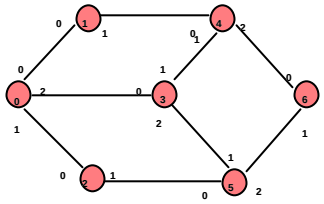
Internetting



Lecture Overview

- internetting concepts
- IPv4 addresses
- IP and MAC addresses
- IPv6

TCP/IP History

U.S. DoD Advanced Research Project Agency (ARPA, sometimes DARPA) created ARPANET starting 1968

ARPANET was first wide-area general purpose packet network, served universities and ARPA contractors

Original developers BBN, UCLA, SRI. Later USC/ISI, MIT.

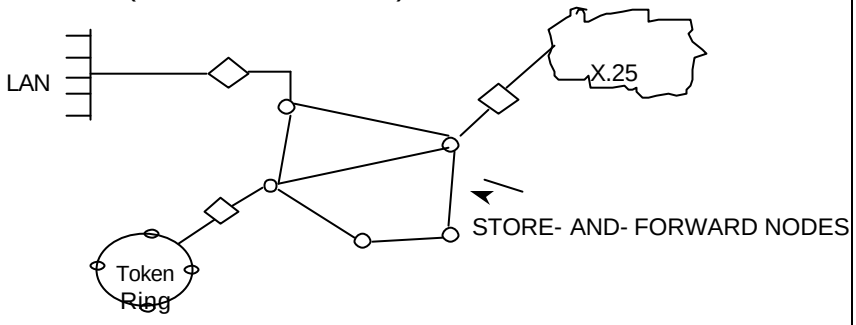
Standards evolved from "Request for Comments" (Now RFC)

In the early 1970's the concept of internetworking (or internetting) was advanced. The ARPANET became the core of the Internet experiment and TCP/IP was developed as its protocol suite

Today the IETF supports dozens of related open protocols that make up the Internet Protocol Suite (IPS)

Internetworking
1 of 3

Subnetworks interconnected by intermediate systems to form an internet (note lower case "i").



Internetworking
2 of 3

Subnet: Physical network over which systems may communicate directly.

- Bridge: Datalink layer
  - Relays frames on same subnet
- Router: Network layer
  - Forwards packets to different subnets

Requirements:

- Network interconnection - PL, DLL, NL
- Routing and data delivery among networks
- Services such that network differences are handled.

Internetworking
3 of 3

- Addressing

  DL address - subnetwork point of attachment (SNPA)

  Different address (and possibly formats) for different subnet

  Assume that subnet physical has local (on subnet) significance only.
- MTU - Maximum transfer unit
  - absolute maximum set by protocol
  - subnets can establish smaller MTU
- Network access mechanism
- Error recovery
- Routing techniques
- Connection-oriented (VC) or connectionless (datagram)

Internet Protocol (IP)

- IP is a “connectionless” protocol
  - each packet is routed independent of all others
- IP packets are called “datagrams”
- IP delivers packets on a “best effort” basis - no guarantee of delivery
- IP does not guarantee ordered delivery
- Created as a protocol to link different networks, IP has has become increasingly popular to run directly on hosts

IP Addresses - Historic

There are four active classes of IP address.
All are 32 bits long

Class A - 128 nets with 16M hosts each
0nnnnnnn hhhhhhhh hhhhhhhh hhhhhhhh

Class B - 16K nets with 64K hosts each
10nnnnnn nnnnnnnn hhhhhhhh hhhhhhhh

Class C - 2M nets with 256 hosts each
110nnnnn nnnnnnnn nnnnnnnn hhhhhhhh

Class D - Multicast, 256M groups
1110gggg gggggggg gggggggg gggggggg

Classless - split in some other way, *e.g.*:
110nnnnn nnnnnnnh hhhhhhhh hhhhhhhh

NOTE: here K = 2^{10} = 1024, M = 2^{20} = 1048576

IP Number Notation

192.	127.	253.	1
11000000	01111111	11111101	00000001

Domain Name System

- - -

cs.gmu.edu = 129.174.40.13

site.gmu.edu = 129.174.40.83

netlab.gmu.edu = 129.174.65.1

cne.gmu.edu = 129.174.120.40

dsigw.gmu.edu = 199.56.135.2

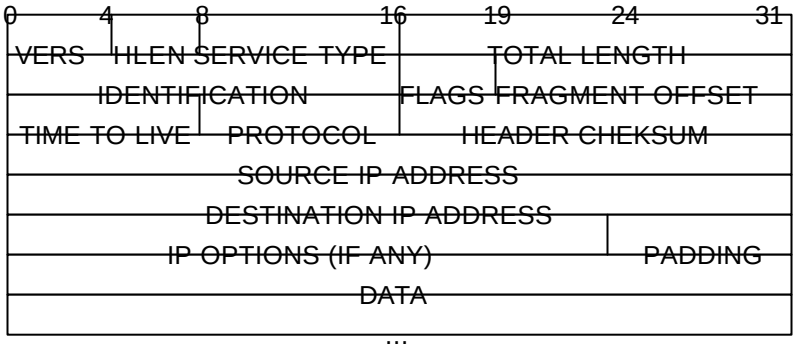
CIDR

- An Internet router must know a path to any network (nnnn above) to which it needs to send a packet
  - when the packet gets to the network hhhh is used to deliver it
  - for an edge router (for example a LAN gateway) this is easy- there is only one path
- Classless InterDomain Routing (CIDR) is an approach to using blocks of addresses in other than classes A,B,C
  - as with classful routing, the whole block must be reachable through one gateway/router

