

Réseaux avancés

Le routage sur les réseaux longue distance

sonnh@soict.hut.edu.vn

Source: <http://duda.imag.fr>
et cours de Jean-Patrick Gelas, Lyon, France

1

Contenu

- Principes de routage
- Algorithmes
 - Vecteur de distance
 - Etat de lien
- Protocole de routage
 - Routage interne: RIP, OSPF
 - Routage externe: BGP

2

Réseau longue distance ?

- Grande couverture géographique
- Hétérogénéité des modes de transmission
- Mélange des réseaux publics et privés
- Agrégation du trafic
- Tarification par des opérateurs

3

Routeur?

Un routeur est une machine possédant **plusieurs interfaces**. Elle est connectée à plusieurs réseaux en même temps et peut faire passer un paquet d'un réseau à un autre



BUFFALO
BHR-4RV



PLANEX
GW-AP54SAG



YAMAHA
RTX-1500



Cisco 2600



Cisco CRS-1



Hitachi
GR2000-1B



Juniper M10



Foundry Networks
NetIron 800



Cisco 3700

[http://www.cisco.com.
vn](http://www.cisco.com.vn)

<http://www.juniper.net/>

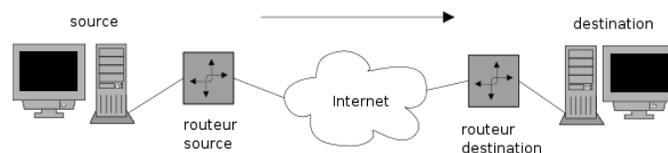
Principe du routage

- Pour transférer des paquets, la couche **réseau** doit déterminer le parcours (ou la route) à emprunter, que ce soit
 - un service à datagramme (parcours différents) ou
 - un service à circuit virtuel (même parcours).
- Cette fonction incombe au **protocole de routage** de la couche **réseau**.

5

Principe du routage

- Un serveur est généralement connecté à un routeur spécifique = le routeur par défaut (ou routeur de premier bond).
- Acheminer un paquet d'une *source* à une *destination* revient à l'acheminer entre un routeur source et un routeur destination.



6

Différents mécanismes de routage

7

Routage centralisé

- Dans le routage centralisé, un processeur central :
 - collecte l'information sur l'état de chaque liaison,
 - établit une table de routage pour chaque noeud et l'envoie à ceux-ci.
- Le routage centralisé n'est pas envisageable sur des réseaux de grande taille :
 - Et si les liaisons au voisinage du processeur central tombe en panne ou que celui-ci tombe en panne ?
 - Quels seraient les temps de calcul ?
 - Les routeurs situés près du processeur central reçoivent les tables de routage bien avant ceux qui sont plus éloignés (inconsistance dans le réseau pendant un certain temps).
 - Les liaisons aboutissant au processeur central peuvent être surchargées si trop de routeurs envoient des informations.

8

Routage distribué

- Dans le routage distribué, les routeurs s'envoient périodiquement des informations pour créer des tables de routage dynamique.
- Le **routage sur Internet** est résolument **distribué**.

9

Routage « à partir de la source »

- L'en-tête d'un paquet contient les adresses de tous les noeuds par lesquels il va devoir passer pour arriver à destination.
- La source doit connaître la **topologie** de tous le réseau.
- Inconvénients :
 - Si une liaison ou un routeur sur le chemin disparaît, le paquet n'atteindra pas la destination.
 - Si le chemin est long, l'en-tête du paquet peut-être très grande.

10

Routage « noeud après noeud »

- L'en-tête d'un paquet contient juste l'adresse de destination.
- C'est aux routeurs de déterminer le noeud suivant.
- Sur Internet on utilise le routage « noeud après noeud » (hop-by-hop).

11

Routage stochastique ou routage déterministe

- Dans le **routage déterministe**, un routeur envoie un paquet vers sa destination selon une et une seule direction.
- Dans le **routage stochastique**, une table de routage contient plusieurs noeuds pour chaque destination possible.
 - les paquets peuvent arriver dans le désordre, avec des délais variables.
- Les réseaux modernes utilisent le routage **déterministe**.

12

Routage « à chemin unique » ou routage « multi-chemins »

- Un routeur à **chemin unique** maintient un seul chemin pour chaque destination.
- Un routeur **multi-chemin** maintient un chemin primaire pour chaque destination, et des chemins alternatifs au cas où le chemin primaire serait **indisponible** (Remarque: dans le routage stochastique, un routeur peut envoyer des paquets sur un chemin alternatif alors que le chemin primaire est disponible).
- Sur Internet on utilise le routage à **chemin unique**. (car maintenir des chemins alternatifs requiert de trop grandes tables de routage).

