IPv6

- **Initial motivation:** 32-bit address space soon to be completely allocated.

- **Additional motivation:**
  - header format helps speed processing/forwarding
  - header changes to facilitate QoS

**IPv6 datagram format:**
- fixed-length 40 byte header
- no fragmentation allowed

**IPv6 Header**

*Priority:* identify priority among datagrams in flow; help distinguish packets that can be flow controlled and those that cannot

*Flow Label:* experimental - identify datagrams in same "flow" and treat them specially. Potential use?

*Next header:* identify upper layer protocol for data or optional IP header
IPv6 Addressing

- BTW, how big is $2^{128}$??
- There are $1,000,000,000,000,000,000,000,000,000$ stars in the observable Universe.
  - $2^{128}$ provides more than 300 billion addresses for every such star!

Other Changes from IPv4

- **Checksum**: removed entirely to reduce processing time at each hop
- **Options**: allowed, but outside of header, indicated by “Next Header” field
- **ICMPv6**: new version of ICMP
  - additional message types, e.g. “Packet Too Big”
  - multicast group management functions
Header Extensions

Example optional headers

- Hop-by-Hop Option - Special options that require hop-by-hop processing
- Routing - Extended routing, like IPv4 loose source route
- Fragmentation - Fragmentation and reassembly (processed at destination)
- Authentication - Integrity and authentication, security
- Encapsulation - Confidentiality
- Destination Options - Optional information to be examined by the destination node
Transition From IPv4 To IPv6

- Not all routers are upgraded simultaneously
  - How will the network operate with mixed IPv4 and IPv6 routers?

- **Tunneling**: IPv6 carried as payload in IPv4 datagram among IPv4 routers. How?

Tunneling

**Logical view:**

- A (IPv6)
- B (IPv6)
- C (IPv6)
- D (IPv6)
- E (IPv6)
- F (IPv6)

**Physical view:**

- A (IPv6)
- B (IPv6)
- C (IPv4)
- D (IPv4)
- E (IPv6)
- F (IPv6)
**Tunneling**

**Logical view:**

A IPv6  B IPv6 | tunnel |
E IPv6  F IPv6

**Physical view:**

A IPv6 | B IPv6 | C IPv4 | D IPv4 | E IPv6 | F IPv6

Flow: X
Src: A
Dest: E
data

Flow: X
Src: A
Dest: F
data

Flow: X
Src: A
Dest: F
data

A-to-B: IPv6

A-to-B: IPv6 inside IPv4

B-to-C: IPv6 inside IPv4

E-to-F: IPv6

**IPv6 Status**

- CIDR introduced in 1993 to stave off exhaustion of IPv4 addresses
- CIDR (plus NAT) has been very effective, but final phase of allocation began in 2011.
- Still, in 2013 only 16% of networks supported IPv6
- In 2014, IPv4 still carried more than 99% of all Internet traffic
- But, by end of 2015, 10% of users reached Google via IPv6...
Chapter 4: Network Layer

- Introduction (forwarding and routing)
- Review of queueing theory
- Routing algorithms
  - Link state, Distance Vector
- Router design and operation
- IP: Internet Protocol
  - IPv4 (datagram format, addressing, ICMP, NAT)
  - IPv6
- Routing in the Internet
  - Autonomous Systems
  - Routing protocols (RIP, OSPF, BGP)

Hierarchical Routing

Our routing study thus far has focused on individual networks
- network is “flat”
- all routers are identical

But, the Internet is a network of networks
- Routing is hierarchical. Why?
- Individual networks have autonomy.
Hierarchical Routing

- aggregate routers into regions, “autonomous systems” (AS)
- routers in same AS run same routing protocol
  - “intra-AS” routing protocol
  - routers in different AS can run different intra-AS routing protocol

Gateway router
- Direct link to router in another AS

Interconnected Autonomous Systems

- forwarding table configured by both intra- and inter-AS routing algorithm
  - intra-AS sets entries for internal destinations
  - inter-AS & intra-AS sets entries for external destinations
**Intra-AS Routing**

- also known as Interior Gateway Protocols (IGP)
- most common Intra-AS routing protocols:
  - RIP: Routing Information Protocol
  - OSPF: Open Shortest Path First
  - IGRP: Interior Gateway Routing Protocol (Cisco proprietary)

**RIP (Routing Information Protocol)**

- distance vector algorithm
- included in BSD-UNIX Distribution in 1982
- distance metric: # of hops (max = 15 hops)

<table>
<thead>
<tr>
<th>destination</th>
<th>hops</th>
</tr>
</thead>
<tbody>
<tr>
<td>u</td>
<td>1</td>
</tr>
<tr>
<td>v</td>
<td>2</td>
</tr>
<tr>
<td>w</td>
<td>2</td>
</tr>
<tr>
<td>x</td>
<td>3</td>
</tr>
<tr>
<td>y</td>
<td>3</td>
</tr>
<tr>
<td>z</td>
<td>2</td>
</tr>
</tbody>
</table>

