

Chapitre IV

La couche réseau

Le plan de contrôle

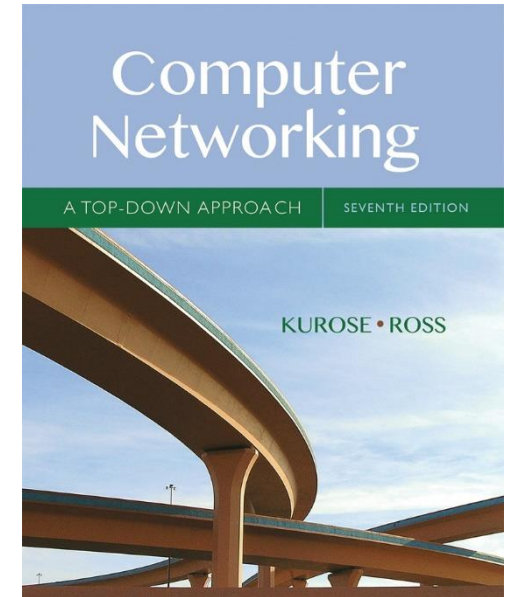
A note on the use of these Powerpoint slides:

We're making these slides freely available to all (faculty, students, readers). They're in PowerPoint form so you see the animations; and can add, modify, and delete slides (including this one) and slide content to suit your needs. They obviously represent a *lot* of work on our part. In return for use, we only ask the following:

- If you use these slides (e.g., in a class) that you mention their source (after all, we'd like people to use our book!)
- If you post any slides on a www site, that you note that they are adapted from (or perhaps identical to) our slides, and note our copyright of this material.

Thanks and enjoy! JFK/KWR

© All material copyright 1996-2016
J.F Kurose and K.W. Ross, All Rights Reserved



*Computer
Networking: A Top
Down Approach*
7ème édition
Jim Kurose, Keith Ross
Addison-Wesley
2017

Routage hiérarchique

le routage étudié est idéalisé

- ❖ tous les routeurs sont identiques
- ... loin de la réalité*

mise en échelle difficile:

avec des milliards de destinations:

- ❖ impossible de les stocker dans un routeur!
- ❖ la quantité des échanges n'est pas tolérable!

autonomie de l'administration

- ❖ internet = réseau des réseaux
- ❖ chaque admin veut contrôler le routage dans son propre réseau

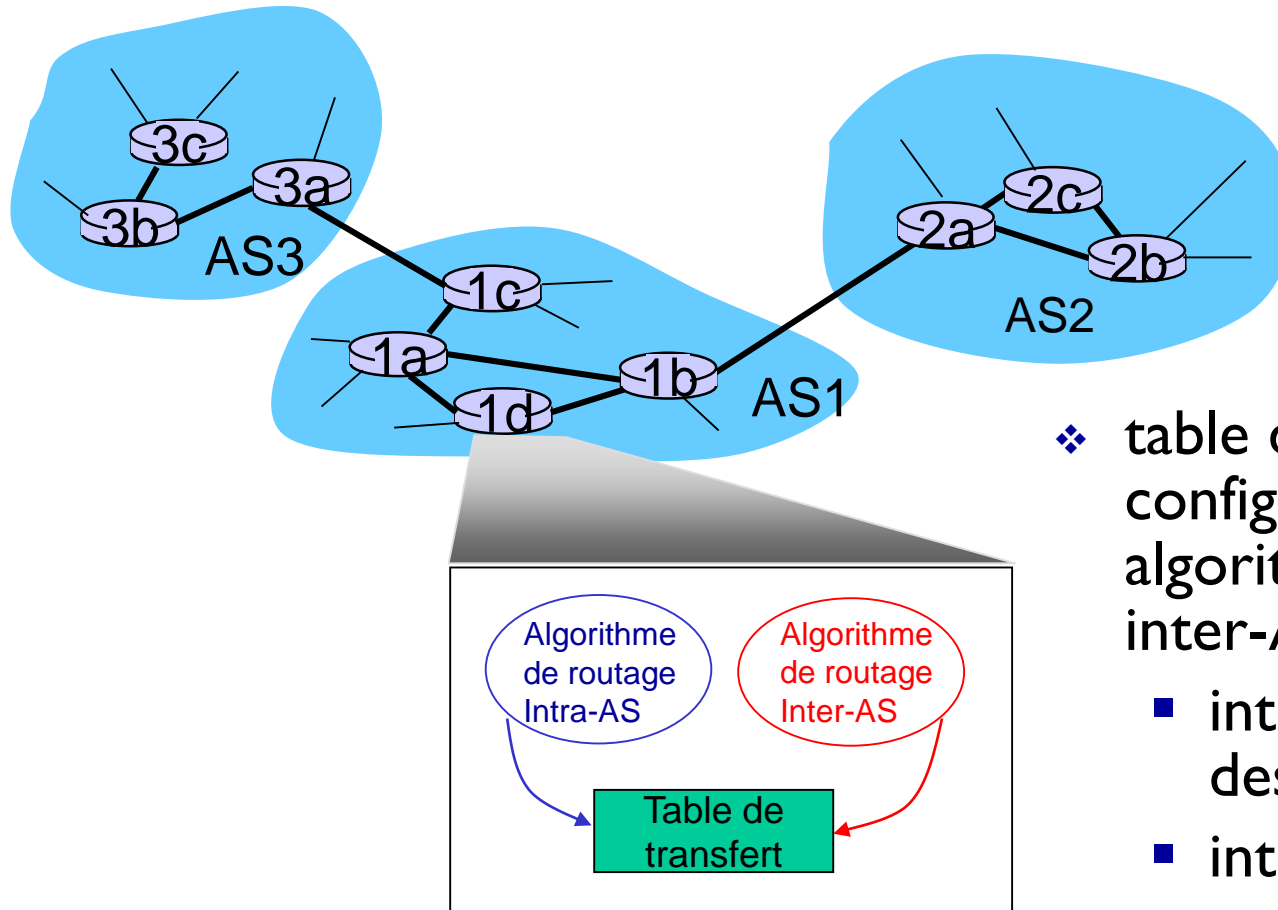
Routage hiérarchique

- ❖ diviser les routeurs en régions, “**systèmes autonomes**” (AS)
- ❖ les routeurs dans le même AS exécutent le même protocole
 - **routage “intra-AS”**
 - les routeurs dans des AS différents peuvent utiliser des protocoles intra-AS différents

routeur passerelle:

- ❖ sur le bord de son propre AS
- ❖ possède un lien au routeur d'un autre AS
- ❖ Les numéros des ASs sont attribués par ICANN

