Università di Roma Tor Vergata
Corso di Laurea triennale in Informatica

Sistemi operativi e reti

A.A. 2016-17

Pietro Frasca

Parte II: Reti di calcolatori

Lezione 19 (43)

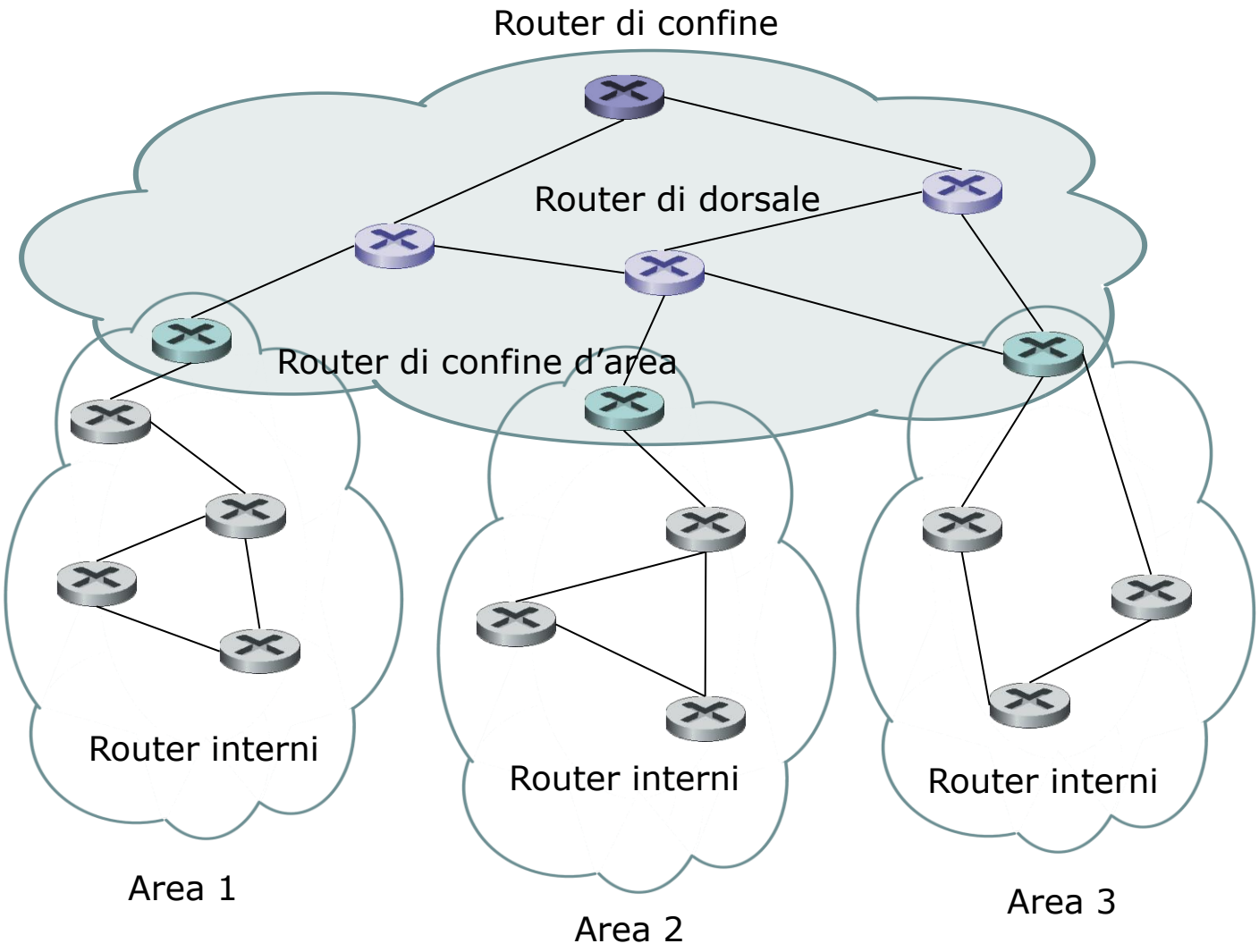
Martedì 16-05-2017

OSPF: Open Shortest Path First

- L'OSPF è un protocollo basato sullo **stato dei link** molto usato nell'instradamento intra-AS.
- Come il RIP, OSPF è un protocollo di pubblico dominio (a differenza del protocollo EIGRP della Cisco).
- L'OSPF, in un router, costruisce un **grafo orientato pesato,** dell'intero sistema autonomo. Tramite l'algoritmo di Dijkstra determina **l'albero dei percorsi a costo minimo** verso tutte le reti di destinazione.
- La tabella di instradamento del router è ottenuta da questo albero dei percorsi minimi.
- Con OSPF i pesi associati ai collegamenti possono essere assegnati con vari criteri. Ad esempio, i pesi possono essere assegnati con valori inversamente proporzionali alle larghezze di banda dei link in modo che i pacchetti siano rinviati principalmente verso linee con maggiore velocità di trasmissione.

- I pesi dei link possono essere anche scelti in base ad altri parametri o posti ad 1, ottenendo così un instradamento a numero minimo di hop (salti).
- Un router OSPF trasmette periodicamente informazioni di instradamento a ***tutti*** **gli altri router del sistema autonomo** ogni volta che si verifica un cambiamento nello stato dei link come ad esempio, un cambiamento del peso o un cambiamento dello stato attivo/inattivo.
- Per rendere il funzionamento più affidabile OSPF, a intervalli di **30 minuti**, invia informazioni anche se lo stato dei link non è cambiato.
- OSPF implementa funzionalità tipiche dello strato di trasporto come il trasferimento affidabile di dati oltre al broadcast dello stato dei link e pertanto, i **messaggi OSPF** sono trasportati direttamente da IP, e sono identificati dal valore **89** del campo IP "**protocollo di strato superiore"**.
- un router OSPF, per verificare che i link siano attivi, invia a ogni vicino il **messaggio HELLO.** Come risposta a tale messaggio il router riceve da un vicino le informazioni dello stato dei link di tutta la rete.

- I messaggi scambiati tra router OSPF possono essere autenticati. Sono supportate due modalità di autenticazione, **basic** e **MD5.** La prima invia i messaggi in chiaro, ed è quindi poco sicura, mentre con l'autenticazione MD5 i messaggi sono cifrati per evitare che hacker possano inviare ad un router false tabelle di instradamento e provocare blocchi critici della rete.
- OSPF è stato progettato per l'instradamento in AS di dimensioni sia piccole che grandi. Tuttavia, per sistemi di grandi dimensioni il traffico di routing, prodotto dai router in base alla tecnica del broadcast delle informazioni inviate, può diventare eccessivo. Per evitare ciò con OSPF un sistema autonomo può essere strutturato gerarchicamente suddividendolo in "**aree**". I router appartenenti ad una stessa area si scambiano tra loro informazioni sullo stato dei link e non inviano informazioni a router appartenenti ad aree diverse.



