Internet Routing Protocols
Lecture 01

Advanced Systems Topics
Lent Term, 2008

Timothy G. Griffin
Computer Lab
Cambridge UK

Common View of the Telco Network
IP routing is the little bit-o-smarts left in the IP network layer

- Dynamic Routing protocols are used to implement and maintain connectivity in the Internet.
- Which protocols are used?
- How do they work?
- How do they behave?
- What are some of the fundamental tradeoffs in the design space of routing protocols?
Outline

- Lecture 1: Routing vs. Forwarding. Internet routing architecture
- Lecture 3: Inter-domain routing. The Border Gateway Protocol (BGP)
- Lecture 4: BGP continued
- Lecture 5: BGP dynamics
- Lecture 6: BGP routing anomalies
Architecture of Dynamic Routing

OSPF, IS-IS, RIP, EIGRP (cisco)


Only one: BGP
The Routing Domain of BGP is the entire Internet

How many ASN are used today?

http://bgp.potaroo.net

Jan 28. 2008
Routers Talking to Routers
(The “control plane”)

- Routing computation is distributed among routers within a routing domain
- Computation of best next hop based on routing information is the most CPU/memory intensive task on a router
- Routing messages are *usually* not routed, but exchanged via layer 2 between physically adjacent routers (internal BGP and multi-hop external BGP are exceptions)

Technology of Distributed Routing

**Link State**
- Topology information is *flooded* within the routing domain
- Best end-to-end paths are computed locally at each router
  - Best end-to-end paths determine next-hops
  - Based on minimizing some notion of distance
  - Works only if policy is *shared* and *uniform*
  - Examples: OSPF, IS-IS

**Vectoring**
- Each router knows little about network topology
- Only best next-hops are chosen by each router for each destination network
  - Best end-to-end paths result from composition of all next-hop choices
- Does not require any notion of distance
- Does not require uniform policies at all routers
- Examples: RIP, BGP
The Gang of Four

<table>
<thead>
<tr>
<th>Link State</th>
<th>Vectoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>OSPF</td>
<td>RIP</td>
</tr>
<tr>
<td>IS-IS</td>
<td>BGP</td>
</tr>
</tbody>
</table>

IGP

EGP

The standard model

data received

Application
Presentation
Session
Transport
Network
DataLink
Physical

data sent

Application
Presentation
Session
Transport
Network
DataLink
Physical
### All Hail the IP Datagram!

```
+---------------------------------------------+------------------+
<table>
<thead>
<tr>
<th>Version</th>
<th>IHL</th>
<th>Service Type</th>
<th>Total Length</th>
</tr>
</thead>
</table>
+---------------------------------------------+------------------+
| Identification | Flags | Fragment Offset |
+---------------------------------------------+------------------+
| Time to Live | Protocol | Header Checksum |
+---------------------------------------------+------------------+
| Source Address |
+---------------------------------------------+------------------+
| Options | Padding |
+---------------------------------------------+------------------+
```

... up to 65,515 octets of data ...

1981, RFC 791

### IP Hour Glass

**Networking Applications**
- Remote Access
- Voice
- Multimedia
- Email
- File Transfer
- Web
- VPN

**Networking Technologies**
- Frame
- ATM
- Ethernet
- SONET
- DWDM
- FDDI
- X.25

**IP**
- Minimalist network layer
Best Effort, Connectionless, Connectivity

The is the fundamental service provided by Internet Service Providers (ISPs)

All other IP services depend on connectivity: DNS, email, VPNs, Web Hosting, ...

IP is a Network Layer Protocol

Separate physical networks glued together into one logical network
Hosts, Networks, and Routers

Network A
Host 7
Host 1
Network B
Host 12
Host 2
Router

Network C
Host 1
Host 2

Unique IP Address = Network Number + Host Number

Actually, IP addresses identify interfaces

Network A
Host 7
Host 1
Network B
Host 12
Host 2

Network C
Host 1
Host 2

Network A, Host 3

Network B, Host 77

Network C, Host 3

Machines can have more than one IP address. All routers do!
### IP Forwarding Table

<table>
<thead>
<tr>
<th>Destination</th>
<th>Next Hop</th>
<th>Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Net A</td>
<td>Router 1</td>
<td>INT 7</td>
</tr>
<tr>
<td>Net B</td>
<td>Direct</td>
<td>INT 4</td>
</tr>
<tr>
<td>Net C, Host 3</td>
<td>Router 2</td>
<td>INT 3</td>
</tr>
<tr>
<td>Net C</td>
<td>Router 1</td>
<td>INT 7</td>
</tr>
</tbody>
</table>

A destination is usually a network. May also be a host, or a “gateway of last resort” (default).

