The Internet Protocol (IP)

John Kristoff
jtk@depaul.edu
+1 312 362-5878
DePaul University
Chicago, IL 60604
Will layer 2 networking suffice?
Layer 3 usually provides

- Internetworking for data link technologies
- Globally unique addresses
- Scalable routing
- A common communications format
- Packet fragmentation capability
- A hardware independent interface
- Packet independence
An IP Router (or gateway)

- Usually a special purpose, dedicated device
- Connects heterogeneous networks
- Directs packets *toward* ultimate destination
- Dynamic routing algorithms often used
  - They make automatic forwarding decisions
  - They can forward based on various metrics
- Official pronunciation is really *rooter*
- Layer 3 switch = router = layer 3 switch
IP Routing

- **Scope**
  - Autonomous system, interior, exterior

- **Dynamic routing**
  - Protocol for route exchange and computation

- **Static routing**
  - Manually configured routes

- **Destination address driven**
Internet protocol (IP)

- Standardized in RFC 791
- Connectionless
- Unreliable
- Fairly simple
- The Internet Glue
IP addresses

- Virtual – not bound to hardware
- 32-bit fixed size
- Unique address for each IP interface
- Global authorities assign a prefix (network)
- Local administrators assign the suffix (host)
- Usually written as dotted decimal notation
  - e.g. 140.192.1.6
IP address types

- Unicast (one-to-one)
  - Source address should always be unicast
- Multicast (one-to-many)
  - Receivers join/listen to multicast group address
- Broadcast (one-to-all)
  - Special case of a multicast, usually best avoided
- Anycast (one-to-one-of-many)
  - Preferably one-to-nearest, defined for IPv6
### IP address notation

<table>
<thead>
<tr>
<th>32-bit Binary Number</th>
<th>Equivalent Dotted Decimal</th>
</tr>
</thead>
<tbody>
<tr>
<td>10000001 00110100 00000110 00000000</td>
<td>129.52.6.0</td>
</tr>
<tr>
<td>11000000 00000101 00110000 00000011</td>
<td>192.5.48.3</td>
</tr>
<tr>
<td>00001010 00000010 00000000 00100101</td>
<td>10.2.0.37</td>
</tr>
<tr>
<td>10000000 00001010 00000010 00000011</td>
<td>128.10.2.3</td>
</tr>
<tr>
<td>10000000 10000000 11111111 00000000</td>
<td>128.128.255.0</td>
</tr>
</tbody>
</table>
# Special IP addresses

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Suffix</th>
<th>Type Of Address</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>all-0s</td>
<td>all-0s</td>
<td>this computer</td>
<td>used during bootstrap</td>
</tr>
<tr>
<td>network</td>
<td>all-0s</td>
<td>network</td>
<td>identifies a network</td>
</tr>
<tr>
<td>network</td>
<td>all-1s</td>
<td>directed broadcast</td>
<td>broadcast on specified net</td>
</tr>
<tr>
<td>all-1s</td>
<td>all-1s</td>
<td>limited broadcast</td>
<td>broadcast on local net</td>
</tr>
<tr>
<td>127</td>
<td>any</td>
<td>loopback</td>
<td>testing</td>
</tr>
</tbody>
</table>
# Classful IP Addressing

<table>
<thead>
<tr>
<th>Class</th>
<th>Prefix</th>
<th>Suffix</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>B</td>
<td>10</td>
<td>24</td>
</tr>
<tr>
<td>C</td>
<td>110</td>
<td>16</td>
</tr>
<tr>
<td>D</td>
<td>1110</td>
<td>8</td>
</tr>
<tr>
<td>E</td>
<td>1111</td>
<td>Reserved for future use</td>
</tr>
</tbody>
</table>

- **Class A** uses 0 to 126.255.255.255.
- **Class B** uses 128.0.0.0 to 191.255.255.255.
- **Class C** uses 192.0.0.0 to 223.255.255.255.
- **Class D** uses 224.0.0.0 to 239.255.255.255.
- **Class E** is reserved for future use.
## Classful address sizes

<table>
<thead>
<tr>
<th>Address Class</th>
<th>Bits In Prefix</th>
<th>Maximum Number of Networks</th>
<th>Bits In Suffix</th>
<th>Maximum Number Of Hosts Per Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>7</td>
<td>128</td>
<td>24</td>
<td>16777216</td>
</tr>
<tr>
<td>B</td>
<td>14</td>
<td>16384</td>
<td>16</td>
<td>65536</td>
</tr>
<tr>
<td>C</td>
<td>21</td>
<td>2097152</td>
<td>8</td>
<td>256</td>
</tr>
</tbody>
</table>
Example IP network
Example IP router addressing
Limitations of classful addressing

- Internet growth
- Route table size
- Address depletion
- Misappropriation of addresses
- Lack of support for different sized networks
  - Class B too big, class C too small
IP addressing solutions