13

Routage adaptatif ou routage statique

- Routage adaptatif :
 - Le choix du chemin dépend de l'état du réseau (trafic, files d'attente, ...).
 - Meilleur choix des chemins, mais peut mener à des oscillations.
 - Besoin d'une charge supplémentaire sur le réseau pour l'établissement de l'état des liaisons.
- Routage statique : on ignore l'état du réseau.
- Internet utilise les deux type de routage.

14

Algorithmes de routage

15

Algorithmes de routage

- Problème
 - trouver le « bon » parcours entre le routeur source et le routeur de destination
- Le bon parcours?
 - Le parcours le « moins onéreux »: le parcours dont la somme des liaisons entre l'expéditeur et le destinataire est la plus faible.
- Métrique pour mesurer le coût d'une liaison
 - bande passante, coût financier, valeur fixé par l'admin., délai, nombre de sauts, charge, MTU, fiabilité, ...
- Optimisation graphique - le plus court chemin
 - Trouver le plus court chemin dans une graphe
 - « le plus court » dans le contexte d'un métrique

16

Algorithmes principales

- Vector de distance (Bellman-Ford)
 - routeurs ne connaissent que leur état local
 - metrique des liens et estimation des voisins
 - Protocole de routage interne (RIP, IGRP)
- Link state
 - connaissance de l'état global
 - metrique de tous les liens
 - Optimization globale (Shortest Path First - Dijkstra)
 - Protocole de routage interne (OSPF, PNNI (ATM))
- Vecteur de chemin
 - connaissance de l'état global
 - chemin: séquence d'AS avec attributs
 - Optimization globale et routage par strategie
 - Protocole de routage externe(BGP)

17

Protocole de Routage

	Internet	ISO
IGP	distance vector: RIP, RIP v2, IGRP	
	link state: OSPF dual: EIGRP	IS-IS
EGP	EGP (obsolete) BGP	IDRP
host	ICMP Redirect	IS-ES

18

Métriques

- Statiques – ne depende pas à l'état du réseau
 - Numero de sauts
 - Capacité, délais
 - cout
- Dynamique - depende pas à l'état du réseau
 - Charge du lien
 - Delais courant

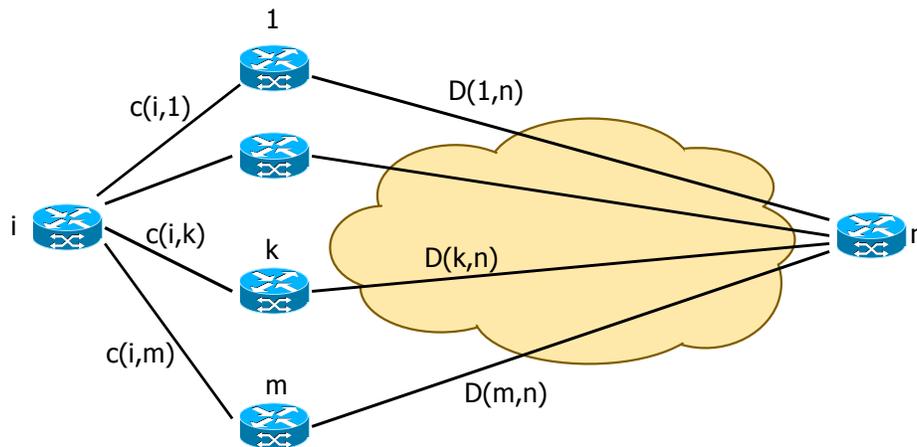
19

Inondation

- Simple et robuste
 - Pas de table de routage
 - duplication de paquet
 - Tolérance au pannes
 - optimale pour certains aspects, e.g 1ere paquet arrive à la destination par le chemin le plus court
- Problème
 - Charge élevés
- Utilise dans OSPF pour diffuser l'information de l'état de lien et protocole de routage ad hoc (AODV, OLSR)

