Chapter 3 Internetworking

Basic Internetworking (IP) (cont’d)

• Datagram forwarding in IP
  ▪ Every datagram contains the IP address of the destination host
  ▪ If directly connected to destination network, then forward to destination host
  ▪ If not directly connected to destination network, then forward to some router
  ▪ Each host has a default router
  ▪ Each router maintains a forwarding table
    ▪ Table contains <NetworkNum, NextHop> pairs.
    ▪ Table may contain a default entry specifying a default router (used if none of the entries in the table matches the destination’s network number).
  
○ Router datagram forwarding algorithm

  if (NetworkNum of destination = NetworkNum of one of my interfaces) then Deliver packet to destination over that interface
else if (NetworkNum of destination is in my forwarding table) then Deliver packet to NextHop router
else Deliver packet to default router

○ Host datagram forwarding algorithm

  if (NetworkNum of destination = my NetworkNum) then Deliver packet to destination directly
else Deliver packet to default router
Hierarchical addressing improves scalability: forwarding table lists only network numbers rather than all the hosts in the network.

- Address Translation (ARP)
  - When node A wants to send an IP datagram to node B on the same network, A needs to know B’s link-layer address (or physical address).
  - Each node maintains a table that maps IP addresses into physical addresses
    - Host can dynamically learn the contents of the table using the Address Resolution Protocol (ARP)
    - Each table entry has an expiration timer, entry removed when timer expires.
  - The Address Resolution Protocol
    - A node broadcasts an ARP query if the target IP address is not in ARP table
      - ARP query contains the sender’s IP address and link-layer address, and the target IP address.
    - Target node responds with its link-layer address; it also adds the sending node’s information in its ARP table.
    - The originator adds the IP address to link-layer address mapping of the target node to its ARP table.
    - All nodes on the network sniff the ARP query. If a node already has an entry for the sending node in its ARP table, it refreshes the entry.
  - ARP packet format
Subnetting

- Assigning one network number per physical network is inefficient
- Add another level to address/routing hierarchy: subnet
  - Allocate a single network number to several physical networks, referred to as subnets.
- An IP address now has three parts: network + subnet + host

<table>
<thead>
<tr>
<th>Network number</th>
<th>Host number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class B address</td>
<td></td>
</tr>
<tr>
<td>11111111111111111111111</td>
<td>00000000</td>
</tr>
</tbody>
</table>

- Subnet mask (255.255.255.0)

<table>
<thead>
<tr>
<th>Network number</th>
<th>Subnet ID</th>
<th>Host ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subnetted address</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A host is configured with both an IP address and a subnet mask for the subnet to which it is attached</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
  - Subnet mask indicates the split between network + subnet part and host part
  - All hosts on a given subnet are configured with the same subnet mask
  - IP address AND subnet mask = subnet number
- Host forwarding algorithm

if (my subnet mask & destination IP address = my subnet number)
  Deliver packet to destination directly
else deliver packet to default router
o Router forwarding algorithm

D = destination IP address
for each forwarding table entry <SubnetNumber, SubnetMask, NextHop>
    if (SubnetMask & D = SubnetNumber)
        if NextHop is an interface
            deliver packet directly to destination
        else deliver packet to NextHop (a router)
    if no matches are found, then deliver packet to default router.

o Subnets are not visible from routers outside the site; outside routers keep only one entry in their forwarding tables.

o Benefits of subnetting
  ▪ Improve address assignment efficiency
  ▪ Help aggregate routing information

• Classless InterDomain Routing (CIDR)
  ▪ CIDR addresses two scaling concerns in the Internet
    ▪ The growth of backbone routing table as more and more network numbers need to be stored in them
    ▪ Potential exhaustion of the 32-bit address space due to address assignment inefficiency
  ▪ Address assignment inefficiency arises because of the IP address structure with class A, B, and C addresses
    ▪ A network with 2 hosts needs a class C address
    ▪ A network with 256 hosts needs a class B address
  ▪ Exhaustion of IP address space centers on exhaustion of the class B network numbers
    ▪ Solution: assign an appropriate number of class C addresses
Problem with this solution: Excessive storage requirement at the routers

- If an organization has 16 class C network numbers assigned to it, every Internet backbone router needs 16 entries in its routing tables for that organization
- If we had assigned a class B address to the organization, the same routing information can be stored in one entry. But efficiency = $16 \times 254 / 65,534 = 6.2\%$

CIDR tries to balance the desire to minimize the number of routes that a router needs to know against the need to hand out addresses efficiently.

CIDR helps us to aggregate routes

- Uses a single entry in the forwarding table to tell the router how to reach a lot of different networks
- This is achieved by breaking the rigid boundaries between address classes

Consider an organization with 16 class C network numbers.

- Instead of handing out 16 addresses at random, hand out a block of contiguous class C addresses
- Suppose we assign the class C network numbers from 192.4.16 through 192.4.31
- Observe that top 20 bits of all the addresses in this range are the same (11000000 00000100 0001)
- We have created a 20-bit network number (which is in between class B network number and class C number)
- CIDR requires to hand out blocks of class C addresses that share a common prefix
  - Each block must contain a number of class C networks that is a power of 2
- To represent a network number, or *prefix*, the convention is to place a /X after the prefix where X is the prefix length in bits
  - Prefixes may be of any length (2 to 32 bits)
- For example, the 20-bit prefix for all the networks 192.4.16 through 192.4.31 is represented as 192.4.16/20
  - By contrast, 192.4.16/24 represents a single class C network number
- IP forwarding revisited
  - Each forwarding table entry is a `<NetNumber/MaskLength, NextHop>` pair
  - When a packet comes in, router scans the forwarding table entry by entry, looking for a match. A match occurs when destination IP address & Mask = NetNumber.
  - If multiple entries match, the entry with the longest match wins.
- Internet Control Message Protocol (ICMP)
  - Defines a collection of error messages that are sent back to the source host whenever a router or host is unable to process an IP datagram successfully
    - Destination host unreachable due to link/node failure
    - Reassembly process failed
    - TTL had reached 0
    - IP header checksum failed
- ICMP-Redirect
  - Sent from a router to a source host
  - With a better route information
- ICMP messages are carried inside IP datagrams