Corso di Laurea in Ingegneria Informatica

Corso di Reti di Calcolatori
(a.a. 2010/11)

Roberto Canonico (roberto.canonico@unina.it)
Giorgio Ventre (giorgio.ventre@unina.it)

Routing interdominio in Internet
BGP

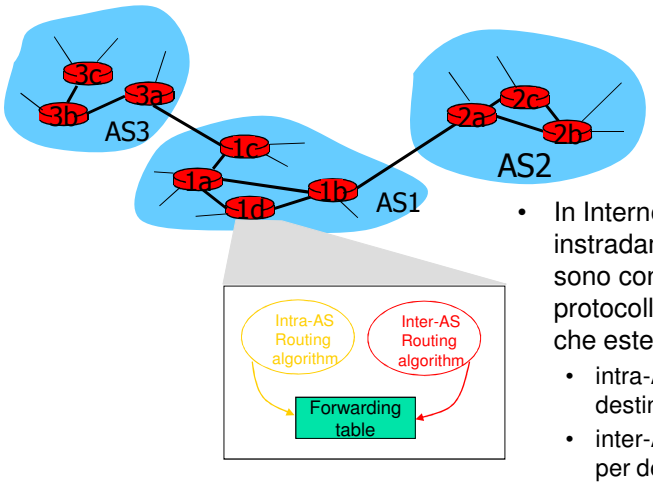
24 novembre 2010

I lucidi presentati al corso sono uno strumento didattico
che NON sostituisce i testi indicati nel programma del corso

Nota di copyright per le slide COMICS

Nota di Copyright

Interconnessione di Autonomous System



- In Internet le tabelle di instradamento dei router sono configurate sia da protocolli di routing interni che esterni
  - intra-AS entry per destinazioni interne
  - inter-As ed intra-AS entry per destinazioni esterne

3

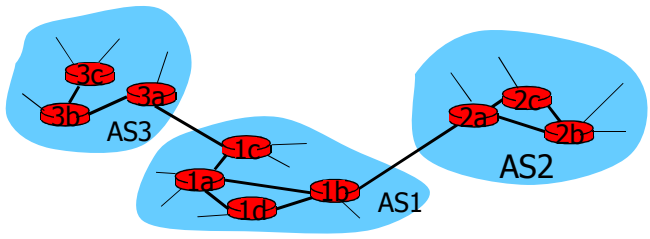
Inter-AS tasks

- suppose router in AS1 receives datagram destined outside of AS1:
  - router should forward packet to gateway router, but which one?

AS1 must:

1. learn which dests are reachable through AS2, which through AS3
2. propagate this reachability info to all routers in AS1

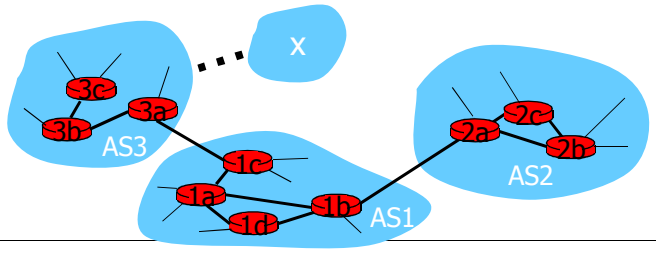
Job of inter-AS routing!



4

Example: Setting forwarding table in router 1d

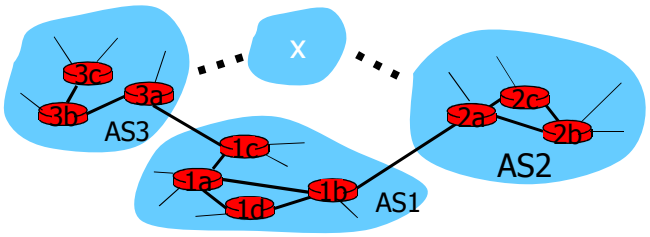
- suppose AS1 learns (via inter-AS protocol) that subnet *x* is reachable via AS3 (gateway 1c) but not via AS2.
- inter-AS protocol propagates reachability info to all internal routers.
- router 1d determines from intra-AS routing info that its interface *I* is on the least cost path to 1c.
  - installs forwarding table entry *(x,I)*



5

Example: Choosing among multiple ASes

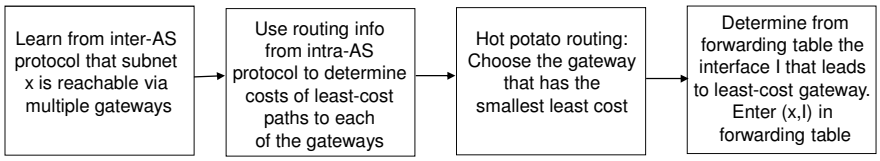
- now suppose AS1 learns from inter-AS protocol that subnet *x* is reachable from AS3 *and* from AS2.
- to configure forwarding table, router 1d must determine towards which gateway it should forward packets for dest x.
  - this is also job of inter-AS routing protocol!