20

Bellman-Ford algorithmme



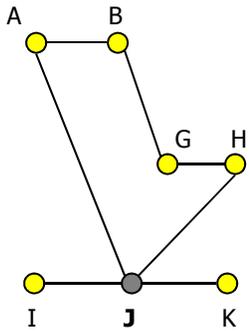
21

Bellman-Ford algorithmme

- Bellman-Ford algorithmme
 - noeud i sait $c(i,k)$ vers ses voisins
 - distance $D(i,n)$ est calcule: $D(i,n) = \min_k (c(i,k) + D(k,n))$
 - convergence après $N-1$ iterations
- algorithme Bellman-Ford distribué
 - Init: $D(i,n) = 0$ si i est connecté à n , si non $D(i,n) = +\infty$
 - noeud i \leftarrow ses voisin k estimation de $D(k,n)$ pour tous n (vecteur de distance)
 - noeud i compute la meilleure estimation
 $D(i,n) = \min_k (c(i,k) + D(k,n))$

22

Example de Bellman-Ford



	A	I	H	K
A	0	24	20	21
B	12	36	31	28
G	18	31	6	31
H	17	20	0	19
I	21	0	14	22
J	9	11	7	10
K	24	22	22	0

Table de J

8	A
20	A
18	H
12	H
10	I
0	-
6	K

computation de G : $18+8=26$, $31+10=41$, $6+12=18$, $6+31=37$
 → choisi de 18, H

