

Corsi di Laurea in Ingegneria Informatica
Ingegneria delle Telecomunicazioni
Ingegneria dell'Automazione
Corso di Reti di Calcolatori



Simon Pietro Romano (spromano@unina.it)
Antonio Pescapè (pescapè@unina.it)
Giorgio Ventre (giorgio@unina.it)

Routing

–

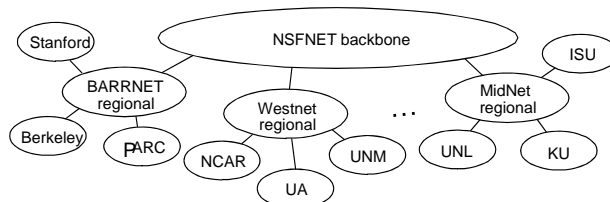
Parte terza:

Internet e il routing gerarchico

Il routing in Internet: com'era



- Negli anni 80 l'architettura di Internet era molto semplice:
 - c'era un'unica rete backbone
 - ogni rete fisica era collegata alla backbone da un core router:
 - ogni core router conosceva le rotte per tutte le reti fisiche



Il routing in Internet: problematiche



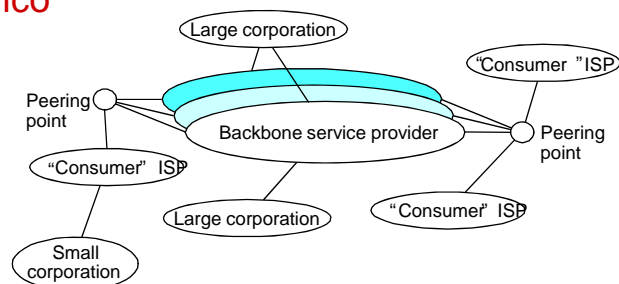
- Non è accettabile che ci sia un unico proprietario per la backbone di tutta la rete
- Non tutte le reti fisiche possono essere collegate direttamente alla backbone
- Soluzione non scalabile:
 - al crescere del numero di core router diventa impossibile mantenerli tutti aggiornati...

3

Il routing in Internet: com'è

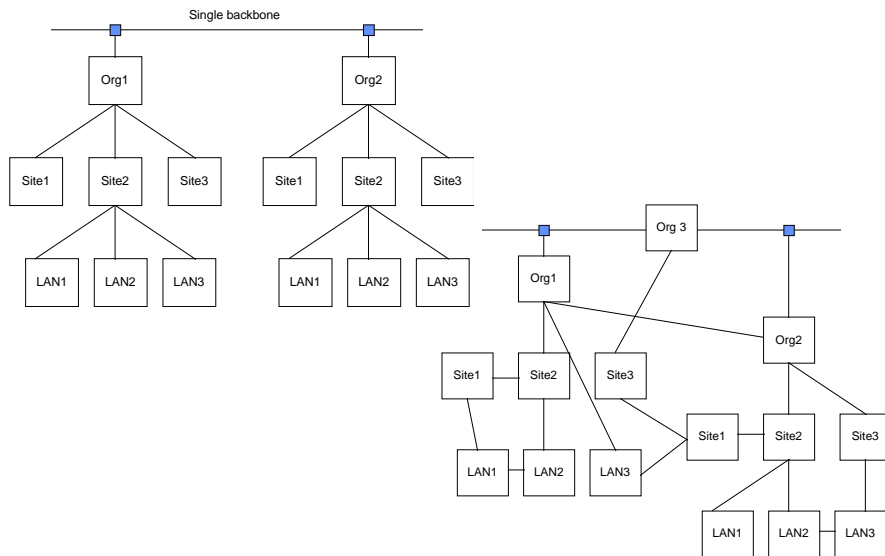


- Reti con *Peer Backbone*:
 - prevedono l'esistenza di diverse dorsali:
 - gli amministratori delle reti backbone devono concordare una politica di routing per evitare la creazione di cicli
 - i core router delle diverse reti devono scambiarsi informazioni sulle rotte
- **Routing Gerarchico**