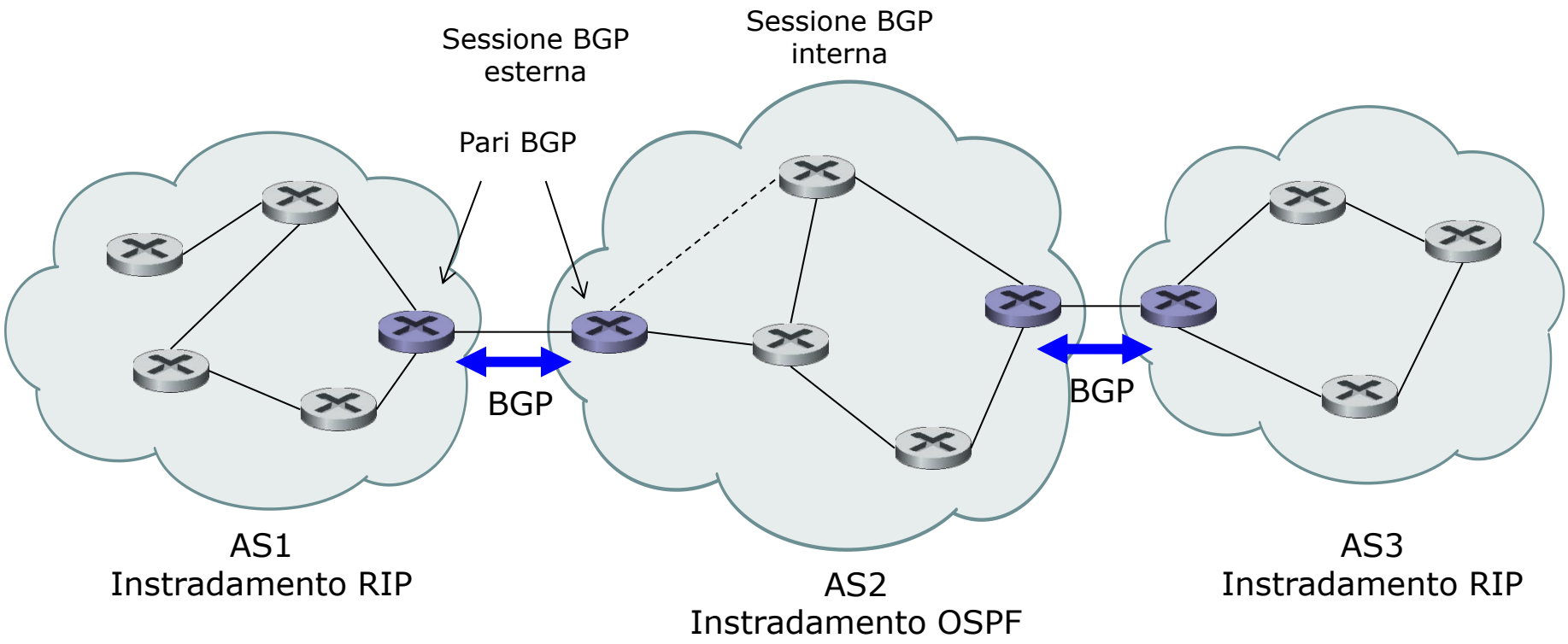
Sistema autonomo OSPF strutturato gerarchicamente con quattro aree

- Per ciascuna area è presente uno o più **router di confine di area *(area border router)*** che si occupano di instradare i pacchetti verso altre aree.
- Una particolare area è configurata per essere l'area **dorsale** (**backbone)** che serve per instradare i pacchetti tra tutte le altre aree dell'AS.
- Nella area dorsale sono presenti tutti i **router di confine** di area delle diverse aree dell'AS e può contenere anche dei router non di confine di area.
- L'instradamento inter-area richiede che il pacchetto sia instradato prima verso un router di confine di area, poi che venga instradato attraverso la dorsale al router di confine di area che si trova nell'area di destinazione, e da qui verso la sua destinazione finale.

- In una rete OSPF strutturata gerarchicamente possiamo identificare quattro tipi di router OSPF:
  - ***Router interni****.* Questi router sono interni ad una determinata area ed eseguono solo l'instradamento intra-AS.
  - ***Router di confine di area****.* Sono router che appartengono sia alla dorsale sia a un'area.
  - ***Router della dorsale (non di confine)****.* Questi router non sono router di confine di area ma eseguono l'instradamento dentro la dorsale. All'interno di un'area diversa dalla dorsale, i router interni apprendono dell'esistenza delle rotte verso altre aree dalle informazioni (essenzialmente dai messaggi sullo stato dei link, ma l'informazione riguarda il costo del percorso verso un'altra area, piuttosto del costo del link) trasmesse entro l'area dai suoi router backbone.
  - ***Router di confine*** *(****boundary router****). Questi* router scambiano informazioni di instradamento con router di altri sistemi autonomi. Usano **BGP** per svolgere l'instradamento inter-AS.

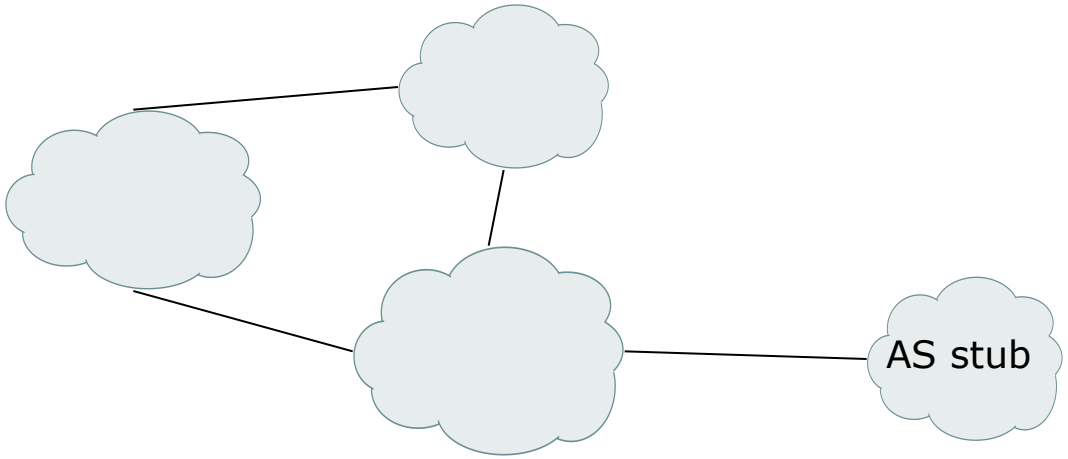
Instradamento tra sistemi autonomi: BGP

- Il protocollo **BGP** *(**border gateway protocol**)*, è il protocollo standard "de facto" per l'instradamento inter-AS di Internet. L'attuale versione è il BGP4 (RFC 1771, 1995).
- E' un protocollo molto complesso sia a livello di progettazione che di amministrazione.
- Il BGP è un protocollo di tipo ***path vector*** (vettore dei percorsi).
- I router BGP di confine, detti **pari BGP**, si scambiano informazioni dettagliate sui percorsi, specificando la **sequenza di AS da attraversare**, che consentono di raggiungere una determinata rete di destinazione.
- I pari BGP utilizzano il TCP e la porta 179 per scambiarsi messaggi (sessione BGP). Ricordiamo invece che, i router RIP usano UDP e OSPF usa un suo proprio protocollo per inviare i messaggi OSPF.