Example: Initialization

A

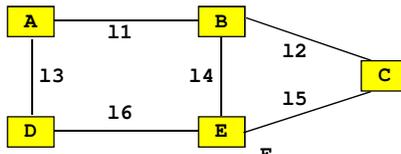
dest	link	cost
A	local	0

B

dest	link	cost
B	local	0

C

dest	link	cost
C	local	0



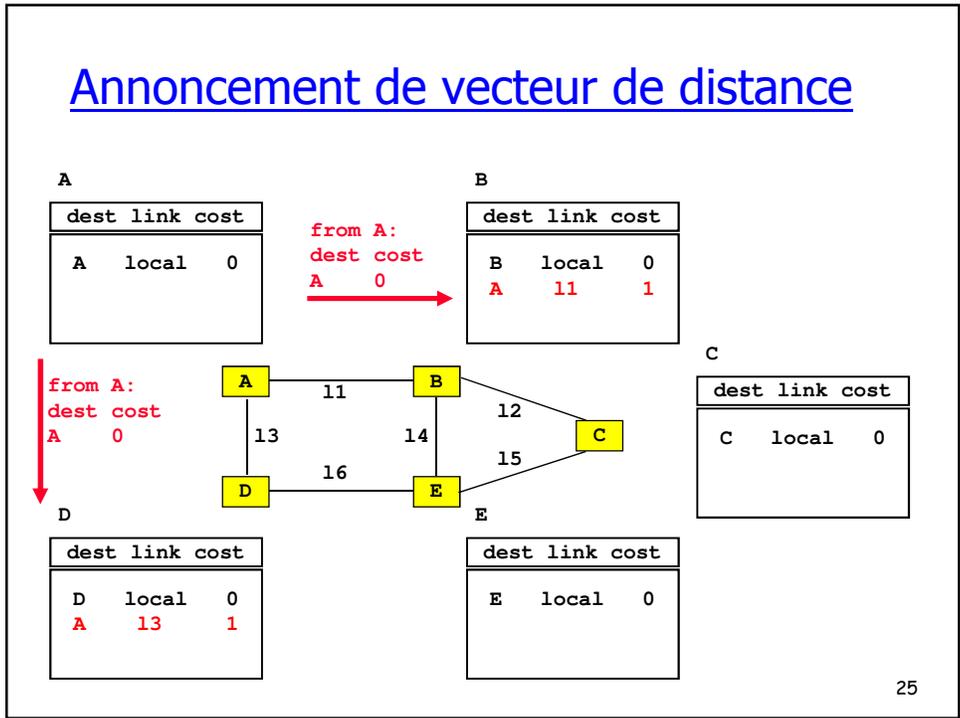
D

dest	link	cost
D	local	0

E

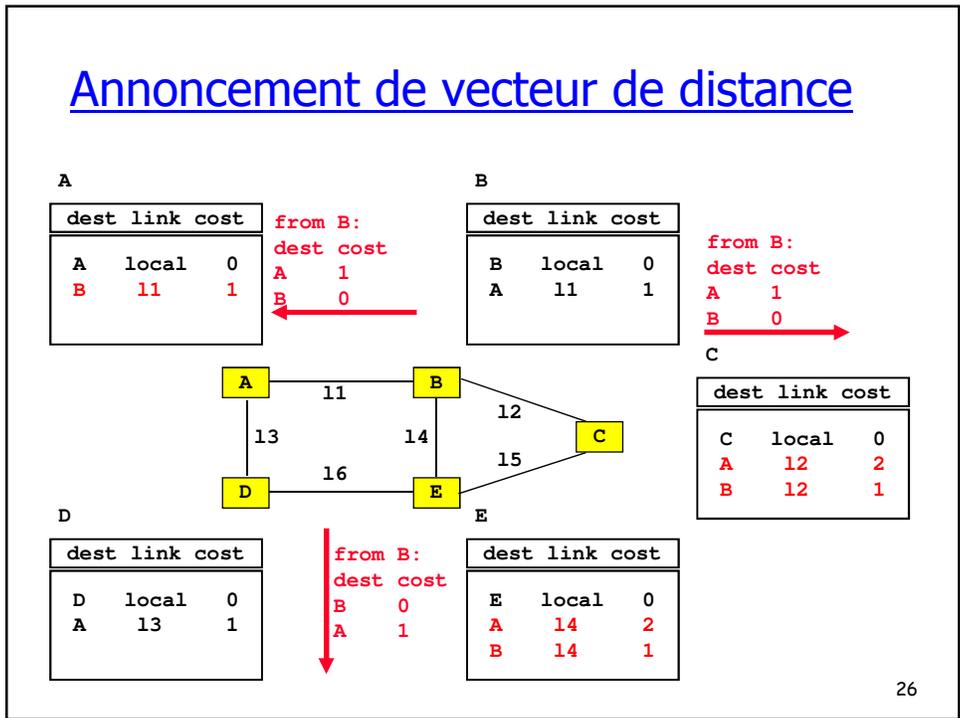
dest	link	cost
E	local	0

Annoncement de vecteur de distance



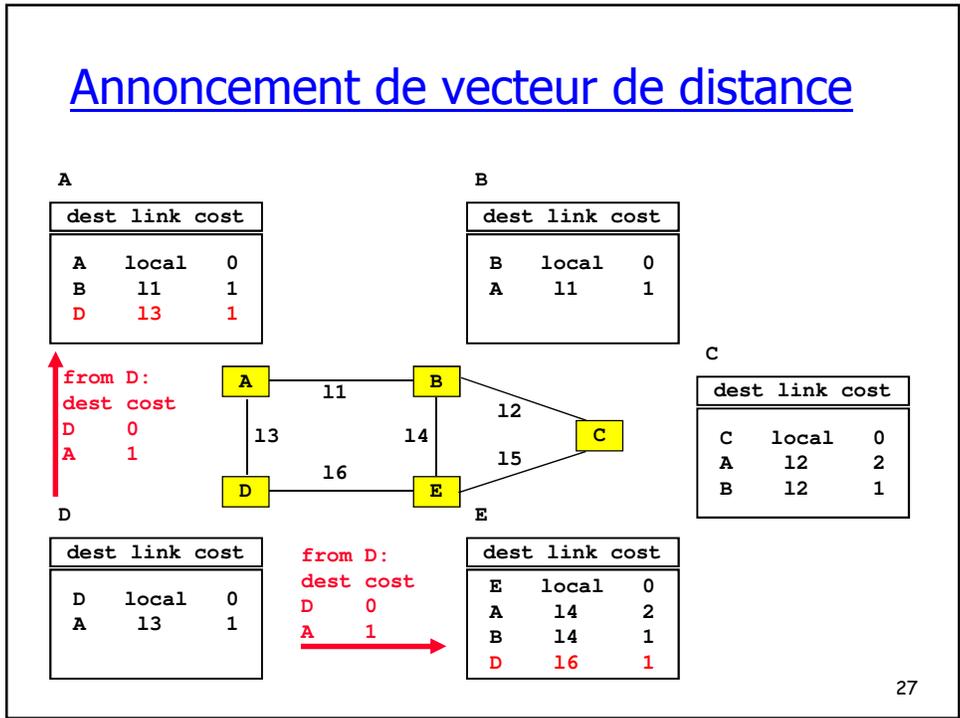
25