- Subnetting
- Supernetting
- Classless interdomain routing (CIDR)
- Variable length subnet masks (VLSM)
Subnetting

network  host

140.192.9.63

network  subnet  host

140.192.9.63
Subnet mask

- The bit length of the prefix (network)
- Prefix (network) is no longer classful
- Dotted decimal or '/' notation
  - 140.192.1.6's subnet mask is 255.255.255.128
  - ...or 140.192.1.6/25
- You may want to convert to binary for clarity
  - A /25 or 255.255.255.128 subnet mask is:
  - 11111111.11111111.11111111.10000000
Example: Using subnet masks

- Given 140.192.50.8/20, what is the:
  - subnet mask in dotted decimal notation?
  - directed broadcast address in dotted decimal
  - total number of hosts that can be addressed?
Supernetting

- Combine smaller blocks into larger aggregate
- If class B too big, class C too small...
- Maybe do this:
  - Combine 199.63.0.0/24 to 199.63.15.0/24
  - Equals 199.63.0.0/20
Example: Using Supernets

- Given that ISP has 128.15.0.0/16:

  - If a customer needs to address 300 hosts, how might the ISP assign them address space?
  - What is the address space assigned in the example above in 'slash' notation?
  - How many, if any, maximum free IP addresses will the customer have at their disposal?
  - Can you think of any reason why the customer might have less than that maximum?
CIDR

- Use supernetting for routing tables
- Routes advertised as smaller CIDR blocks
- So instead of advertising:
  - 199.5.6.0/24, 199.5.6.1/24, 199.5.6.2/24 and 199.5.6.1/24 separately
- Advertise:
  - 199.5.6.0/22 one time
- Internet CIDR report
Example: Using CIDR

- Given that an ISP announces netblocks 64.5.0.0/20, 64.5.16.0/20, 192.0.2.0/25 and 192.0.2.192/26 and 192.0.2.128/26:
  
- What is the smallest number of CIDR announcements that this ISP can make?

- If these routes are received from another provider, can you think of any reason why they might not be able to be CIDR-ized?
VLSM

- Multiple subnet sizes in a single AS
- Allows efficient use of address space
- Can be used to build internal hierarchy
- External view of AS does not change
- An organization may have 140.192.0.0/16
  - But internally may use 140.192.0.0/17, 140.192.128.0/24, 140.192.129.0/24 and so on.
Example: Using VLSM

- Given an address space of 140.192.0.0/16 to work with, assign netblocks and addresses based on the following network:
  - 6 satellite sites and 1 main office center
  - About 7000 hosts exist on entire network today
  - Main site uses approximately 50% of addresses
  - Satellites vary from 200 to 700 total addresses
  - Overall growth for organization is 500 hosts/year
Obtaining IP addresses

- IANA has global authority for assignment
  - Regional registries delegate (ARIN/RIPE/APNIC)
  - ISPs assign addresses to end users
- RFC 1918 defines private address netblocks
  - Not globally unique
  - Must not appear on the public Internet
  - 10.0.0.0/8, 172.16.0.0/12, 192.168.0.0/16
IP datagram layout
Inside an IPv4 datagram

- Version field
  - Binary 0100 (equals what in decimal?)
- Header length
  - Length of the IP header in 32 bit words
  - Will usually be equal 5 (in decimal)
- Type of Service (now DiffServ field)
  - An indication of quality/class of service
  - Rarely used, but if so usually within single AS
Inside an IPv4 datagram [cont.]

- **Total length**
  - Total IP datagram length in octets
  - Maximum value is 65535, but rarely > 1500

- **Identification**
  - Used for to identify fragmented packets
  - Experimental use for tracing (D)DoS attacks

- **Flags**
  - Bit 0 reserved, others control fragmentation
Inside an IPv4 datagram [cont.]

- Fragment offset
  - Helps piece together fragment datagrams
- Time to live (TTL)
  - Bounds time/hops of IP datagram in network
  - Counts down to zero and stops being forwarded
- Protocol type
  - Indicates next level protocol in data portion
Inside an IPv4 datagram [cont.]

- Header checksum
  - Used to verify header validity at each hop
- Source address
  - 32-bit IP address
- Destination address
  - 32-bit IP address
- Options
  - Variable, not oftenly used
Demo: Understanding PING

• Setup packet capture session using tcpdump
  • tcpdump -n -s 1500 -w ping.cap icmp and ( dst host <my-ip> or src host <my-ip> )

• Ping remote host
  • ping <remote-ip>

• View capture using Ethereal
Demo: Understanding traceroute

• Setup packet capture session using tcpdump
  • tcpdump -n -s 1500 -w traceroute.cap (udp or icmp) and ( dst <my-ip> or src <my-ip> )

• Trace remote host
  • traceroute -n <remote-ip>

• View capture using Ethereal
Other tools and references

- Find contacts for IP address or netblocks
  - `whois <ip-address-or-network>`

- View network path from external sites
  - `http://www.traceroute.org`

- Verify DNS entry to IP address or vice versa
  - `nslookup <ip-address>`

- `http://www.iana.org`
- `http://www.arin.org`
Supporting protocols

- ARP
- BOOTP/DHCP
- DNS
- ICMP
- SNMP
Final thoughts

- IP is unreliable
- IP addressing can be a pain
- IPv6 doesn’t make it any easier
- IP address is both a who and a where
- IP addresses provide little security
- Private IPs and NAT are best avoided
- IP fragmentation is generally best avoided