Lecture 10 Overview

• IP Protocol

• Fragmentation
  • MTU and IP ID
  • Path MTU discovery

• IP Addresses
  • Class-based addressing

• Subnetting
  • Classless addressing
Routers

- A router is a store-and-forward device
  - Routers are connected to multiple networks
  - On each network, looks just like another host
  - A lot like a switch, but supports multiple datalink layers and makes decisions at the network layer

- Must be explicitly addressed by incoming frames (L2)
  - Not at all like a switch, which is transparent
  - Removes link-layer header, parses IP header (L3)

- Looks up next hop, forwards on appropriate network
  - Each router need only get one step closer to destination
Impose few demands on network
- Make few assumptions about what network can do
- No QoS, no reliability, no ordering, no large packets
- No persistent state about communications; no connections

Manage heterogeneity at hosts (not in network)
- Adapt to underlying network heterogeneity
- Re-order packets, detect errors, retransmit lost messages…
- Persistent network state only kept in hosts (fate-sharing)

Service model: best effort, a.k.a. send and pray
# IP Packet Header

The IP packet header comprises the following fields:

- **Version (ver)**: 4 bits, indicating the IP version (IPv4 is 0).
- **Header Length (HL)**: 4 bits, specifying the length of the header in 32-bit words (max 20 bytes).
- **Type of Service (TOS)**: 8 bits, indicating the service type.
- **Total Length (length)**: 16 bits, indicating the total length of the packet in bytes.
- **Identification (identification)**: 16 bits, used to identify fragments of a single packet.
- **Flag (offset)**: 13 bits, composed of two flags (MF and DF) and an offset field.
- **Time to Live (TTL)**: 8 bits, indicating the number of hops a packet can take before it is discarded.
- **Protocol (protocol)**: 8 bits, specifying the protocol used in the upper layer.
- **Header Checksum (header checksum)**: 16 bits, a checksum for the header.
- **Source Address (source address)**: 32 bits, the source IP address.
- **Destination Address (destination address)**: 32 bits, the destination IP address.
- **Options (if any)**
- **Data (if any)**
**Version field**

- Which version of IP is this?
  - Plan for change
  - Very important to be at the beginning (why?)

- Current versions
  - 4: most of Internet today
  - 6: new protocol with larger addresses
  - What happened to 5?
    Standards body politics.
How big is IP header?
- Counted in 32-bit words
- Variable length header
  - Options
- Engineering consequences of variable length...
  - Hardware can’t always assume fixed length

Most IP packet headers are 20 bytes long
How should this packet be treated?
- Care/don’t care for delay, throughput, reliability, cost
- How to interpret, how to apply on underlying net?
- Largely unused until 2000 (hijacked for new purposes, ECN & Diffserv)
**Length**

- How long is whole packet in bytes?
  - Includes header
  - Limits total packet to 64K
  - Redundant?

<table>
<thead>
<tr>
<th>ver</th>
<th>HL</th>
<th>TOS</th>
<th>offset</th>
<th>length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

20 bytes
TTL (Time-to-Live)

- How many more routers can this packet pass through?
  - Designed to limit packet from looping forever
- Each router decrements TTL field
- If TTL is 0 then router discards packet
Protocol

- Which transport protocol is the data using?
  - i.e. how should a host interpret the data

- TCP = 6
- UDP = 17
IP Checksum

- Header contains simple checksum
  - Validates content of header only

- Recalculated at each hop
  - Routers need to update TTL
  - Hence straightforward to modify

- Ensures *correct* destination receives packet
So what does IP do?

- Addressing
- Fragmentation
  - E.g. FDDI’s maximum packet is 4500 bytes while Ethernet is 1500 bytes, how to manage this?
- Some error detection
- Routers only forward packets to next hop
  - They do not:
    » Detect packet loss, packet duplication
    » Reassemble or retransmit packets

Today we’ll talk about fragmentation
Fragmentation

- Different networks may have different maximum frame sizes
  - Maximum Transmission Unit (MTUs)
  - Ethernet 1500, WiFi 2,346
- Router breaks up single IP packet into two or more smaller IP packets
  - Each fragment is labeled so it can be correctly reassembled
  - End host reassembles them into original packet
IP ID and Bitflags

- Source inserts unique value in identification field
  - Also known as the IPID
  - If packet is fragmented, the router copies this value into any fragments

- Offset field indicates position of current fragment (in bytes/8)
  - Zero for non-fragmented packet

- Bitflags provide additional information
  - More Fragments bit helps identify last fragment
  - Don’t Fragment bit prohibits (further) fragmentation
  - Note recursive fragmentation easily supported—just requires care with More Fragments bit
Fragmentation Example

One large datagram becomes several smaller datagrams

<table>
<thead>
<tr>
<th>length=4000</th>
<th>ID=x</th>
<th>MF=0</th>
<th>offset=0</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>length=1500</th>
<th>ID=x</th>
<th>MF=1</th>
<th>offset=0</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>length=1500</th>
<th>ID=x</th>
<th>MF=1</th>
<th>offset=1480</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>length=1040</th>
<th>ID=x</th>
<th>MF=0</th>
<th>offset=2960</th>
</tr>
</thead>
</table>

(Offset actually encoded as bytes/8)

CSE 123 – Lecture 10: IP Protocol
Costs of Fragmentation

- Interplay between fragmentation and retransmission
  - A single lost fragment may trigger retransmission
  - Any retransmission will be of entire packet (why?)

- Packet must be completely reassembled before it can be consumed on the receiving host
  - Takes up buffer space in the mean time
  - When can it be garbage collected?

- Why not reassemble at each router?
Discovering MTU on path

- Path MTU is the smallest MTU along path
  - Packets less than this size don’t get fragmented

- Fragmentation is a burden for routers
  - We already avoid reassembling at routers
  - Avoid fragmentation too by having hosts learn path MTUs

- Hosts send packets, routers return error if too large
  - Hosts can set “don’t fragment” flag
  - Hosts discover limits, can size packets at source
    » ICMP protocol: special IP packet format for sending error msgs
  - Reassembly at destination as before
For Next Time

- Read 4.1.3
- Homework 2 out tonight by 10pm
- Good time to start reading spec for Project 2!