Annoncement de vecteur de distance



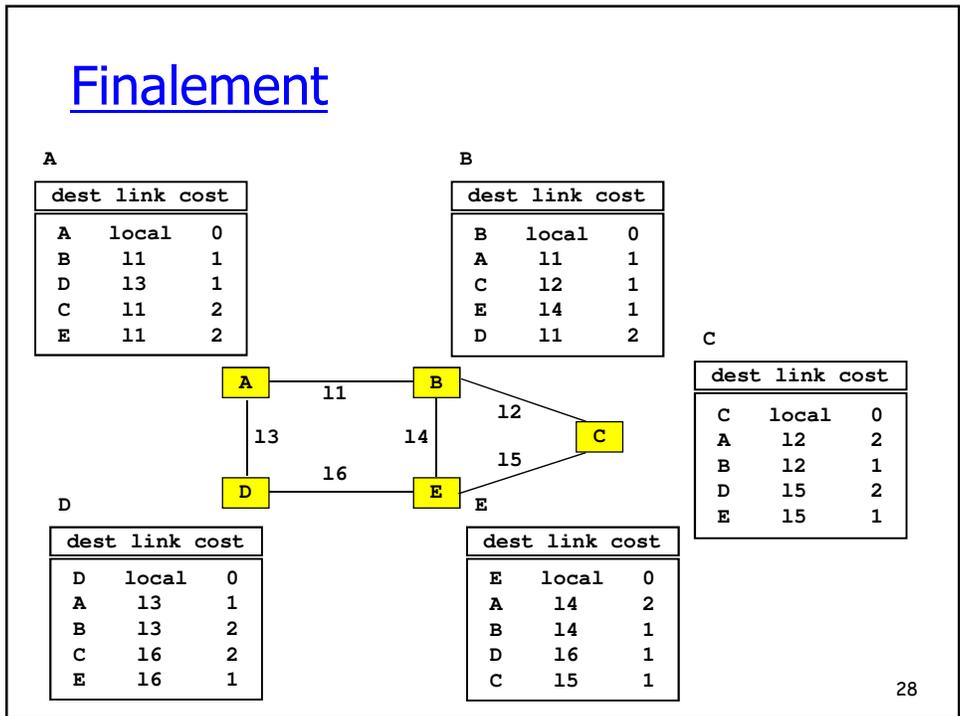
26

Annoncement de vecteur de distance



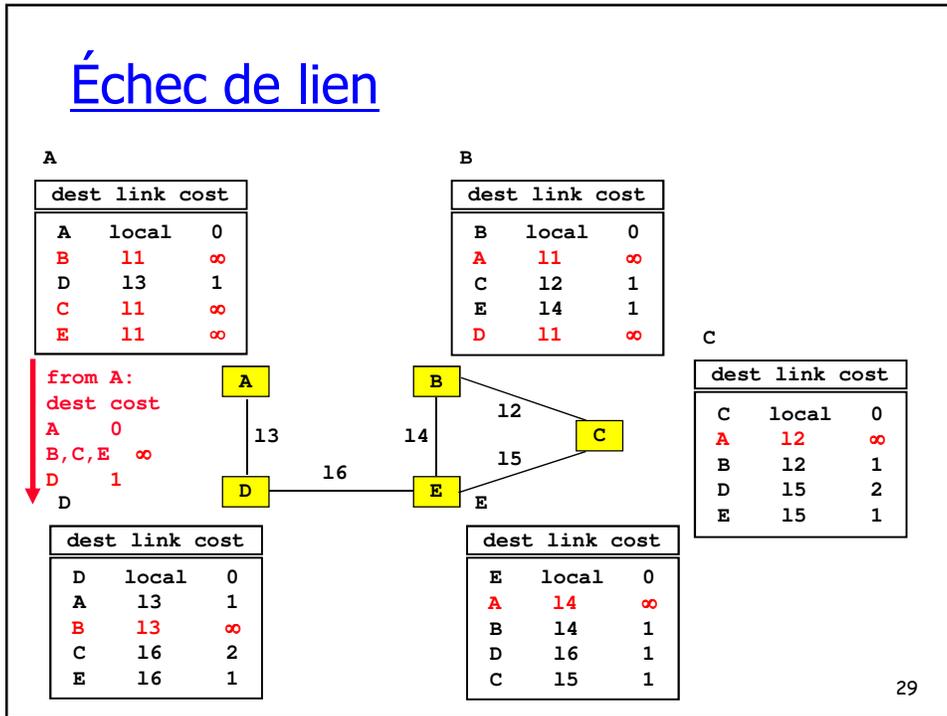
27

Finalement



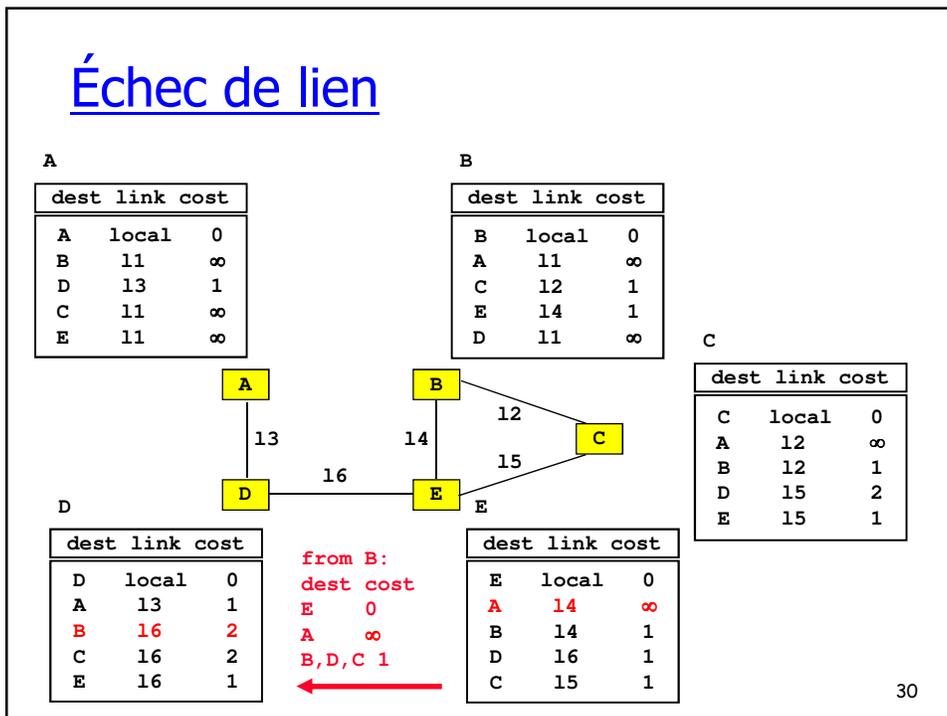
28

Échec de lien



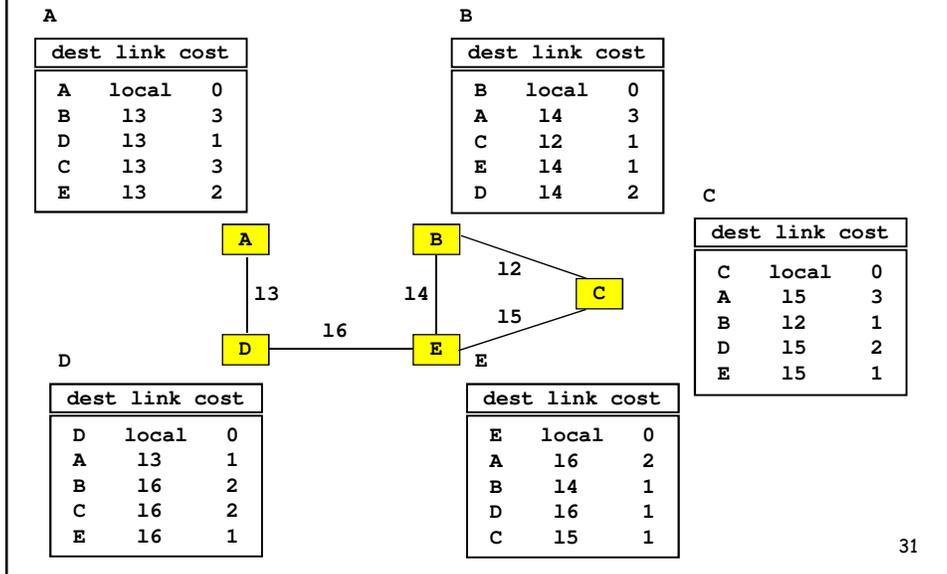
29

Échec de lien



30

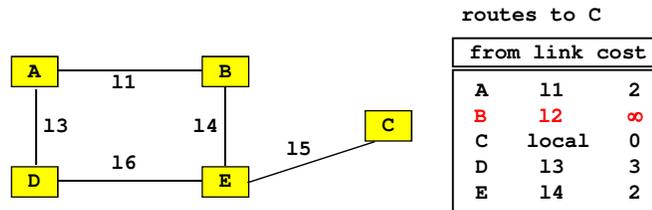
État final après la failure



31

Échec de lien

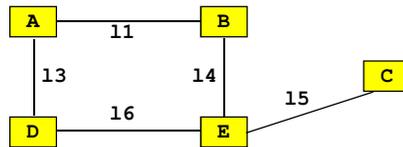
- Lien 2 échoue et B met à jour sa table



32

Échec de lien

- Avant B mise à jour sa table, A diffuse sa table avec cout 2 à C
- B accepte - lien 1, cout 3 et envoie le nouveau vecteur à A et E



routes to C

	from	link	cost
A	11		4
B	11		3
C	local		0
D	13		3
E	14		2

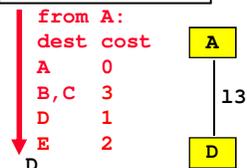
- Boucle entre A et B!

