Network Layer – Forwarding

Kurose & Ross, Chapter 4 (5th ed.)

Many slides adapted from:
J. Kurose & K. Ross
Computer Networking: A Top Down Approach (5th ed.)
Addison-Wesley, April 2009.
Copyright 1996-2010, J.F Kurose and K.W. Ross, All Rights Reserved.

Network layer
- transport segment from sending to receiving host
- on sending side, encapsulates segments into datagrams
- on receiving side, delivers segments to transport layer
- network layer protocols in every host, router
- router examines header fields in all IP datagrams passing through it

Two Key Network-Layer Functions
- forwarding: move packets from router’s input to appropriate router output
- routing: determine route taken by packets from source to dest.
  - routing algorithms

Interplay between routing and forwarding

Network service model

Q: What service model for “channel” transporting datagrams from sender to receiver?

example services for individual datagrams:
- guaranteed delivery
- guaranteed delivery with less than 40 msec delay

example services for a flow of datagrams:
- in-order datagram delivery
- guaranteed minimum bandwidth to flow
- restrictions on changes in inter-packet spacing
Network layer service models:

<table>
<thead>
<tr>
<th>Network Architecture</th>
<th>Service Model</th>
<th>Bandwidth</th>
<th>Loss</th>
<th>Order</th>
<th>Timing</th>
<th>Congestion feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internet</td>
<td>best effort</td>
<td>none</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no (interred via loss)</td>
</tr>
<tr>
<td>ATM</td>
<td>CBR</td>
<td>constant rate</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no congestion</td>
</tr>
<tr>
<td>ATM</td>
<td>VBR</td>
<td>guaranteed rate</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no congestion</td>
</tr>
<tr>
<td>ATM</td>
<td>ABR</td>
<td>guaranteed minimum</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>ATM</td>
<td>UBR</td>
<td>none</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td></td>
</tr>
</tbody>
</table>

**IP provides best-effort service**

Virtual circuits

*source-to-dest path behaves much like telephone circuit*  
- performance-wise  
- network actions along source-to-dest path

- call setup, teardown for each call before data can flow  
- each packet carries VC identifier (not destination host address)  
- every router on source-dest path maintains "state" for each passing connection  
- link, router resources (bandwidth, buffers) may be allocated to VC (dedicated resources = predictable service)

VC implementation

A VC consists of:
1. path from source to destination
2. VC numbers, one number for each link along path
3. entries in forwarding tables in routers along path
- packet belonging to VC carries VC number (rather than dest address)
- VC number can be changed on each link.
  - New VC number comes from forwarding table

VC Forwarding table

<table>
<thead>
<tr>
<th>Incoming interface</th>
<th>Incoming VC #</th>
<th>Outgoing interface</th>
<th>Outgoing VC #</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12</td>
<td>3</td>
<td>22</td>
</tr>
<tr>
<td>2</td>
<td>63</td>
<td>1</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>2</td>
<td>17</td>
</tr>
<tr>
<td>1</td>
<td>97</td>
<td>3</td>
<td>87</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

**Routers maintain connection state information!**

Virtual circuits: signaling protocols

- used to setup, maintain, teardown VC
- used in ATM, frame-relay, X.25
- not used in today's Internet

Datagram networks

- no call setup at network layer
- routers: no state about end-to-end connections
  - no network-level concept of "connection"
- packets forwarded using destination host address
  - packets between same source-dest pair may take different paths
Datagram or VC network: why?

**Internet (datagram)**
- data exchange among computers
- “elastic” service, no strict timing req.
- “smart” end systems (computers)
- can adapt, perform control, error recovery
- simple inside network, complexity at “edge”
- many link types
- different characteristics
- uniform service difficult

**ATM (VC)**
- evolved from telephony
- human conversation:
  - strict timing, reliability requirements
  - need for guaranteed service
- “dumb” end systems
- telephones
- complexity inside network

IP datagram format

- IP protocol version number
- header length (bytes)
- “type” of data
- max number of hops (decremented at each router)
- upper layer protocol to deliver payload
- how much overhead with TCP?
  - 20 bytes of TCP
  - 20 bytes of IP
  - 40 bytes + app layer overhead

The Internet Network layer

Host, router network layer functions:

- Routing protocols
- RIP, OSPF, BGP
- IP protocol addressing conventions
- datagram format
- packet handling conventions
- ICMP protocol error reporting
- router “signaling”

Dynamic Host Configuration Protocol

- How can a newly connected host get an IP address?
  - (And other useful information: its network mask, an IP address of DNS server etc.)
  - Manual configuration by a system administrator
  - DHCP
    - Plug and play – host obtains this information automatically
    - Defined in RFCs 2131 & 2132.
    - Sent over UDP. Server’s port: 67, client’s port: 68.
    - As usual, we’ll give an overview and not go into all the technical details.

DHCP client-server scenario

- Arriving client sends a DHCP discover message
  - Src IP: 0.0.0.0 (“this host”)
  - Dst IP: 255.255.255.255 (“broadcast”)
  - Transaction ID: some number x.
  - Message is broadcast to all nodes on the subnet
DHCP client-server scenario 2

- Server replies with a DHCP offer message
  - Src IP: server’s IP
  - Dst IP: 255.255.255.255 ("broadcast")
    - Why? The client still can’t receive direct messages – it has no IP address.
  - Transaction ID: x
  - Message contains the proposed IP address.