Prefix Notation

- Latest IP address block notation is /k where k is the number of n bits used to deliver the packet from the Internet
  - the destination may subdivide the block for internal delivery
  - for example, GMU class B is 129.174/16
  - NETLAB subnet is 129.174.65/24
  - another notation: mask with 1 for n, 0 for h, for example (class C)
  - 11111111 11111111 11111111 00000000
    = 255.255.255.0

IPv4 Header



IPv4 Header Information
1 of 3

VERS - protocol version (current = 4) (new = 6)

HLEN - header length in 32 - bit words
max. header length = 60 bytes

SERVICE TYPE - precedence and handling options
Telnet - low delay
FTP data - high throughput
??? - high reliability

3 bits: Precedence field
4 bits: TOS - min delay, max. throughput, max reliability
1 bit : unused.

IPv4 Header Information
2 of 3

TOTAL LENGTH - packet length in bytes
16 bits ➢ max = 65,535 bytes
Field needed because link layer may add padding

IDENTIFICATION, FLAGS, FRAGMENT OFFSET - used in fragmentation, i.e. passing a packet through a net where its length exceeds that net's Maximum Transmission Unit (MTU)

TIME TO LIVE (TTL)- max number of routers through which datagram may be forwarded. Set by sender (e.g. 32 or 64) and decremented by each router that processes datagram.

IP Header Information
3 of 3

PROTOCOL - code for transport protocol

HEADER CHEKSUM - error check for header
(assuming zeros in this field).
- Sender: Divide header into 16 bit pieces and calculate 1's complement of sum)
- Receiver: 1's complement sum over entire header, including checksum. ALL 1's -> OK.

SOURCE IP, DESTINATION IP - addresses of from/to hosts (32 bits each)

OPTIONS - Variable length; little used

IP Subnet Mechanisms

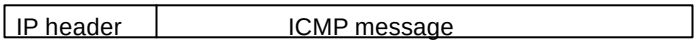
Using IP, devices on a physically contiguous (or bridged) LAN will share the same high-order address bits
  – typically, a class C network address

- Address Resolution Protocol (ARP) - broadcast to LAN "does anybody out there use this IP address?" builds table of Ethernet addresses corresponding to IP addresses.
- Reverse ARP (RARP) - used by diskless client to find its IP address at bootup, designated RARP host replies
  – protocol "BOOTP" is used
- Dynamic Host Configuration Protocol (DHCP) allows IP addresses to be assigned from a server when needed rather than statically configured
  – selected devices can still always be assigned the same IP address

IP Control Message Protocol (ICMP)

1 of 4

- ICMP passes data among routers:
  - error statistics
  - reachability tests (“ping”)
  - “source quench” congestion control (not widely used)
  - routing loops
  - clock synchronization
- Part of Internet Layer
- Used to communicate errors and other conditions not handled by IP or higher layer protocols
- Carried in IP datagrams with Protocol = 0x01

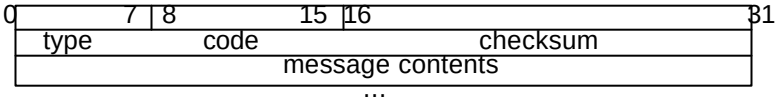


20 bytes

IP Control Message Protocol

2 of 4

- ICMP Message Format



- Types include

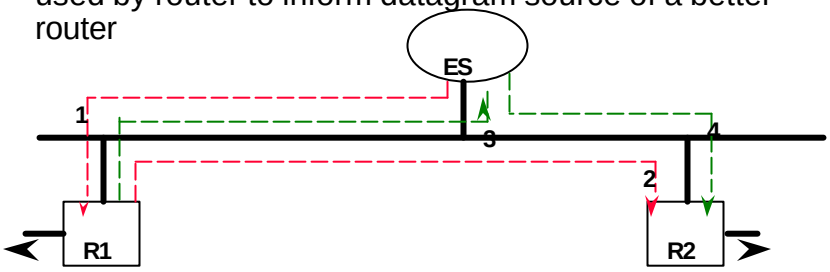
0 Echo (ping) reply	Query response
3 Destination unreachable	Error
Codes include:	
0 Network unreachable	
1 Host unreachable	
5 Redirect	Error
8 Echo request	Query

IP Control Message Protocol
3 of 4

- Checksum
  Calculated over entire ICMP message
  (same algorithm as IP)
- Message contents
  Query: Depends on type
  Error: IP header and first 8 bytes data of diagram that caused ICMP message to be generated.
- ICMP error message never generated in response to
  ICMP error message
  Datagram destined to IP broadcast or IP multicast addr.
  Fragment other than first
  Datagram with source address <> single host

IP Control Message Protocol
4 of 4

ICMP Redirect
used by router to inform datagram source of a better router



Allows good routing on subnet with minimal configuration in end systems.