Interconnecter les AS



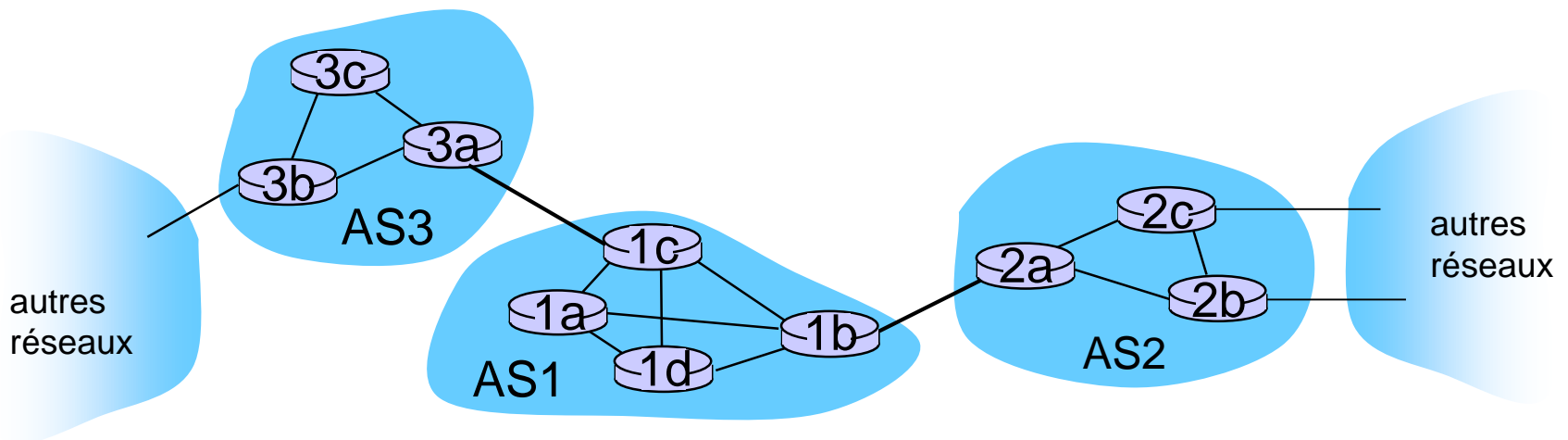
- ❖ table de transfert configurée par les algorithmes de routage inter-AS et intra-AS
 - intra-AS pour les destinations internes
 - inter-AS & intra-AS pour les destinations externes

Routing inter-AS

- ❖ un routeur de AS1 reçoit un datagramme destiné hors AS1:
 - routeur transfère le paquet au routeur passerelle, mais lequel?

AS1 doit:

1. connaître quelles dests sont accessibles via AS2, et lesquelles via AS3
2. propager cette info à tous les routeurs de AS1



Routage Intra-AS

- ❖ en anglais *interior gateway protocols (IGP)*
- ❖ Les protocoles les plus connus:
 - RIP: *Routing Information Protocol*
 - OSPF: *Open Shortest Path First*
 - IGRP: *Interior Gateway Routing Protocol*
(propriété de Cisco)
 - IS-IS: *Intermediate System to Intermediate System*

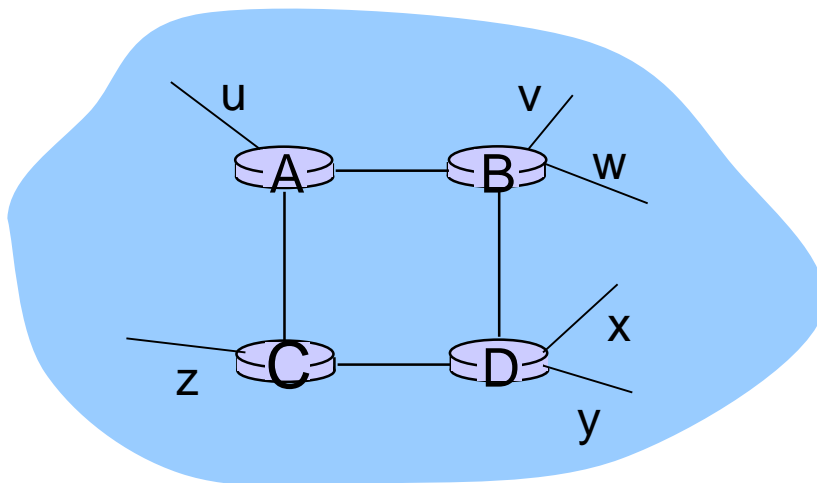
Chapitre IV (suite): plan

4.5 protocoles de routage dans l'Internet

- RIP
- OSPF
- BGP

RIP (Routing Information Protocol)

- ❖ depuis 1982 dans les distributions BSD-UNIX
- ❖ algorithme à vecteurs de distance
 - métrique de distance: # sauts (max = 15 sauts), le coût de chaque lien 1
 - vecteurs échangés avec les voisins chaque 30 sec dans les “messages réponses” (alias *advertisement ou annonce*)
 - chaque annonce: liste d'au plus 25 adresses de *sous-réseau* destination



du routeur A au sous-réseau:

<u>sous-réseau</u>	<u>saut</u>
u	1
v	2
w	2
x	3
y	3
z	2

RIP: exemple

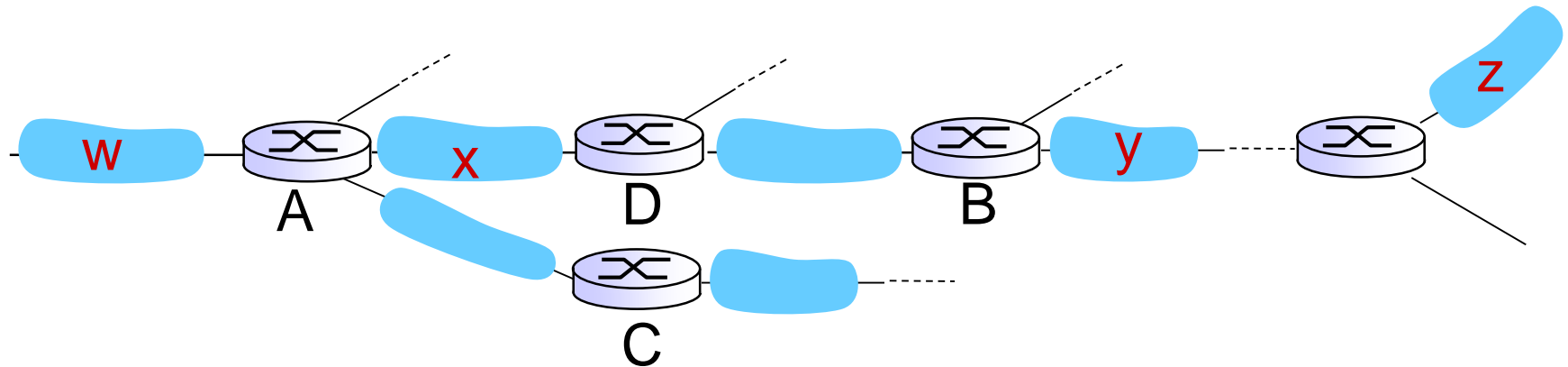


table de routage dans D

sous-réseau dest	prochain routeur	# saut à dest
W	A	2
y	B	2
z	B	7
x	--	1
....

RIP: exemple

A-à-D annonce

dest	prochain	sauts
w	-	1
x	-	1
z	C	4
...

