (link state) maintenue dans les routeurs des stub areas-> uniquement une route par défault
- Route optimale pas assurée!

OSPF route summarization



- Plusieurs routes dans un seul paquet (advertisement)
- But: réduction de la table de routage et de la base de données « link state »
  1. Inter-area summary: done by ABR
  2. External route summary: done by ASBR

TOS et métriques OSPF

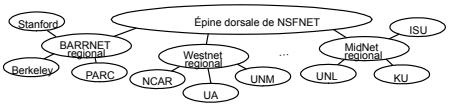
- TOS
  - Mapping des 4 bits de TOS en un chiffre décimal
  - 0 – normal service
  - 2 – minimize monetary cost
  - 4 – minimize reliability
  - 8 – maximize throughput
  - 16 – minimize delay
- Métrique
  - Temps pour envoyer 100 Mb/s sur l'interface
  - C = 10^8/largeur de bande
  - 1 si plus grand que 100 Mb/s
  - Peut être configurée par l'administrateur

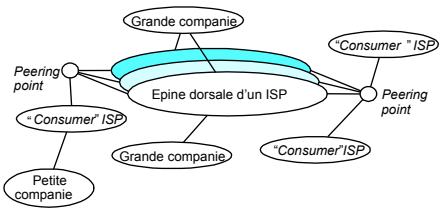
Comparaison RIP-OSPF

- OSPF est bien plus compliqué mais présente des avantages:
  - Pas de comptage à l'infini
  - Pas de limite sur le nombre de sauts (hops)
  - Moins de trafic de signalisation (mise à jour LS toutes les 30 minutes)
  - Métrique évoluée
  - Routage hiérarchique pour les grands réseaux
  - Génère du trafic surtout quand on a un changement de topologie (bien qu'on ait des paquets Hello envoyés périodiquement)
- Désavantage
  - Difficile à configurer

Protocoles de routage inter-domaine (Exterior Gateway Protocols)

- Protocoles de routage entre Systèmes Autonomes
  - Premier protocole : EGP
    - Nécessitait une topologie en arborescence simple
    - N'est plus utilisé
  - Protocole actuelle : BGP (Border Gateway Protocol)
    - A remplacé EGP dans Internet
    - Permet une topologie arbitraire





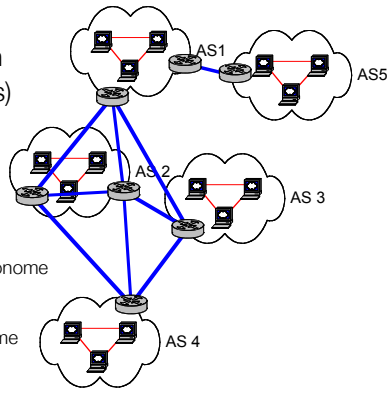
Exercice 50

BGP

- Principes de conception
  - Doit pouvoir gérer les routes dans l'Internet global
    - Actuellement, un routeur BGP connaît environ 90'000 routes
  - Ne peut pas se baser sur les métriques utilisées dans les AS
    - Chaque AS est libre de choisir sa stratégie de routage
    - La notion du plus court chemin n'est pas applicable
  - Utilise des stratégies de routage pour filtrer les routes acceptables
    - Exemple Sun : ne pas utiliser une route qui travers l'AS de Microsoft
- BGP ne cherche pas le meilleur chemin mais un chemin quelconque
  - Échange des informations d'accessibilité
  - Évite des boucles de routage
  - Configuration locale d'une stratégie de routage

Systèmes Autonomes (AS)

- Sous le contrôle d'une seule administration
- Peut comprendre plusieurs réseaux (NetIds)
  - Exemples:
    - Réseau d'un ISP et de ses clients
    - Réseau d'une grande entreprise
- Types d'AS
  - Le bout de AS : (stub AS)
    - A une seule connexion avec un autre système autonome
    - Transporte du trafic local seulement
  - AS multi-ports: (multi-homed AS)
    - A des connexions avec plus d'un système autonome
    - Refuse de transporter le trafic de transit
  - AS de transit : (transit AS)
    - A des connexions avec plusieurs autres AS
    - Transporte le trafic local et le trafic de transit