From router A to subnets:
RIP vectors

- **distance vectors**: exchanged among neighbors every 30 sec via Response Message (also called advertisement)
- each advertisement: list of up to 25 destination subnets within AS

### RIP: Example

<table>
<thead>
<tr>
<th>Dest</th>
<th>Next hops</th>
</tr>
</thead>
<tbody>
<tr>
<td>w</td>
<td>- 1</td>
</tr>
<tr>
<td>x</td>
<td>- 1</td>
</tr>
<tr>
<td>z</td>
<td>c 4</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Advertisement from A to D

![Routing/Forwarding table in D](image)

<table>
<thead>
<tr>
<th>Destination Network</th>
<th>Next Router</th>
<th>Num. of hops to dest.</th>
</tr>
</thead>
<tbody>
<tr>
<td>w</td>
<td>A</td>
<td>2</td>
</tr>
<tr>
<td>y</td>
<td>B</td>
<td>2</td>
</tr>
<tr>
<td>z</td>
<td>X A</td>
<td>X 5</td>
</tr>
<tr>
<td>x</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
RIP: Link Failure and Recovery

If no advertisement heard after 180 sec --
neighbor/link declared dead
- routes via neighbor invalidated
- new advertisements sent to neighbors
- neighbors in turn send out new advertisements (if tables changed)
- link failure info quickly (?) propagates to entire net
- poison reverse used to prevent ping-pong loops
  (infinite distance = 16 hops = “you can’t get there from here”)

RIP Table processing

- RIP routing tables managed by application-level process called route-d (daemon)
- advertisements sent in UDP packets, periodically repeated

![Diagram of network layer components]
**OSPF (Open Shortest Path First)**

- uses Link State algorithm
  - LS packet dissemination
  - topology map at each node
  - route computation using Dijkstra’s algorithm

- OSPF advertisement carries one entry per neighbor router
- advertisements disseminated to entire AS
  - via flooding
  - carried in OSPF messages directly over IP, rather than TCP or UDP (how?)

**Hierarchical OSPF**

- **two-level hierarchy**: local area, backbone.
  - Link-state advertisements only in area
  - each node has detailed area topology; only know direction (shortest path) to nets in other areas.
- **area border routers**: “summarize” distances to nets in own area, advertise to other Area Border routers.
- **backbone routers**: run OSPF routing limited to backbone.
- **boundary routers**: connect to other AS’ s.
Hierarchical OSPF

Internet inter-AS routing: BGP

- **BGP (Border Gateway Protocol):** the de facto standard

- BGP provides each AS a means to:
  1. Obtain subnet reachability information from neighboring ASs.
  2. Propagate reachability information to all AS-internal routers.
  3. Determine “good” routes to subnets based on reachability information and policy.

- allows subnet to advertise its existence to rest of Internet: “I am here”

- Some resemblance to distance-vector, but based on paths not simply distance
**BGP basics**

- Pairs of routers (BGP peers) exchange routing info over semi-permanent TCP connections: **BGP sessions**
  - BGP sessions need not correspond to physical links.
- When AS2 advertises a **CIDRized prefix** to AS1:
  - AS2 promises it will forward datagrams towards that prefix.
  - AS2 can aggregate prefixes in its advertisement

**Distributing reachability info**

- Using eBGP session between 3a and 1c, AS3 sends prefix reachability info to AS1.
  - 1c can then use iBGP to distribute new prefix info to all routers in AS1
  - 1b can then **re-advertise** new reachability info to AS2 over 1b-to-2a eBGP session
- When router learns of new prefix, it creates entry for prefix in its forwarding table.
Path attributes & BGP routes

- advertised prefix includes BGP attributes.
  - prefix + attributes = “route”

- two important attributes:
  - AS-PATH: contains ASs through which prefix advertisement has passed: e.g., AS 67, AS 17
  - NEXT-HOP: indicates specific internal-AS router that begins the AS-PATH. (enables configuration of routing table entries)

BGP routing policy

- A, B, C are provider networks
- X, W, Y are customers (of provider networks)
- X is dual-homed: attached to two networks
  - X does not want to route from B via X to C
  - .. so X will not advertise to B a route to C
BGP routing policy (cont.)

- A advertises path AW to B
- B advertises path BAW to X
- Should B advertise path BAW to C?
  - No, B gets no “revenue” for routing CBAW since neither W nor C are B’s customers
  - B wants to force C to route to W via A
  - B wants to route only to/from its customers.

Chapter 4 Summary

- Introduction (forwarding and routing)
- Review of queueing theory
- Routing algorithms
  - Link state, Distance Vector
- Router design and operation
- IP: Internet Protocol
  - IPv4 (datagram format, addressing, ICMP, NAT)
  - IPv6
- Routing in the Internet
  - Autonomous Systems
  - Routing protocols (RIP, OSPF, BGP)