DHCP client-server scenario 3

- Why aren’t the previous messages enough?
  - The network might contain more than one DHCP server; a client may receive multiple DHCP offers.
- Client sends a DHCP request message
  - Src IP: 0.0.0.0 ("this host")
  - Dst IP: 255.255.255.255 ("broadcast")
  - Why? To allow other servers know their offer was declined
  - DHCP Server ID: the IP of the server whose offer the client wish to accept
  - Transaction ID: some number y

DHCP client-server scenario 4

- Server responds with a DHCP ACK message
  - Src IP: Server’s IP
  - Dst IP: 255.255.255.255 ("broadcast")
    - The client still doesn’t have an IP address
  - Transaction ID: y

Lease times

- DHCP servers assigns a lease-time for each IP address allocation
  - A client may renew its allocation when it is about to expire
  - A client may relinquish its allocation

Internet Control Message Protocol

- ICMP – defined in RFC 792.
- Carried directly over IP
  - No transport protocol used
- Used by hosts and routers to communicate network layer information – usually report errors.
ICMP (cont.)

- For instance: “Destination unreachable”:
  - A router was unable to find a path to host B specified in host A’s request
  - The router sends an ICMP message (type 3) to A to indicate the error
- Another common use: ping
  - Host A sends an ICMP message (type 1) to host B
  - Host B sees this, and replies to A with another type 1 ICMP message.

Traceroute and ICMP

- Source sends series of UDP segments to dest
  - first has TTL = 1
  - second has TTL = 2, etc.
  - unlikely port number
- When nth datagram arrives to nth router:
  - router discards datagram and sends to source an ICMP message (type 11, code 0)
  - ICMP message includes name of router & IP address
- When ICMP message arrives, source calculates RTT
- Traceroute does this 3 times

Stopping criterion
- UDP segment eventually arrives at destination host
- destination returns ICMP “port unreachable” packet (type 3, code 3)
- when source gets this ICMP, stops.

IP Addressing: introduction

- **IP address**: 32-bit identifier for host or router interface
- **interface**: connection between host/router and physical link
  - router’s typically have multiple interfaces
  - a host has typically a single interface
- IP addresses associated with interface, not host, or router

What is a network in IP view?

**IP network terminology:**

- **a Subnet is**:
  - a set of devices that can physically reach each other without intervening router(s)
  - e.g. a LAN
- **a Network is**:
  - a subnet, or:
  - the union of several subnets that are interconnected by links
  - three subnets (LANs) 223.11.*, 223.12.*, 223.13.*, together they form a larger network with prefix 223.1 (16 bits) (OR MORE bits?)

Qn: What is the router’s IP address in the drawing we see?
IP Addresses

Given the notion of “network”, let’s re-examine IP addresses:

Classful addressing:

- "Classful" addressing:
  - Does not need mask or /K indicator

<table>
<thead>
<tr>
<th>Class</th>
<th>Notation</th>
<th>Subnet Part</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>10</td>
<td>127.255.255.255</td>
</tr>
<tr>
<td>B</td>
<td>110</td>
<td>128.0.0 to 191.255.255.255</td>
</tr>
<tr>
<td>C</td>
<td>110</td>
<td>192.0.0 to 233.255.255.255</td>
</tr>
<tr>
<td>D</td>
<td>1100</td>
<td>224.0.0.0 to 239.255.255.255</td>
</tr>
</tbody>
</table>

(*) This range used as multicast also in CIDR method.

IP addressing: CIDR

- Classful addressing:
  - Inefficient use of address space, address space exhaustion
  - E.g., Class B not allocated enough addresses for 65K hosts, even if only 2K hosts in that network

CIDR: Classless Inter-Domain Routing:

- Network portion of address of arbitrary length
- Address format: a.b.c.d/x, where x is # bits in network portion of address
- Requires inclusion of mask or /K in routing table

Hierarchical addressing: route aggregation

Hierarchical addressing allows efficient advertisement of routing information:

ISP’s R-Us has a more specific route to Organization 1.

Hierarchical addressing: more specific routes

IP addresses: how to get one?

Q: How does network get subnet part of IP addr?
A: ISP gets allocated portion of its provider ISP’s address space

ISP’s block: 11010000 00010111 00010000 00000000 200.23.16.0/23

Organization 0
  - 11010000 00010111 00010000 00000000 200.23.16.0/23

Organization 1
  - 11010000 00010111 00010010 00000000 200.23.16.0/23

Organization 2
  - 11010000 00010111 00010010 00000000 200.23.20.0/23

Organization 3
  - 11010000 00010111 00010010 00000000 200.23.30.0/23

Organization 4
  - 11010000 00010111 00010010 00000000 200.23.30.0/23

Organization 5
  - 11010000 00010111 00010010 00000000 200.23.30.0/23

Organization 6
  - 11010000 00010111 00010010 00000000 200.23.30.0/23

Organization 7
  - 11010000 00010111 00011110 00000000 200.23.30.0/23

IP addressing: the last word...

Q: How does an ISP get block of addresses?
A: ICANN: Internet Corporation for Assigned Names and Numbers
- Allocates addresses
- Manages DNS
- Assigns domain names, resolves disputes