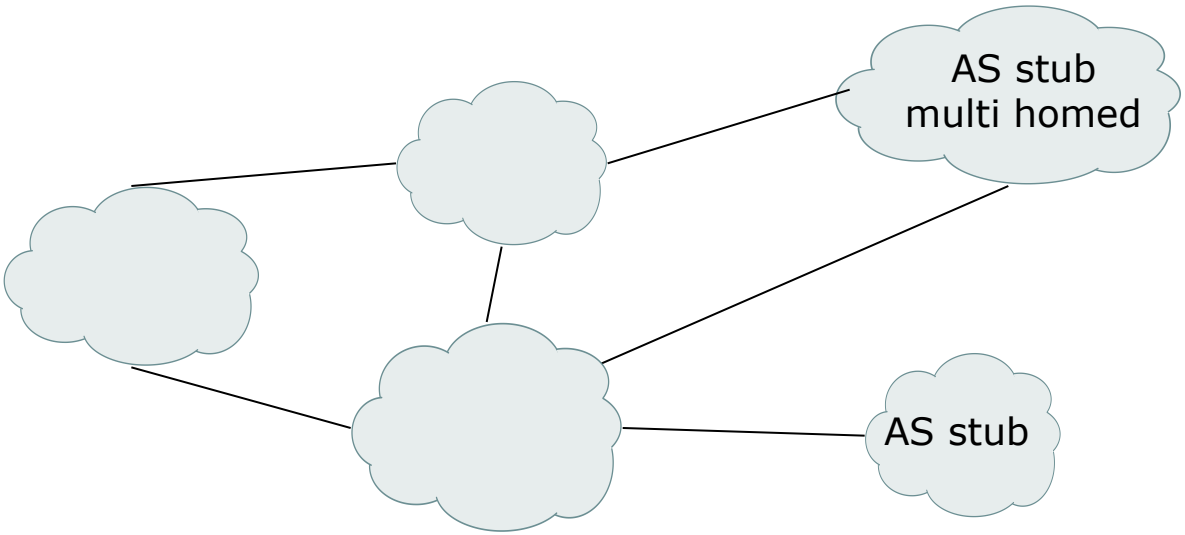


Uso di BGP per l'instradamento inter-AS

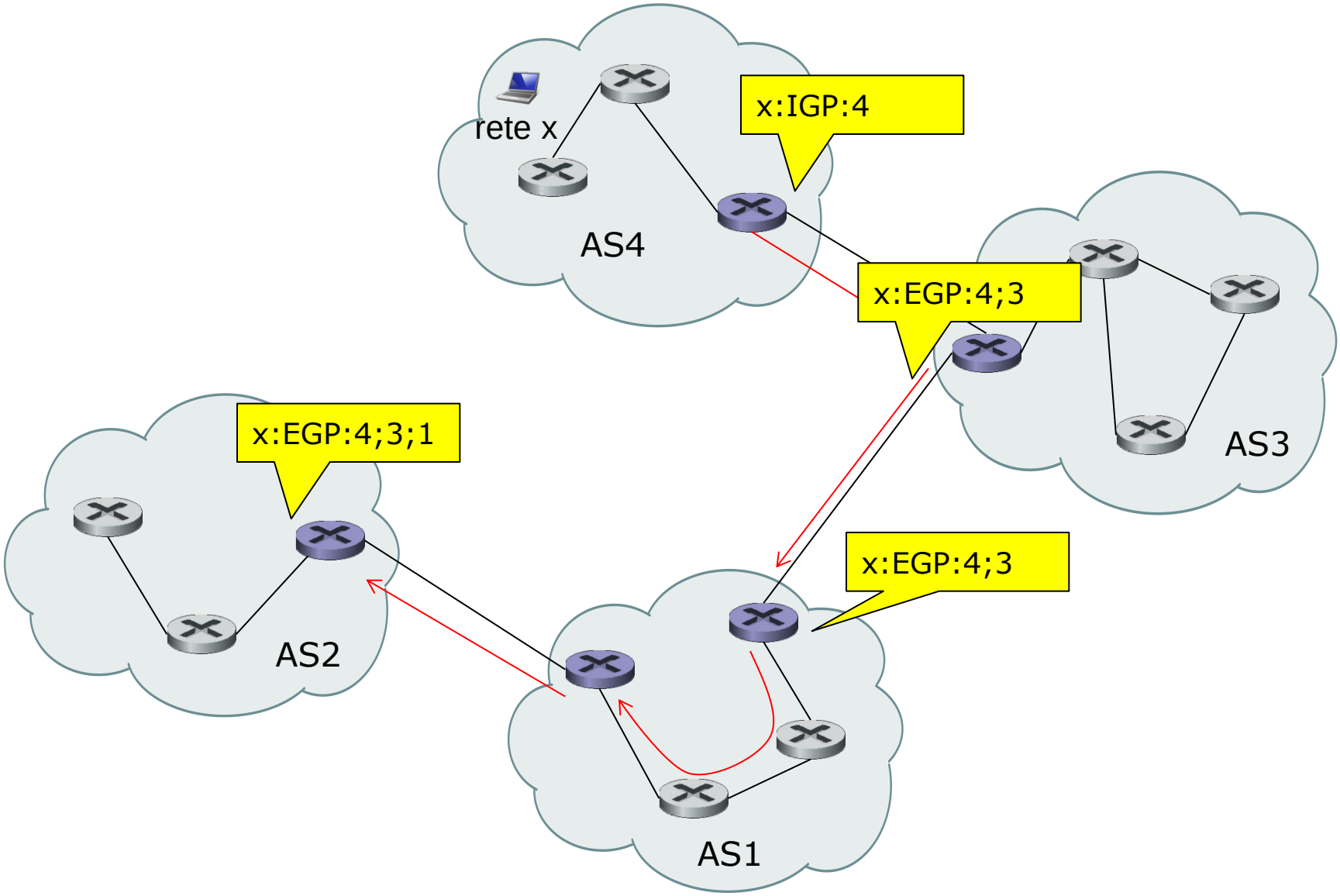
- **In BGP, un sistema autonomo è identificato da un numero di sistema autonomo** (**ASN**, ***Autonomous System Number)*** **globalmente unico (16 bit).**
- I numeri di AS, come gli indirizzi IP, sono assegnati dai registri regionali ICANN (**I**nternet **C**orporation for **A**ssigned **N**ames and **N**umbers).
- Non tutti gli AS hanno un ASN. In particolare, gli AS privi di ASN sono chiamati ***AS stub*** (troncone).
- Una rete è detta **rete stub** quando tutto il traffico in ingresso è indirizzato solo agli host di quella rete, e tutto il traffico in uscita ha origine dagli host di quella rete.



- Una rete stub è detta **rete stub multi homed** (**ad appartenenza multipla**), se è connessa al resto della rete Internet tramite due o più diversi ISP.

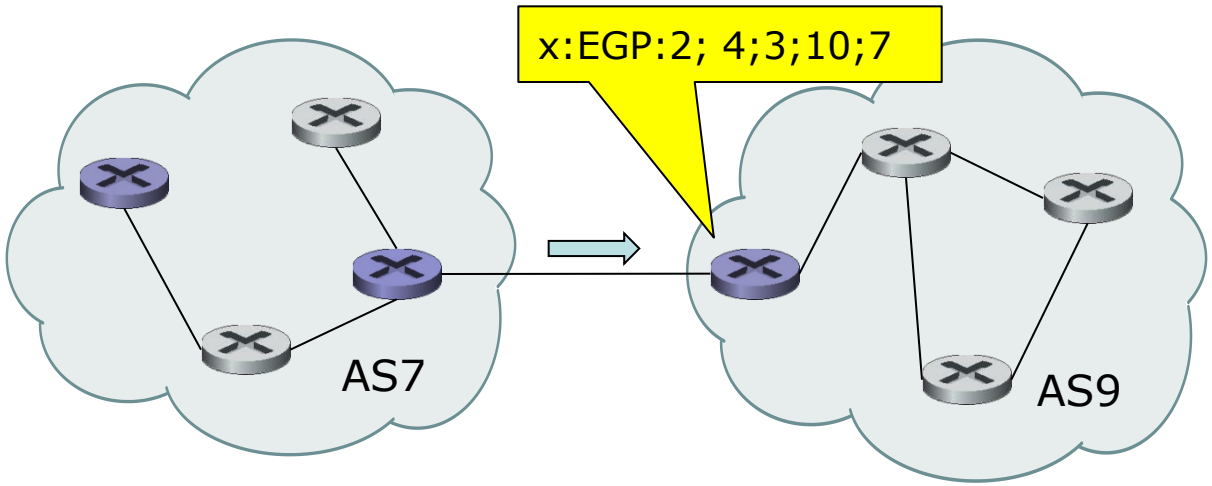


- I pari BGP si scambiano **annunci sui percorsi** su una connessione TCP.
- Un annuncio consiste in:
  - **un indirizzo di rete di destinazione** in forma CIDR (per esempio, 128.119.40/24) e
  - **un insieme di attributi** associati al percorso verso quella rete di destinazione.
- Alcuni attributi sono sempre presenti, altri sono opzionali. Quest'ultimi possono anche non essere interpretati allo stesso modo da tutti i router e possono essere o meno propagati.
- Due degli attributi più importanti sono
  - **AS-PATH** (**percorso)** una lista di tutti gli AS (identificati con gli ASN) sul percorso verso la specifica rete di destinazione e
  - **NEXT-HOP.** L'identità del prossimo router lungo il percorso verso la rete di destinazione.



