

Livello di Rete:

Protocolli di instradamento: RIP, OSPF, BGP

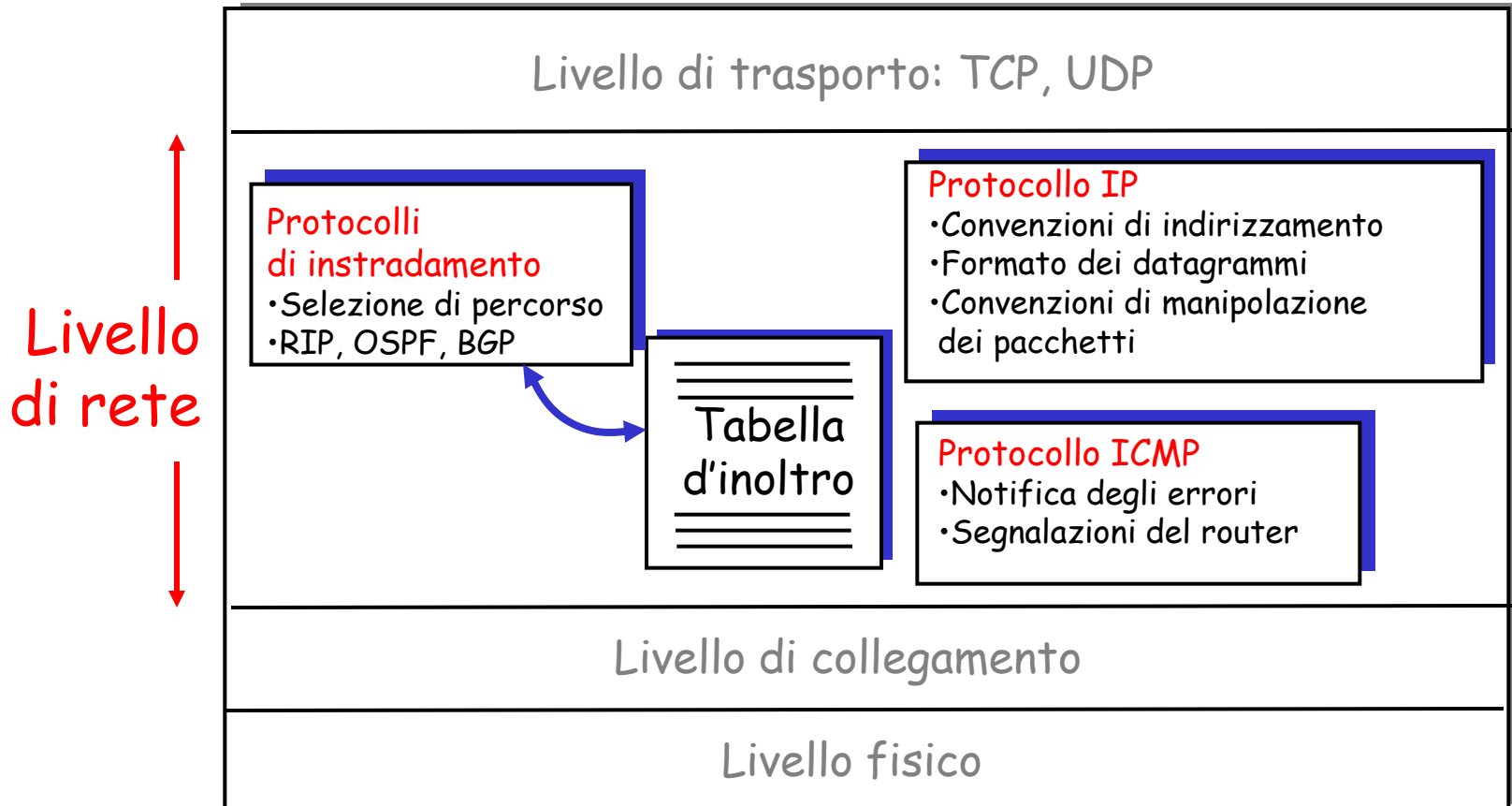
Gaia Maselli

maselli@di.uniroma1.it

Queste slide sono un adattamento delle slide fornite dal libro di testo e pertanto protette da copyright.

All material copyright 1996-2007 J.F Kurose and K.W. Ross, All Rights Reserved

# Livello di rete



# Livello di rete

Introduzione

Reti a circuito virtuale e  
a datagramma

Che cosa si trova  
all'interno di un  
router?

Protocollo Internet (IP)

- Formato dei datagrammi
- Indirizzamento IPv4
- ICMP
- IPv6

Algoritmi di  
instradamento

- Stato del collegamento
- Vettore distanza
- Instradamento gerarchico

Instradamento in  
Internet

- RIP
- OSPF
- BGP

Instradamento  
broadcast e multicast

# Instradamento gerarchico

Abbiamo fin qui visto la rete come una collezione di router interconnessi

- ❑ Ciascun router era indistinguibile dagli altri
- ❑ Visione omogenea della rete

*... nella pratica le cose non sono così semplici*

**Scala:** con 200 milioni di destinazioni:

- ❑ Archiviare le informazioni d'instradamento su ciascun host richiederebbe un'enorme quantità di memoria.
- ❑ Il traffico generato dagli aggiornamenti LS non lascerebbero banda per i pacchetti di dati!
- ❑ DV non convergerebbe mai

**Autonomia amministrativa:**

- ❑ Internet = la rete delle reti
- ❑ Da un punto di vista ideale, ciascuno dovrebbe essere in grado di amministrare la propria rete nel modo desiderato, pur mantenendo la possibilità di connetterla alle reti esterne.

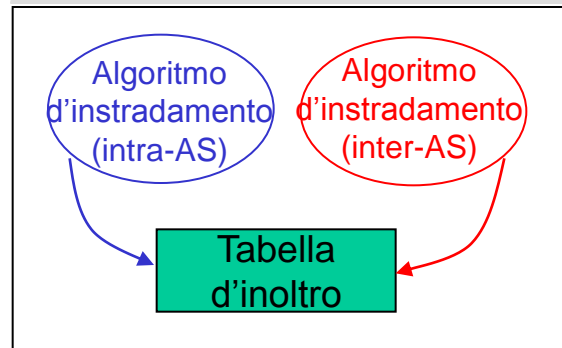
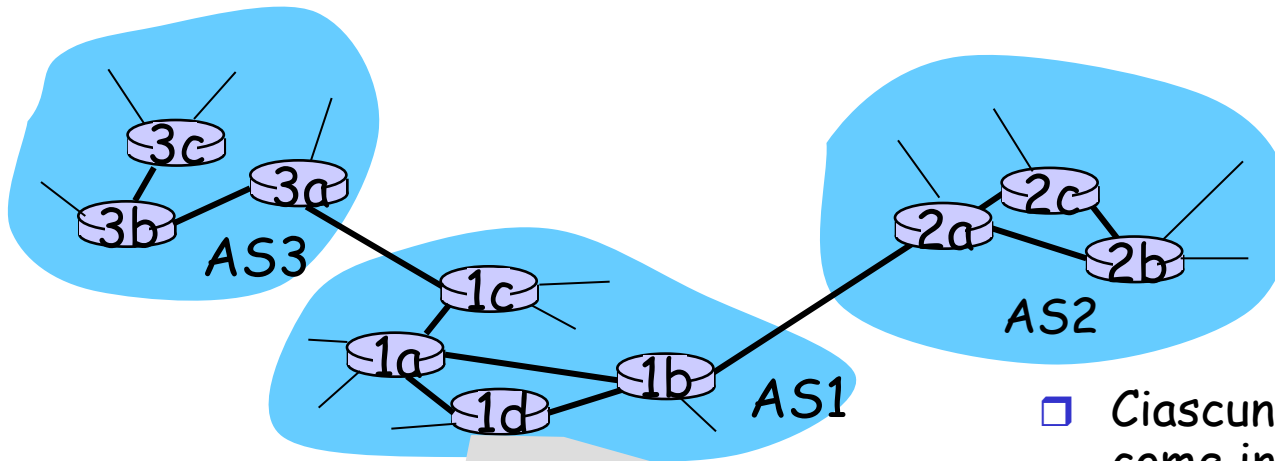
# Instradamento gerarchico

