Lecture 8 (Feb 12, 2004)

Outline
ICMP
RARP
DHCP
NAT

Internet Control Message Protocol (ICMP)
- Echo (ping)
- Redirect (from router to source host)
- Destination unreachable (protocol, port, or host)
- TTL exceeded (so datagrams don’t cycle forever)
- Checksum failed
- Reassembly failed
- Cannot fragment

ICMP
- Uses IP but is a separate protocol in the network layer

ICMP HEADER
- IP HEADER
  - PROTOCOL = 1
  - TYPE, CODE, CHECKSUM
- IP DATA
  - IDENTIFIER
  - SEQUENCE #
  - REMAINDER OF ICMP MESSAGE (FORMAT IS TYPE SPECIFIC)

Echo and Echo Reply
- TYPE: 8 = ECHO, 0 = ECHO REPLY CODE; CODE = 0
- IDENTIFIER
  - An identifier to aid in matching echoes and replies
- SEQUENCE #
  - Same use as for IDENTIFIER
- UNIX “ping” uses echo/echo reply

Redirect when no route to Destination
- TYPE = 5
- CODE =
  - 0 = Network redirect
  - 1 = Host redirect
  - 2 = Network redirect for specific TOS
  - 3 = Host redirect for specific TOS
Use the network to obtain an IP from a remote server

How does a machine without permanent storage determine its IP address?

- How does a machine without permanent storage determine its IP address?
  - Critical for network appliances or embedded systems
  - System must use its physical address to communicate
  - Requests address from server which maintains table of IP’s
  - System doesn’t know the server - sends broadcast request for address

Determine an IP Address at Startup

- How does a machine without permanent storage determine its IP address?
  - Critical for network appliances or embedded systems
  - System must use its physical address to communicate
  - Requests address from server which maintains table of IP’s
  - System doesn’t know the server - sends broadcast request for address

Traceroute

- UNIX utility - displays router used to get to a specified Internet Host
- Operation
  - Router sends ICMP Time Exceeded message to source if TTL is decremented to 0
  - If TTL starts at 5, source host will receive Time Exceeded message from router that is 5 hops away
- Traceroute sends a series of probes with different TTL values... and records the source address of the ICMP Time Exceeded message for each
- Probes are formatted to that the destination host will send an ICMP Port Unreachable message

TraceRoute Example

tenacious-traceroute www.set.dle.ad.jp

Destination Unreachable

<table>
<thead>
<tr>
<th>TYPE CODE CHECKSUM</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP HEADER + 64 bits data from original DG</td>
</tr>
</tbody>
</table>

- UNUSED

- TYPE = 3
  - CODE 0 = Network unreachable
  - 1 = Host unreachable
  - 2 = Protocol unreachable
  - 3 = Port unreachable
  - 4 = Fragmentation needed but DF set
  - 5 = Source route failed

Source Quench

<table>
<thead>
<tr>
<th>TYPE CODE CHECKSUM</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP HEADER + 64 bits data from original DG</td>
</tr>
</tbody>
</table>

- UNUSED

- TYPE = 4; CODE = 0
  - Indicates that a router has dropped the original DG or may indicate that a router is approaching its capacity limit.

Correct behavior for source host is not defined.

Reverse Address Resolution Protocol

- RARP is part of the TCP/IP specification
- RARP operates much like ARP
  - A requestor broadcasts is RARP request
  - Server responds by sending response directly to requestor
  - Requestor keeps IP delivered by first responder
  - Requestor keeps sending requests until it gets an IP
- Clearly there is a need for redundant RARP servers for reliability
  - Timeouts can be used to activate backup RARP servers
    - Backup servers reply to a RARP request if they don’t hear the RARP response from the primary server after some time

Determining an IP Address at Startup

- How does a machine without permanent storage determine its IP address?
  - Critical for network appliances or embedded systems
  - System must use its physical address to communicate
  - Requests address from server which maintains table of IP’s
  - System doesn’t know the server - sends broadcast request for address
Alternatives to RARP

- RARP has shortcomings
  - Most are subtle and all deal with fact that RARP operates at physical level
- BOOTstrap Protocol (BOOTP) was developed as an alternative to RARP – moves process to network level
  - Uses UDP/IP packets to carry messages
  - Hosts are still identified by MAC address
  - How can UDP running over IP be used by a computer to discover its IP address?
    - Uses special case IP address 255.255.255.255 – limited broadcast – not forwarded by routers
    - Forces IP to broadcast on LAN before host IP is known
    - BOOTP server responds using limited broadcast
    - Request transmission via random timeout to avoid synchronization

Dynamic Host Configuration Protocol

- DHCP extends BOOTP
  - Still supports static allocation
  - Supports automatic configuration where addresses are permanent but assigned by DHCP
  - Supports temporary allocation
- Relies on existence of a DHCP server
  - Repository for host configuration information
  - Maintains a pool of available IP’s for use on demand
  - Considerably reduces administration overhead
    - Autoconfiguration of course depends on administrative policy
  - Uses UDP to send messages
    - Uses a relay agent to communicate with servers of LAN (same as BOOTP)
      - Relay agent is statically configured with DHCP server address

DHCP Implementation

- State machine (6 states) which determines DHCP operation
  - Host boots into INITIALIZE state
  - To contact the DHCP server(s) a client sends DHCPDISCOVER message to IP broadcast address and moves to SELECT state
    - Unique header format with variable length options field
      - UDP packet sent to well known BOOTP port 67
  - Server(s) respond with DHCPOFFER message
    - Client can receive 0 or more responses and responds to one
  - Client moves to REQUEST state to negotiate IP lease with 1 server
    - Sends DHCPREQUEST message to server which responds with DHCPACK
  - Client is then in BOUND (normal) state

DHCP Implementation contd.

- From BOUND, client can issue DHCPRELEASE and return to INITIALIZE state
  - This is simply client deciding it no longer needs the IP
- When lease reaches 50% of lease expiration time, it issues DHCPREQUEST to extend lease of current IP with server and moves to RENEW state
  - Receipt of DHCPACK moves client back to BOUND state
  - Receipt of DHCPNACK moves client back to INITIALIZE state
- If no response is received by 87.5% of lease expiration time, the client resends the DHCPREQUEST and moves to REBIND state
  - Receipt of DHCPACK moves client back to BOUND state
  - Receipt of DHCPNACK or timeout moves client back to INITIALIZE state

DHCP Details

- Without relay agent, DHCP would not scale since it would require large number of servers (one per LAN)
- Addresses which are leased over a given period of time and must be updated
  - This means that DHCP requests might have to be made multiple times by the same system (RENEW requests)
Network Address Translation

- Maps an internal \( <\text{address, port}> \) to an external \( <\text{address, port}> \)
- Source address, port of outgoing packet changed
- Destination address, port of incoming packet changed

Maintains a table to translate \( <\text{IP addr, port}> \) pairs

- Used to temporarily tide over IP address space depletion
- Also avoids re-numbering of IP addresses when customer changes provider
- But breaks end-to-end properties