4

Il routing in Internet: com'era e com'è



5

Il routing in Internet: com'è



- Ai nostri giorni Internet è strutturata come un insieme di **Autonomous System (AS)**:
 - un AS è una collezione di reti amministrate da un'unica autorità
- Ogni AS contiene un numero limitato di reti:
 - la gestione delle informazioni di routing all'interno dell'AS è più semplice
- In teoria, un AS per poter raggiungere un qualsiasi punto di Internet devo concordare ed eventualmente pagare diritti di accesso con altri AS

6

Il routing in Internet: i Tier



- Gli AS quindi devono essere connessi ad altri AS secondo relazioni
 - Paritarie: io concedo a te il passaggio e viceversa (peering).
 - Queste reti sono dette Tier 1
 - Di acquisto: io pago ad un AS il diritto di inviare e ricevere traffico da Internet attraverso di lui (transit)
 - Queste reti sono dette Tier 3
 - Ibride: quando ho sia accordi paritari che di acquisto
 - Queste reti sono dette Tier 2

7

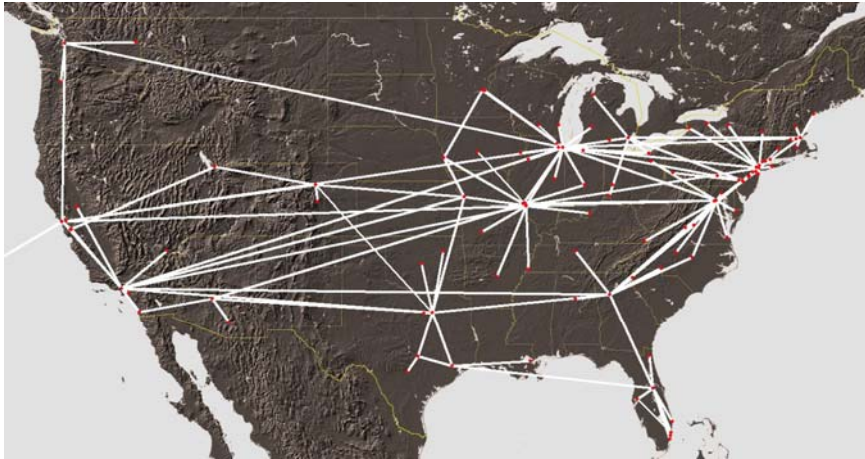
Il Routing in presenza di Autonomous System



- Ogni AS è responsabile del routing all'interno delle sue reti:
 - **routing interno**
- Gli AS devono scambiarsi informazioni di raggiungibilità:
 - **routing esterno**
 - garantisce la correttezza e la consistenza delle informazioni memorizzate nelle tabelle dei router
- Ogni AS deve essere identificato da un nome:
 - AS number (16 bit)

8

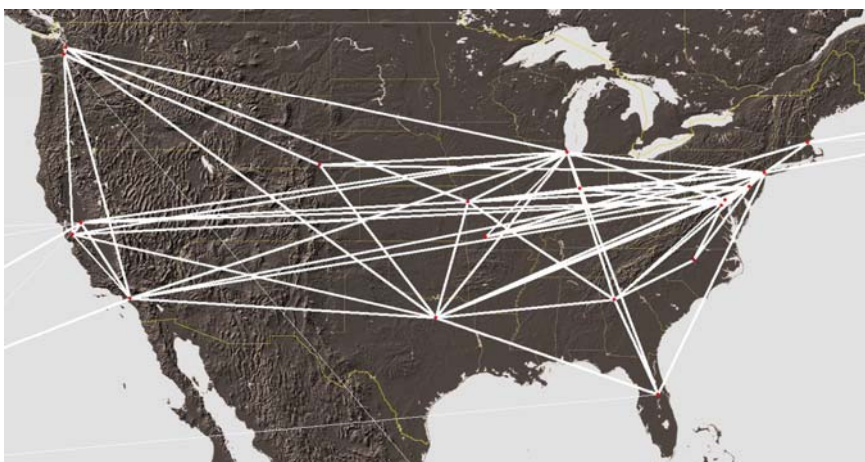
AT&T (AS-7018)