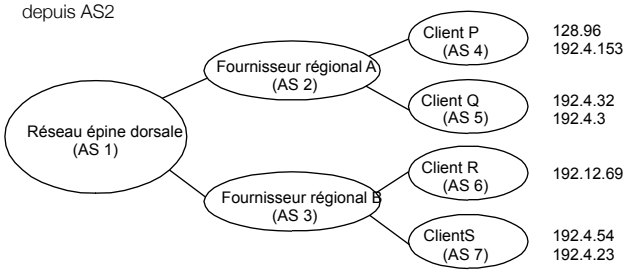


Introduction au fonctionnement de BGP

- Chaque système autonome a un ou plusieurs routeurs BGP
- Un routeur BGP annonce vers l'extérieur :
  - les réseaux à l'intérieur de l'AS
  - les réseaux externes atteignables à travers l'AS
  - Communique le chemin entier pour atteindre chaque réseau
    - Exemple : AS A annonce une route vers un réseau n:
      - n : A-B-C-F (chemin des AS traversés)
    - Permet de filtrer les routes
    - Permet de détecter facilement des boucles de routage
  - ➢ External BGP
- Un routeur BGP communique avec les routeurs internes
  - Diffuse quelques routes apprises vers l'intérieur du AS
  - S'assure de la connectivité à d'autres AS
  - ➢ Internal BGP

Exemple EBGP

- Router BGP pour AS2 annonce l'accessibilité de P et Q
  - Les réseaux 128.96, 192.4.153, 192.4.32, et 192.4.3, peuvent être atteints directement depuis AS2



- Le routeur BGP de l'épine dorsale annonce
  - Les réseaux 128.96, 192.4.153, 192.4.32, et 192.4.3 peuvent être atteints le long du chemin (AS1, AS2).
- Le routeur BGP peut supprimer des chemins annoncés précédemment

Exercices 44, 46

Routages sans classes

- Problème actuelle de BGP
  - Les adresses classe B sont épuisées
  - Une grande entreprise ayant plus de 255 hôtes doit utiliser plusieurs adresses classe C
  - Un routeurs BGP doit connaître et annoncer des routes pour chaque réseau classe C
    - En théorie jusqu'à 2 mio de routes vers des réseaux classe C
- Solution
  - Allocation de blocs de taille variable d'adresses classe C
    - Site à 2000 hôtes : allocation de 8 adresses classe C contiguës
      - Meilleure efficacité que l'allocation d'une adresse classe B
  - Un bloc d'adresses est alloué de telle manière qu'il forme un super-réseau avec un préfix d'identificateur de réseau commun

Exemple

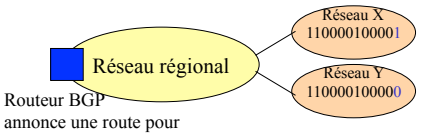
- Site ayant besoin de 4000 adresses
  - Allocation de 16 adresses classe C:
    - 192.4.16 – 192.4.31
    - Structure normale des adresses

  - Ces adresses on le même préfixe binaire
    - Premiers 20 bits = 11000000 00000100 0001
    - Elles peuvent être agrégées dans un seul 'super-réseau' ayant un identificateur de réseau sur 20 bits

- Contrainte
  - Les blocs d'adresses doivent avoir une taille de 2^x d'adresses classe C

Classless Inter-Domain Routing

- CIDR
  - Implémenté dans la nouvelle version BGP-4
    - Définit des identificateurs de réseau de longueur variable
    - Compromis entre efficacité et complexité du routage
  - Agrégation de routes avec CIDR
    - Si un routeur utilise la même route pour plusieurs blocs d'adresses contigus, il peut annoncer une seule route



    - Exemple
      - Adresses classe C 194.0.0.0 – 195.255.255.255 --> Europe
      - Un routeur BGP américain considère les premiers 8 bit pour le routage
      - Un routeur BGP européen doit considérer des préfixes plus longs

Exercices 40, 41, 42, 43

Mobilité dans IP

- Objectif
  - Rendre possible le déplacement d'un ordinateur mobile (PC portable, agenda électronique, …) d'un réseau à un autre de manière transparente pour les applications
- Types
  - Nomadicité (portability)
    - Déplacement 'off-line', mais sans re-paramétrage manuelle
    - Nécessite l'interruption de toutes les connexions en cours
    - --> DHCP
  - Mobilité d'un ordinateur
    - Un ordinateur mobile peut changer sont point d'attachement sans interrompre les communications en cours
  - Réseau mobile
    - Réseau ad-hoc sans infrastructure