- ❑ Organizzazione di router in *sistemi autonomi (AS, autonomous system)*.
- ❑ I router di un gruppo autonomo eseguono lo stesso algoritmo d'instradamento.
  - Protocollo d'instradamento interno al sistema autonomo (intra-AS).
  - I router appartenenti a differenti AS possono eseguire protocolli d'instradamento intra-AS diversi

## Router gateway

- ❑ Connettono gli AS tra loro
- ❑ Hanno il compito aggiuntivo d'inoltrare pacchetti a destinazioni esterne.

# Sistemi autonomi interconnessi



- Ciascun sistema autonomo sa come inoltrare pacchetti lungo il percorso ottimo verso qualsiasi destinazione interna al gruppo
  - Il sistema AS1 ha quattro router
  - I sistemi AS2 e AS3 hanno tre router ciascuno
  - I protocolli d'instradamento dei tre sistemi autonomi non sono necessariamente gli stessi
  - I router 1b, 1c, 2a e 3a sono gateway

# Instradamento tra sistemi autonomi

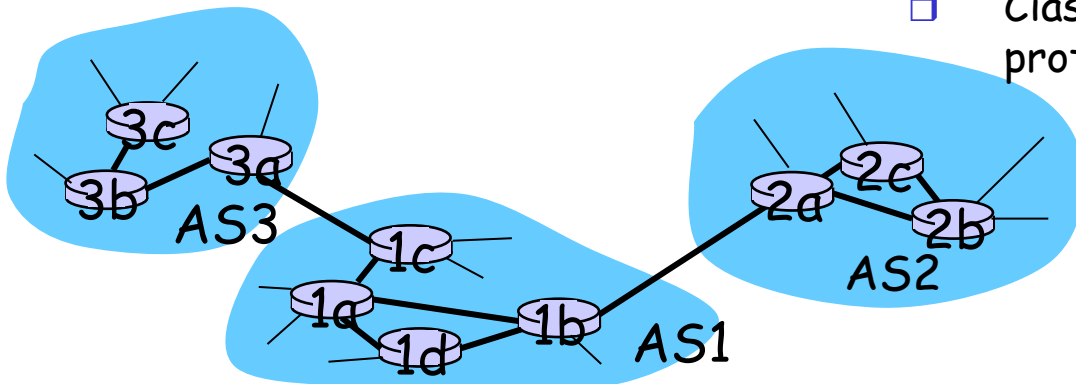
- Semplice se c'è un solo collegamento verso l'esterno
- Ma supponiamo che un router in AS1 riceva un datagramma la cui destinazione ricade al di fuori di AS1
  - Il router dovrebbe inoltrare il pacchetto verso uno dei due gateway. Ma quale??

## AS1 deve:

1. Sapere quali destinazioni sono raggiungibili attraverso AS2 e quali attraverso AS3
2. Informare tutti i router all'interno del sistema AS1 in modo che ciascuno possa configurare la propria tabella d'inoltro per gestire destinazioni esterne

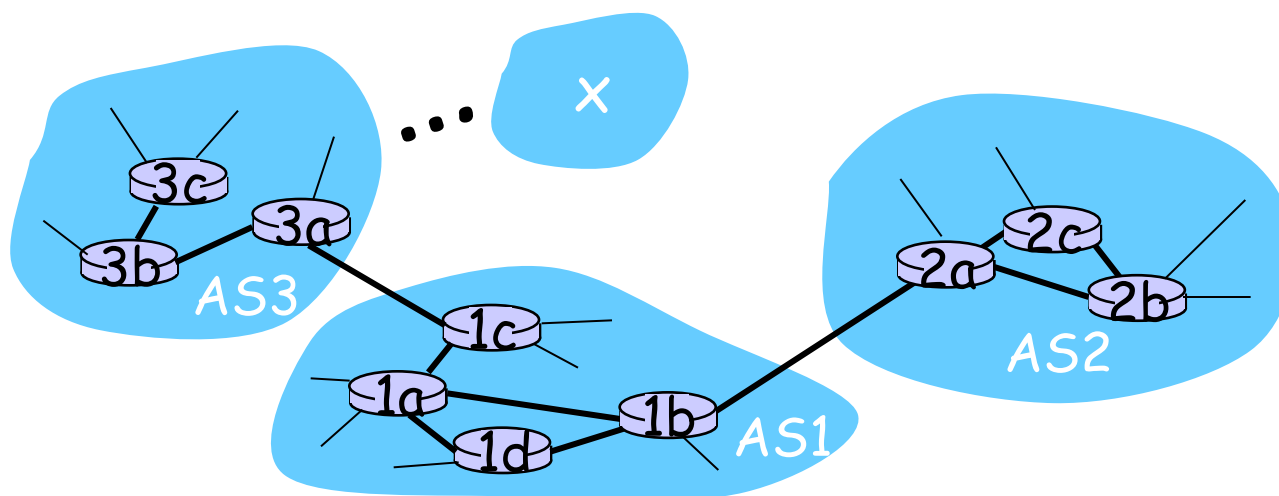
## Questo è l'instradamento inter-AS!

- Stesso protocollo tra sistemi autonomi
- Ciascun router riceve informazioni dai protocolli inter e intra AS e configura la propria tabella di routing