Background image courtesy JHU, applied physics labs

9

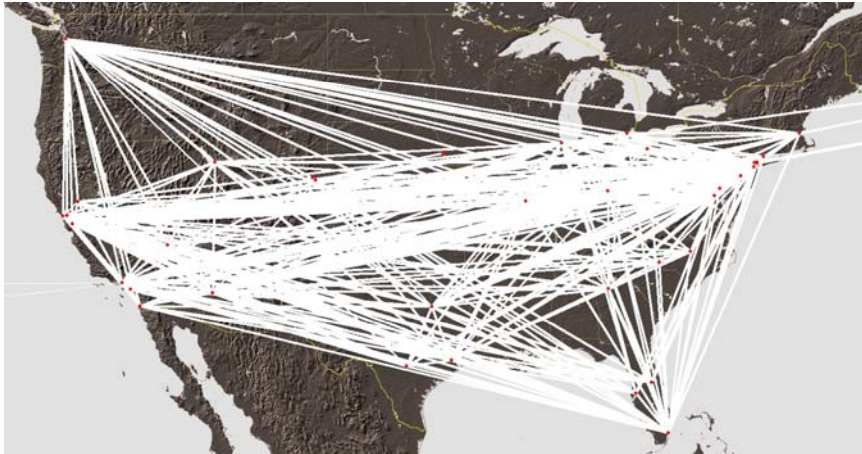
Sprint (AS-1239)



Background image courtesy JHU, applied physics labs

10

Level3 (AS-3356)



Background image courtesy JHU, applied physics labs

11

Routing interno e routing esterno



- Le tabelle di routing interne di un AS sono mantenute dall'**Interior Gateway Protocol** (IGP):
 - i messaggi IGP sono scambiati tra router appartenenti al medesimo AS
 - contengono solo informazioni sulle reti dell'AS
 - RIP (distance vector)
 - OSPF (link state)
 - IGRP (Interior Gateway Routing Protocol – Cisco)
- Le tabelle di routing esterne di un AS sono mantenute dall'**Exterior Gateway Protocol** (EGP):
 - i messaggi EGP sono scambiati tra router designati dai rispettivi AS (border router)
 - contengono informazioni sulle rotte conosciute dai due AS
 - EGP (Exterior Gateway Protocol), ormai obsoleto
 - BGP (Border Gateway Protocol): approccio *path vector*

12

Interior and exterior protocols

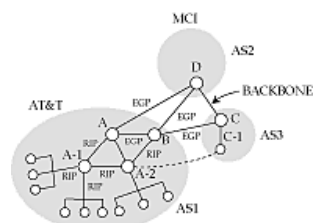


- Internet has three levels of routing
 - highest is at *backbone* level, connecting *autonomous systems (AS)*
 - next level is within AS
 - lowest is within a LAN
- Protocol between AS gateways:
 - exterior gateway protocol (EGP)
- Protocol within AS:
 - interior gateway protocol (IGP)

Exterior gateway protocol (EGP)



- Between untrusted routers
 - mutually suspicious
- Must tell a *border gateway* who can be trusted and what paths are allowed



- *Transit over backdoors* is a problem

Interior Gateway Protocol (IGP)



- Much easier to implement
- Typically partition an AS into *areas*
- Exterior and summary records used between areas

RIP



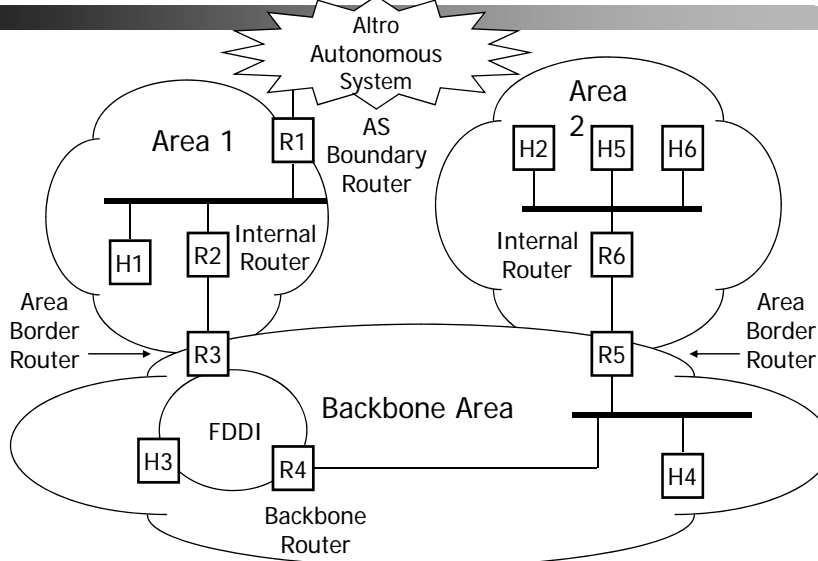
- Distance vector
- Cost metric is hop count
- Infinity = 16
- Exchange distance vectors every 30 s
- Split horizon
- Useful for small subnets
 - easy to install