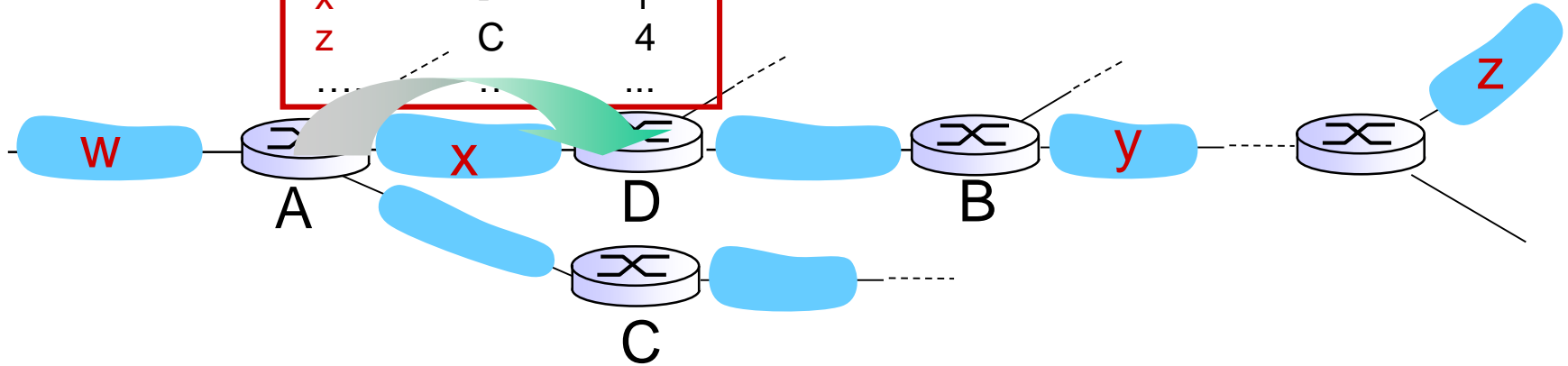


table de routage dans D

sous-réseau dest	prochain routeur	# saut à dest
w	A	2
y	B	2
z	B → A	7 → 5
x	--	1
....

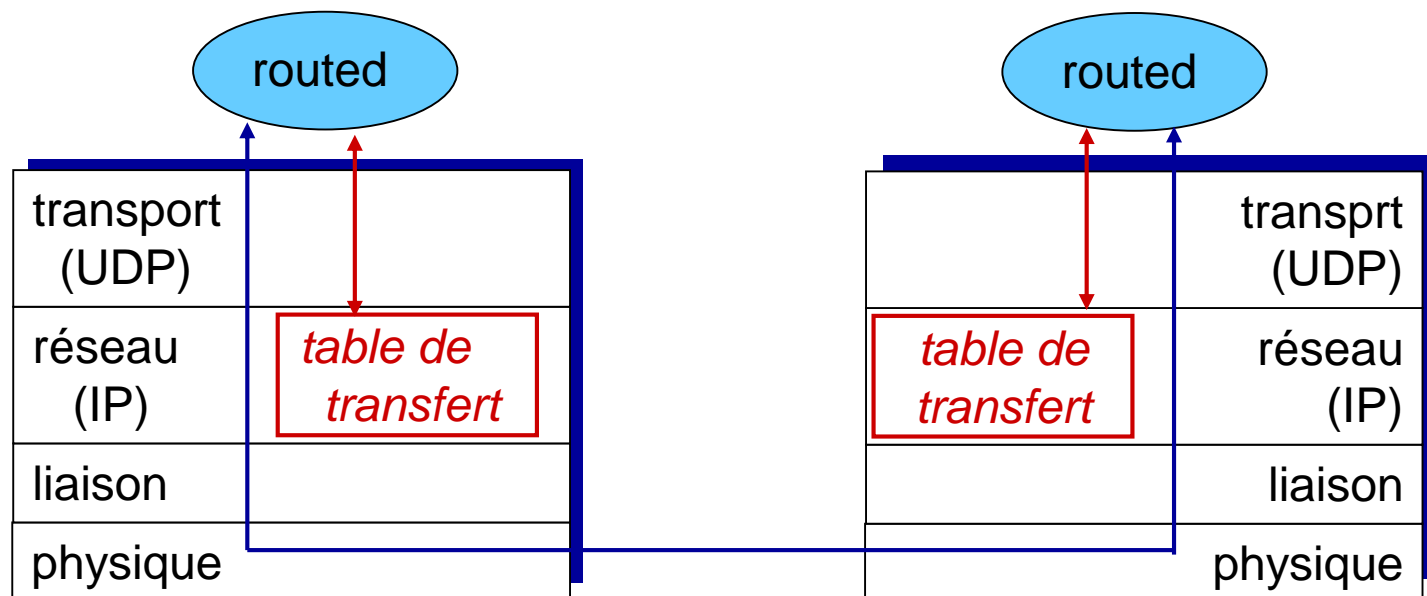
RIP: défaillance de lien, récupération

si aucune annonce n'est entendu après 180 sec →
voisin/liens déclaré mort

- les routes via le voisin considérées non-valides
- envoyer des annonces aux voisins
- les voisins envoient à leur tour des annonces (si leurs tables changent)
- l'information se propage sur la totalité du réseau
- l'utilisation du “*poison reverse*” (distance infinie = 16 sauts)

Gestion des tables RIP

- ❖ les tables de routage RIP sont gérées par un processus (couche application) route-d (daemon)
- ❖ les “annonces” envoyés dans des paquets UDP



Chapitre IV (suite): plan

4.5 protocoles de routage dans l'Internet

- RIP
- OSPF
- BGP

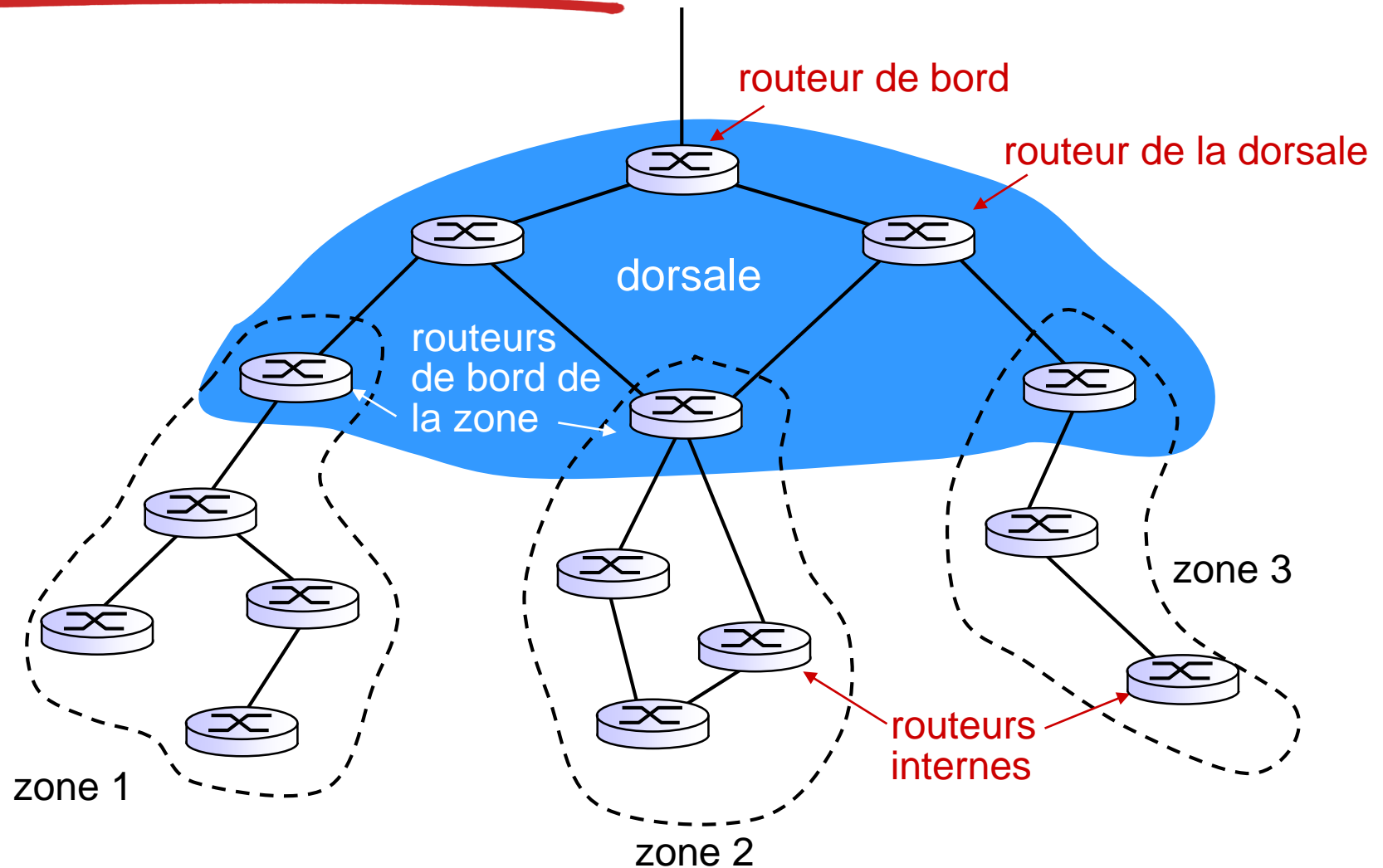
OSPF (Open Shortest Path First)