Aggiornamento dell'AS path

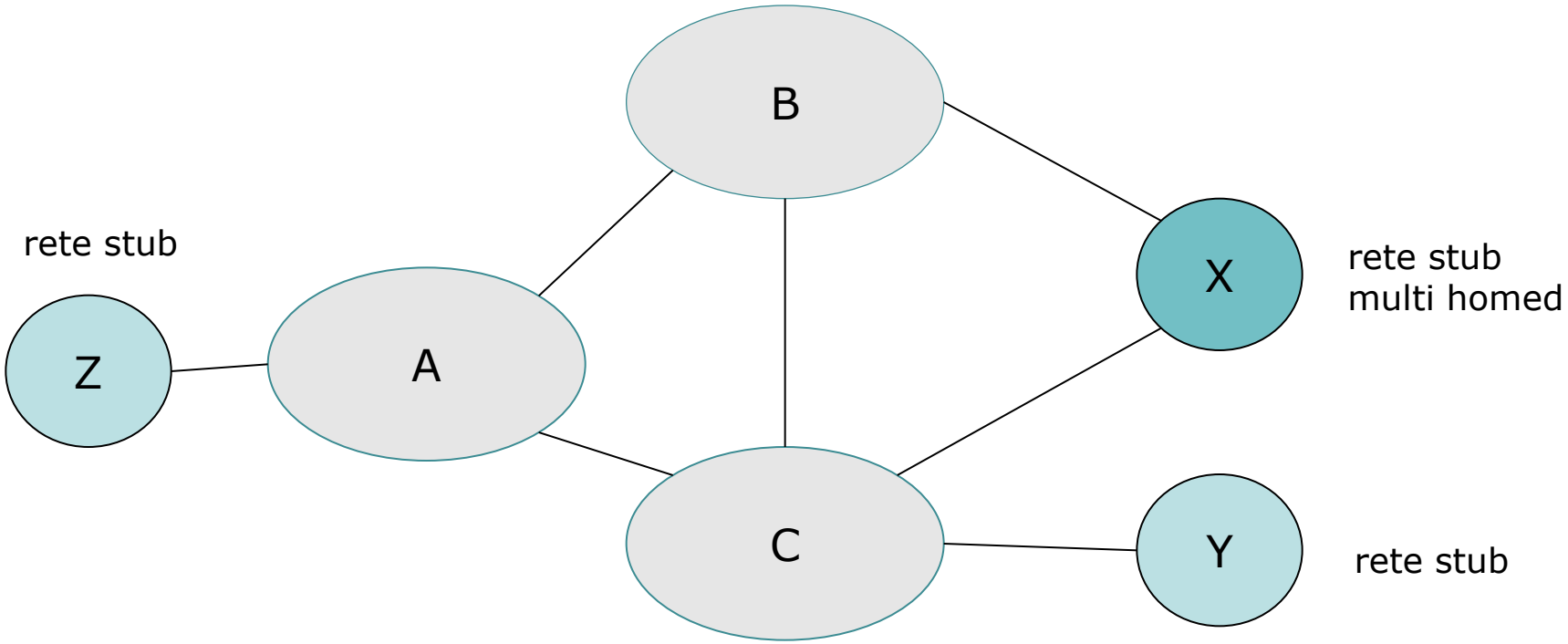
- Il funzionamento di BGP è basato principalmente su tre tipi di operazione:
  - ***Ricezione e filtraggio di annunci sui percorsi***. *Quando* un router BGP riceve gli annunci sui percorsi da un suo pari, può anche filtrare (scartare) gli annunci sui percorsi ricevuti. Poiché gli annunci specificano l'intero percorso per raggiungere la rete annunciata, come lista di AS attraversati, un amministratore può decidere quale sarà l'instradamento seguito dai pacchetti. Ad esempio nella figura seguente, se il gestore di AS9 non volesse rinviare traffico proveniente dall'AS4 configurerebbe BGP in modo tale che siano filtrati i datagram provenienti dalla rete x.



- ***Selezione del percorso.*** Un router BGP può ricevere diversi annunci sui percorsi verso la stessa rete di destinazione, e deve scegliere quale percorso usare tra quelli annunciati. La rete di destinazione e il prossimo router per il percorso scelto devono quindi essere inseriti nelle tabelle di instradamento del router. Un router BGP può conoscere diversi percorsi verso una certa rete di destinazione, ma inserirà un solo router di next-hop per quella destinazione nella tabella di instradamento. BGP consente di scegliere un percorso tra quelli annunciati in **modalità manuale o automatica**. Nel primo caso potrebbe essere una decisione ***politica o commerciale*** che viene presa dall'amministratore di rete dell'AS. Un amministratore di rete può specificare delle ***preferenze locali***, per esempio, indicando che l'instradamento attraverso uno specifico AS confinante è da preferire rispetto all'instradamento attraverso altri AS confinanti. In assenza di preferenze locali, BGP seleziona automaticamente il percorso che attraversa il minor numero di AS per arrivare a una determinata destinazione.

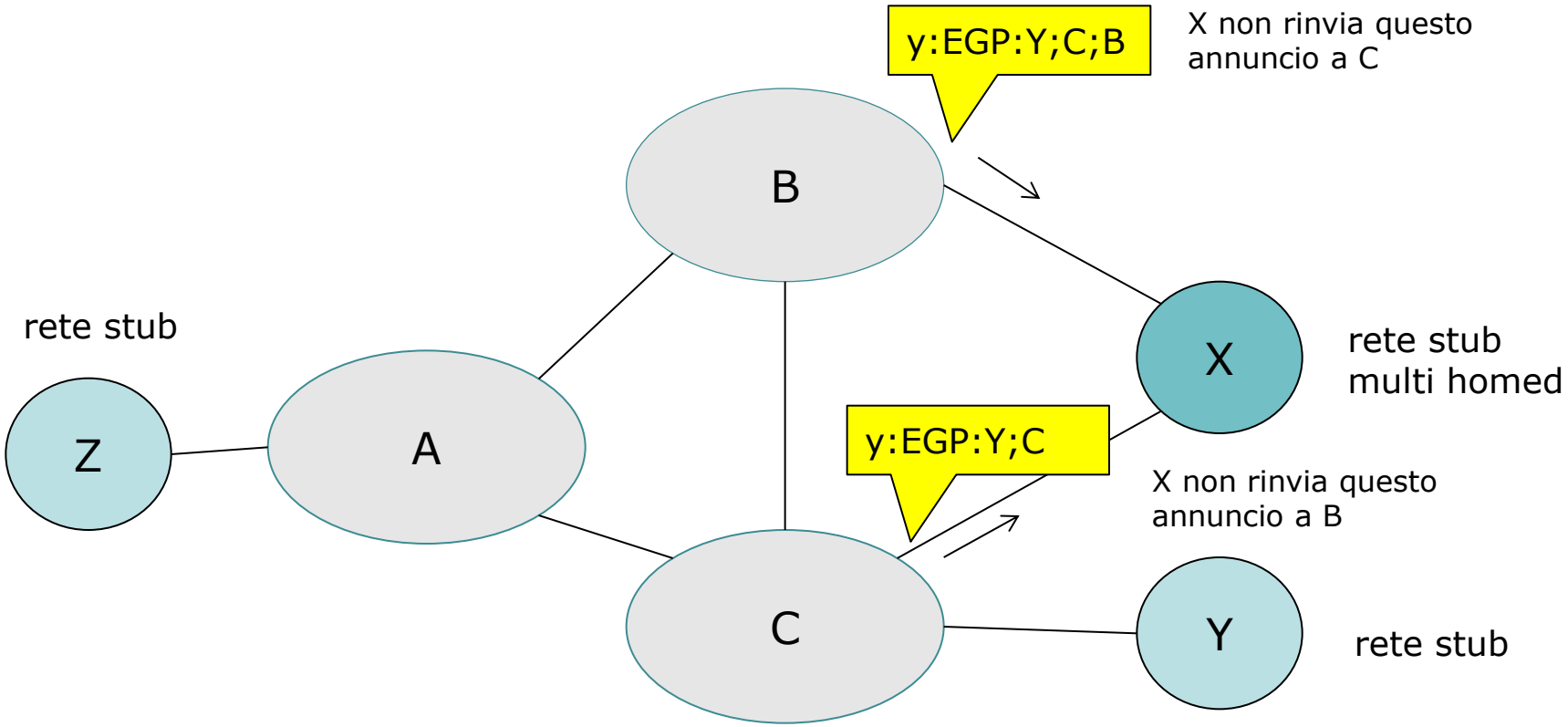
- ***Invio di annunci sui percorsi ai vicini***. Il BGP permette all'amministratore di rete un notevole grado di controllo del traffico che sarà instradato attraverso la sua rete, inviando ai router BGP vicini alcune informazioni che vuole comunicare e nascondendole altre, in base a qualche politica.