OSPF



- Link-state
- Uses areas to route packets hierarchically within AS
- Complex
 - LSP databases to be protected
- Uses *designated routers* to reduce number of endpoints

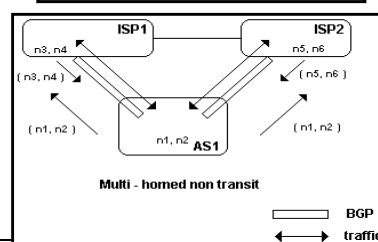
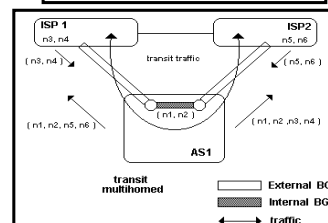
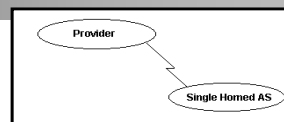
OSPF: scenario operativo



Tipi di AS



- Un solo border router:
 - *stub o single-homed*:
 - (piccole corporate)
- Più border router:
 - *multi-homed*:
 - *transit* (provider)
 - accetta di essere attraversato da traffico diretto ad altri AS
 - *non-transit* (grandi corporate)
 - non accetta di essere attraversato da traffico diretto ad altri AS



19

I gateway router



Sono speciali router dell'AS, che:

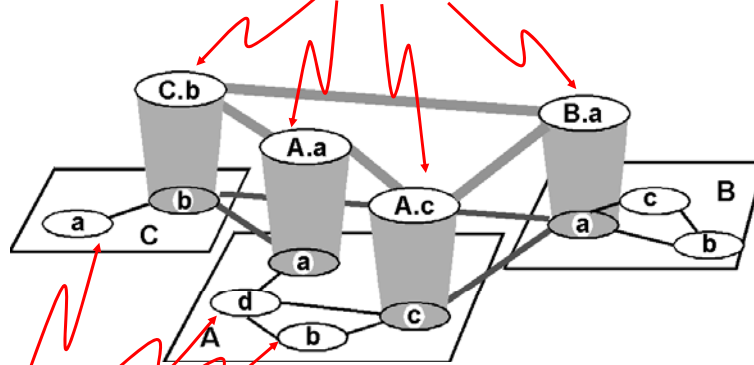
- eseguono protocolli di *routing intra-AS* con altri router appartenenti all'AS
- sono, inoltre, responsabili del routing verso destinazioni esterne al proprio AS:
 - a tal fine, eseguono un protocollo di *routing inter-AS* con altri gateway router
- Su questi router sono pertanto attivi contemporaneamente sia protocolli di routing IGP (ad es. OSPF) e protocolli di routing EGP (ad es. BGP)

20

Instradamento gerarchico in Internet (1/4)



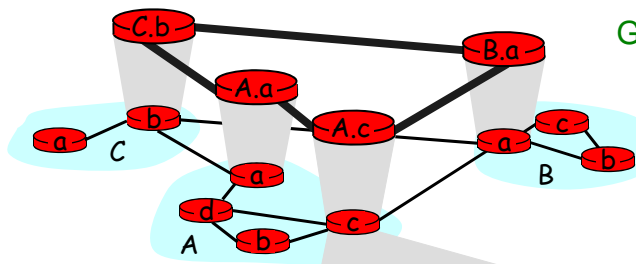
Inter-AS border (exterior gateway) routers