- ❖ “open”: disponible au public
- ❖ utilise un algorithme à états de lien
 - diffusion de paquets LS
 - la topologie est disponible dans chaque nœud
 - utilise l'algorithme de Dijkstra
- ❖ les annonces des états de lien inondent l'AS en entier
 - les messages OSPF utilisent directement IP (pas de TCP ou UDP)
- ❖ Les annonces si changement d'un état de lien ou chaque 30 min.

Caractéristiques de OSPF (pas dans RIP)

- ❖ **sécurité**: tous les messages sont authentifiés
- ❖ **multiple chemins** ayant le même coût sont permis (un seul chemin dans RIP)
- ❖ utilise le champ **TOS** (*type of service*) pour supporter plusieurs services en calculant différentes routes
- ❖ **supporte** le routage hiérarchique

OSPF hiérarchique



OSPF hiérarchique

- ❖ *hiérarchie à deux niveaux*: local, dorsale.
 - les annonces uniquement dans la zone
 - chaque nœud connaît la topologie de la zone; connaît les chemins les moins onéreux vers les autres zones.
- ❖ *routeur du bord de la zone*: connaît les distances vers les routeurs de sa propre zone et les annonce aux autres « routeurs de bord de zone »
- ❖ *routeurs de la dorsale*: utilise OSPF pour le routage à l'intérieur de la dorsale.
- ❖ *routeurs du bord*: se connecte aux autres AS.

Chapitre IV (suite): plan

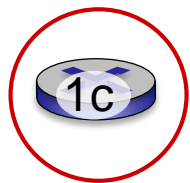
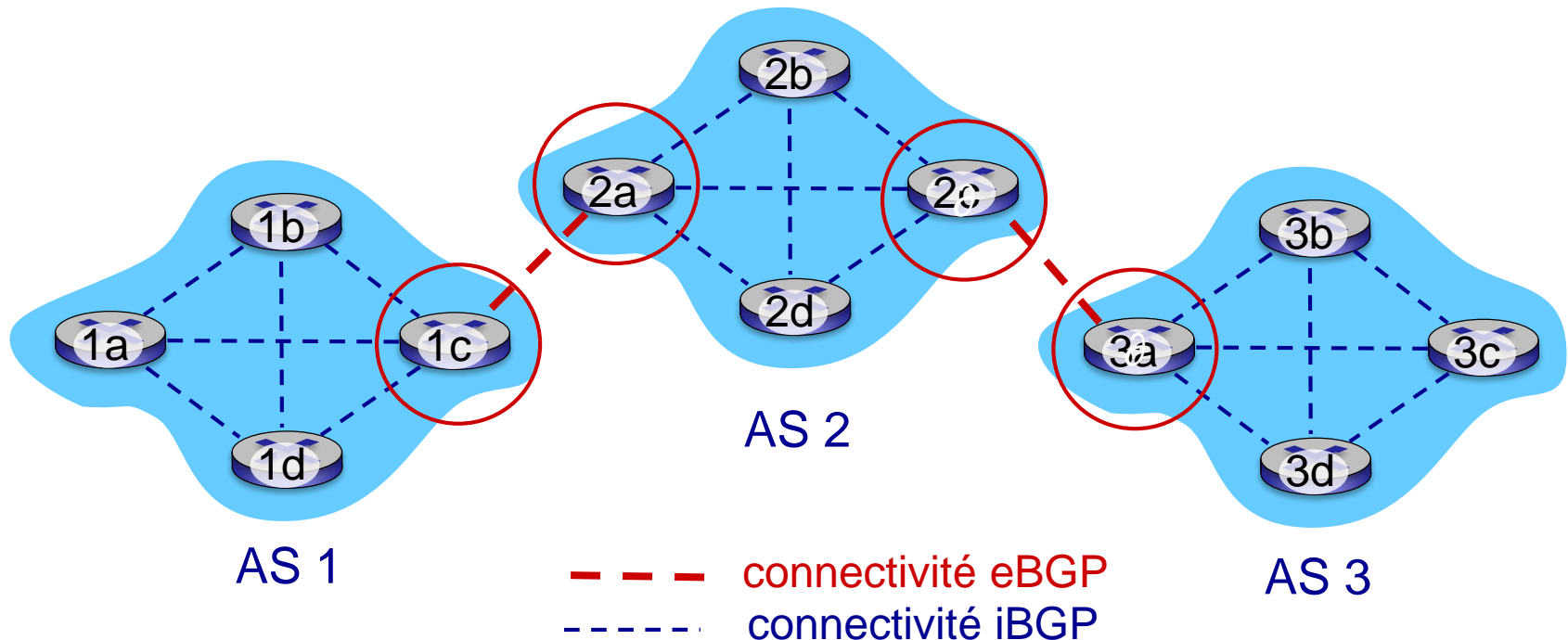
4.5 protocoles de routage dans l'Internet

- RIP
- OSPF
- BGP

Routage inter-AS: BGP

- ❖ **BGP (Border Gateway Protocol):** *le Protocole de routage inter-AS*
 - “La colle qui rassemble les différents morceaux d’Internet”
- ❖ BGP fournit à chaque AS les moyens de:
 - **eBGP:** échanger d’information entre ASs voisins.
 - **iBGP:** propager cette information à tous les routeurs internes de l’AS.
 - déterminer les bonnes routes vers les autres réseaux en se basant sur l’information échangé.
- ❖ BGP permet à chaque sous-réseau de dire aux autres : “*je suis là*”