The next hop is either a directly connected network or a router on a directly connected network.

A physical interface

---

### IP Forwarding Process

1. Remove a packet from an input queue
2. Check for sanity, decrement TTL field
3. Match packet’s destination to a table entry
4. Place packet on correct output queue

If queues get full, just drop packets!
IPv4 Addresses are 32 Bit Values

IPv6 addresses have 128 bits

IP Addresses come in two parts

Where is this dividing line? Well, that depends....
Classful Addresses

Class A
0nnnnnnn hhhhhhhh hhhhhhhh hhhhhhhh

Class B
10nnnnnn nnnnnnnn hhhhhhhh hhhhhhhh

Class C
110nnnnn nnnnnnnn nnnnnnnn hhhhhhh

n = network address bit  h = host identifier bit

Leads to a rigid, flat, inefficient use of address space …

RFC 1519: Classless Inter-Domain Routing (CIDR)

Use two 32 bit numbers to represent a network.
Network number = IP address + Mask

IP Address : 12.4.0.0   IP Mask: 255.254.0.0

Address 00001100 00000100 00000000 00000000

Mask 11111111 11111110 00000000 00000000

Network Prefix for hosts

Usually written as 12.4.0.0/15
**Which IP Addresses are Covered by a Prefix?**

12.5.9.16 is covered by prefix 12.4.0.0/15

<table>
<thead>
<tr>
<th>12.5.9.16</th>
<th>00001100 00000101 00001011 00010000</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.4.0.0/15</td>
<td>00001100 00000100 00000000 00000000</td>
</tr>
<tr>
<td>12.7.9.16</td>
<td>00001100 00000111 00001001 00010000</td>
</tr>
</tbody>
</table>

12.7.9.16 is not covered by prefix 12.4.0.0/15

**CIDR allows Hierarchy in Addressing**

12.0.0.0/8

- 12.0.0.0/16
- 12.1.0.0/16
- 12.2.0.0/16
- 12.3.0.0/16

- 12.253.0.0/16
- 12.254.0.0/16

- 12.253.0.0/19
- 12.253.32.0/19
- 12.253.64.0/19
- 12.253.96.0/19
- 12.253.128.0/19
- 12.253.160.0/19
- 12.253.192.0/19
Classless Forwarding

Destination = 12.5.9.16
-------------------------------
payload

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Next Hop</th>
<th>Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0.0.0/0</td>
<td>10.14.11.33</td>
<td>ATM 5/0/9</td>
</tr>
<tr>
<td>12.0.0.0/8</td>
<td>10.14.22.19</td>
<td>ATM 5/0/8</td>
</tr>
<tr>
<td>12.4.0.0/15</td>
<td>10.13.77</td>
<td>Ethernet 0/1/3</td>
</tr>
<tr>
<td>12.5.8.0/23</td>
<td>attached</td>
<td>Serial 1/0/7</td>
</tr>
</tbody>
</table>

IP Forwarding Table

How Are Forwarding Tables Populated to implement Routing?

Statically
Administrator manually configures forwarding table entries
+ More control
+ Not restricted to destination-based forwarding
- Doesn’t scale
- Slow to adapt to network failures

Dynamically
Routers exchange network reachability information using ROUTING PROTOCOLS.
Routers use this to compute best routes
+ Can rapidly adapt to changes in network topology
+ Can be made to scale well
- Complex distributed algorithms
- Consume CPU, Bandwidth, Memory
- Debugging can be difficult
- Current protocols are destination-based

In practice: a mix of these. Static routing mostly at the “edge”
Routing vs. Forwarding

Routing: establish end-to-end paths
Forwarding: determine next hop

Default to upstream router

Forwarding always works
Routing can be badly broken

Happy Packets: The Internet Does Not Exist Only to Populated Routing Tables

RIP Process
BGP Process
OSPF Process
Forwarding Table Manager

RIP Domain
OSPF Domain

RIP Routing tables
BGP Routing tables
OSPF Routing tables
Forwarding Table

20
Before We Go Any Further

IP ROUTING PROTOCOLS DO NOT DYNAMICALLY ROUTE AROUND NETWORK CONGESTION

- IP traffic can be very bursty
- Dynamic adjustments in routing typically operate more slowly than fluctuations in traffic load
- Dynamically adapting routing to account for traffic load can lead to wild, unstable oscillations of routing system

Next Lecture: Shortest Path Routing

This is what IS-IS, OSPF, and RIP do, more or less.