Lecture 10: Addressing

CSE 123: Computer Networks
Aaron Schulman
Lecture 10 Overview

• Finish fragmentation
  • MTU and IP ID
  • Path MTU discovery

• IP Addresses
  • Class-based addressing

• Subnetting
  ◆ Classless addressing
Fragmentation

- Different networks may have different maximum frame sizes
  - Maximum Transmission Unit (MTUs)
  - Ethernet 1.5K, FDDI 4.5K
- 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)

<table>
<thead>
<tr>
<th>length</th>
<th>ID</th>
<th>MF</th>
<th>offset</th>
</tr>
</thead>
<tbody>
<tr>
<td>4000</td>
<td>x</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1500</td>
<td>x</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1500</td>
<td>x</td>
<td>1</td>
<td>1480</td>
</tr>
<tr>
<td>1040</td>
<td>x</td>
<td>0</td>
<td>2960</td>
</tr>
</tbody>
</table>

CSE 123 – Lecture 9: Internetworking
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?
Path MTU Discovery

- 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
Aside: ICMP

- What happens when things go wrong?
  - Need a way to test/debug a large, widely distributed system

- ICMP = Internet Control Message Protocol (RFC792)
  - Companion to IP – required functionality

- Used for error and information reporting:
  - Errors that occur during IP forwarding
  - Queries about the status of the network
ICMP Error Message Generation

Error during forwarding!

source

IP packet

dest

ICMP

IP packet
Common ICMP Messages

- **Destination unreachable**
  - “Destination” can be host, network, port, or protocol
- **Redirect**
  - To shortcut circuitous routing
- **TTL Expired**
  - Used by the “traceroute” program
    - traceroute traces packet routes through Internet
- **Echo request/reply**
  - Used by the “ping” program
    - ping just tests for host liveness
- **ICMP messages include portion of IP packet that triggered the error (if applicable) in their payload**
ICMP Restrictions

- The generation of error messages is limited to avoid cascades … error causes error that causes error…

- Don’t generate ICMP error in response to:
  - An ICMP error
  - Broadcast/multicast messages (link or IP level)
  - IP header that is corrupt or has bogus source address
  - Fragments, except the first

- ICMP messages are often rate-limited too
  - Don’t waste valuable bandwidth sending tons of ICMP messages
Addressing Considerations

- Fixed length or variable length addresses?

- Issues:
  - Flexibility
  - Processing costs
  - Header size

- Engineering choice: IP uses fixed length addresses
IP Addresses

- 32-bits in an IPv4 address
  - Dotted decimal format a.b.c.d
  - Each represent 8 bits of address

- Hierarchical: Network part and host part
  - E.g. IP address 128.54.70.238
    - 128.54 refers to the UCSD campus network
    - 70.238 refers to the host ieng6.ucsd.edu

- Which part is network vs. host?
Class-based Addressing

- Most significant bits determines “class” of address
  - Class A: 0 Network Host
    - 127 nets, 16M hosts
  - Class B: 1 0 Network Host
    - 16K nets, 64K hosts
  - Class C: 1 1 0 Network Host
    - 2M nets, 254 hosts

- Special addresses
  - Class D (1110) for multicast, Class E (1111) experimental
  - 127.0.0.1: local host (a.k.a. the loopback address)
  - Host bits all set to 0: network address
  - Host bits all set to 1: broadcast address

CSE 123 – Lecture 10: Addressing
IP Forwarding Tables

- Router needs to know where to forward a packet
- Forwarding table contains:
  - List of network names and next hop routers
  - Local networks have entries specifying which interface
    » Link-local hosts can be delivered with Layer-2 forwarding
- E.g. cseweb.ucsd.edu address is 132.239.8.30
  - Class B address – class + network is 132.239
  - Lookup 132.239 in forwarding table
  - Prefix – part of address that really matters for routing
Subnetting

- Individual networks may be composed of several LANs
  - Only want traffic destined to local hosts on physical network
  - Routers need a way to know which hosts on which LAN

- Networks can be arbitrarily decomposed into subnets
  - Each subnet is simply a prefix of the host address portion
  - Subnet prefix can be of any length, specified with netmask
Subnet Addresses

- Every (sub)network has an address and a **netmask**
  - Netmask tells which bits of the network address is important
  - Convention suggests it be a proper prefix

- Netmask written as an all-ones IP address
  - E.g., Class B netmask is 255.255.0.0
  - Sometimes expressed in terms of number of 1s, e.g., /16

- Need to size subnet appropriately for each LAN
  - Only have remaining bits to specify host addresses
IP Address Problem (1991)

- Address space depletion
  - In danger of running out of classes A and B

- Why?
  - Class C too small for most organizations (only ~250 addresses)
  - Very few class A – very careful about giving them out (who has 16M hosts anyway?)
  - Class B – greatest problem
Classless Inter-Domain Routing (1993)

- Networks described by variable-length prefix and length
- Allows arbitrary allocation between network and host address

E.g. 10.95.1.2 contained within 10.0.0.0/8:
  - 10.0.0.0 is network and remainder (95.1.2) is host

- Pro: Finer grained allocation; aggregation
- Con: More expensive lookup: longest prefix match

CIDR

Network  Host

Prefix

Mask = # significant bits representing prefix

e.g. 10.95.1.2 contained within 10.0.0.0/8:
  - 10.0.0.0 is network and remainder (95.1.2) is host
For Next Time

- Read 4.1.3
- Homework 2 out tonight by 10pm
- Finish up Project 1 only around 1 week left!