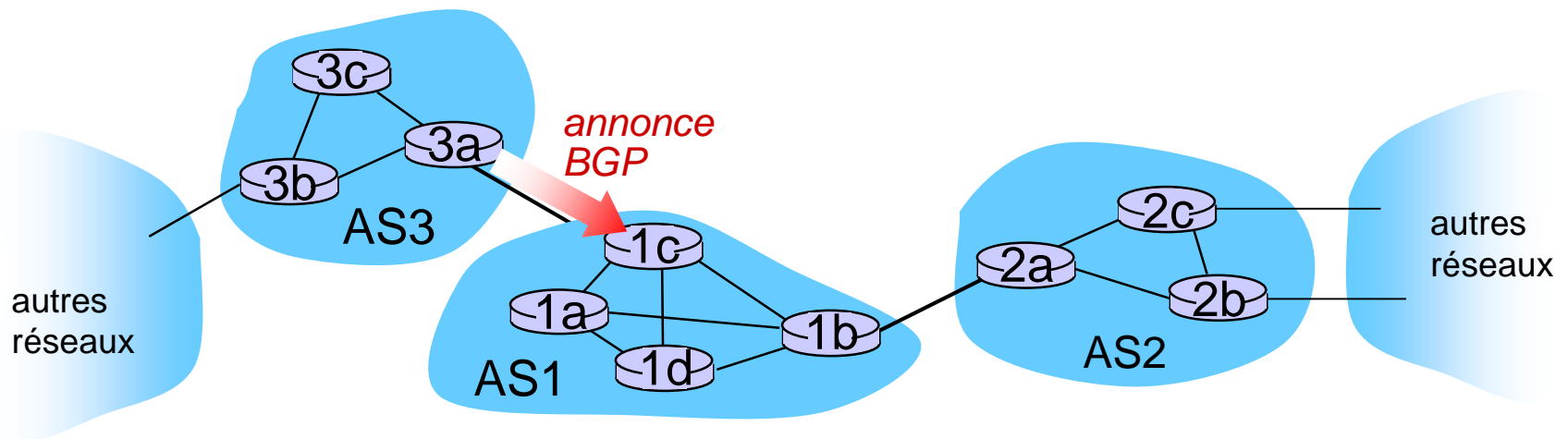
Connexions eBGP, iBGP



Les routeurs de passerelle utilise les protocoles eBGP et iBGP

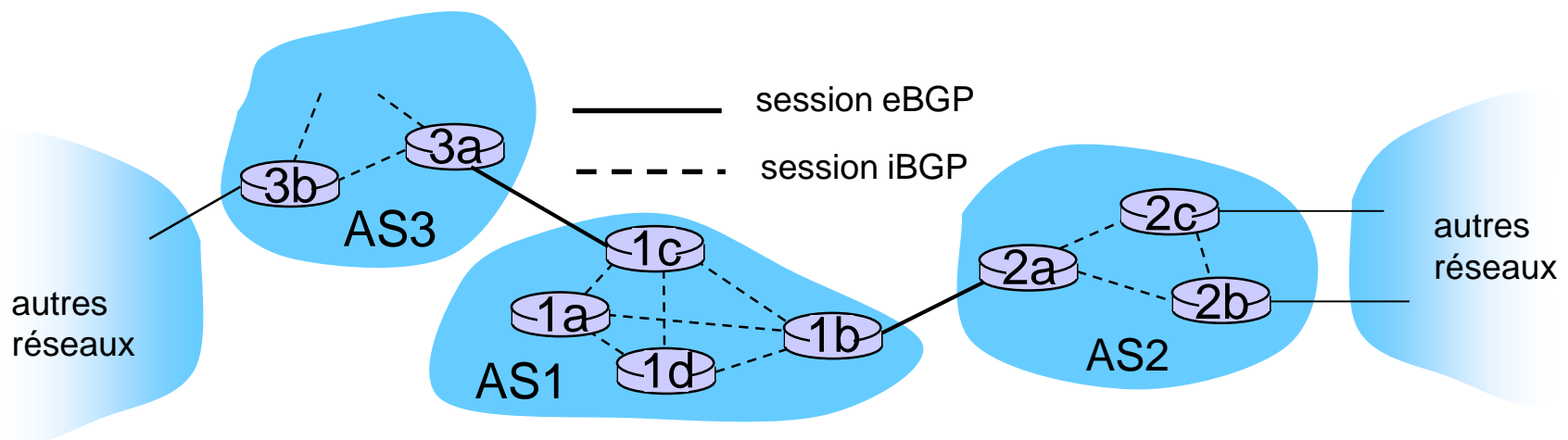
Bases de BGP

- ❖ **une session BGP** : deux routeurs BGP (« *peers* ») échangent des messages BGP :
 - annonces des *chemins* vers les différents préfixes des réseaux
 - utilisation de connections semi-permanentes TCP (port 179)
- ❖ Quand AS3 annonce un préfixe à AS1:
 - AS3 *promet* qu'il s'occupera des paquets vers ce préfixe
 - AS3 peut utiliser des agrégations de préfixes



Bases de BGP

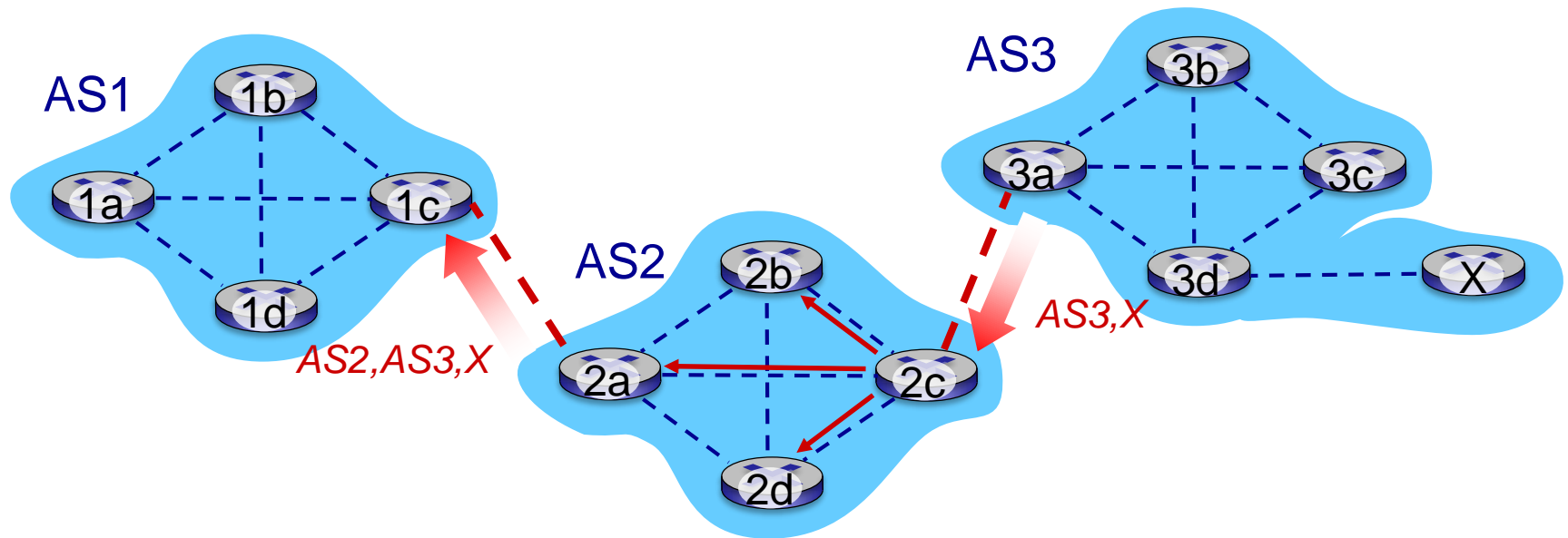
- ❖ en utilisant une session eBGP entre 3a et 1c, AS3 envoie l'info sur l'accessibilité d'un préfixe à AS1.
 - 1c peut utiliser iBGP pour distribuer cette nouvelle info à tous les routeurs dans AS1
 - 1b peut annoncer cette info à AS2 via 1b-à-2a (session eBGP)
- ❖ quand un routeur apprend un nouveau préfixe, il crée une entrée correspondante dans sa table.