## Esempio: impostare la tabella d'inoltro nel router 1d

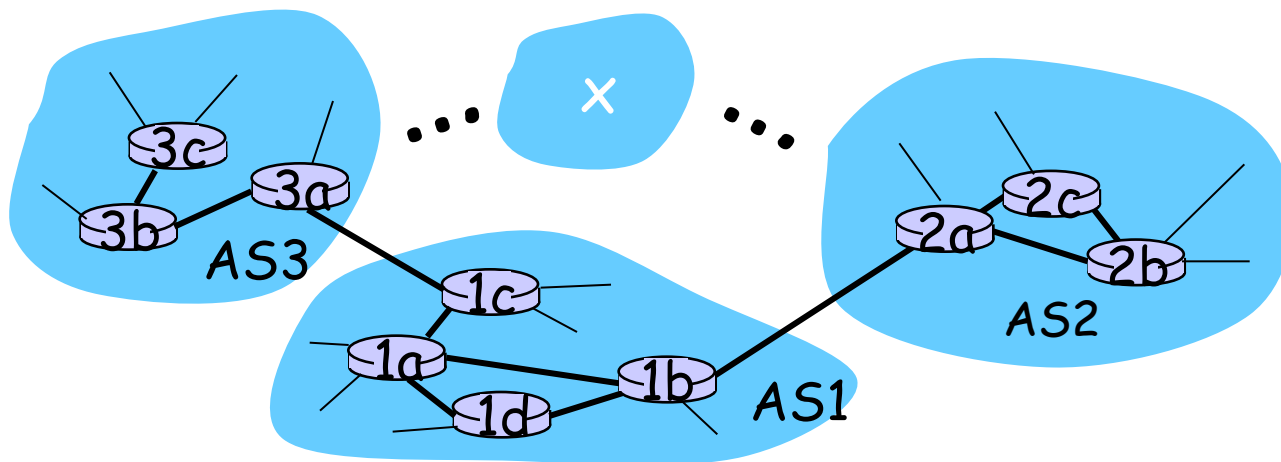
- Supponiamo che AS1 apprenda dal proprio protocollo d'instradamento inter-AS che la sottorete  $x$  è raggiungibile da AS3 (gateway 1c), ma non da AS2.
- Il protocollo inter-AS propaga questa informazione a tutti i propri router.
- Il router 1d determina, partendo dall'informazione fornita dal protocollo intra-AS, l'interfaccia  $I$  del router sul percorso a costo minimo dal router 1d al gateway 1c.
- Il router 1d può inserire la riga  $(x, I)$  nella propria tabella d'inoltro.





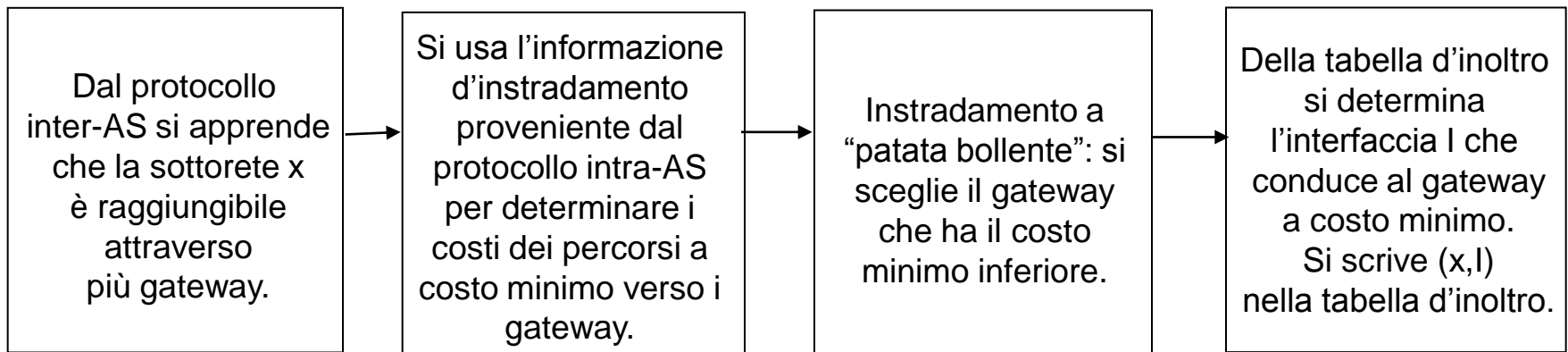
# Esempio: scegliere fra più AS

- Supponiamo inoltre che AS1 apprenda dal protocollo d'instradamento tra sistemi autonomi che la sottorete  $x$  è raggiungibile da AS2 (tramite 1b) e da AS3 (tramite 1c).
- AS1 propaga tale informazione ai suoi router, compreso 1d
- Al fine di configurare la propria tabella d'inoltro, il router 1d dovrebbe determinare a quale gateway, 1b o 1c, indirizzare i pacchetti destinati alla sottorete  $x$ .
- Anche questo è un compito che spetta al protocollo d'instradamento inter-AS!



# Esempio: scegliere fra più AS

- ❑ **Instradamento a patata bollente:** il sistema autonomo si sbarazza del pacchetto (patata bollente) non appena possibile (ovvero nel modo meno costoso possibile).
- ❑ Si sceglie il gateway (che ha un percorso verso la destinazione) con il minor costo.



# Instradamento tra sistemi autonomi

- ❑ Protocollo BGP (Border Gateway Protocol)
- ❑ Un sistema autonomo presenta una certa flessibilità nel decidere quali destinazioni pubblicizzare ai propri vicini
- ❑ Decisione prevalentemente "politica" che dipende da questioni economiche

# Livello di rete

Introduzione

Reti a circuito virtuale e  
a datagramma

Che cosa si trova  
all'interno di un  
router?

Protocollo Internet (IP)

- Formato dei datagrammi
- Indirizzamento IPv4
- ICMP
- IPv6

Algoritmi di  
instradamento

- Stato del collegamento
- Vettore distanza
- Instradamento gerarchico

Instradamento in  
Internet

- RIP } Intra-AS
- OSPF } Intra-AS
- BGP } Inter-AS

Instradamento  
broadcast e multicast

# Instradamento in Internet

- ❑ I protocolli d'instradamento intra-AS sono noti anche come **protocolli gateway interni (IGP)**
- ❑ I protocolli intra-AS più comuni sono:
  - RIP: routing information protocol
  - OSPF: open shortest path first
  - IGRP: Interior Gateway Routing Protocol (di proprietà Cisco)

# Livello di rete

Introduzione

Reti a circuito virtuale e  
a datagramma

Che cosa si trova  
all'interno di un  
router?

Protocollo Internet (IP)

- Formato dei datagrammi
- Indirizzamento IPv4
- ICMP
- IPv6

Algoritmi di  
instradamento

- Stato del collegamento
- Vettore distanza
- Instradamento gerarchico

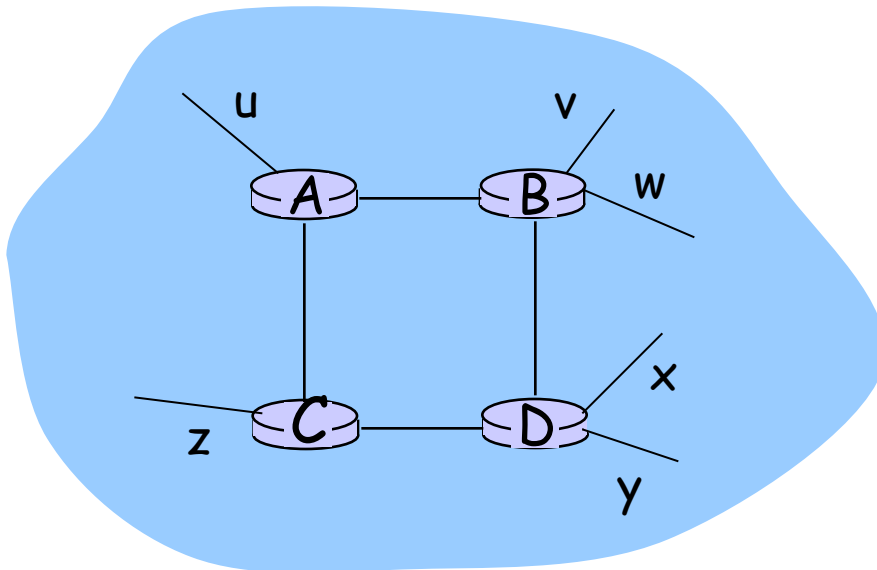
Instradamento in  
Internet

- RIP
- OSPF
- BGP

Instradamento  
broadcast e multicast

# RIP (Routing Information Protocol)

- ❑ È un protocollo a vettore distanza.
- ❑ È tipicamente incluso in UNIX BSD dal 1982.
- ❑ metrica di costo: distanza misurata in hop (max = 15 hop)
  - ❑ Ogni link ha costo unitario



Dal router A alle varie sottoreti:

destinazione	hop
u	1
v	2
w	2
x	3
y	3
z	2

# RIP protocol

## □ ***RIP Routing Information and Route Distance Metric***

- Each router in an RIP internetwork keeps track in its routing table of all networks (and possibly individual hosts) in the internetwork. For each network or host, the device includes a variety of information, of which the following are the most important:
  1. The address of the network or host.
  2. The distance from that router to the network or host (# hop).
  3. The first hop for the route: the device to which datagrams must first be sent to eventually get to the network or host.
  