Problème de l'adressage IP

- Une adresse IP
  - identifie un système terminal
    - Correspondance statique entre le nom de domaine et l'adresse IP par DNS
  - détermine la route vers un système terminal
    - L'adresse IP comprend un NetId et un HostId
    - Le routage utilise le NetId pour trouver le chemin vers le système terminal
- Contradiction
  - Un ordinateur mobile nécessite une adresse fixe pour être joignable
  - Une adresse fixe implique un routage fixe
- Idée : Utiliser deux adresses
  - Une adresse qui identifie un ordinateur
  - Une adresse qui permet de joindre l'ordinateur

Mobile IP

- Défini dans la RFC 2002 (3344)
  - Permet une mobilité globale dans Internet
  - Solution au niveau de la couche Réseau
- Principes de conception
  1. Transparence pour les applications existantes
     - Un ordinateur mobile doit être capable de communiquer avec un autre ordinateur qui n'implémente pas IP Mobile
     - Un ordinateur mobile doit être joignable en utilisant uniquement son adresse IP (normale)
  2. Transparence pour les routeurs existants
     - Aucune modification de la méthode de routage
     - Sans modification permanente des tables de routage
  3. Sécurité
     - Un ordinateur mobile ne doit pas être plus exposé qu'une autre machine

Terminologie

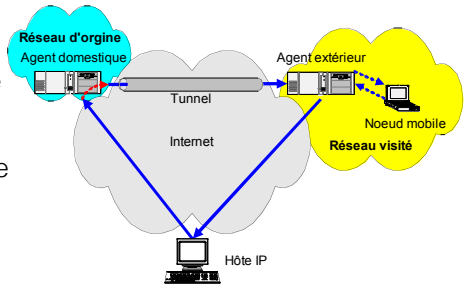
- Nœud mobile
  - Nœud qui peut changer de points d'attachement sur l'Internet tout en maintenant les communications en cours
- Réseau d'origine (Home Network)
  - Réseau auquel appartient l'adresse IP du nœud mobile
- Réseau extérieur (Foreign Network)
  - Réseau visité par le nœud mobile
- Agent domestique (Home Agent)
  - Routeur avec une interface sur le réseau d'origine du nœud mobile
- Agent extérieur (Foreign Agent)
  - Routeur situé dans le réseau visité par le mobile

Adresses

1. Adresse de domiciliation (Home Address)
   - Adresse principale de l'ordinateur mobile dans son réseau d'origine
   - Adresse sur laquelle le mobile est contacté par d'autres machines
2. Adresse de réexpédition (Care-of Address, c/o address)
   - Adresse faisant partie du réseau visité
   - Utilisée par l'agent domestique et l'agent extérieur pour acheminer des messages
   - Deux types
     - « Adresse de réexpédition par agent extérieur »
       - Adresse de l'agent extérieur
     - « Adresse de réexpédition par colocataire »
       - Adresse assignée de manière temporaire à l'ordinateur mobile

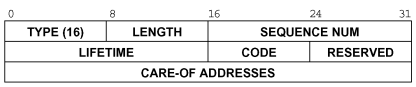
Principe du protocole

1. Un nœud mobile obtient une adresse de réexpédition du réseau visité
2. Le nœud mobile enregistre son adresse c/o auprès de son agent domestique
3. L'agent domestique intercepte tous les paquets destinés au nœud mobile et les transmet à travers un tunnel vers l'adresse c/o
4. Le nœud mobile envoie ses paquets directement aux correspondants



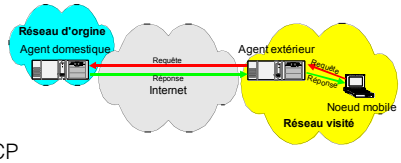
Découverte des agents

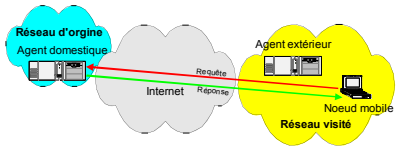
- Permet à un nœud mobile de savoir s'il se trouve dans son réseau d'origine ou dans un autre réseau
- Principe
  - Les agents diffusent périodiquement des messages « Agent advertisement » sur le LAN auquel ils sont attachés
  - Un nœud mobile peut aussi solliciter une réponse d'un agent présent en envoyant un message de découverte d'agent
- Format des messages
  - Extension des messages ICMP standard
  - Le message « Agent advertisement » contient
    - Une adresse de réexpédition que le nœud mobile peut utiliser
    - Une durée de vie pendant laquelle l'agent prend en compte l'enregistrement du mobile