- Illustriamo con degli esempi alcuni concetti base degli annunci sui percorsi BGP. La figura mostra sei sistemi autonomi interconnessi: A, B, C, X, Y e Z.



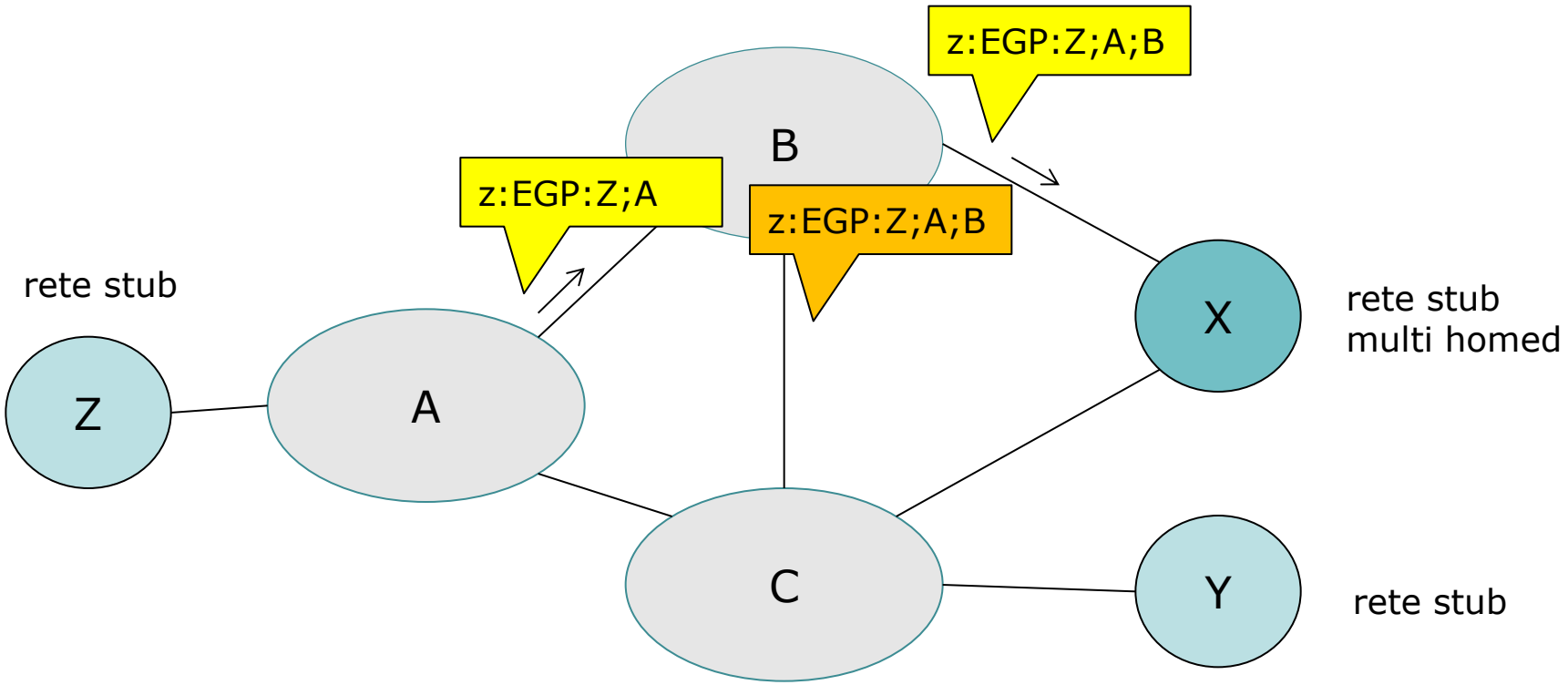
A, B e C: reti di ISP

- Come primo esempio vediamo come **l'annuncio selettivo** di percorsi può essere usato per implementare relazioni di instradamento cliente/fornitore.
- Assumiamo che i sistemi autonomi (clienti) X, Y e Z siano **reti stub** e che A, B e C siano reti di ISP (fornitori).
- E' evidente che Y e Z sono reti stub.
- L'AS X è connesso con gli AS B e C ma dato che è una rete stub deve essere la sorgente/destinazione di tutto il traffico uscente/entrante in X e pertanto non deve rinviare traffico tra B e C.
- Questo può essere ottenuto controllando il modo in cui i percorsi BGP vengono annunciati. Nell'esempio, X funzionerà come rete stub se annuncia ai suoi vicini B e C che non ha percorsi verso altre reti. Cioè, anche se X può conoscere un percorso, ad esempio XCY, che raggiunge la rete Y, esso ***non*** annuncerà questo percorso a B.
- Dato che B ignora che X ha un percorso verso Y, B non rilancerà il traffico destinato a Y o a C attraverso X.



A, B e C: reti di ISP

- Consideriamo ora la rete di un ISP, ad esempio l'AS B. Supponiamo che B abbia ricevuto da A un annuncio del percorso AZ verso Z. B può quindi registrare il percorso BAZ nella sua tabella dei percorsi.
- Chiaramente, B vuole annunciare il percorso BAZ al suo cliente, X, in modo che X sappia che può instradare i suoi pacchetti verso B per giungere a Z.
- Ma non è detto che B debba annunciare il percorso BAZ a C. Se lo annuncia, allora C potrebbe instradare il traffico verso Z attraverso CBAZ.
- Se A, B e C sono tutti ISP della dorsale, allora B potrebbe non essere d'accordo a trasportare il traffico di transito tra A e C. B potrebbe giustamente ritenere che è compito (e costo) di A e C assicurarsi che C possa instradare ai/dai clienti di A attraverso una connessione diretta tra A e C.
- Attualmente non ci sono standard ufficiali che regolano come gli ISP della dorsale instradino tra loro.
- Accordi individuali da pari a pari (che regolamentano questioni come quella sollevata prima) vengono tipicamente sottoscritti tra coppie di ISP.

