Internet Protocols
Fall 2005

Lectures 11-12
Inter-domain routing, mobility support, multicast routing
Andreas Terzis
Outline

• Inter-domain Internet Routing
  – BGP
• Routing for mobile nodes
• Multicast routing
• **AS (Autonomous System)** - a collection of routers under the same technical and administrative domain.
• **EGP (External Gateway Protocol)** - used between two AS’s to allow them to exchange routing information so that traffic can be forwarded across AS borders. Example: BGP
Questions

• How do we “tie” the two routing levels together?
• What are the special constraints of inter-domain routing?
• What is the routing algorithm used at the inter-domain level?
Two-Tier Routing

AS1

AS2

AS3

R1

R2

R3

R4

R5
So Many Choices

Which route should Hopkins pick to 13.13.0.0./16?

AS 1
AS 2
AS 3
AS 4

Hopkins

13.13.0.0/16
Choice of Routing Algorithm

• Constraints:
  - scaling
  - Support policy-based route selection

• Link-state?
  - requires sharing of complete network information
  - information exchanges doesn’t scale
  - can’t express policy

• Distance Vector?
  - scales and retains privacy
  - can’t implement policy
  - can’t avoid loops if shortest paths not taken
Path Vector Protocol

• Distance vector algorithm with extra information
  - For each route, store the complete path (ASs)
  - No extra computation, just extra storage

• Advantages:
  - Can make policy choices based on set of ASs in path
  - Can easily avoid loops
BGP Operations (Simplified)

Establish session on TCP port 179

Exchange all active routes

Exchange incremental updates

While connection is ALIVE exchange route UPDATE messages
Four Types of BGP Messages

- Open: Establish a peering session.
- Keep Alive: Handshake at regular intervals.
- Notification: Shuts down a peering session.
- Update: Announcing new routes or withdrawing previously announced routes.

**Announcement**

=  
\text{prefix} + \text{attributes values}
<table>
<thead>
<tr>
<th>Value</th>
<th>Code</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ORIGIN</td>
<td>[RFC1771]</td>
</tr>
<tr>
<td>2</td>
<td>AS_PATH</td>
<td>[RFC1771]</td>
</tr>
<tr>
<td>3</td>
<td>NEXT_HOP</td>
<td>[RFC1771]</td>
</tr>
<tr>
<td>4</td>
<td>MULTI_EXIT_DISC</td>
<td>[RFC1771]</td>
</tr>
<tr>
<td>5</td>
<td>LOCAL_PREF</td>
<td>[RFC1771]</td>
</tr>
<tr>
<td>6</td>
<td>ATOMIC_AGGREGATE</td>
<td>[RFC1771]</td>
</tr>
<tr>
<td>7</td>
<td>AGGREGATOR</td>
<td>[RFC1771]</td>
</tr>
<tr>
<td>8</td>
<td>COMMUNITY</td>
<td>[RFC1997]</td>
</tr>
<tr>
<td>9</td>
<td>ORIGINATOR_ID</td>
<td>[RFC2796]</td>
</tr>
<tr>
<td>10</td>
<td>CLUSTER_LIST</td>
<td>[RFC2796]</td>
</tr>
<tr>
<td>11</td>
<td>DPA</td>
<td>[Chen]</td>
</tr>
<tr>
<td>12</td>
<td>ADVERTI SER</td>
<td>[RFC1863]</td>
</tr>
<tr>
<td>13</td>
<td>RCID_PATH / CLUSTER_ID</td>
<td>[RFC1863]</td>
</tr>
<tr>
<td>14</td>
<td>MP_REACH_NLRI</td>
<td>[RFC2283]</td>
</tr>
<tr>
<td>15</td>
<td>MP_UNREACH_NLRI</td>
<td>[RFC2283]</td>
</tr>
<tr>
<td>16</td>
<td>EXTENDED COMMUNITIES</td>
<td>[Rosen]</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>255</td>
<td>reserved for development</td>
<td></td>
</tr>
</tbody>
</table>
Attributes are Used to Select Best Routes

Given multiple routes to the same prefix, a BGP speaker must pick at most one best route (Note: it could reject them all!)
Every time a route announcement crosses an AS boundary, the Next Hop attribute is changed to the IP address of the border router that announced the route.
Interdomain Loop Prevention

- BGP at AS YYY will never accept a route with ASPATH containing YYY

AS 7018

Don't Accept!