Attributs des chemins

- ❖ Les annonces BGP comprennent des attributs
 - préfixe + attributs = “route”
- ❖ deux attributs importants :
 - **AS-PATH**: contient les numéros des ASs par lesquels l’annonce est passée: ex., AS 67, AS 17
 - **NEXT-HOP**: l’adresse de l’interface du routeur passerelle (de l’autre AS) responsable du transfert des paquets vers le préfixe
- ❖ un routeur passerelle qui reçoit une annonce peut l’accepter ou la rejeter
 - selon sa politique

Annonce de chemin BGP

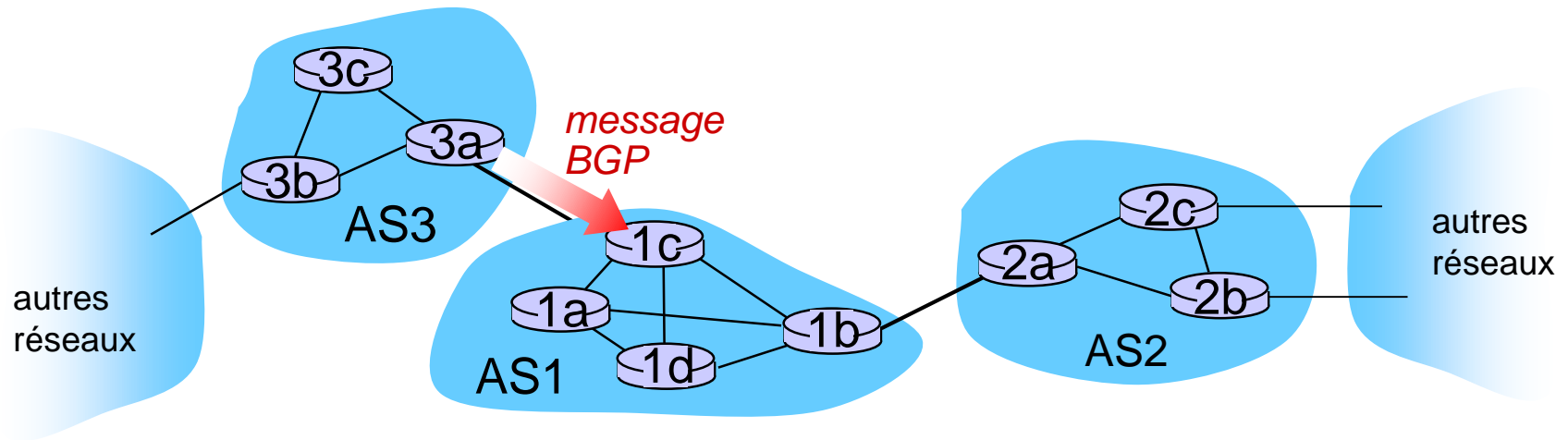


- Le routeur 2c de AS2 reçoit une annonce de chemin **AS3,X** (via eBGP) du routeur 3a de AS3
- Selon sa politique, le routeur 2c de AS2 accepte le chemin **AS3,X**, le propage (via iBGP) à tous les routeurs AS2
- Selon sa politique, le routeur 2a de AS2 annonce (via eBGP) le chemin **AS2,AS3,X** au routeur 1c de AS1

messages BGP

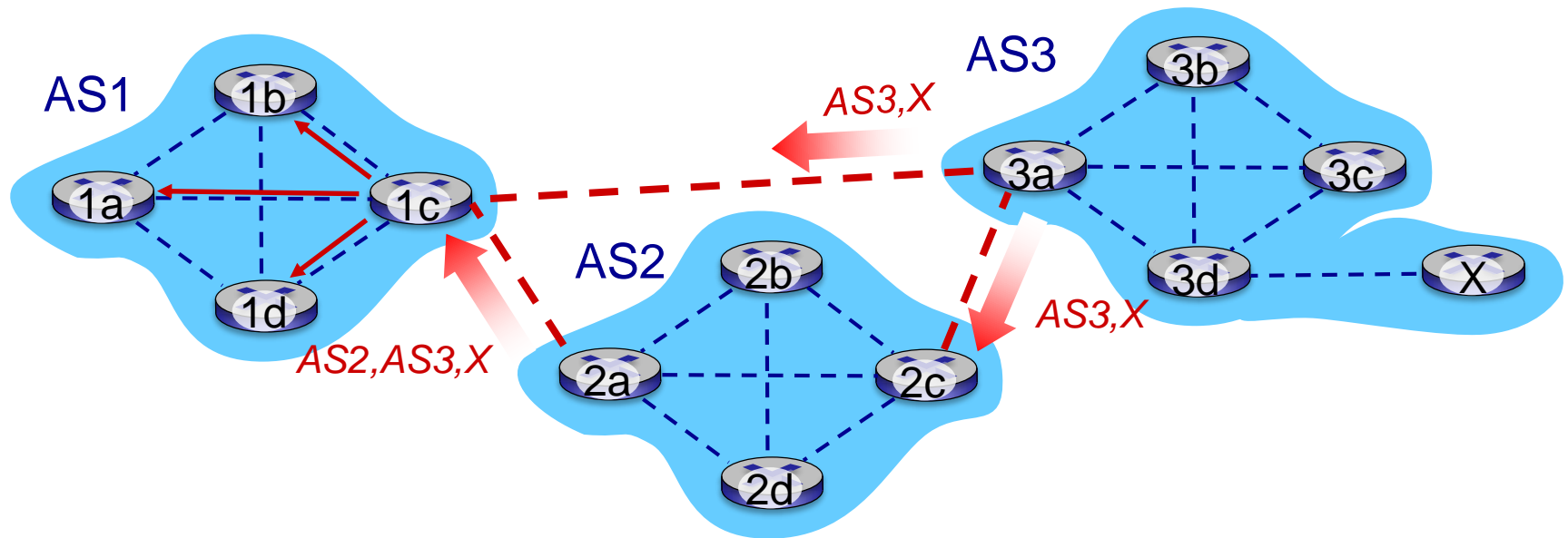
- ❖ les messages BGP échangés entre deux routeurs passerelles *peers* sur une connexion TCP
- ❖ Les messages BGP :
 - **OPEN**: ouvrir une connexion TCP vers un routeur BGP *peer* lointain et authentifier le BGP *peer* d'émission
 - **UPDATE**: annoncer un nouveau chemin (ou éliminer un ancien)
 - **KEEPALIVE**: garder la connexion en absence de UPDATES; aussi acquitter les requêtes ACKs OPEN
 - **NOTIFICATION**: reporter les erreurs dans les messages précédents; aussi utilisé pour fermer la connexion

Le routeur apprend le préfixe



- ❖ message BGP contient les “routes”
- ❖ rappel: “route” est un préfixe et attributs: AS-PATH, NEXT-HOP,...
- ❖ Exemple: route:
 - ❖ Prefix:138.16.64/22 ; AS-PATH: AS3 AS131 ; NEXT-HOP: 201.44.13.125