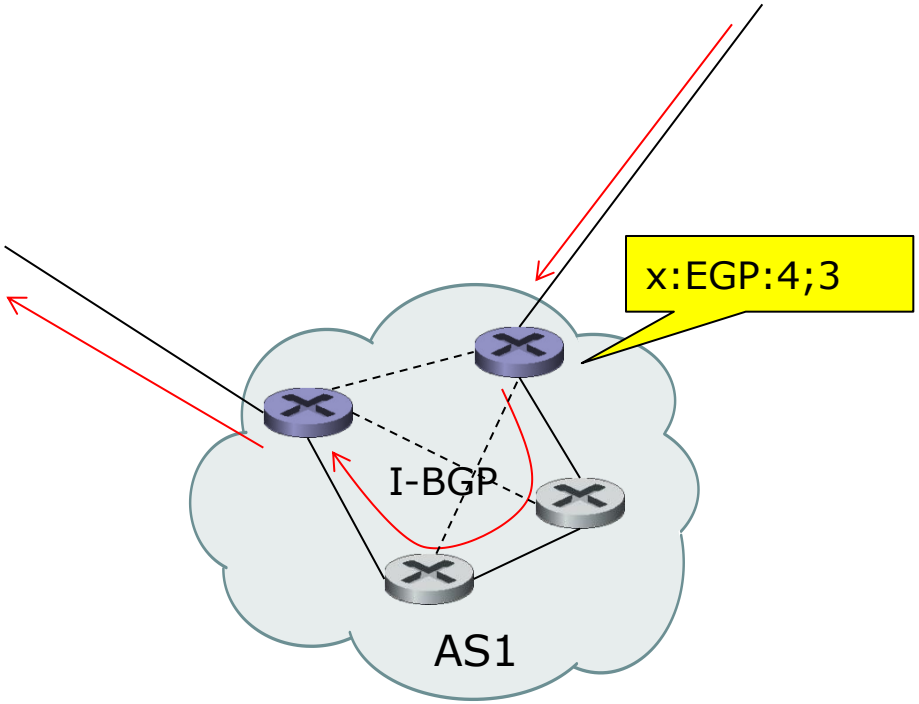


A, B e C: reti di ISP

- Dopo aver descritto alcune funzioni di controllo svolte da BGP, torniamo all'aspetto tecnico del funzionamento di BGP.
- Il protocollo BGP definisce quattro tipi di messaggi: OPEN, KEEPALIVE, UPDATE, e NOTIFICATION.
  - ***OPEN***. Quando un router BGP vuole connettersi per la prima volta con un pari BGP, viene inviato un messaggio OPEN al pari. Questo messaggio permette al router BGP di identificarsi e autenticarsi. Se il messaggio OPEN è accettato dal pari, esso risponderà con un messaggio KEEPALIVE.
  - ***KEEPALIVE***. Questo messaggio oltre che come riscontro di un messaggio OPEN ricevuto, è usato per far conoscere a un pari che il mittente è attivo ma che non ha altre informazioni da spedire.
  - ***UPDATE***. Un router BGP usa il messaggio UPDATE per annunciare un percorso verso una data destinazione al pari BGP. Il messaggio UPDATE può anche essere usato per eliminare un percorso che era stato precedentemente annunciato, cioè, per informare un pari che un percorso che aveva precedentemente annunciato non è più valido.

- ***NOTIFICATION***. Questo messaggio è usato per informare un pari che è stato rilevato un errore , per esempio, in un messaggio BGP trasmesso in precedenza o che il mittente sta per chiudere la sessione BGP.

- Quanto finora descritto è riferito esclusivamente alla versione di BGP nota come **E-BGP** (***BGP esterno***, *External BGP)* di BGP, funzionante tra router in differenti *AS.*
- C'è un'altra versione di BGP, detta **I-BGP** (**BGP interno***),* che è usata per distribuire informazioni di instradamento ai router all'interno di un AS riguardo ad AS di destinazione al suo esterno.



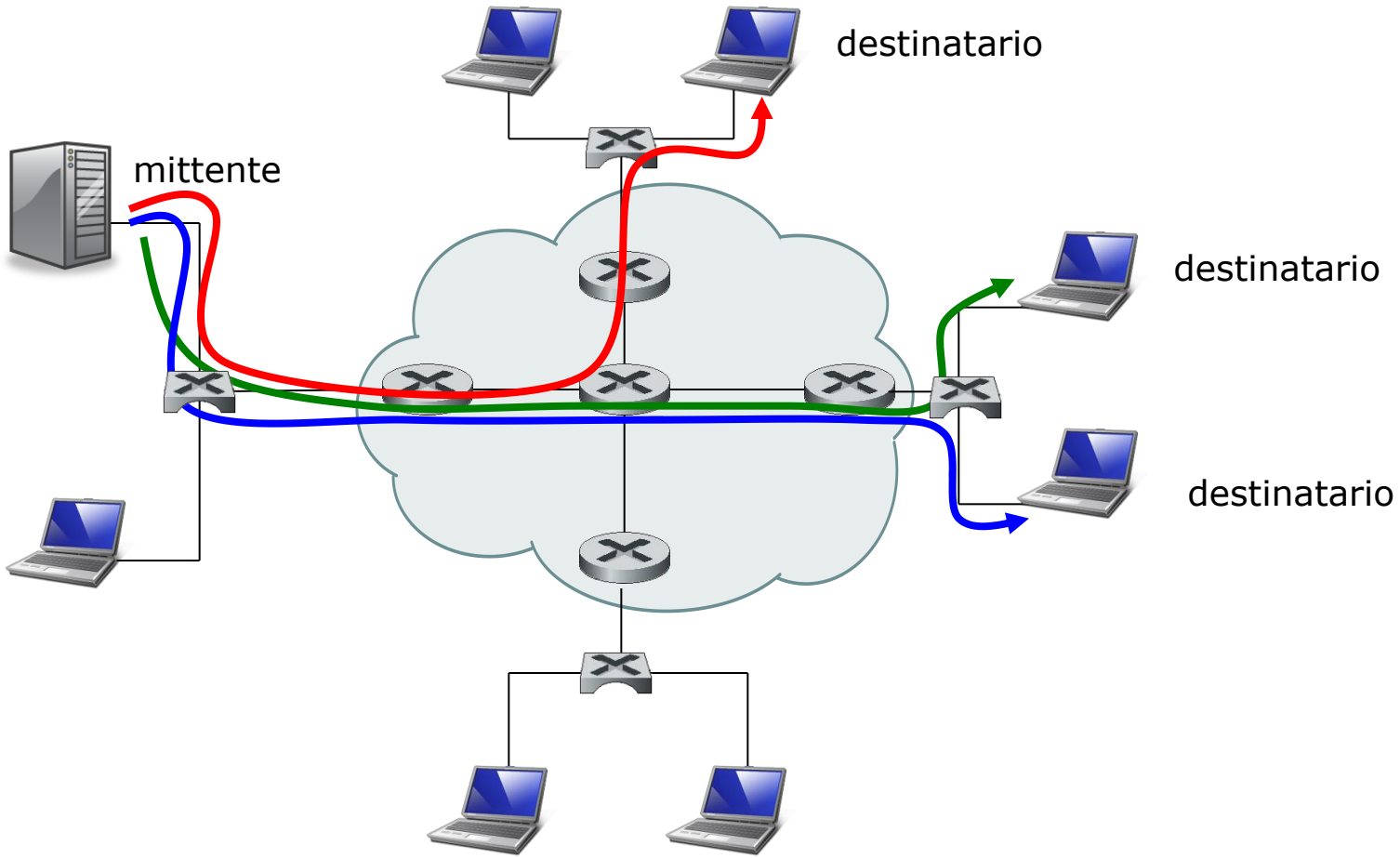
- Abbiamo visto in precedenza che è possibile specificare un percorso di default (next-hop) mediante un protocollo di instradamento intra-AS come RIP, in modo che i datagram che non sono destinati a una rete esplicitamente elencata nella tabella di instradamento saranno rinviati lungo il percorso di default.
- Un'altra possibilità è di usare il protocollo **I-BGP** per distribuire informazioni di instradamento all'interno di un AS riguardo a AS remoti.
- I router I-BGP all'interno di un AS sono tutti logicamente connessi tra loro. In altri termini, i router I-BGP in un AS sono considerati tutti vicini l'uno l'altro.