Intra-AS interior (gateway) routers

21

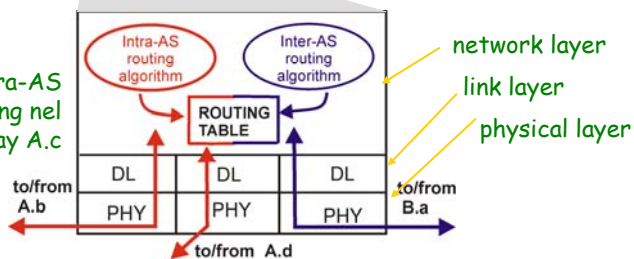
Instradamento gerarchico in Internet (2/4)



Gateway:

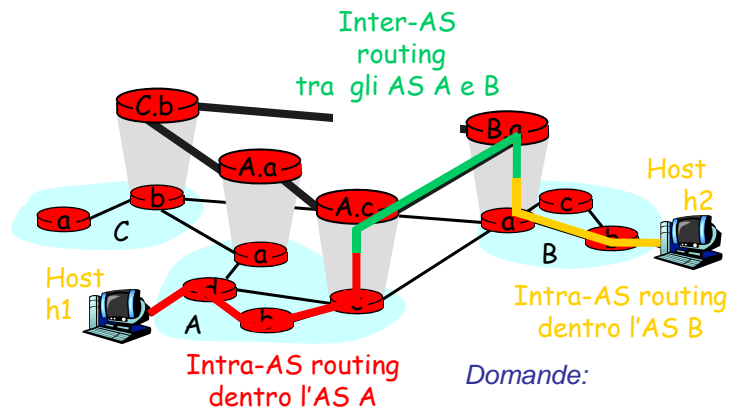
- eseguono inter-AS routing fra loro
- eseguono intra-AS routing con altri router nel loro AS

inter-AS, intra-AS routing nel gateway A.c



22

Instradamento gerarchico in Internet (3/4)



Domande:

- Cosa sa il router A.d ?
- Cosa sa il router A.c ?
- Cosa sa il router B.a ?

23

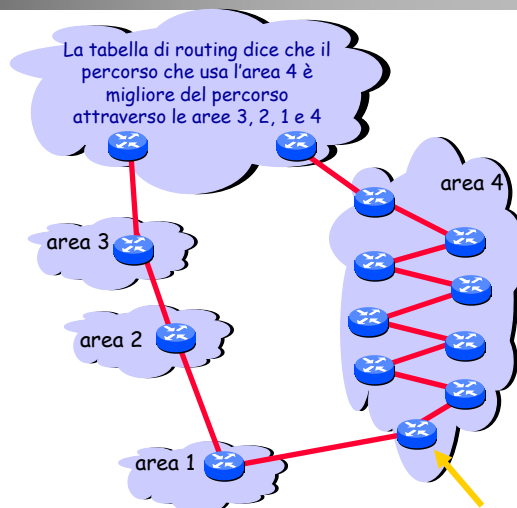
Routing Gerarchico vs Routing Piatto



- Il routing gerarchico è usato per migliorare la scalabilità:

- con 150 milioni di destinazioni:
 - non è possibile memorizzare tutte le destinazioni nelle routing table
 - lo scambio di tabelle di routing così grandi diminuisce notevolmente la banda utilizzata

Ma...



24

Border Gateway Protocol (BGP)



- Uno standard *de facto* (RFC 1772,1773)
- Il più diffuso protocollo EGP
 - sviluppato nell' '89
 - attualmente arrivato alla versione 4
- Utilizza la tecnica *path vector*
 - generalizzazione della tecnica distance vector
 - ogni messaggio contiene una lista di percorsi
- Protocollo TCP, porto 179
- Ogni Border Gateway comunica a tutti i vicini l'intero cammino (cioè la sequenza di AS) verso una specifica destinazione