Useful ICMP Utilities

- ping internet-address
  - responds with reachability
  - variation ping-s pings once per second - shows packet loss
- traceroute internet-address
  - on Windows tracert internet-address
  - prints routers on path to the address and delay to reach them

Ping Example

```
/home/mpullen >ping macaw.nlm.nih.gov
macaw.nlm.nih.gov is alive
/home/mpullen >ping -s macaw.nlm.nih.gov
PING macaw.nlm.nih.gov: 56 data bytes
64 bytes from macaw.nlm.nih.gov (130.14.42.1): icmp_seq=0. time=4. ms
64 bytes from macaw.nlm.nih.gov (130.14.42.1): icmp_seq=1. time=6. ms
64 bytes from macaw.nlm.nih.gov (130.14.42.1): icmp_seq=2. time=5. ms
64 bytes from macaw.nlm.nih.gov (130.14.42.1): icmp_seq=3. time=5. ms
64 bytes from macaw.nlm.nih.gov (130.14.42.1): icmp_seq=4. time=4. ms
64 bytes from macaw.nlm.nih.gov (130.14.42.1): icmp_seq=5. time=4. ms
64 bytes from macaw.nlm.nih.gov (130.14.42.1): icmp_seq=6. time=4. ms
^C
----macaw.nlm.nih.gov PING Statistics----
7 packets transmitted, 7 packets received, 0% packet loss
round-trip (ms)  min/avg/max = 4/4/6
```

Traceroute Example

```
/home/mpullen >traceroute 137.53.87.128
traceroute to 137.53.87.128 (137.53.87.128), 30 hops max, 40 byte packets
 1  129.174.65.2 (129.174.65.2)  2.382 ms  2.259 ms  2.210 ms
 2  129.174.247.118 (129.174.247.118)  0.434 ms  0.433 ms  0.580 ms
 3  206.197.101.66 (206.197.101.66)  3.259 ms  4.037 ms  3.415 ms
 4  192.70.138.22 (192.70.138.22)  3.806 ms  4.270 ms  4.517 ms
 5  nycmng-washng.abilene.ucaid.edu (198.32.8.84)  8.936 ms  8.670 ms
       21.744 ms
 6  chinng-nycmng.abilene.ucaid.edu (198.32.8.82)  27.560 ms  27.980 ms
       31.745 ms
   - - -
12  sttl-sttlng.abilene.ucaid.edu (198.32.11.125)  80.248 ms  79.927 ms
       80.125
13  hnsp1-wes-so-5-0-0-0.pnw-gigapop.net (198.48.91.77)  79.970 ms
       80.581 ms  80.141 ms
  - - -
19  137.53.254.21 (137.53.254.21)  84.383 ms *  84.282 ms
20  137.53.223.11 (137.53.223.11)  84.702 ms  85.175 ms  84.450 ms
21  * * *
22  137.53.87.128 (137.53.87.128)  85.450 ms  84.689 ms  84.756 ms
```


IP Fragmentation

Transparent to higher layers

When IP recovers datagram, if MTU for outgoing interface < datagram (data) size, fragment

Rough Algorithm:

- Divide data based on MTU
  - Max size pieces from front
  - Multiples of 8 except last.
- Copy original IP header to IP header of each frag.
  - If options in original, all copied to first frag, only some copied to other frags.
- In each frag: If last, set more fragments bit to 1, else 0.
  - Set fragment offset to position (in 8 byte units) in data area of original datagram.
  - Set total length of fragment, datagram.
  - Compute IP checksum (over header)

More on IP Fragmentation

Fragment may be further fragmented later.
If don't fragment bit = 1 and fragmentation required, drop datagram and send ICMP error message.

Reassembly
Performed at destination
Straightforward using *fragment offset* and *more* flags.

IPv6

- The 32-bit address of the current IP is becoming a limiting factor
- IETF created the "IPng" working group to devise a replacement
- They came up with a 128-bit address and several other innovations
  - Self-identification so hosts do not need to be configured with an IP address
  - Flow identifier for expedited real-time routing
  - MTU discovery replaces IPv4 fragmentation
  - Mandatory security
- Internet industry has been slow to move to v6
  - Until recently there was no good way to stay compatible with v4
  - In 1999 IETF defined a more practical transition path based on tunneling IPv6 through IPv4 networks

Internet Gateway Functions

- Gateway is another name for a router serving a subnet
- As such it performs normal router functions
  - routing protocol to determine path to external addresses
  - ARP to match local MAC addresses to IP addresses
- Plus dealing with differences in subnets
  - IPv4 fragmentation due to MTU mismatch
- Also may serve as border gateway to advertise Internet reachability