COSC 301 Network Management and Security

Lecture 23: Exterior Routing and BGP

Today's Focus



How routing between different administrative domains works in Internet?

Idealized View of the Internet



- This abstraction is quite misleading
- The TRUE story
 - The Internet service is provided by a large number of commercial enterprises, generally in competition with each other.
 - Global connectivity is achieved through cooperation between profitable commercial enterprises.

The Real Internet Structure



- Internet is composed of over 50000 Autonomous Systems (ASs) – Superlinear growth
- Each public AS is identified a globally unique number
 - IANA allocates AS Numbers to Regional Internet Registries (RIRs)
 - The RIRs further allocate or assign AS Numbers to network operators
 - 16 bit values

Inter-AS Relationships (1)

• **Peering**: two or more ASs interconnect directly with each other to exchange traffic

- often done without charging for the interconnection or the traffic.



Inter-AS Relationships (2)

• **Transit**: one AS agrees to carry the traffic that flows between ASs

– The transit provider receives a "transit fee" for the service



Routing between ASs

• Distance vector or link state?

– No universal routing metric

- Problems with distance vector
 - Bellman-Ford algorithm may not converge
- Problems with link state
 - Metric used by routers not the same
 - Database too large entire Internet
 - May expose policies to other ASs

BGP

- <u>B</u>order <u>G</u>ateway <u>P</u>rotocol
 - Is a **Policy-Based** routing protocol
 - Is the <u>de facto inter-domain routing protocol</u> of today's global Internet (current version BGP4)
- BGP is classified as a path vector routing protocol
 - makes routing decisions based on paths, network policies, or rule-sets configured by a network administrator
 - defines a route as a pairing between a destination and the attributes of the path to that destination.



BGP Operations



Four Types of BGP Messages

- Open
 - Establishes a peering session on port 179.
- Keep Alive
 - sends 19-byte keep-alive messages every 60 seconds to maintain the connection
- Notification
 - Used for error notification
 - Shuts down a peering session

Update

Announces new routes or withdraws previously announced routes.

Two Types of BGP Neighbour Relationships

- eBGP: External Neighbor in a different AS
- iBGP: Internal Neighbor in the same AS



- The main difference is the way to propagate routes
 - New routes from an eBGP is typical redistributed to all other iBGP and eBGP peers.
 - New routes learned by an iBGP are advertised to only iBGP peers.

BGP Route Processing



Import/Export Routes

- Import Routes
 - When a router hears many possible routes to a destination network, it needs to decide which route to install in its forwarding tables.
 - Order of route preference: customer > peer > provider

Export Routes

- Each AS needs to make decisions on which routes to export to its neighboring ISPs using BGP.
 - No ISP wants to act as transit for packets that it isn't somehow making money on.
- Transit customer routes: Highest priority
- Transit provider routes: Most likely not (no money earned)
- Peer routes: only selected routes to other peering ISPs.

Best Route Selection

Use BGP attrbutes



Given multiple routes to the same prefix, a BGP speaker must pick at most <u>one</u> best route

BGP Attributes

Value	Code	Reference	
1 2	ORIGIN AS_PATH	[RFC1771] [RFC1771]	
4	MULTI_EXIT_DISC	[RFC1771]	
5	LOCAL_PREF	[RFC1771]	
6	ATOMIC_AGGREGATE	[RFC1771]	Most
7	AGGREGATOR	[RFC1771]	important
8	COMMUNITY	[RFC1997]	
9	ORIGINATOR_ID	[RFC2796]	attributes
10	CLUSTER_LIST	[RFC2796]	attinbutee
11	DPA	[Chen]	
12	ADVERTISER	[RFC1863]	
13	RCID_PATH / CLUSTER_ID	[RFC1863]	
14	MP_REACH_NLRI	[RFC2283]	
15	MP_UNREACH_NLRI	[RFC2283]	
16	EXTENDED COMMUNITIES	[Rosen]	
255	reserved for development		
	· · · · · · · · · · · · · · · · · · ·		
From IANA: http://www.lana.org/assignments/bgp-barameters			

AS-Path and Loop Detection

AS_PATH: sequence of AS identifiers that the route advertisement has traversed.



Next_Hop

IP address of the next-hop router along the path to the destination.

- On eBGP sessions, the next hop is set to the IP address of the border router.
- On iBGP sessions, the next hop is not modified.



Local Preference

- The first criteria used to select routes
- Not attached on routes learned via eBGP sessions, but assigned by the import policy of these sessions.



Multi-Exit Discriminator (MED)

- Neighboring AS sets the MED values to indicate which router it prefers to receive traffic for destination.
- Used for comparing two or more routes from the same neighboring AS.



Best Route Selection Criteria



BGP Problem and Migration

- Internal BGP scalability
 - Full mesh connectivity
 - Route reflectors and confederations.
- Instability
 - The routing tables are adjusted continually to reflect actual changes
 - Route flap damping
- Routing table growth – Route summarization
- Load-balancing
 - Locater/Identifier Separation Protocol

Summary

- What is BGP?
- How does BGP process possible routes?
- How does BGP choose best routes?

Resources

- Hari Balakrishnan, and Nick Feamster, Interdomain Internet Routing, http://nms.csail.mit.edu/6.829-f05/lectures/L4routing.pdf
- Mike Pennington, BGP Deployment & Scalability, <u>www.pennington.net/tutorial/bgp_001/BGP_Overview.ppt</u>
- Timothy G. Griffin, An Introduction to Interdomain Routing and the Border Gateway Protocol (BGP), http://www.research.att.com/~griffin