Instradamento multicast

- I protocolli degli strati di trasporto e di rete che abbiamo descritto finora, consentono la trasmissione di pacchetti da una singola sorgente a una singola destinazione e sono quindi detti protocolli **unicast**.
- Un numero sempre maggiore di applicazioni di rete esegue la trasmissione di pacchetti da un mittente a un **gru*ppo* di destinatari***.*
- Alcuni esempi di queste applicazioni sono: la trasmissione dell'aggiornamento di un software dallo sviluppatore agli utenti, lo streaming audio e video e testi per lettura diretta a un gruppo di lettori etc.
- Per queste applicazioni, un'astrazione molto utile è il concetto di **multicast**: **l'invio di un pacchetto da un mittente a molti destinatari con una singola operazione di spedizione.**

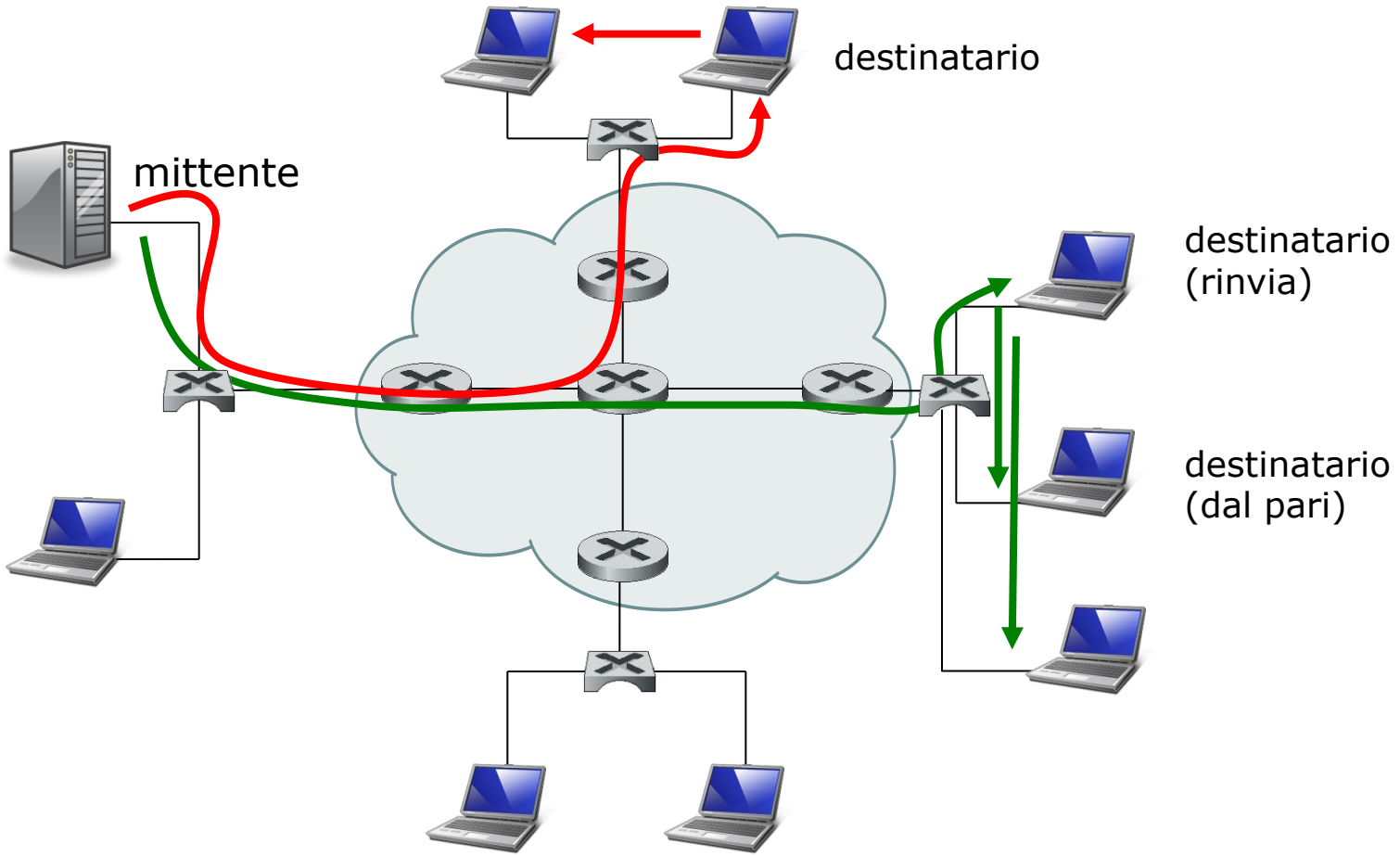
Multicast di Internet e i gruppi multicast

- La comunicazione multicast, può essere implementata in vari modi:
- ***Unicast da-uno-a-molti***. il mittente instaura una diversa connessione unicast per ciascuno dei destinatari. I dati dell'applicazione nel mittente sono trasmessi su ciascuna delle connessioni individuali. Questa soluzione non richiede alcun supporto multicast aggiuntivo nello strato di rete. Questa soluzione è mostrata nella figura seguente, con i router che non sono svolgono alcuna operazione per supportare il multicast.
- In questo esempio, il mittente multicast usa tre connessioni unicast *separate* per raggiungere i tre destinatari.



