Advanced Internetworking
Problems

- How do we build a routing system that can handle hundreds of thousands of networks and billions of end nodes?
- How to handle address space exhaustion of IPV4?
- How to enhance the functionalities of Internet?
The Global Internet

The tree structure of the Internet in 1990
The Global Internet

A simple multi-provider Internet
Interdomain Routing (BGP)

- Internet is organized as autonomous systems (AS) each of which is under the control of a single administrative entity

- Autonomous System (AS)
  - corresponds to an administrative domain
  - examples: University, company, backbone network

- A corporation’s internal network might be a single AS, as may the network of a single Internet service provider
Interdomain Routing

A network with two autonomous system
Route Propagation

- Idea: Provide an additional way to hierarchically aggregate routing information is a large internet.
  - Improves scalability

- Divide the routing problem in two parts:
  - Routing within a single autonomous system
  - Routing between autonomous systems

- Another name for autonomous systems in the Internet is routing domains
  - Two-level route propagation hierarchy
    - Inter-domain routing protocol (Internet-wide standard)
    - Intra-domain routing protocol (each AS selects its own)
EGP and BGP

- Inter-domain Routing Protocols
  - Exterior Gateway Protocol (EGP)
    - Forced a tree-like topology onto the Internet
    - Did not allow for the topology to become general
      - Tree like structure: there is a single backbone and autonomous systems are connected only as parents and children and not as peers
  - Border Gateway Protocol (BGP)
    - Assumes that the Internet is an arbitrarily interconnected set of ASs.
    - Today’s Internet consists of an interconnection of multiple backbone networks (they are usually called service provider networks, and they are operated by private companies rather than the government)
    - Sites are connected to each other in arbitrary ways
BGP

- Some large corporations connect directly to one or more of the backbone, while others connect to smaller, non-backbone service providers.

- Many service providers exist mainly to provide service to “consumers” (individuals with PCs in their homes), and these providers must connect to the backbone providers.

- Often many providers arrange to interconnect with each other at a single “peering point”
BGP-4: Border Gateway Protocol

- Assumes the Internet is an arbitrarily interconnected set of AS's.

- Define *local traffic* as traffic that originates at or terminates on nodes within an AS, and *transit traffic* as traffic that passes through an AS.

- We can classify AS's into three types:
  - *Stub AS*: an AS that has only a single connection to one other AS; such an AS will only carry local traffic (*small corporation in the figure of the previous page*).
  
  - *Multihomed AS*: an AS that has connections to more than one other AS, but refuses to carry transit traffic (*large corporation at the top in the figure of the previous page*).
  
  - *Transit AS*: an AS that has connections to more than one other AS, and is designed to carry both transit and local traffic (*backbone providers in the figure of the previous page*).
The goal of Inter-domain routing is to find any path to the intended destination that is loop free.

- We are concerned with reachability than optimality.
- Finding path anywhere close to optimal is considered to be a great achievement.

Why?
Scalability: An Internet backbone router must be able to forward any packet destined anywhere in the Internet
- Having a routing table that will provide a match for any valid IP address

Autonomous nature of the domains
- It is impossible to calculate meaningful path costs for a path that crosses multiple ASs
- A cost of 1000 across one provider might imply a great path but it might mean an unacceptable bad one from another provider

Issues of trust
- Provider A might be unwilling to believe certain advertisements from provider B
BGP

Each AS has:

- One BGP *speaker* that advertises:
  - local networks
  - other reachable networks (transit AS only)
  - gives *path* information

- In addition to the BGP speakers, the AS has one or more border “gateways” which need not be the same as the speakers

- The border gateways are the routers through which packets enter and leave the AS
BGP

- BGP does not belong to either of the two main classes of routing protocols (distance vectors and link-state protocols)

- BGP advertises *complete paths* as an enumerated lists of ASs to reach a particular network
BGP Example

Example of a network running BGP
BGP Example

- Speaker for AS 2 advertises reachability to P and Q
  - Network 128.96, 192.4.153, 192.4.32, and 192.4.3, can be reached directly from AS 2.