- Distance is measured in *hops*. The RIP distance between a router and a network measures the number of routers that the datagram must pass through to get to the network. If a router connects to a network directly, then the distance is 1 hop. If it goes through a single router, the distance is 2 hops, and so on. In RIP, a maximum of 15 hops are allowed for any network or host. The value 16 is defined as *infinity*, so an entry with 16 in it means "this network or host is not reachable".



# RIP protocol

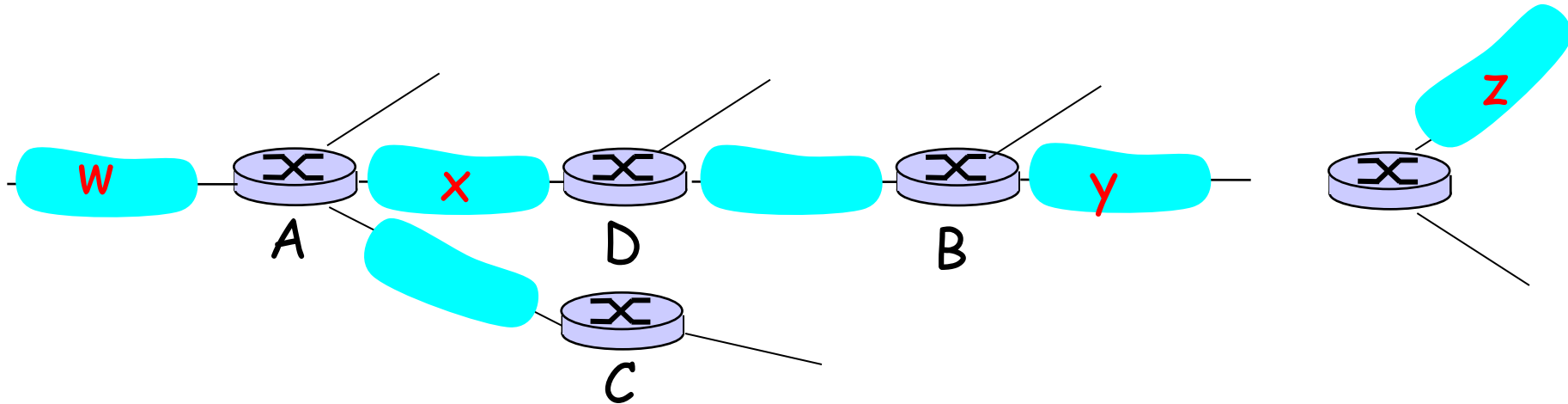
## □ *RIP Route Determination Algorithm*

On a regular basis (= periodically), each router running RIP will send out its routing table entries (distance vector) to provide information to other routers about the networks and hosts it knows how to reach. Any routers on the same network as the one sending out this information will be able to update their own tables based on the information they receive. Any router that receives a message from another router on the same network saying it can reach network  $X$  at a cost of  $N$ , knows it can reach network  $X$  at a cost of  $N+1$  by sending to the router it received the message from.

# Messaggi RIP

- ❑ **RIP Request:**
  - Quando un nuovo router viene inserito nella rete invia una RIP Request per ricevere immediatamente informazioni di routing
  - Fini diagnostici
- ❑ **RIP Response (o advertisements):**
  - In risposta a una Request
  - Periodicamente ogni 30 sec (timer)
  
- ❑ Ogni messaggio contiene un elenco comprendente fino a 25 sottoreti di destinazione all'interno del sistema autonomo nonché la distanza del mittente rispetto a ciascuna di tali sottoreti.

# RIP: esempio



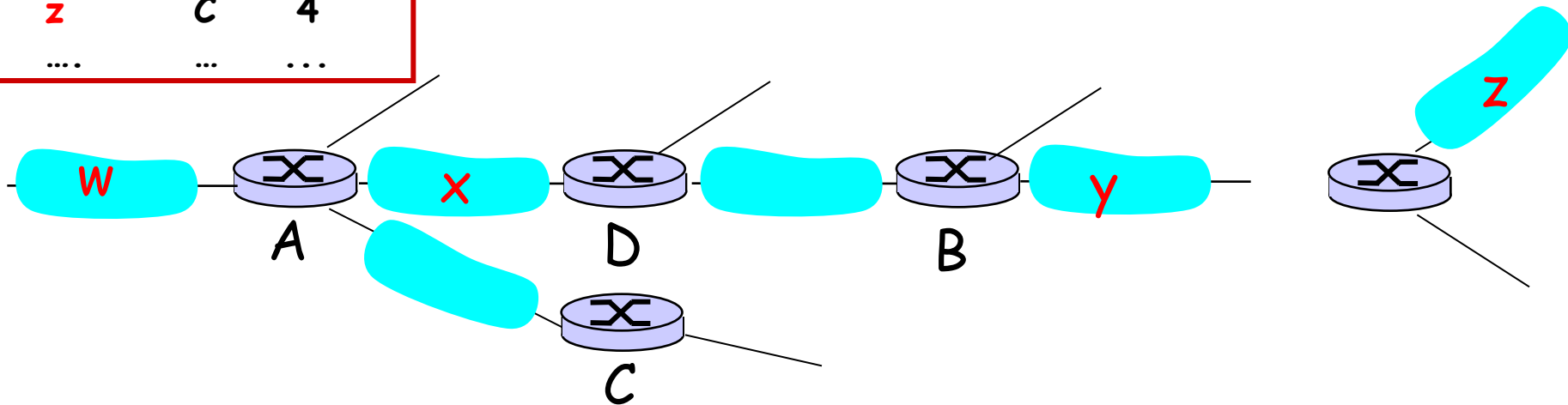
Sottorete destin.	Router successivo	Numero di hop verso la dest.
W	A	2
Y	B	2
Z	B	7
X	--	1
...	...	....

Tabella d'instradamento nel router D.

# RIP: esempio

Dest	Next	hop
w	-	1
x	-	1
z	C	4
....	...	...

Notifica dal router A a D



Sottorete destin.	Router successivo	Numero di hop verso la dest.
w	A	2
y	B	2
z	<del>B</del> A	<del>7</del> 5
x	--	1
....	....	....

Tabella d'instradamento nel router D.

# RIP: guasto sul collegamento e recupero

Se un router non riceve notizie dal suo vicino per 180 sec --> il nodo adiacente/il collegamento viene considerato spento o guasto.

