Inter-domain Routing and Connectivity

Presentation: Vish & Joseph
Scribe: Kirill
CS 538 - February 11, 2016
Abstract

- BGP started as a simple path vector protocol
- Today’s internet: Simple shortest path routing is insufficient
- BGP incrementally evolved factors: operational, economic, and political factors
Today’s discussion

- How BGP and its routing policies evolved
- BGP Characteristics
- Routing in single AS
- BGP decision process and policies
- Taxonomy of BGP routing policy
Early days of Internet

- Small
- Operated by NSFNET
- Shortest-path routing
Result

• Large, commercialized and privatized

• ISPs and vested interests in controlling traffic flow
  • Economic and political reasons

• Incremental modifications to BGP

• Complex and overlapping policies
BGP characteristics

• One time complete exchange of routing table

• Incremental exchanges
  New route advertisements, withdrawals and changes to route attributes

• Prefix level route advertisement
BGP Routing

Exchanging routing state

First, BGP is a relatively simple protocol with a few salient features. It is used to exchange routing information to construct a forwarding table that maps each destination address to a particular path through the network to the chosen border router. Each router combines the BGP and IGP routing tables to learn the internal network topology and compute paths from one router to another. In addition, the routers exchange reachability information using BGP, which is used to advertise routes between neighboring Autonomous Systems (ASes). Border (internal) routers as well as other border routers in neighboring ASes are used to exchange routes between neighboring ASes. These sessions are established between neighboring ASes for exchanging routes, where each router in the set is equivalent according to the BGP protocol. The collective results of the decision process across route attributes are used to produce a set of routes that are equally good, for example, Local Preference (LocalPref).

Second, BGP is a protocol where advertisements contain a list of ASes used to reach the destination network, including a list of prefixes being advertised, a list of prefixes advertised from other ASes, and attributes. These changes may be new route advertisements, route withdrawals, or changes to routes that information are exchanged. These changes may be new to a BGP session, where each router in the set is equivalent according to the BGP protocol. However, in order to give operators greater control over route selection, several additional attributes are used to advertise routes, which are included in advertisements, allowing a router to alter its decision process. For example, the Multi-Exit Discriminator (MED) is typically used by two ASes connected by multiple links to indicate which peering link should be used to send traffic to the destination network.

Third, routes are advertised at the border routers of an AS and border routers in neighboring ASes. These sessions are established between neighboring ASes for exchanging routes, where each router in the set is equivalent according to the BGP protocol. The routes advertised are used by the routers to implement policy.

Fourth, BGP update messages may contain several fields, including a list of prefixes being advertised, a list of prefixes advertised from other ASes, and attributes. These changes may be new to a BGP session, where each router in the set is equivalent according to the BGP protocol. However, in order to give operators greater control over route selection, several additional attributes are used to advertise routes, which are included in advertisements, allowing a router to alter its decision process. For example, the Multi-Exit Discriminator (MED) is typically used by two ASes connected by multiple links to indicate which peering link should be used to send traffic to the destination network.

2.2 Selecting a route at a router

There are different locations where a route attribute can be set at and propagated throughout the local AS and filtered before sending to neighboring ISPs. This ordering (e.g. ignore AS path length, or first choose low MED then highest LocalPref) may require various hacks which cannot be changed. The ordering of attributes allows the operator to predict the outcome of making configuration changes. While considering the ordering of attributes, the operator can complicate router configuration and lead to unforeseen side effects.

- Lowest router ID (to break ties)
- Lowest IGP cost to border router
- Lowest MED
- Lowest origin type
- Lowest AS path length
- Highest LocalPref

Step 1: Locally
Step 2: Neighbor
Step 3: Neighbor
Step 4: Neighbor
Step 5: Neighbor
Step 6: Neighbor

Table 1: Steps in the BGP decision process.
BGP Routing

Selecting a route at a router – decision process

<table>
<thead>
<tr>
<th>Step</th>
<th>Attribute</th>
<th>Controlled by local or neighbor AS?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Highest LocalPref</td>
<td>local</td>
</tr>
<tr>
<td>2.</td>
<td>Lowest AS path length</td>
<td>neighbor</td>
</tr>
<tr>
<td>3.</td>
<td>Lowest origin type</td>
<td>neither</td>
</tr>
<tr>
<td>4.</td>
<td>Lowest MED</td>
<td>neighbor</td>
</tr>
<tr>
<td>5.</td>
<td>eBGP-learned over iBGP-learned</td>
<td>neither</td>
</tr>
<tr>
<td>6.</td>
<td>Lowest IGP cost to border router</td>
<td>local</td>
</tr>
<tr>
<td>7.</td>
<td>Lowest router ID (to break ties)</td>
<td>neither</td>
</tr>
</tbody>
</table>
BGP Routing

Controlling import and export of policies

• **Preference** – decide which BGP route to use
  Done by adding/deleting/modifying route attributes in BGP advertisements

• **Filtering** – eliminate certain routes from consideration
  Done by instructing routers to ignore advertisements with matching attribute value range