12.22.0.0/16
ASPATH = 1 333 7018 877

AS 1
BGP Route Processing

Receive BGP Updates

Apply Policy = filter routes & tweak attributes

Based on Attribute Values

Best Routes

Apply Policy = filter routes & tweak attributes

Transmit BGP Updates

Apply Import Policies

Best Route Selection

Best Route Table

Apply Export Policies

Install forwarding Entries for best Routes.

IP Forwarding Table
Route Selection Summary

- Highest Local Preference
- Enforce relationships
- Shortest ASPATH
- Lowest MED
- i-BGP < e-BGP
- Lowest IGP cost to BGP egress
- Lowest router ID
- traffic engineering
- Throw up hands and break ties
Local Preference

Higher Local preference values are more preferred
Shortest AS Path

- Route Selection Algorithm
  1. Local Pref
  2. Shortest AS Path
- Shortest AS Path != shortest path

BGP says that path 4 1 is better than path 3 2 1
Outline

• Inter-domain Internet Routing
  – BGP

• Routing for mobile nodes

• Multicast routing
Routing for Mobile Hosts

• Internet routing does not support mobile hosts
  - IP address is divided into network part and host part
  - Rest of the Internet knows how to forward packets to the network end-host is connected to
  - If end-host moves packets are delivered in the wrong network
Partial Solution

• As mobile node moves it acquires different IP addresses
  - DHCP

• This is adequate for some applications
  - Web browsing, etc

• Does not work for other applications
  - Mobile host is the server
  - Long lasting connections
Mobile IP

• Solution proposed to solve the problem of mobility
• Does not require changes in most IP infrastructure
• Introduces three new entities
  – Home Agent (HA)
  – Foreign Agent (FA)
  – Mobile Host (MH)
Example
Sub-problems

• How does the home agent intercept packets that are destined to the mobile node?
• How does the home agent deliver the packet to the foreign agent?
• How does the foreign agent deliver the packet to the mobile node?
Home Agent Intercept

- Packets from Corresponding Agents (CAs) arrive at the home network
- Home agent issues proxy ARP
  - Replies with its own MAC to queries for the Mobile Hosts IP
  - Packets are delivered to the Home Agent
- When Mobile Host returns to the home network
  - Gratuitous ARP
Tunneling

• A *tunnel* exists between HA and FA
  - Home agent encapsulates IP packets for the Mobile Host using the FA’s IP address
  - Rest of the Internet forwards packet as if it was destined for the FA
  - When packet arrives at the FA, it decapsulates the packet
FA Delivery

- FA bypasses regular IP routing for registered MHs and delivers directly to the MAC address
- MH can be the FA
- What about packets sourced by the MH?
  - Sent from the remote network with the MH’s permanent address
  - Reverse tunnel
Route Optimization

• Packets from CA -> HA -> FA follow suboptimal route

• If CA is upgraded it can create tunnel directly to the FA
  – HA sends binding update to the CA when packets arrive to it
  – FA can also send binding updates

• Security implications?
Outline

• Inter-domain Internet Routing
  – BGP

• Routing for mobile nodes

• Multicast routing
Multicast Routing

- Applications that want to send same data to multiple destinations
  - Audio, Video, Stock Updates
- Multicast model
  - Sender sends to Multicast Address
  - Receivers join Multicast Address
    - IGMP
  - Packets are delivered to all hosts that joined the multicast address
- How do the receivers know which multicast group to join?
Distance-Vector Multicast

• DVMRP uses a two step approach
  – Flooding mechanism that delivers packets to the whole network
  – *Pruning* networks that have no receivers
Reverse-Path Broadcast

- Flood packets from S to everyone
  - Router forwards packets to all interfaces (minus incoming) if packet arrives on correct interface
    - Interface R would use to send packets to S
  - Packet will be sent multiple times if multiple routers connected on same LAN
    - Designate parent for each network
    - Router with shortest distance to S
Reverse-Path Multicast

• RPB delivers packets to all networks
  – Obviously undesirable
• Leaf networks know if any interested receivers are present
• Routers propagate this information to upstream neighbors
  – Tree is pruned
• What happens if a new receiver joins at a pruned network?
Alternatives

• DVMRP not appropriate for *sparse* groups

• Protocol Independent Multicast (PIM) solves this issue
  – Rendezvous Point (RP)