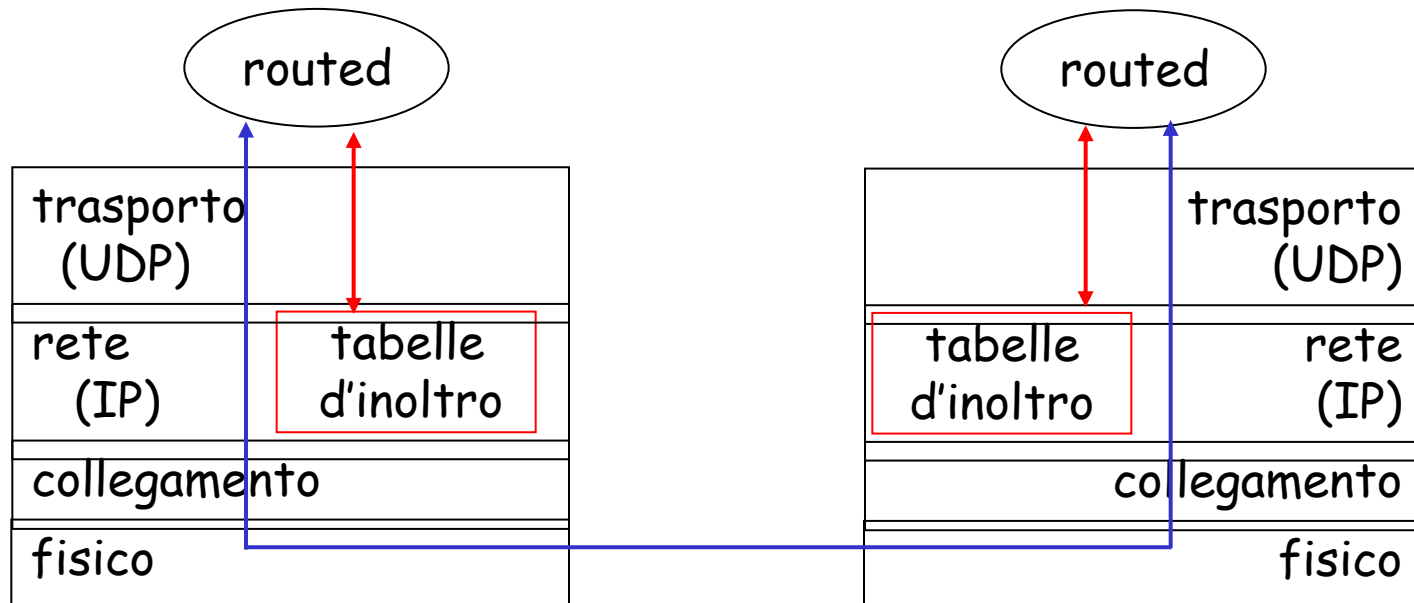
- RIP modifica la tabella d'instradamento locale
- Propaga l'informazione mandando annunci ai router vicini.
- I vicini inviano nuovi messaggi (se la loro tabella d'instradamento è cambiata).
- L'informazione che il collegamento è fallito si propaga rapidamente su tutta la rete.
- L'utilizzo dell'**inversione avvelenata** evita i loop (distanza infinita = 16 hop)

# Caratteristiche di RIP

- ❑ *Split horizon with poisoned reverse (inversione avvelenata)*
  - Serve per evitare che un router invii rotte non valide al router da cui ha imparato la rotta (evitare cicli).
  - Si mette a infinito (16) il costo della rotta che passa attraverso il vicino a cui si manda advertisement
- ❑ *Triggered updates*
  - Riduce il problema della convergenza lenta
  - Quando cambia una rotta si inviano immediatamente informazioni ai vicini senza attendere il timeout.
- ❑ *Hold-down*
  - Fornisce robustezza
  - Quando si riceve una informazione di una rotta non più valida, si avvia un timer e tutti gli advertisement riguardanti quella rotta che arrivano entro il timeout vengono tralasciati

# Tabella d'instradamento RIP

- ❑ Implementato come applicazione sopra UDP porta 520 per RIP-1 e RIP-2, porta 521 per RIPng
- ❑ Un processo chiamato *routed* esegue RIP, ossia mantiene le informazioni d'instradamento e scambia messaggi con i processi *routed* nei router vicini.
- ❑ Poiché RIP viene implementato come un processo a livello di applicazione, può inviare e ricevere messaggi su una socket standard e utilizzare un protocollo di trasporto standard.



# Livello di rete

Introduzione

Reti a circuito virtuale e  
a datagramma

Che cosa si trova  
all'interno di un  
router?

Protocollo Internet (IP)

- Formato dei datagrammi
- Indirizzamento IPv4
- ICMP
- IPv6

Algoritmi di  
instradamento

- Stato del collegamento
- Vettore distanza
- Instradamento gerarchico

Instradamento in  
Internet

- RIP
- OSPF
- BGP

Instradamento  
broadcast e multicast



# OSPF (Open Shortest Path First)

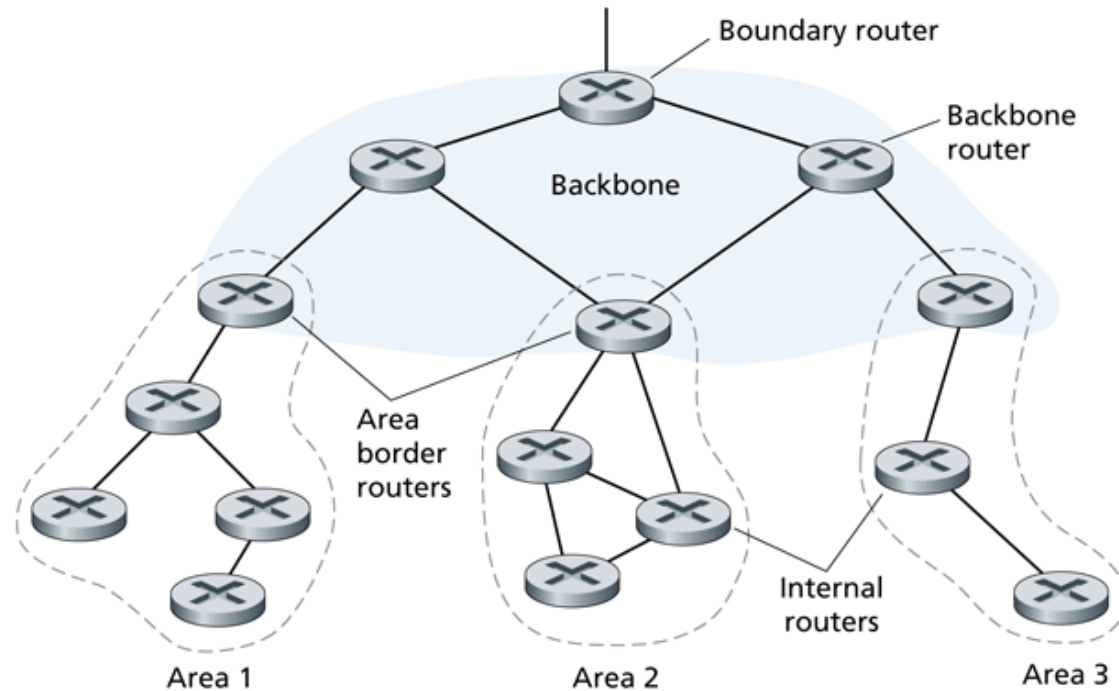
- ❑ “open”: le specifiche del protocollo sono pubblicamente disponibili.
- ❑ È un protocollo a stato del collegamento:
  1. Utilizza il **flooding** di informazioni di stato del collegamento
  2. Utilizza l'algoritmo di Dijkstra per la determinazione del percorso a costo minimo.
- ❑ Flooding
  - ❑ Con OSPF, ogni volta che si verifica un cambiamento nello stato di un collegamento, il router manda informazioni d'instradamento a *tutti* gli altri router.
  - ❑ Invia periodicamente (ogni 30 minuti) messaggi OSPF all'intero sistema autonomo, utilizzando il flooding.
- ❑ I messaggi OSPF vengono trasportati direttamente in datagrammi IP (e non da TCP o UDP), usando il numero di protocollo 89 nel campo IP protocol