- Speaker for backbone network then advertises
  - Networks 128.96, 192.4.153, 192.4.32, and 192.4.3 can be reached along the path <AS 1, AS 2>.

- Speaker can also cancel previously advertised paths
BGP Issues

- It should be apparent that the AS numbers carried in BGP need to be unique.
- For example, AS 2 can only recognize itself in the AS path in the example if no other AS identifies itself in the same way.
- AS numbers are 16-bit numbers assigned by a central authority.
Integrating Interdomain and Intradomain Routing

All routers run iBGP and an intradomain routing protocol. Border routers (A, D, E) also run eBGP to other ASs.
Integrating Interdomain and Intradomain Routing

<table>
<thead>
<tr>
<th>Prefix</th>
<th>BGP Next Hop</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.0/16</td>
<td>E</td>
</tr>
<tr>
<td>12.5.5/24</td>
<td>A</td>
</tr>
<tr>
<td>128.34/16</td>
<td>D</td>
</tr>
<tr>
<td>128.69./16</td>
<td>A</td>
</tr>
</tbody>
</table>

BGP table for the AS

<table>
<thead>
<tr>
<th>Router</th>
<th>IGP Path</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>D</td>
<td>C</td>
</tr>
<tr>
<td>E</td>
<td>C</td>
</tr>
</tbody>
</table>

IGP table for router B

<table>
<thead>
<tr>
<th>Prefix</th>
<th>IGP Path</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.0/16</td>
<td>C</td>
</tr>
<tr>
<td>12.5.5/24</td>
<td>A</td>
</tr>
<tr>
<td>128.34/16</td>
<td>C</td>
</tr>
<tr>
<td>128.69./16</td>
<td>A</td>
</tr>
</tbody>
</table>

Combined table for router B

BGP routing table, IGP routing table, and combined table at router B
Routing Areas

A domain divided into areas

Backbone area
Area border router (ABR)
Next Generation IP (IPv6)
Major Features

- 128-bit addresses
- Multicast
- Real-time service
- Authentication and security
- Auto-configuration
- End-to-end fragmentation
- Enhanced routing functionality, including support for mobile hosts
IPv6 Addresses

- Classless addressing/routing (similar to CIDR)
- Notation: x:x:x:x:x:x:x:x (x = 16-bit hex number)
  - contiguous 0s are compressed: 47CD::A456:0124
  - IPv6 compatible IPv4 address: ::128.42.1.87

- Address assignment
  - provider-based
  - geographic
IPv6 Header

- 40-byte “base” header
- Extension headers (fixed order, mostly fixed length)
  - fragmentation
  - source routing
  - authentication and security
  - other options
Internet Multicast
Overview

- IPv4
  - class D addresses
  - demonstrated with MBone
  - uses tunneling

- Integral part of IPv6
  - problem is making it scale
Overview

- **One-to-many**
  - Radio station broadcast
  - Transmitting news, stock-price
  - Software updates to multiple hosts

- **Many-to-many**
  - Multimedia teleconferencing
  - Online multi-player games
  - Distributed simulations
Overview

- Without support for multicast
  - A source needs to send a separate packet with the identical data to each member of the group
    - This redundancy consumes more bandwidth
    - Redundant traffic is not evenly distributed, concentrated near the sending host
  - Source needs to keep track of the IP address of each member in the group
    - Group may be dynamic

- To support many-to-many and one-to-many IP provides an IP-level multicast
Overview

- Basic IP multicast model is many-to-many based on multicast groups
  - Each group has its own IP multicast address
  - Hosts that are members of a group receive copies of any packets sent to that group’s multicast address
  - A host can be in multiple groups
  - A host can join and leave groups
Overview

- Using IP multicast to send the identical packet to each member of the group
  - A host sends a single copy of the packet addressed to the group’s multicast address
  - The sending host does not need to know the individual unicast IP address of each member
  - Sending host does not send multiple copies of the packet
Overview

- IP’s original many-to-many multicast has been supplemented with support for a form of one-to-many multicast