33

Comptage jusqu'à l'infini. (Counting to infinity)

A

dest	link	cost
A	local	0
B	13	3
D	13	1
C	13	3
E	13	2



dest	link	cost
D	local	0
A	13	1
B	13	4
C	13	4
E	13	3

- Boucle entre A et D
- Exchange de routes, cout augmente 2 par une cycle
- Convergence vers état stable
 - ∞ = grande valeur
 - e.g. RIP: $\infty = 16$

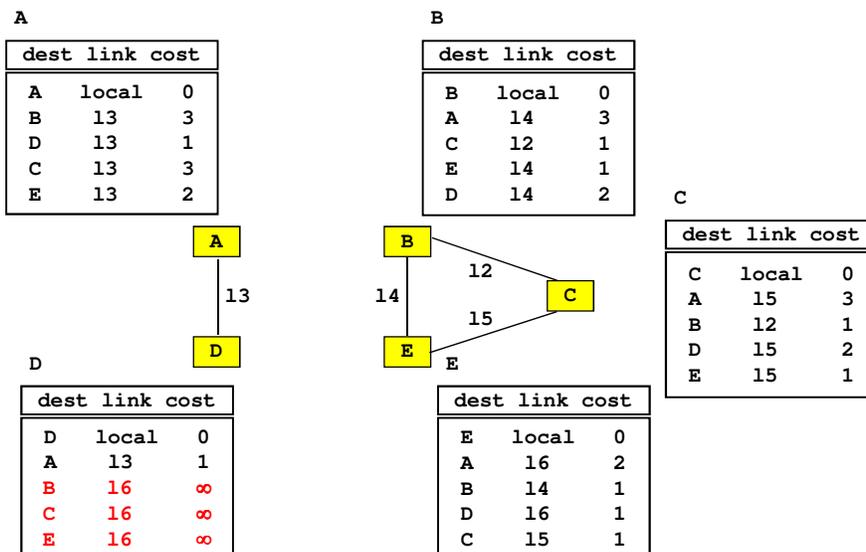
34

Split horizon (horizon éclaté)

- Règle
 - Si A route un paquet à X par B, A ne passe pas ce route à B

35

Example de "split horizon"



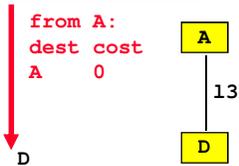
36

Split horizon

A

dest link cost		
A	local	0
B	13	3
D	13	1
C	13	3
E	13	2

"Split horizon" peut résoudre le problème de comptage jusqu'à l'infini.



dest link cost		
D	local	0
A	13	1
B	16	∞
C	16	∞
E	16	∞

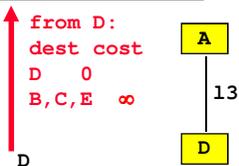
37

Split horizon

A

dest link cost		
A	local	0
B	13	∞
D	13	1
C	13	∞
E	13	∞

- "Split horizon" peut résoudre le problème de comptage jusqu'à l'infini.