# Vantaggi di OSPF (non in RIP)

- ❑ **Multipath:** quando più percorsi verso una destinazione hanno lo stesso costo, OSPF consente di usarli senza doverne scegliere uno, come invece avveniva in RIP
- ❑ **Costo:** Su ciascun collegamento, vi possono essere più metriche di costo
  - ❑ #hop
  - ❑ Capacità del collegamento (inversamente proporzionale)
- ❑ **Supporto integrato per l'instradamento unicast e multicast.**
  - Per consentire l'instradamento multicast viene impiegato MOSPF (OSPF *multicast*) che utilizza il database di collegamenti OSPF.
- ❑ **Supporto alle gerarchie** in un dominio d'instradamento
  - ❑ Possibilità di strutturare i sistemi autonomi in modo gerarchico

# OSPF strutturato gerarchicamente

Rete con 4 aree



- ❑ OSPF può essere configurato in aree che eseguono diversi algoritmi OSPF
- ❑ Ciascun router in un'area invia lo stato dei collegamenti agli altri router dell'area
- ❑ L'instradamento intra-area coinvolge solo i router dell'area
- ❑ Se la sorgente e la destinazione sono in aree differenti, allora il datagramma viene inviato dalla sorgente al suo router di confine, da questi attraverso la dorsale al router di confine dell'area destinazione, che provvede a farlo arrivare alla destinazione interna all'area.

# OSPF strutturato gerarchicamente

- **Gerarchia su due livelli:** aree locale, area dorsale.
  - Messaggio di link-state solo all'interno dell'area
  - Ciascun nodo ha una sua area; conosce solo la direzione (shortest path) verso le reti nelle altre aree
  
- **Router interno:** mantiene informazioni solo per la sua area e non ha conoscenza della topologia delle altre aree
- **Router di confine:** mantengono informazioni LS per ogni area a cui appartengono (partecipano anche al backbone).
- **Router di dorsale:** effettuano l'instradamento all'interno della dorsale, ma non sono router di confine.
  
- Only the backbone routers (which include at least one area border router within each area) need to know the details of the entire autonomous system. These backbone routers condense information about the areas so that only a "summary" of each area's topology needs to be advertised on the backbone.

# Livello di rete

Introduzione

Reti a circuito virtuale e  
a datagramma

Che cosa si trova  
all'interno di un  
router?

Protocollo Internet (IP)

- Formato dei datagrammi
- Indirizzamento IPv4
- ICMP
- IPv6

Algoritmi di  
instradamento

- Stato del collegamento
- Vettore distanza
- Instradamento gerarchico

Instradamento in  
Internet

- RIP
- OSPF
- BGP

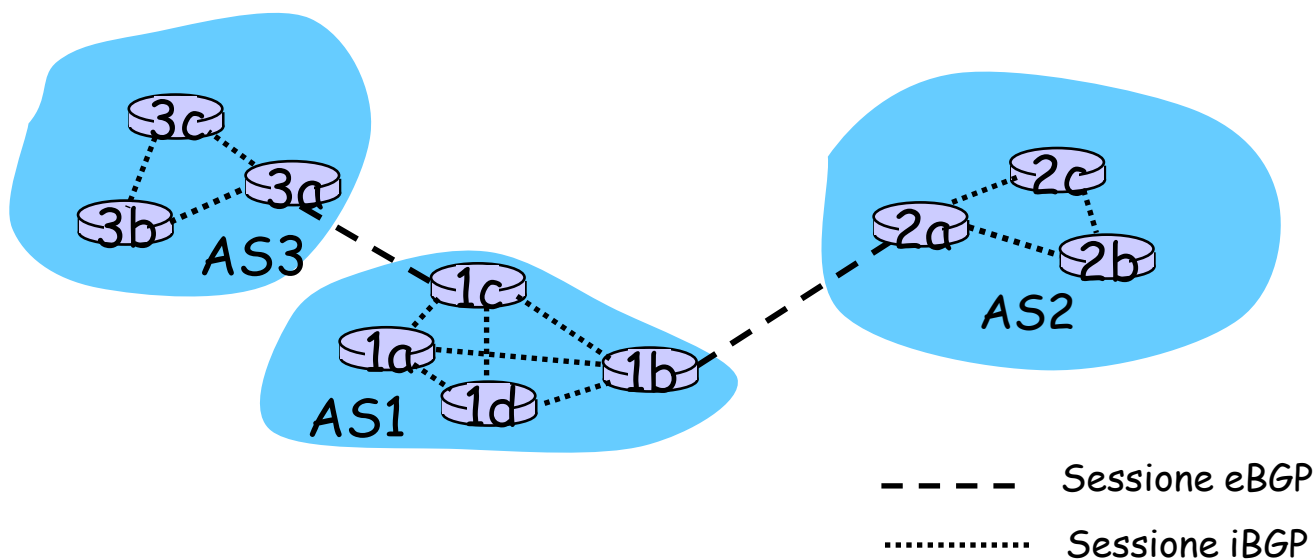
Instradamento  
broadcast e multicast

# Instradamento inter-AS in Internet: BGP

- ❑ RIP e OSPF vengono utilizzati per determinare i percorsi ottimali per le coppie origine-destinazione interne a un sistema autonomo
- ❑ **BGP (Border Gateway Protocol)**
  - ❑ Usato per determinare percorsi per le coppie origine-destinazione che interessano più sistemi autonomi
  - ❑ rappresenta l'attuale standard *de facto*.
- ❑ BGP è un protocollo distance vector che invia più informazioni sulla rotta invece che il solo costo
- ❑ BGP mette a disposizione di ciascun AS un modo per:
  1. ottenere informazioni sulla raggiungibilità delle sottoreti da parte di AS confinanti
  2. propagare le informazioni di raggiungibilità a tutti i router interni di un AS
  3. determinare percorsi "buoni" verso le sottoreti sulla base delle informazioni di raggiungibilità e delle politiche dell'AS
- ❑ BGP consente a ciascuna sottorete di comunicare la propria esistenza al resto di Internet.
- ❑ In BGP le destinazioni non sono host ma prefissi CIDR che rappresentano una sottorete o una collezione di sottoreti