- One-to-many multicast
  - Source specific multicast (SSM)
  - A receiving host specifies both a multicast group and a specific sending host

- Many-to-many model
  - Any source multicast (ASM)
Overview

- A host signals its desire to join or leave a multicast group by communicating with its local router using a special protocol
  - In IPv4, the protocol is Internet Group Management Protocol (IGMP)
  - In IPv6, the protocol is Multicast Listener Discovery (MLD)

- The router has the responsibility for making multicast behave correctly with regard to the host
Multicast Routing

- A router’s unicast forwarding tables indicate for any IP address, which link to use to forward the unicast packet.
- To support multicast, a router must additionally have multicast forwarding tables that indicate, based on multicast address, which links to use to forward the multicast packet.
- Unicast forwarding tables collectively specify a set of paths.
- Multicast forwarding tables collectively specify a set of trees:
  - Multicast distribution trees.
Multicast Routing

- To support source specific multicast, the multicast forwarding tables must indicate which links to use based on the combination of multicast address and the unicast IP address of the source.

- Multicast routing is the process by which multicast distribution trees are determined.
Distance-Vector Multicast

- Each router already knows that shortest path to source S goes through router N.
- When receive multicast packet from S, forward on all outgoing links (except the one on which the packet arrived), iff packet arrived from N.
- Eliminate duplicate broadcast packets by only letting
  - “parent” for LAN (relative to S) forward
    - shortest path to S (learn via distance vector)
    - smallest address to break ties
Reverse Path Broadcast (RPB)

- **Goal:** Prune networks that have no hosts in group G
- **Step 1:** Determine if LAN is a leaf with no members in G
  - leaf if parent is only router on the LAN
  - determine if any hosts are members of G using IGMP

- **Step 2:** Propagate “no members of G here” information
  - augment \( <\text{Destination}, \text{Cost}> \) update sent to neighbors with set of groups for which this network is interested in receiving multicast packets.
  - only happens when multicast address becomes active.

Distance-Vector Multicast
Protocol Independent Multicast (PIM)

Shared Tree

Source specific tree

RP = Rendezvous point
--- Shared tree
--- Source-specific tree for source R1
Protocol Independent Multicast (PIM)

Delivery of a packet along a shared tree. R1 tunnels the packet to the RP, which forwards it along the shared tree to R4 and R5.
Inter-domain Multicast

Multicast Source Discovery Protocol (MSDP)

(a) Domain A

1: Register

SR

RP1

2a: Join

3: Join

Domain B

RP2

2b: MSDP Source active

(b) Domain A

Shared tree

Source-specific tree for source SR
Routing for Mobile Hosts

- Mobile IP
  - home agent
    - Router located on the home network of the mobile hosts
  - home address
    - The permanent IP address of the mobile host.
    - Has a network number equal to that of the home network and thus of the home agent
  - foreign agent
    - Router located on a network to which the mobile node attaches itself when it is away from its home network
Routing for Mobile Hosts

- Problem of delivering a packet to the mobile node

  - How does the home agent intercept a packet that is destined for the mobile node?
    - Proxy ARP

  - How does the home agent then deliver the packet to the foreign agent?
    - IP tunnel
    - Care-of-address

  - How does the foreign agent deliver the packet to the mobile node?
Routing for Mobile Hosts

- Route optimization in Mobile IP
  - The route from the sending node to mobile node can be significantly sub-optimal
  - One extreme example
    - The mobile node and the sending node are on the same network, but the home network for the mobile node is on the far side of the Internet
      - Triangle Routing Problem
  - Solution
    - Let the sending node know the care-of-address of the mobile node. The sending node can create its own tunnel to the foreign agent
    - Home agent sends binding update message
    - The sending node creates an entry in the binding cache
    - The binding cache may become out-of-date
      - The mobile node moved to a different network
      - Foreign agent sends a binding warning message
Summary

- We have looked at the issues of scalability in routing in the Internet
- We have discussed IPV6
- We have discussed Multicasting
- We have discussed Mobile IP