dest link cost		
D	local	0
A	13	1
B	16	∞
C	16	∞
E	16	∞

38

Routage interne- protocole de vecteur de distance

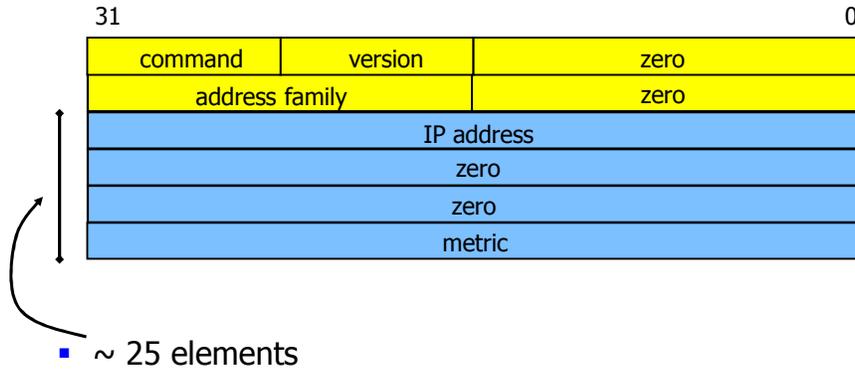
39

RIP v1

- protocole à **vecteur de distance** utilisant une métrique statique.
- Métrique – sauts, limitation: 15
 - $\infty = 16$
- Utilise une variante du *split horizon* pour éviter le problème du comptage jusqu'à l'infini.
- ID de réseau de destination par l'adresse IP
 - Pas de prefix/subnet
- Encapsulé par UDP, porte 520
- Version variée (**routed** sur Unix)
- les routeurs s'échangent leur vecteurs de distance toutes les 30 secondes

40

Format de message



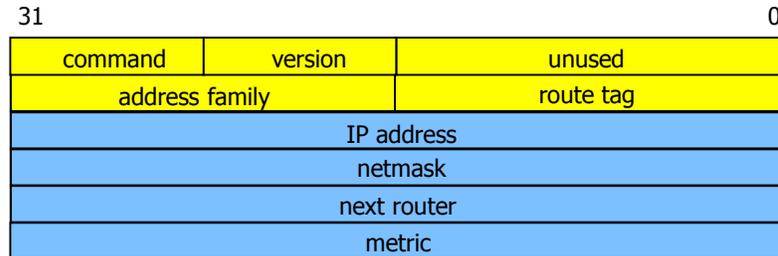
41

RIP v2 (RFC 2453)

- Sous reseaux
 - Supporte CIDR, prefixe réseau et masque de sous réseau
- Authentication
- Multicast
 - multicast adresse: 224.0.0.9

42

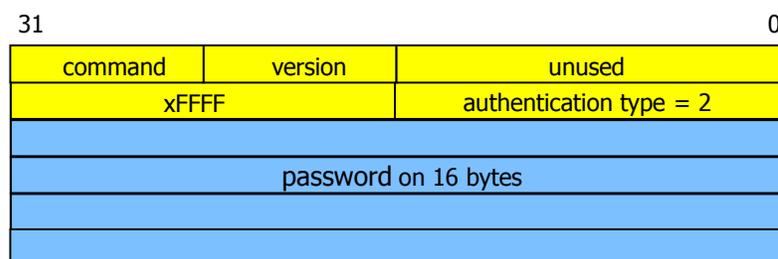
Format de message



- *Command, version* ~ v1.0
- *Address family* - authentication
- *Next router*
 - Utilise à la frontière de domain)
- *Route tag*
 - Pour route externe (par BGP)

43

Authentication simple



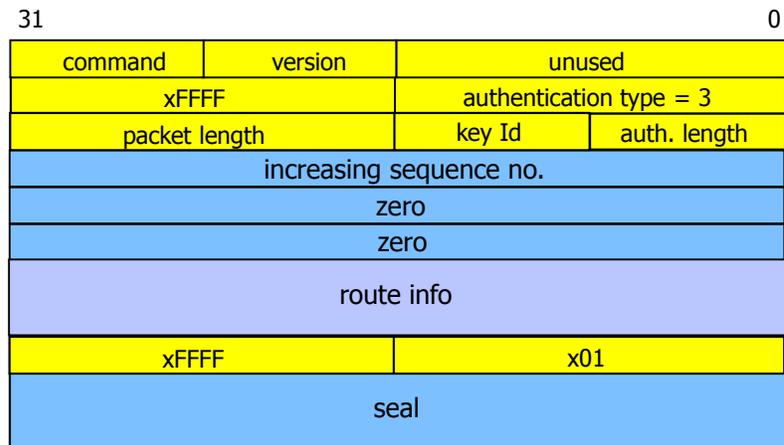
- Configuration de port (/etc/gated.conf)


```

rip yes {
    interface all
    version 2 multicast
    authentication simple "qptszwmz"
}
      
```

44

MD5 authentication



45

MD5 authentication

- *Seal*
 - MD5 digest
- Configuration de port (/etc/gated.conf)


```
rip yes {
    interface all
    version 2 multicast
    authentication md5 "qptszwmz"
}
```

46

Conclusion de RIP

- Protocole à vecteur de distance
 - Unix: `routed`
- Simple
- Convergence: lente
- fonctionne bien sur des systèmes de petites taille...