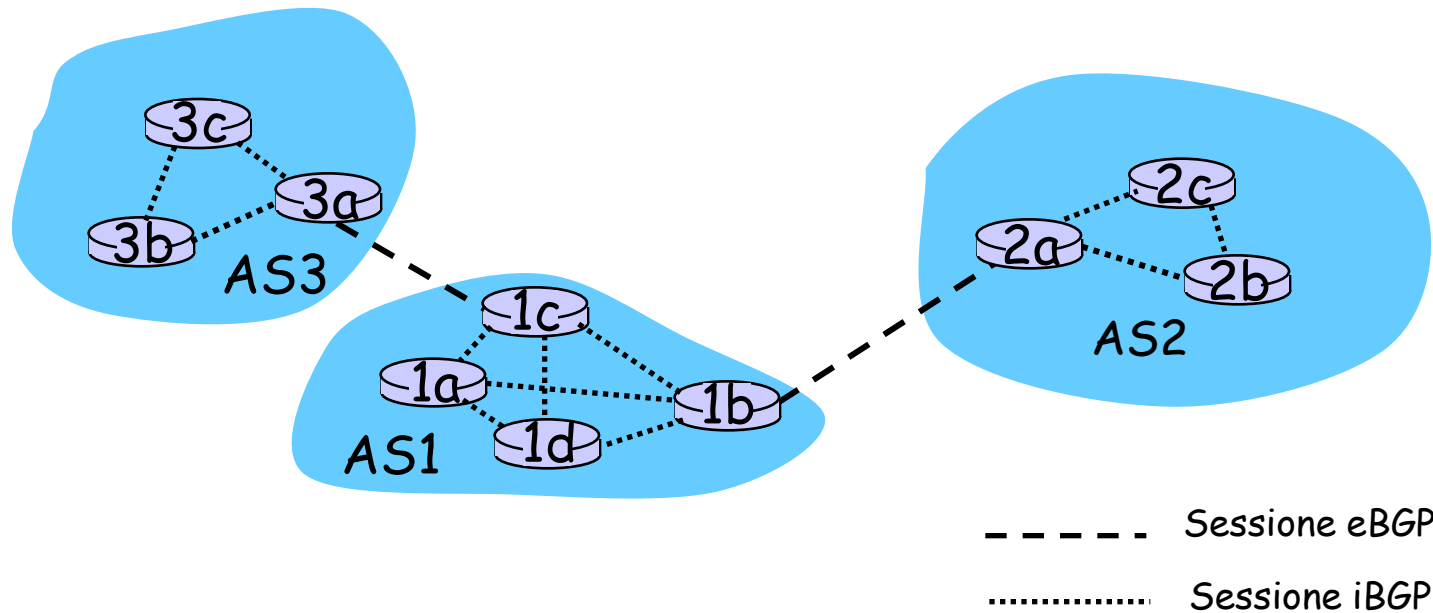
# Fondamenti di BGP

- ❑ Coppie di router si scambiano informazioni di instradamento su connessioni TCP usando la porta 179
- ❑ I router ai capi di una connessione TCP sono chiamati **peer BGP**, e la connessione TCP con tutti i messaggi BGP che vi vengono inviati è detta **sessione BGP**.
- ❑ Notiamo che le linee di sessione BGP non sempre corrispondono ai collegamenti fisici.
- ❑ Quando AS2 annuncia un prefisso a AS1, AS2 sta in realtà **promettendo** che inoltrerà i datagrammi su un percorso verso il prefisso cui sono destinati.
  - AS2 può aggregare più prefissi nel suo annuncio



# Distribuzione delle informazioni di raggiungibilità

- In una sessione eBGP tra i gateway 3a e 1c, AS3 invia ad AS1 la lista di prefissi raggiungibili.
  - 1c utilizza le proprie sessioni iBGP per distribuire i prefissi agli altri router del sistema autonomo.
  - Anche AS1 e AS2 si scambiano informazioni sulla raggiungibilità dei prefissi attraverso i propri gateway 1b e 2a.
- Quando un router viene a conoscenza di un nuovo prefisso, lo memorizza in una nuova riga della propria tabella d'inoltro.

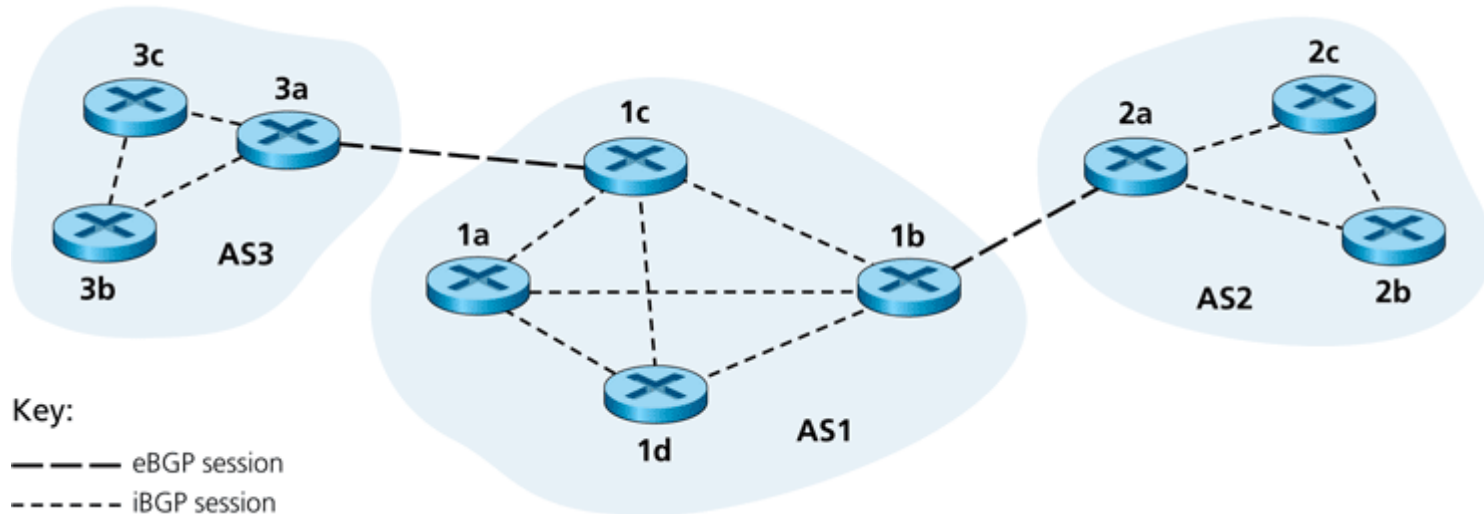




# Attributi del percorso e rotte BGP

- ❑ Quando un router annuncia una rotta per un prefisso (di rete) per una sessione BGP, include anche un certo numero di **attributi BGP**.
  - prefisso + attributi = "rotta"
- ❑ Due dei più importanti attributi sono:
  - **AS-PATH**: elenca i sistemi autonomi attraverso i quali è passato l'annuncio del prefisso (e quindi gli hop intermedi della rotta). Ogni sistema autonomo non stub ha un identificativo univoco (in questo modo si evitano cicli)
  - **NEXT-HOP**: indirizzo IP dell'interfaccia su cui viene inviato il prefisso. (Un router ha più indirizzi IP, uno per ogni interfaccia)
- ❑ Quando un router gateway riceve un annuncio di rotta, utilizza le proprie **politiche d'importazione** per decidere se accettare o filtrare la rotta.
  - ❑ Il sistema autonomo può non voler inviare traffico su uno degli AS presenti nel AS-PATH
  - ❑ Router conosce rotta migliore

