Lecture 10 Overview

• IP Protocol

• Fragmentation
  • MTU and IP ID
  • Path MTU discovery
Internetworking

- Cerf & Kahn74, “A Protocol for Packet Network Intercommunication”
  - Foundation for the modern Internet

- Routers forward packets from source to destination
  - May cross many separate networks along the way

- All packets use a common Internet Protocol
  - Any underlying data link protocol
  - Any higher layer transport protocol
  - This is the “thin waist”
IP Networking

Router

WiFi

data packet

WiFi  IP  TCP  HTTP

Ethernet

data packet

ETH  IP  TCP  HTTP

CSE 123 – Lecture 10: IP Protocol
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
IP Philosophy

- 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 the 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 (version 4) Packet Header

<table>
<thead>
<tr>
<th>0</th>
<th>15</th>
<th>16</th>
<th>31</th>
</tr>
</thead>
<tbody>
<tr>
<td>ver</td>
<td>HL</td>
<td>TOS</td>
<td>length</td>
</tr>
<tr>
<td>identification</td>
<td></td>
<td>R M</td>
<td>D</td>
</tr>
<tr>
<td>TTL</td>
<td>protocol</td>
<td>header checksum</td>
<td></td>
</tr>
<tr>
<td>source address</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>destination address</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>options (if any)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>data (if any)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Description**: An IP packet header consists of several fields, each occupying a specific position in the header. The header is 20 bytes long. Each field has a specific purpose:

- **version (ver)**: Specifies the version of IP protocol, currently 4.
- **header length (length)**: Specifies the length of the header in 32-bit words.
- **type of service (TOS)**: Specifies the type of service for the packet.
- **total length**: Indicates the total length of the packet in bytes, including the header.
- **identification**: A field used for packet identification and fragmentation.
- **flags**: Indicates if the packet is fragmentable and the fragment offset.
- **fragment offset**: Specifies the offset of the fragment within the original packet.
- **time to live (TTL)**: Specifies the maximum number of hops a packet can travel before being discarded.
- **protocol**: Specifies the protocol for the packet content.
- **source address**: Indicates the source IP address.
- **destination address**: Indicates the destination IP address.
- **options (if any)**: Contains any additional options that may be present in the packet.
- **data (if any)**: Contains the actual data payload of the packet.
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, widely adopted today
  - What happened to 5? Standards body politics.
Header length

- 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
Type-of-Service

- 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>length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>20 bytes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>identification</th>
<th>offset</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTL</td>
<td>protocol</td>
</tr>
<tr>
<td></td>
<td>header checksum</td>
</tr>
<tr>
<td>source address</td>
<td>destination address</td>
</tr>
<tr>
<td>options (if any)</td>
<td>data (if any)</td>
</tr>
</tbody>
</table>
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
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 to/from hosts and networks
- Fragmentation (handling different link layer protocols)
  - 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

(Offset actually encoded as bytes/8)
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 receiver in the meantime
  - When can it be garbage collected?

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

- Path Maximum Transmission Unit (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 today

• Good time to start reading spec for Project 2!