Enregistrement

- Lorsqu'un mobile est hors de son réseau d'origine il enregistre son adresse temporaire auprès de son agent domestique
- Deux cas
  1. Le mobile utilise l'adresse de réexpédition de l'agent extérieur
     - Le mobile envoie une requête d'enregistrement à l'agent extérieur qui la fait suivre à l'agent domestique
     - L'agent domestique revoie la réponse d'enregistrement à l'agent extérieur qui la passe au mobile
  2. Le mobile utilise une adresse de réexpédition par colocataire, obtenue p.ex. à l'aide de DHCP
     - La requête et la réponse sont envoyées directement entre le mobile et l'agent domestique





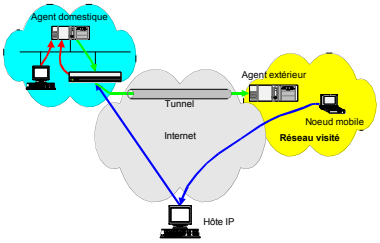
Contenu des messages d'enregistrement

0	8	16	31
TYPE (1 or 3)	FLAGS	LIFETIME	
HOME ADDRESS			
HOME AGENT			
CARE-OF ADDRESS			
IDENTIFICATION			
EXTENSIONS . . .			

- **Type**: requête ou réponse
- **Durée de vie**: les agents peuvent limiter la durée de validité de l'enregistrement
- Adresse de domiciliation et adresse de réexpédition
- Agent domestique
- **Identification**: valeur générée par le mobile pour identifier les requêtes et réponse et pour des raisons de sécurité
- **Drapeaux/Code**: indique le succès de la requête ou des options supplémentaires
- **Extensions**: p.ex. authentification

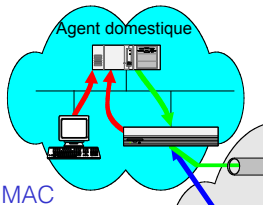
Transmission des datagrammes

- Mobile --> Correspondant
  - Le mobile envoie des datagramme avec son adresse de domiciliation comme source
  - Le datagramme utilise le routage habituel pour arriver au destinataire
- Correspondant --> Mobile
  - Les datagrammes envoyés par une machine quelconque au mobile sont routés vers le réseau d'origine du mobile
  - L'agent domestique doit
    - Intercepter les datagrammes destinés au mobile
    - Les réexpédier vers le mobile en contournant le routage normal



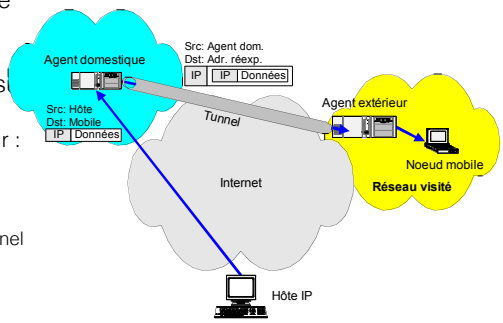
Interception des messages

- L'agent domestique doit intercepter les messages provenant
  - De l'extérieur (d'un ordinateur se trouvant dans un autre réseau)
  - De l'intérieur (d'un ordinateur local connecté au même réseau)
- L'agent domestique ne se trouve pas nécessairement dans le chemin des datagrammes
- Technique : proxy ARP
  - L'agent mobile répond aux requêtes ARP concernant le nœud mobile avec son adresse MAC
    - ➢ Toutes les trames destinées au nœud mobile arrivent à l'agent domestique



Tunnel

- Technique souvent utilisée pour contourner le routage habituel
- Principe
  - L'agent domestique encapsule le datagramme intercepté dans un autre datagramme, ayant comme destinataire l'adresse de réexpédition
  - La terminaison du tunnel décapsule le datagramme original
    - Réexpédition par agent extérieur :
      - L'agent extérieur sert de terminaison de tunnel
    - Réexpédition par colocation :
      - Le nœud mobile termine le tunnel



Aspects problématiques de Mobile IP

- Sécurité
  - Le mécanisme d'enregistration permet à un intrus de dévier/ interrompre des communications
    - ➢ Mécanisme d'authentification
  - Configuration des firewalls des réseaux visités/d'origine
- Inefficacité de routage (triangle routing)
  - Une solution est de communiquer l'adresse de réexpédition au correspondant pour établir un tunnel direct vers le nœud mobile
    - Le correspondant doit implémenter IP Mobile
- Fast handoff
  - Pendant la transition d'un réseau à un autre, des paquets peuvent arriver à la mauvaise adresse de réexpédition
  - « A better than nothing fast handover », Doswald, Robert