COMP/ELEC 429/556
Introduction to Computer Networks

Inter-domain routing

Some slides used with permissions from Edward W. Knightly, T. S. Eugene Ng, Ion Stoica, Hui Zhang
Previous focus: Intra-Domain Routing

Intra-domain routing protocol aka Interior Gateway Protocol (IGP)
Today’s focus: Inter-Domain Routing
Inter-Domain Routing Considerations

• Global connectivity is at stake
• Inevitably leads to one single protocol that everyone must speak
  – Unlike many choices in intra-domain routing
• What are the requirements?
  – Scalability
  – Flexibility in choosing routes

• Border Gateway Protocol (BGP)
  – A hybrid between link state and distance vector
  – “Path vector”
Two types of routers
- Border router, Interior router
Border Gateway Protocol Part II: I-BGP

- Two types of routers
  - Border router, Interior router
BGP Operations (Simplified)

Establish session on TCP port 179

Exchange all active routes

Exchange incremental updates

While connection is ALIVE exchange route UPDATE messages
BGP Update Messages

• **Update**: Announcing new routes or withdrawing previously announced routes.

\[
\text{Update} = \text{Destination IP address prefix + attributes values (e.g. a routing path)}
\]
Part I: E-BGP, Share connectivity information across ASs

- You can reach network prefix A via addr and the path is "AS2"
Part II: I-BGP, Carrying Info within an AS

- I-BGP used to disseminate learned routes to all routers in AS

E-BGP update

I-BGP updates
Part II: I-BGP, Carrying Info within an AS

you can reach net B via addr1 and the path is “AS3”

you can reach net B via addr2 and the path is “AS3 AS2”

you can reach net B via addr1 and the path is “AS3”
Attributes are Used to Select Best Routes

Given multiple routes to the same prefix, a BGP speaker must pick at most one best route
Example: Multiple AS Paths

Default choice: Pick shortest path
Shorter Doesn’t Always Mean better

Is path 4 1 better than path 3 2 1?

AS can use custom policies other than shortest path
Benefits of BGP Design

• Path Vector style routing
  – Distance vector algorithm with extra information
  – For each route, store the complete path (ASs)

• Advantages:
  – can make policy choices (choose among many possible learned paths) based on set of ASs in path
  – can easily avoid loops
Announcing and Choosing Routes

• BGP may learn many different paths for a destination network
• Learns only reachability information, no performance metrics
  – Not about optimizing anything
  – All about policy (business and politics)
• What a BGP speaker announces or not announces to a neighbor determines what routes may get used by that neighbor
• Router chooses among paths based on policy
Nontransit vs. Transit ASes

Internet Service providers (often) are transit networks.

Traffic NEVER flows from ISP 1 through NET A to ISP 2 (At least not intentionally!)

Nontransit AS might be a corporate or campus network.
Selective Transit

NET A provides transit between NET B and NET C and between NET D and NET C.

NET A DOES NOT provide transit between NET D and NET B.

Most transit networks transit in a selective manner...

T. S. Eugene Ng
eugeneng at cs.rice.edu
Rice University
Customers and Providers

Customer pays provider for access to the Internet

Customer pays provider for access to the Internet
The Peering Relationship

Peers provide transit between their respective customers

Peers do not provide transit between peers

Peers (often) do not exchange $$$

Traffic allowed

Traffic NOT allowed
Peering Provides Shortcuts

peer
peer
customer
provider
Import Routes

- Provider route
- Peer route
- Customer route
- ISP route

Diagram showing routes from provider, peer, and customer.
## Export Routes

<table>
<thead>
<tr>
<th>Diamond</th>
<th>Red Plus</th>
<th>Heart Plus</th>
<th>Circle</th>
<th>Plus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provider route</td>
<td>Peer route</td>
<td>Customer route</td>
<td>ISP route</td>
<td>Filters block</td>
</tr>
</tbody>
</table>

**Diagram:**
- **To provider**
- **From provider**
- **To peer**
- **To customer**

**Filters block**
- Diamond
- Red Plus
How can routes be marked as “provider”, “peer”, “customer”, “isp”? 

Use “Community Attribute” in route announcement

A community attribute is 32 bits

By convention, first 16 bits is ASN indicating who is giving it an interpretation

Used for signaling within and between ASs

Very flexible BECAUSE it has no predefined meaning

Community number
BGP Issues

• BGP designed for policy not performance

• Susceptible to router misconfiguration
  – Blackholes: announce a route you cannot reach

• Slow convergence time
  – Rate limiting and route flap dampening
Combining IGP and BGP

- IGP Process
  - IGP Routing tables

- BGP Process
  - BGP Routing tables

Forwarding Table Manager

Forwarding Table

OS kernel
Combine BGP and IGP Tables to Create Forwarding Table

I-BGP: 135.207.0.0/16
Next Hop = 192.0.2.1

IGP

<table>
<thead>
<tr>
<th>destination</th>
<th>next hop</th>
</tr>
</thead>
<tbody>
<tr>
<td>192.0.2.0/30</td>
<td>10.10.10.10</td>
</tr>
</tbody>
</table>

IGP + BGP

<table>
<thead>
<tr>
<th>destination</th>
<th>next hop</th>
</tr>
</thead>
<tbody>
<tr>
<td>135.207.0.0/16</td>
<td>192.0.2.1</td>
</tr>
</tbody>
</table>

Forwarding Table

<table>
<thead>
<tr>
<th>destination</th>
<th>next hop</th>
</tr>
</thead>
<tbody>
<tr>
<td>135.207.0.0/16</td>
<td>10.10.10.10</td>
</tr>
<tr>
<td>192.0.2.0/30</td>
<td>10.10.10.10</td>
</tr>
</tbody>
</table>
Life Cycle of a Packet in the Internet

- Address Resolution Protocol (ARP)
  - On the same subnet, need to map IP address to MAC (e.g. Ethernet) address
  - Host and router have ARP cache to store the IP-MAC pairs
  - In case of no match in ARP cache, broadcast an ARP request with the IP address in question and the device with the IP address will reply with its MAC address
Life Cycle of a Packet in the Internet

For each hop in the network, do the following steps:

1. Decapsulate the Ethernet frame to get the IP header (except no need to do this at the source)
2. Check routing table by the destination IP address, get the next-hop IP address and the network interface
3. Learn the MAC address of the next hop (look up in ARP cache or broadcast an ARP request)
4. Encapsulate the IP packet into an Ethernet frame with the destination MAC address
5. Send the Ethernet frame out from the next-hop network interface