25

BGP: un esempio



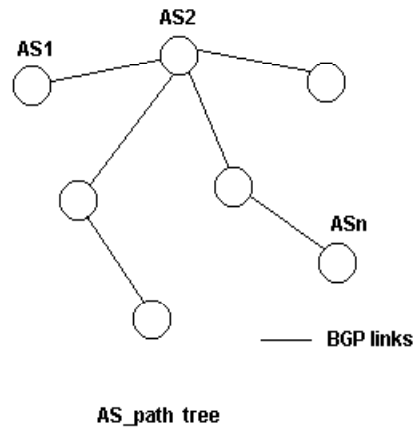
- Il gateway X può memorizzare, per la destinazione Z, il seguente cammino:
$$\text{Path}(X,Z) = X, Y1, Y2, Y3, \dots, Z$$
- Il gateway X manda il suo cammino al peer gateway W
- Il gateway W può scegliere se selezionare il cammino offerto dal gateway X, in base, ad esempio:
 - al costo
 - a questioni politico/economiche
- Se W seleziona il cammino annunciato da X:
$$\text{Path}(W,Z) = W, \text{Path}(X,Z)$$
 - la selezione del cammino è basata più su aspetti politici ed amministrativi (ad es. non passare attraverso concorrenti) che sul costo (ad es. # di AS attraversati)

26

BGP: analisi



- BGP utilizza i messaggi scambiati tra i border router per costruire un grafo di AS
- In genere si costruisce un albero:
 - **AS path tree**

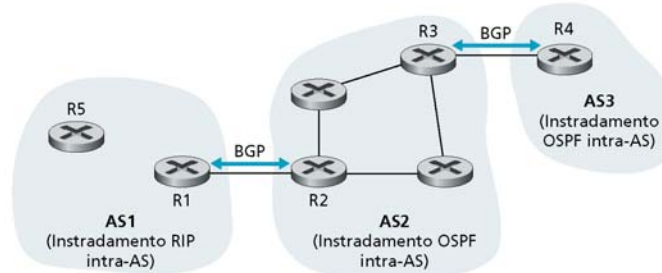


27

BGP: funzioni dei peer



- I BGP peer svolgono tre funzioni principali:
 - stabiliscono la connessione e concordano i parametri di comunicazione
 - si scambiano informazioni di raggiungibilità
 - effettuano un monitoraggio periodico dello stato degli altri peer



28

BGP: politiche di instradamento



- BGP consente solo di pubblicizzare dei percorsi verso altri AS:
 - non associa nessuna metrica ai percorsi
- Il border router “esporta” solo le informazioni consentite dalla politica di routing dell’AS
- Non è possibile considerare BGP come un classico algoritmo di routing

29

BGP: attività principali (I)



1. Ricezione e filtraggio di annunci sui percorsi da parte di vicini direttamente attaccati

- Un pari BGP che annuncia un percorso verso un AS di destinazione promette che se un AS confinante gli rilancerà un pacchetto destinato a quell’AS di destinazione, esso sarà in grado di inoltrare quel pacchetto lungo un percorso verso quella destinazione
- Un router BGP ignorerà gli annunci che contengono il proprio numero di AS nell’AS-PATH, dato che quel percorso darebbe luogo a un loop di instradamento, se usato.
- Poiché viene specificato l’intero percorso verso l’AS, un amministratore di rete può esercitare un notevole controllo sull’instradamento seguito dai pacchetti.

30

BGP: attività principali (II)



2. Selezione del percorso

- Un router BGP può ricevere diversi annunci sui percorsi verso lo stesso AS di destinazione, e deve scegliere quale percorso usare tra quelli annunciati.
- BGP fa una distinzione chiara tra **meccanismo di instradamento** e **politica di instradamento**.
- In particolare, BGP non specifica come un AS deve scegliere un percorso tra quelli annunciati. Questa è una decisione politica che viene lasciata all'amministratore di rete dell'AS
- In assenza di preferenze locali, il percorso selezionato è spesso il più breve percorso di AS (cioè, quello che attraversa il minor numero di AS nel percorso verso la destinazione).

3. Invio di annunci sui percorsi ai vicini

- Così come un router BGP riceverà annunci sui percorsi dai suoi vicini, anche lui annuncerà percorsi ai suoi vicini.