• **Tagging** – associate additional state with a route
  Done by using the community attribute
Taxonomy of BGP routing policy

1. Business Relationships
2. Traffic Engineering
3. Scalability
4. Security
Business Relationships

ISPs relationships

1. Customer-provider
2. Peer – peer
3. Backup relationships
Business Relationships

• Influencing decision process (assigning LocalPrefs)
  Customer learned routes over peers and providers learned routes

• Controlling route export (by using community attribute)
Traffic engineering

Important goal of ISPs apart from business goals and when several routes have equal preference

Ways to control:

1. Outbound traffic control (By changing LocalPref and IGP costs)

2. Inbound traffic control (By AS prepending and MED)

3. Remote control (by changing community attribute)
Scalability

- Limiting routing table size
  - Protection from other ISPs
    - Filtering long prefixes
    - Fixed per-session prefix limit
    - Default routing
  - Protecting other ISPs – using route aggregation
- Limit the number of routing changes
  - Routing changes cause jitter and packet loss
  - Accomplished via flap damping (route suppression based on a penalty system)
Security

- Discarding invalid routes
  - Unallocated address spaces
  - AS paths with equal Tier level
  - Unpublished configurations
- Protect integrity of routing policies
  - Delete attributes or overwrite them (during import)
  - Set next-hop attribute to specific IP address
- Securing the network infrastructure
  - Not exporting internal backbone IP addresses
  - Export filtering of invalid routes (courtesy)
- Blocking denial-of-service attacks
  - Maximum acceptable prefixes
  - Discard advertisements for offending prefixes
  - Blackhole routes (offenders and victims)
Key BGP Research Areas

- Configuration Checking
  - Consistency criteria checking tools
  - Side-effect predicting modeling tools
  - Routing Arbiter
  - Coordinate route policy selection
- Language Design - Routing Policy Specification Language (RPSL)
- New Architectures
  - Hybrid link-state path-vector (HLP) – improved scalability and convergence than eBGP
  - Routing Control Platform (RCP) – centralizes routes within ISP, improving configurations
Lessons From 10 Years of Studying the Internet AS

1. Too many meanings for “inter-domain topology of the Internet.”

2. The internal structure of ASes is critical to understanding features (route diversity, policy diversity, or multi-connectivity) of real world ASes.

3. Studies that depict the AS-level internet as a digraph are lacking data of how ASes are “really” connected.
Typical AS Graph

\[ G = (V, E) \]

\( V \) – ASes

\( E \) – Connection between Ases

(Connected if they can exchange routing data and presumably IP traffic without the help of intermediary AS that provides transit)
AS Graph Is Not A Di-Graph

ASN-1

ASN-2
AS Graph Is Not A Di-Graph

Prefix in ASN-DEST
AS PATH: ASN-1, -2, -DEST

Prefix in ASN-DEST
AS PATH: ASN-1, -3, -4, -DEST
AS Graph Should Be

• Multigraph – two ASes are connected more than once

• Hypergraph – Single edge can connect multiple Ases

• ASes are NOT atomic
AS Should Be

• Edges and nodes must be labeled by their policies
• The intent of a study is critical in determining how an AS graph’s edges are labeled
  • Business Relationships
  • Physical Link-Levels
  • Connectivity
  • BGP Routing
  • Policy
  • Traffic
AS Graph Should Be

- Economic construct
- Constrained by socio-technological factors
- Driven by economic incentives and business decisions
- Decisions made by the major players (e.g. service and content providers, large corporations, governments)
Lessons continued..

4. BGP hides how ASes are physically connected. It only shows how ASes prefer (and exported) to route. (RouteViews, RIPE RIS)

5. Traceroute was not designed for Internet topology discovery/mapping. (CADA, DIMES, iPlane)

6. The publicly available and widely-used measurement data is not revealing enough to prove new routing protocols and perform sound simulation studies
Lessons continued..

7. AS connectivity is of little importance when examining the vulnerability of the Internet

8. The scale of the Internet’s ASes make strictly “observational” studies hard to interpret. The focus should be on active tests that can isolate behavior.

9. Focused studies are more successful than “fishing expeditions”

10. Robustness is key to analyze the Internet’s topology.
Valid BGP Data Uses

- Discovery of slow convergence and persistence oscillation in routing protocols
- Understanding of the impacts (positive and negative) of route flap dampening
- Determining how much address space and how many ASNs are being actively used
- Looking for routing “Bogons” often related to Internet address hijacking
- Debugging network problems
Improving View of AS Topology

- Heisenberg-like Uncertainty Principle
- Combining datasets (RIPs, looking glasses, IXP data) to find missing edges
- Beacons
- Route Poisoning
- Quality Measurements
Discussion

1. What are main issues having so much flexibility of BGP policies?
2. Is BGP and its policies still evolving?
3. What are other Internet characteristics hidden by BGP?
4. Is it acceptable for the chosen inter-domain routing protocol to hide topology characteristics?
5. Should ISPs, large corporations, and/or governments have vested interest in data currently hidden by BGP?