6

Example: Choosing among multiple ASes

- now suppose AS1 learns from inter-AS protocol that subnet *x* is reachable from AS3 *and* from AS2.
- to configure forwarding table, router 1d must determine towards which gateway it should forward packets for dest x.
  - this is also job of inter-AS routing protocol!
- hot potato routing: send packet towards closest of two routers



7

Internet inter-AS routing: BGP

- BGP (Border Gateway Protocol): *the* de facto standard
- BGP provides each AS a means to:
  1. Obtain subnet reachability information from neighboring ASs.
  2. Propagate reachability information to all AS-internal routers.
  3. Determine "good" routes to subnets based on reachability information and policy.
- allows subnet to advertise its existence to rest of Internet: *"I am here"*

8

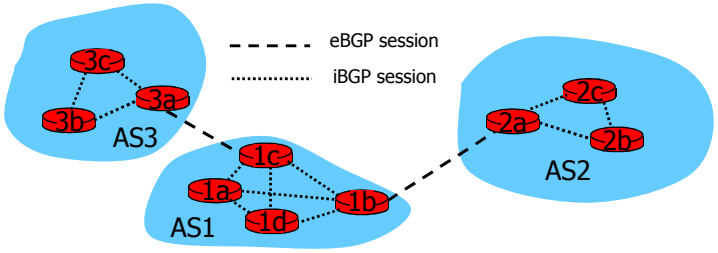
Border Gateway Protocol (BGP)

- 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

9

BGP basics

- pairs of routers (BGP peers) exchange routing info over semi-permanent TCP connections (port 179): BGP sessions
  - BGP sessions need not correspond to physical links.
- when AS2 advertises a prefix to AS1:
  - AS2 *promises* it will forward datagrams towards that prefix.
  - AS2 can aggregate prefixes in its advertisement



10

BGP: cammini

- Il gateway X può memorizzare, per la destinazione Z, il seguente cammino:

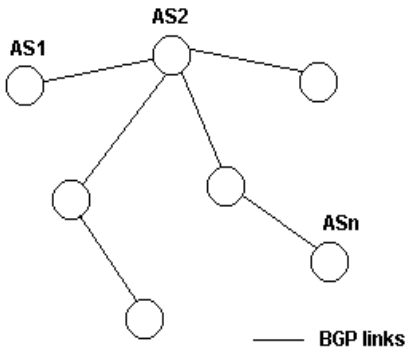
  Path (X,Z) = X,Y1,Y2,Y3,…,Z
- Il gateway X manda il suo cammino al peer gateway W
- Il gateway W può scegliere se selezionare il cammino offerto dal gateway X, in base, ad esempio:
  - al costo
  - a questioni politico/economiche
- Se W seleziona il cammino annunciato da X:

  Path (W,Z) = W, 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)

11

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

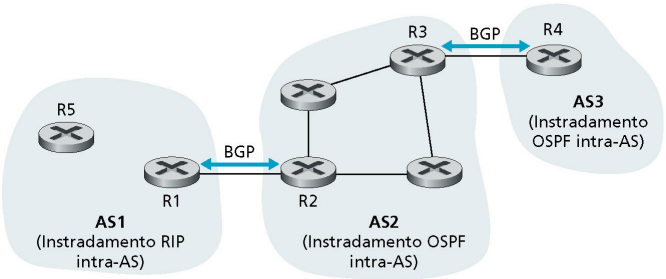


AS_path tree

12

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



13

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

14

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.

15

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.

16

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’annoucement del prefisso: e.g., AS 67, AS 17
  - NEXT-HOP: l’indentità del prossimo router sul percorso
- Quando un router BGP riceve un annuncio di rotta, usa le “import policy” per accettarla o rifiutarla.

17

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

18

BGP: scambi di messaggi tra router

- Due router BGP neighbors inizialmente si scambieranno le intere routing table, dopodiché 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.

19

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

20

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

BGP: bgp summary IPv4 (looking glass)

```
BGP router identifier 195.28.164.125, local AS number 196614
RIB entries 465673, 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  260742     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  159373      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
89.105.65.65    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  597134      329        0    0    0 2405h52m   232704
195.28.164.1    4 31569  269540     3226        0    0    0 07:28:22   232774
195.28.165.1    4 31569  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 28315       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
```

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).

23

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

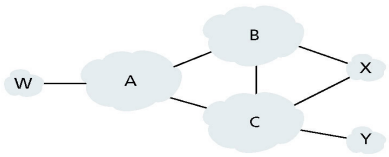
24

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

25

BGP: esempio di routing policy



- X non dice a B che sa raggiungere C
- B apprende da A che ha un percorso AW
- B installa il percorso BAW sulla sua RIB
- B annuncia il suo percorso BAW al proprio cliente X
- Deve annunciarlo anche a C ?
  - No, perché non avrebbe alcun vantaggio economico dal momento che né W né C sono suoi clienti
  - B preferisce che il traffico da C a W passi per A

26

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

27