31

BGP: attributi del path e route BGP



- I prefissi annunciati da BGP includono anche degli "attributes"
 - prefix + attributes = "route"
- Due attributi importanti sono:
 - **AS-PATH**: contiene gli ASs attraversati durante l'annuncio del prefisso: e.g., AS 67, AS 17
 - **NEXT-HOP**: l'identità del prossimo router sul percorso
- Quando un router BGP riceve un annuncio di rotta, usa le "**import policy**" per accettarla o rifiutarla.

32

BGP: tipi di messaggio



- **OPEN**
 - inizializza la connessione tra peer:
 - apre connessione TCP
 - autentica il mittente
- **UPDATE**
 - aggiornamento delle informazioni di raggiungibilità
 - annuncio di un nuovo cammino
 - eliminazione di un cammino preesistente
- **NOTIFICATION**
 - risposta ad un messaggio errato
 - chiusura di una connessione
- **KEEPALIVE**
 - verifica che il peer sia ancora attivo
 - si tratta di messaggi che mantengono la connessione attiva in assenza di UPDATE
 - serve a:
 - tenere attiva la connessione TCP
 - dare l'ACK ad una richiesta di OPEN

33

Riassumendo



- Due routers BGP neighbors inizialmente si scambieranno le intere routing tables, dopodichè solo le modifiche attraverso messaggi **UPDATE**.
- Dopo la connessione il primo messaggio ad essere spedito è quello **OPEN** che l'interlocutore confermerà con un messaggio **KEEPALIVE**.
- I messaggi **KEEPALIVE** sono trasmessi periodicamente per mantenere attiva la connessione.
- Il messaggio **NOTIFICATION** viene trasmesso quando si rileva un errore nella trasmissione o per speciali condizioni.

BGP: funzionamento



- Due peer periodicamente si scambiano informazioni di raggiungibilità:
 - nuove rotte
 - vecchie rotte non più valide
- Le informazioni di raggiungibilità vengono trasmesse tramite il messaggio UPDATE
- Tipi di UPDATE:
 - **WITHDRAWN**
 - percorsi non più disponibili
 - **PATH**
 - nuovi percorsi:
 - lista delle reti raggiungibili, con relativi attributi

35

BGP: routing



- BGP consente solo di pubblicizzare informazioni di raggiungibilità:
 - non garantisce la consistenza delle informazioni nelle tabelle di routing
 - **non è un algoritmo di routing**
- Per implementare un sistema di routing inter-AS è necessario che gli AS si fidino l'uno dell'altro
 - il demone **gated** implementa un'interfaccia tra AS distinti:
 - supporta politiche di routing basate su vari tipi di metriche
 - è in grado di integrare il routing interno con quello esterno:
 - può usare un protocollo IGP su un'interfaccia e BGP su un'altra

36

BGP: bgp summary IPv4 (looking glass)



BGP router identifier 195.28.164.125, local AS number 196614
 RIB entries 465670, using 27 MiB of memory
 Peers 32, using 79 KiB of memory
 Peer groups 1, using 16 bytes of memory
 Dampening enabled.