Annonce de chemin BGP



Le routeur passerelle peut apprendre plusieurs chemins vers une destination:

- ❖ Le routeur passerelle 1c de AS1 apprend le chemin **AS2,AS3,X** de 2a
 - Le routeur passerelle 1c de AS1 apprend le chemin **AS3,X** de 3a
 - Selon sa politique, Le routeur passerelle 1c de AS1 choisit le chemin **AS3,X**, et *annonce le chemin dans AS1 via iBGP*

Choisir la meilleure route BGP

- ❖ Le routeur choisit la route en se basant sur le AS-PATH le moins onéreux

- ❖ Exemple:

❖ AS2 AS17 to 138.16.64/22

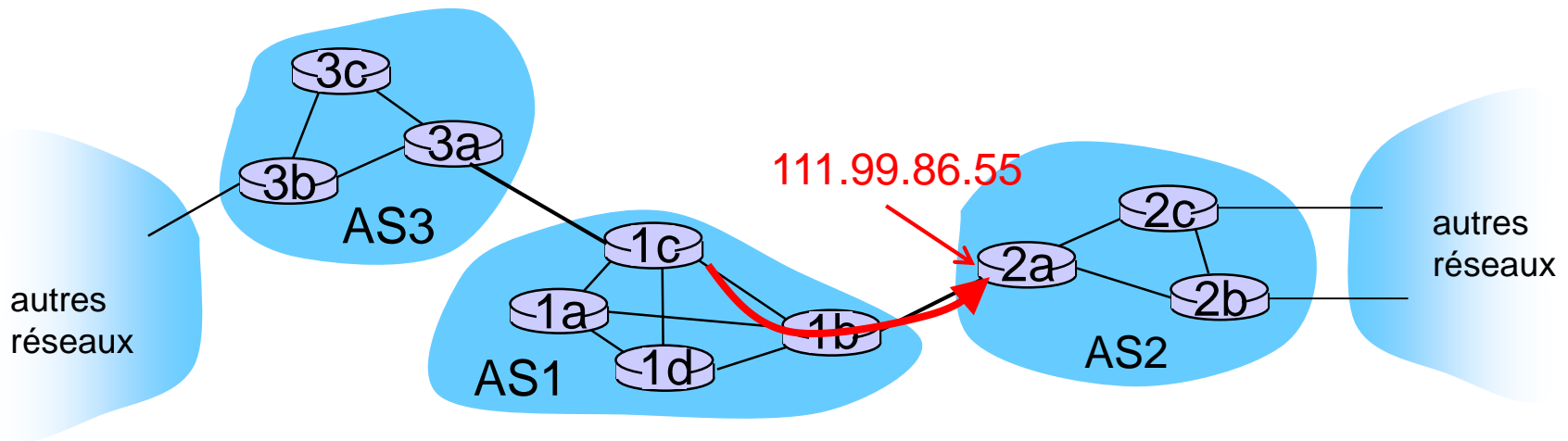
❖ AS3 AS131 AS201 to 138.16.64/22

choisit

- ❖ Et s'ils sont égaux?

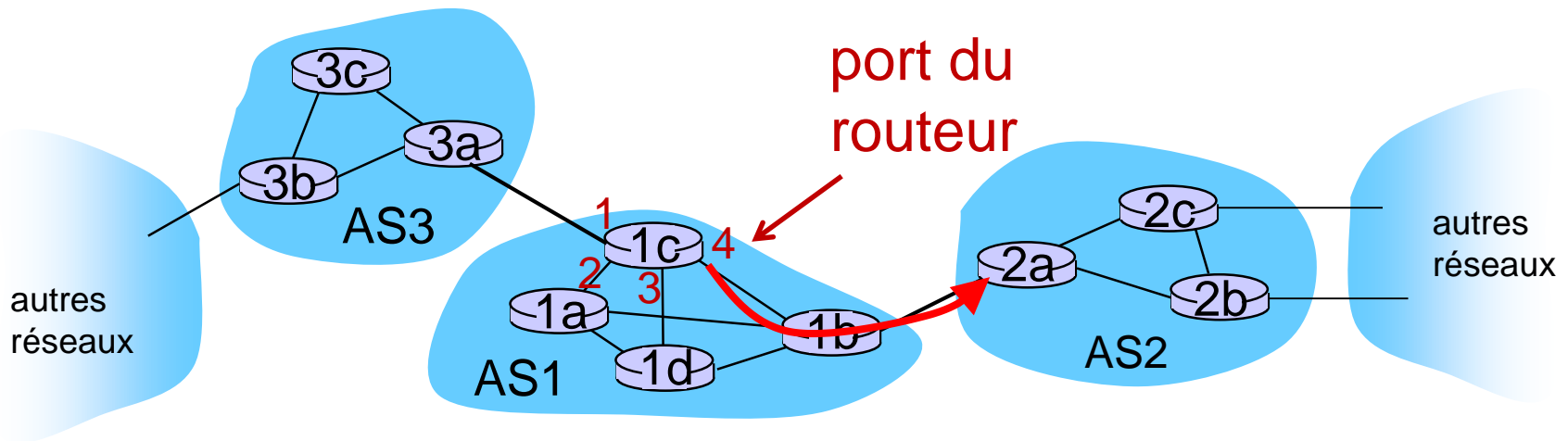
Trouver la meilleure route intra-AS

- ❖ Utiliser l'attribut NEXT-HOP
 - NEXT-HOP est l'adresse IP de l'interface du premier routeur dans AS PATH.
- ❖ Exemple:
 - ❖ AS-PATH: AS2 AS17 ; NEXT-HOP: 111.99.86.55
- ❖ Routeur utilise OSPF pour trouver le chemin de 1c à 111.99.86.55



Routeur identifie le bon port

- ❖ Identifie le port qui permet de prendre le chemin OSPF
- ❖ Ajoute l'entrée préfixe-port dans sa table:
 - (138.16.64/22 , port 4)