# Esempio su utilizzo NEXT-HOP



- ❑ Il router 3a annuncia una rotta a 1c con eBGP
  - (x, AS-PATH, interfaccia 3a-1c)
- ❑ 1d apprende questa rotta con iBGP
- ❑ 1d per includere la rotta nella sua tabella di routing dovrà trovare l'interfaccia l, che inizia il suo cammino minimo verso 1c
- ❑ Per trovare l, 1d calcola il suo cammino minimo verso l'interfaccia indicata nel NEXT-HOP del messaggio (3a-1c)
- ❑ 1d inserisce (x,l) nella sua tabella di routing

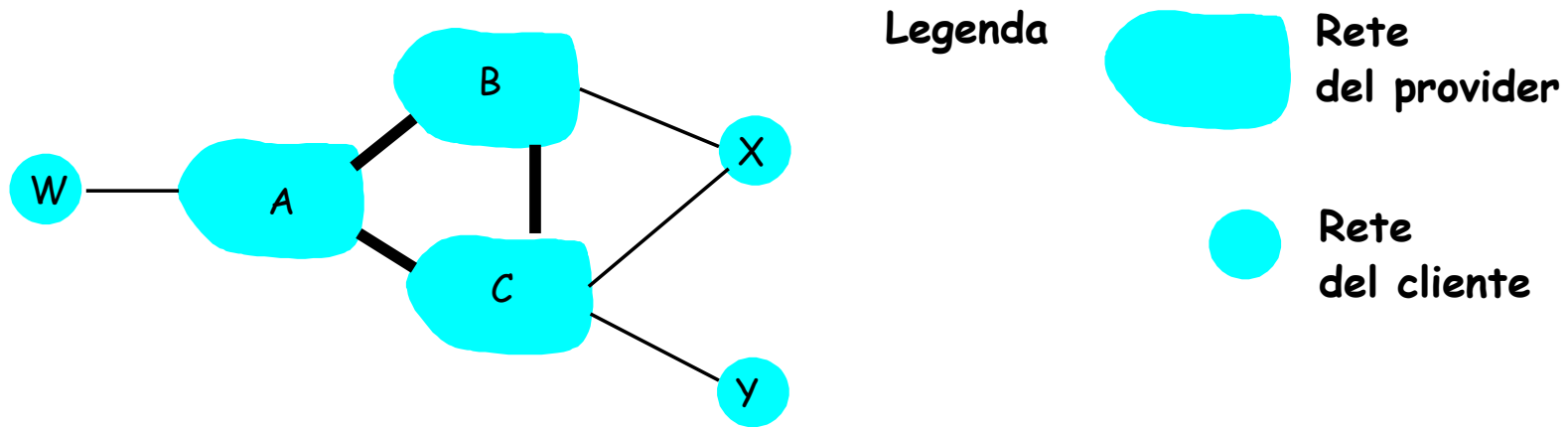
# Selezione dei percorsi BGP

- ❑ Un router può ricavare più di una rotta verso un determinato prefisso, e deve quindi sceglierne una.
- ❑ Regole di eliminazione:
  1. Alle rotte viene assegnato come attributo un valore di preferenza locale. Si selezionano quindi le rotte con i più alti valori di preferenza locale.
  2. Si seleziona la rotta con valore AS-PATH più breve.
  3. Si seleziona quella il cui router di NEXT-HOP è più vicino: instradamento *a patata bollente*.
  4. Se rimane ancora più di una rotta, il router si basa sugli identificatori BGP.

# Messaggi BGP

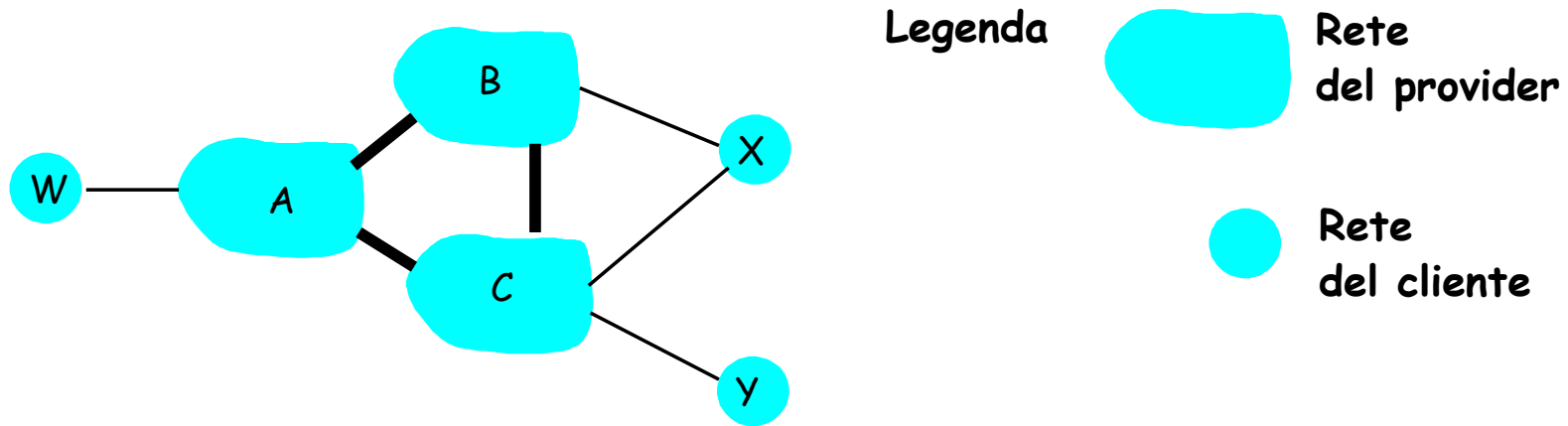
- ❑ I messaggi BGP vengono scambiati attraverso TCP.
- ❑ Messaggi BGP:
  - **OPEN**: apre la connessione TCP e autentica il mittente
  - **UPDATE**: annuncia il nuovo percorso (o cancella quello vecchio)
  - **KEEPALIVE** mantiene la connessione attiva in mancanza di UPDATE
  - **NOTIFICATION**: riporta gli errori del precedente messaggio; usato anche per chiudere il collegamento.

# Politiche d'instradamento BGP



- ❑ A, B, C sono reti di provider di dorsale.
- ❑ X, W, Y sono reti stub
- ❑ X è una rete stub a più domicili
  - X non vuole che il traffico da B a C le passi attraverso
  - ... e così X non annuncerà a B la rotta verso C

# Politiche d'instradamento BGP (2)



- ❑ A annuncia a B del percorso AW.
- ❑ B annuncia a X del percorso BAW.
- ❑ B deve annunciare a C del percorso BAW?
  - Certo che no! B non ha nessun "interesse" nella rotta CBAW poiché né W né C sono clienti di B
  - B vuole costringere C ad instradare verso W attraverso A
  - B vuole instradare *solo* da/verso i suoi clienti!

# Perché i protocolli d'instradamento inter-AS sono diversi da quelli intra-AS?

## Politiche:

- ❑ Inter-AS: il controllo amministrativo desidera avere il controllo su come il traffico viene instradato e su chi instrada attraverso le sue reti.
- ❑ Intra-AS: unico controllo amministrativo, e di conseguenza le questioni di politica hanno un ruolo molto meno importante nello scegliere le rotte interne al sistema

## Scala:

- ❑ L'instradamento gerarchico fa "risparmiare" sulle tabelle d'instradamento, e riduce il traffico dovuto al loro aggiornamento

## Prestazioni:

- ❑ Intra-AS: orientato alle prestazioni
- ❑ Inter-AS: le politiche possono prevalere sulle prestazioni