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
64.71.255.61	4	812	268147	3227	0	0	0	07:29:59	230826
70.47.139.3	4	22400	332004	326	0	0	0	2405h52m	233703
70.47.139.4	4	22400	375815	326	0	0	0	2405h52m	233697
80.81.196.177	4	6695	260743	3229	0	0	0	07:29:42	228719
80.81.196.178	4	6695	259443	3229	0	0	0	07:29:32	228710
81.24.16.116	4	43474	6847	3233	0	0	0	07:35:15	2799
84.232.0.242	4	42493	163973	327	0	0	0	2405h52m	235104
87.99.32.1	4	38944	483327	3231	0	0	0	07:30:24	232865
87.99.32.2	4	38944	279617	3232	0	0	0	07:29:46	232942
88.81.250.1	4	31445	179005	327	0	0	0	2405h52m	237294
89.106.65.65	4	31530	238411	3228	0	0	0	07:29:24	232952
91.193.239.1	4	42916	225083	9651	0	0	0	07:27:03	233368
91.194.224.253	4	43366	306988	3206	0	0	0	07:30:16	240007
91.194.224.254	4	43366	299665	3204	0	0	0	07:30:19	240015
193.0.4.28	4	12654	0	400	0	0	0	never	OpenSent
193.26.26.1	4	34024	88322	649	0	0	0	2405h52m	239944
193.142.245.150	4	41495	173947	326	0	0	0	2405h52m	233816
193.242.111.224	4	2128	161149	971	0	0	0	2405h52m	238753
193.242.111.225	4	2128	0	0	0	0	0	never	Connect
194.0.217.1	4	42542	194471	326	0	0	0	2405h52m	235032
194.140.246.253	4	41153	51822	1276	0	0	0	04:17:31	240052
194.140.246.254	4	41153	51699	1284	0	0	0	04:09:31	240057
195.14.247.111	4	8422	697134	329	0	0	0	2405h52m	232704
195.28.164.1	4	31669	269640	3226	0	0	0	07:28:22	232774
195.28.165.1	4	31669	221759	3227	0	0	0	07:29:20	232771
200.152.255.2	4	25933	203667	327	0	0	0	2405h52m	231580
200.152.255.7	4	25933	1594	1594	0	0	0	never	Idle
200.194.237.1	4	11844	203048	329	0	0	0	2405h51m	226420
201.84.224.132	4	83915	0	0	0	0	0	never	Connect
203.84.155.5	4	24523	1594	1594	0	0	0	never	Idle
208.81.207.1	4	19866	553012	3229	0	0	0	07:27:48	233417
217.79.79.12	4	16154	176511	326	0	0	0	2405h52m	234148

Total number of neighbors 32

37

BGP: RIB



- I percorsi vengono immagazzinati nel **RIB** (Routing Information Base), suddiviso come segue:
 - **ADJ-RIB-IN**: contiene tutti i percorsi appresi da messaggi UPDATE, che vengono dati in input al processo decisionale.
 - **LOC-RIB**: contiene le informazioni di routing locale, cioè all'interno dell'AS, che il BGP speaker ha selezionato in base alla politica locale che viene stabilita dall'amministratore.
 - **ADJ-RIB-OUT**: contiene le informazioni di routing che il BGP speaker locale ha selezionato e che sono annunciate ai suoi interlocutori (peers).

Il Routing Arbiter System



- Un meccanismo per coordinare il routing a livello globale
- Un database distribuito ed autenticato che mantiene tutte le informazioni di raggiungibilità
- Sostituisce il core network

39

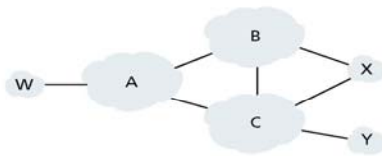
Route Server



- L'architettura di Internet è basata sui **Network Access Point (NAP)**:
 - punti di interconnessione di tutti gli ISP di un'area geografica
- Ogni NAP ha un **route server (RS)**, che mantiene una copia del Routing Arbiter Database
- Ogni ISP ha un border gateway che usa BGP per comunicare con il route server

40

Un semplice esempio



Legenda:



1. X non dice a B che sa raggiungere C
1. B conosce da A che ha un percorso AW
2. B installa il percorso BAW sulla sua RIB
3. B annuncia il suo percorso al proprio cliente X
4. Deve annunciarlo anche a C ?

QUALUNQUE TRAFFICO CHE ATTRAVERSA LA DORSALE DI UN ISP DEVE AVERE UNA SORGENTE O UNA DESTINAZIONE IN UNA RETE CLIENTE DI QUELL'ISP

41

Inter-AS vs Intra-AS routing



- **Politica:**
 - Inter-AS
 - si concentra su aspetti politici (es: quale provider scegliere o evitare)
 - Intra-AS
 - si applica in una singola organizzazione:
 - all'interno dell'organizzazione, la politica di routing applicata è coerente
- **Dimensioni:**
 - si realizza un routing gerarchico
 - si diminuisce il traffico per aggiornare le tabelle di routing
- **Prestazioni:**
 - Inter-AS
 - gli aspetti politico-amministrativi sono prevalenti
 - Intra-AS
 - si concentra sull'ottimizzazione delle prestazioni

42