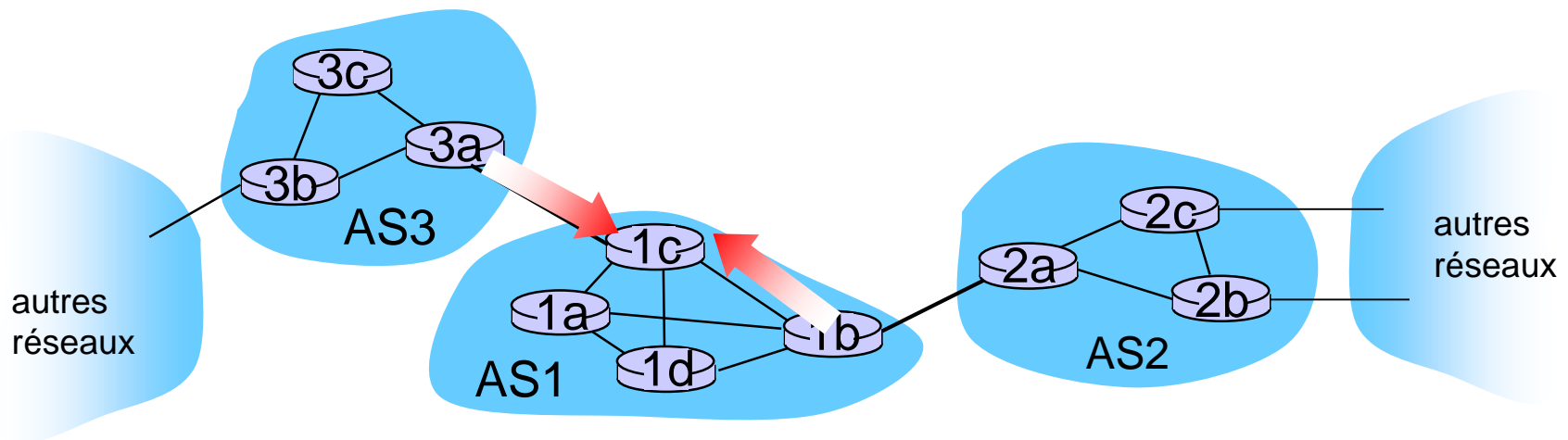


Sélection des routes dans BGP

- ❖ un routeur peut recevoir plusieurs annonces de route vers la même destination, sa sélection est basée (en ordre de priorité) sur:
 1. politique de préférence (travail de l'admin)
 2. le AS-PATH le moins onéreux (le nombre de AS dans le AS-PATH)
 3. routeur NEXT-HOP le plus proche: routage hot potato
 4. d'autres critères

Routage Hot Potato

- ❖ Supposons plusieurs meilleures inter-routes.
- ❖ Choisir la route avec NEXT-HOP le plus proche
 - Utiliser OSPF pour déterminer la plus proche passerelle
 - Q: de 1c, choisir AS3 AS131 ou AS2 AS17?
 - R: route AS3 AS201

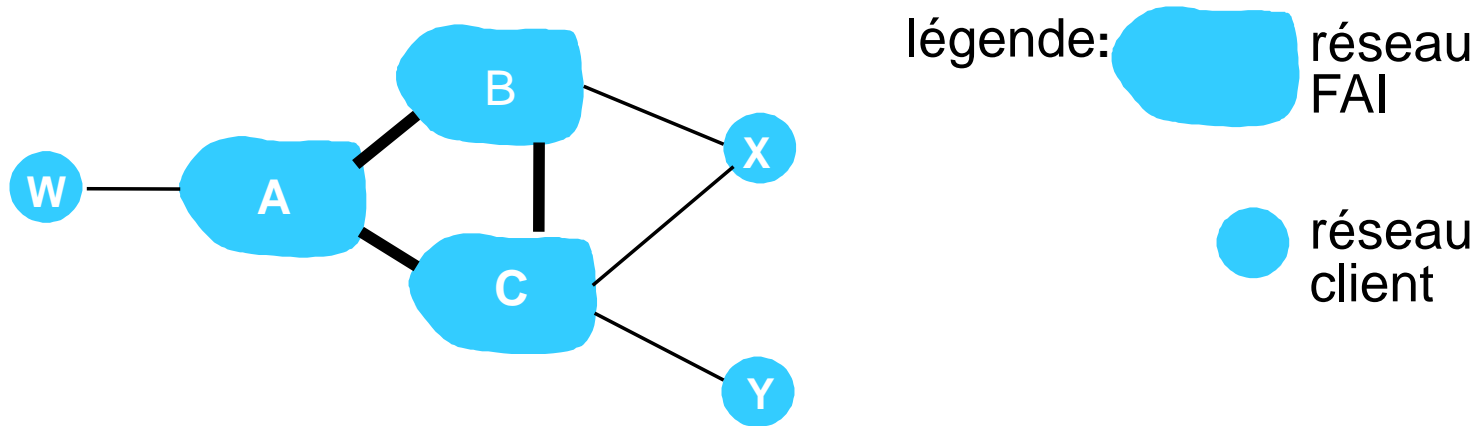


Résumé: Comment ajouter une nouvelle entrée?

Trois étapes

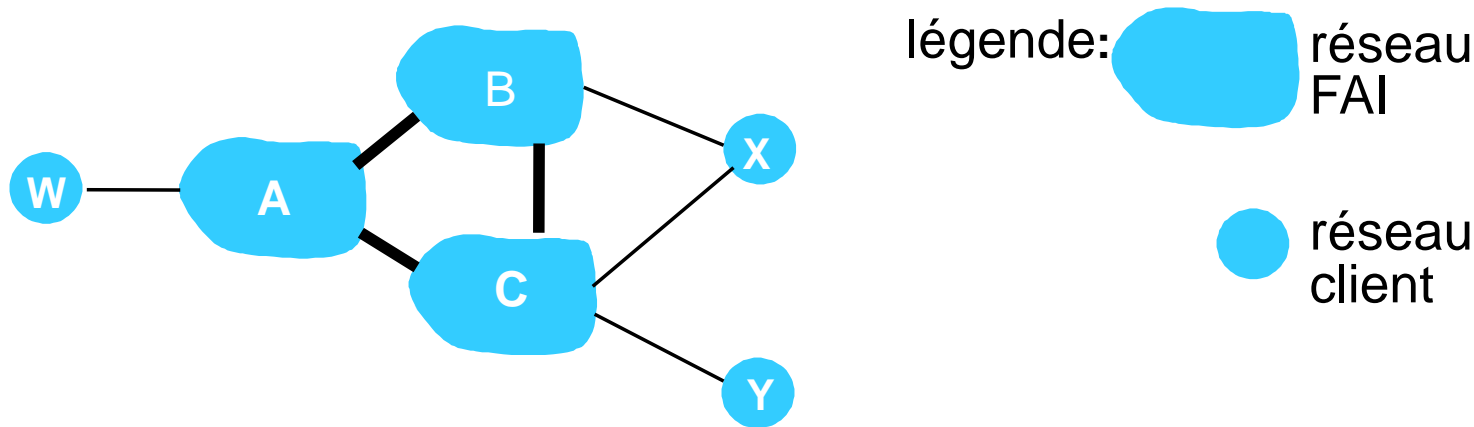
1. Le routeur apprend le préfixe
 - via les annonces BGP envoyées par les autres routeurs
2. Le routeur détermine le port de sortie pour ce préfixe
 - utilise la sélection de route BGP pour trouver la meilleure route inter-AS
 - utilise OSPF pour trouver la meilleure route intra-AS
 - identifie le port pour cette route
3. Le routeur ajoute préfixe-port dans sa table

Politique de routage BGP



- ❖ A,B,C *réseaux FAI*
- ❖ X,W,Y clients
- ❖ X attaché à deux FAIs
 - X ne veut pas router de B via X à C
 - ..alors X ne fait pas l'annonce à B concernant la route vers C

Politique de routage BGP (2)



- ❖ A annonce le chemin AW à B
- ❖ B annonce le chemin BAW à X
- ❖ Est ce que B doit annoncer BAW à C?
 - Non, pas profitable (question d'argent)
 - B veut forcer C à servir w via A
 - B veut servir seulement ses clients!

Pourquoi des routages Intra et Inter différent ?

politique:

- ❖ inter-AS: admin veut contrôler comment son trafic est routé, qui passe par son réseau.
- ❖ intra-AS: un seul admin, pas de grandes décisions (pas de politique)

mise-en-échelle:

- ❖ le routage hiérarchique réduit les tailles des tables et le volume des échanges

performance:

- ❖ intra-AS: performance est le premier objectif
- ❖ inter-AS: politique plus importante que les performances