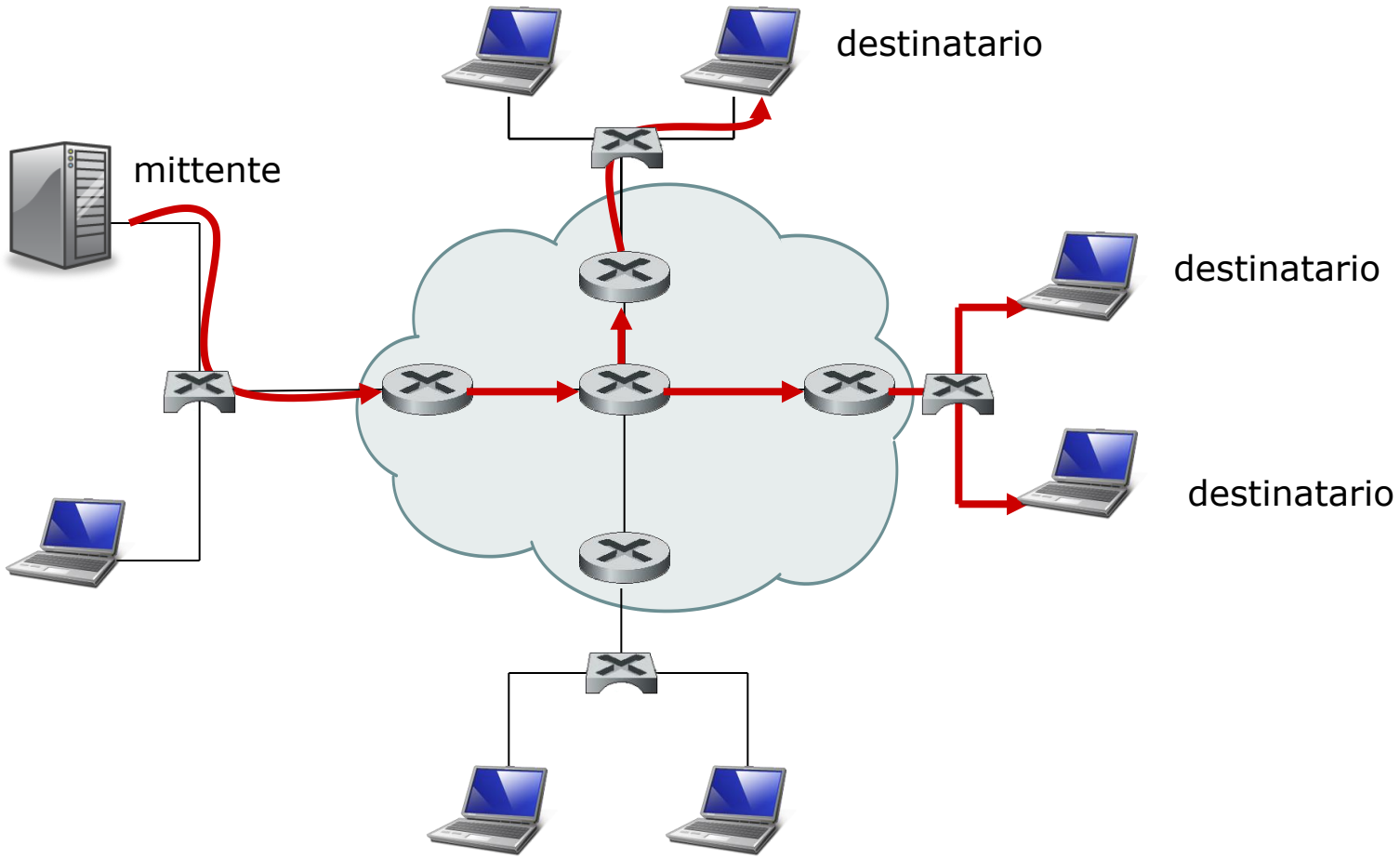
Unicast da uno a molti

- ***Multicast* a *livello applicativo***. Una seconda soluzione usa anch'essa connessioni unicast ma nella trasmissione parte dei riceventi partecipano nella duplicazione e rinvio dei dati.
- Il mittente trasmette i dati a un numero minore di destinatari, i quali fanno delle copie dei dati ricevuti e le rilanciano ad altri riceventi, che possono quindi duplicare e rinviare copie a ulteriori riceventi, e così via.
- Questa soluzione è più efficiente dell'unicast da-uno-a-molti ma richiede la realizzazione di un'infrastruttura di distribuzione a livello applicativo molto complessa. Nell'esempio seguente, i datagram del flusso in verde sono inviati in unicast dal mittente al ricevente a destra in alto. Il ricevente fa, per ciascun datagram ricevuto, due copie, e rilancia in unicast le copie agli altri riceventi sulla sua LAN.



Multicast a livello applicativo

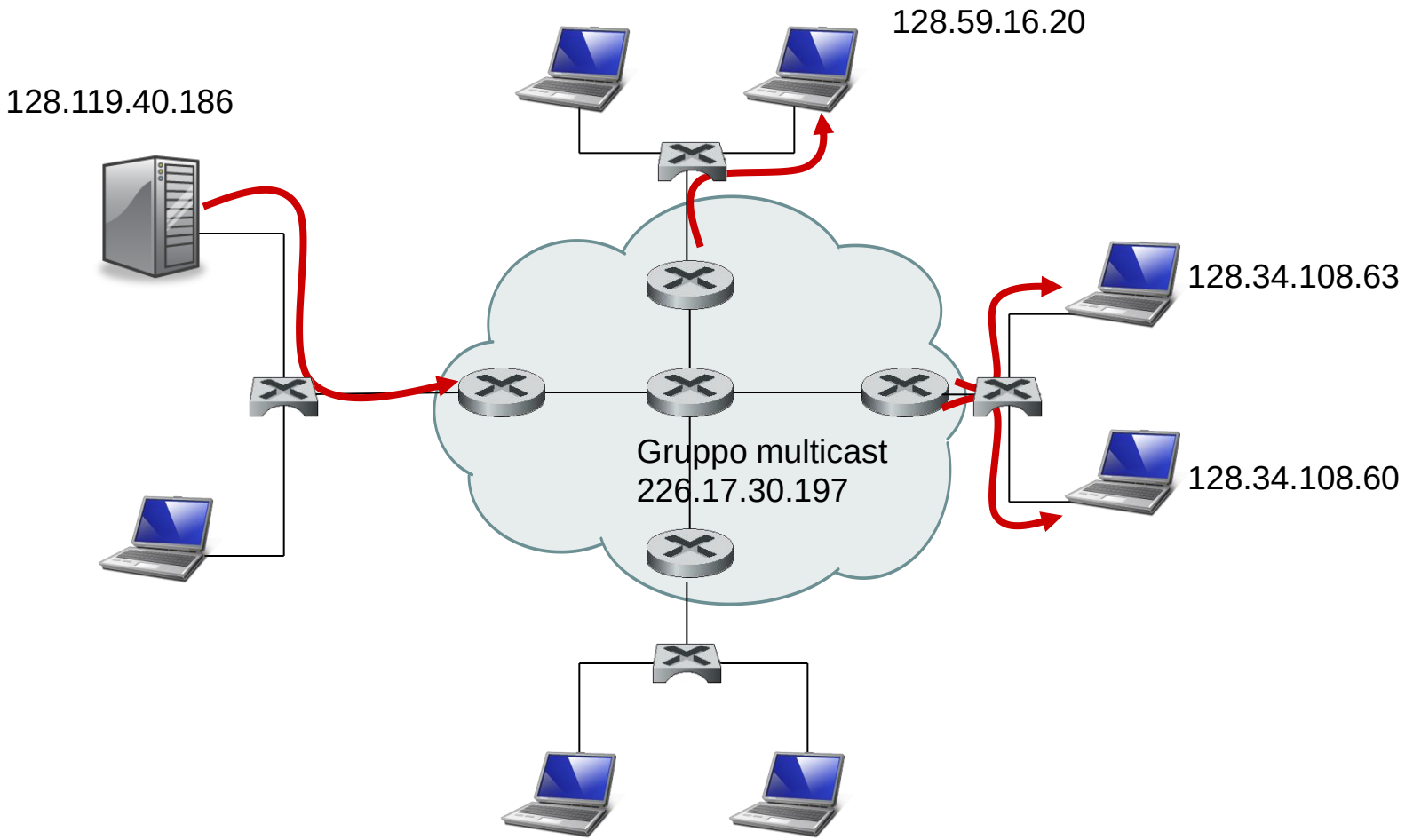
- ***Multicast esplicito***. Una terza soluzione è che lo strato di rete fornisca un **supporto aggiuntivo multicast**.
- In questo modo, i **datagram** sono trasmessi dall'host mittente e quindi rinviati da un router della rete ogni volta che devono essere rinviati su più link in uscita per poter raggiungere i destinatari.
- Nella figura seguente, una copia è rinviata al destinatario in alto e un'altra è rispedita ai destinatari a destra.
- Chiaramente, la terza soluzione è più efficiente in termini di occupazione di larghezza di banda, in quanto solo una *singola* copia di un datagram attraverserà un link. D'altra parte, **allo strato di rete è richiesto un supporto complesso per implementare il multicast.**



Multicast esplicito

- Il multicast, introduce due problemi che sono molto più complessi rispetto al caso dell'unicast:
  - **Identificazione dei destinatari e**
  - **Indirizzamento di un datagram spedito ai destinatari**.
- Nel caso della comunicazione unicast, l'indirizzo IP del destinatario è un campo del datagram IP e identifica un singolo destinatario. Ma nel caso del multicast, i destinatari sono molti.
- Un datagram multicast è indirizzato usando **un indirizzo indiretto**, in altre parole, si usa un unico indirizzo per un gruppo di destinazioni.

- L’indirizzo che rappresenta un gruppo di destinazioni è un **indirizzo multicast di classe D**.
- Gli indirizzi della classe D sono riservati per gli indirizzi multicast e sono compresi da **224.0.0.0** a **239.255.255.255**.
- Un gruppo di destinatari associati ad un indirizzo in classe D prende il nome di **gruppo multicast *(multicast group*)**.
- Uno schema di un gruppo multicast è illustrata nella figura seguente. In questo esempio, quattro host sono associati con il gruppo multicast che ha indirizzo **226.17.30.197** e riceveranno tutti i datagram indirizzati a questo indirizzo multicast.
- Vediamo ora in che modo un host può essere associato ad un numero multicast. Tale compito è svolto dal protocollo IGMP.



Gruppo multicast