Introduction to TCP/IP

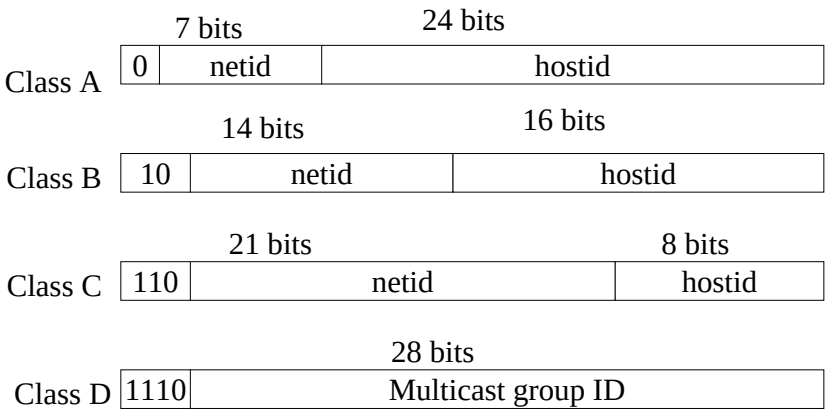
TCP/IP

- Four layer Architecture
- Developed in 1960's
- *Open* System
- Not just one protocol, whole family.
- Many programming interfaces available.
- Standardised protocol set.

IP Addressing Scheme

- Need capability of mapping addresses of one type onto another.
- LAN address, Network Point of Attachment NPA, must be mapped onto an IP address.
- NPA formats differ from one LAN standard to another.
- IP addresses are homogenous within single IP version.

IP Address Format



IP Address Format (cont.)

- Different size networks may use different address classes, defined by the first few bits in the address. 0 for Class A, 10 for Class B, 110 for Class C, etc. etc.
- Networks with large numbers of hosts may use Class A, while Class C may have many subnets with a small number of attached hosts.

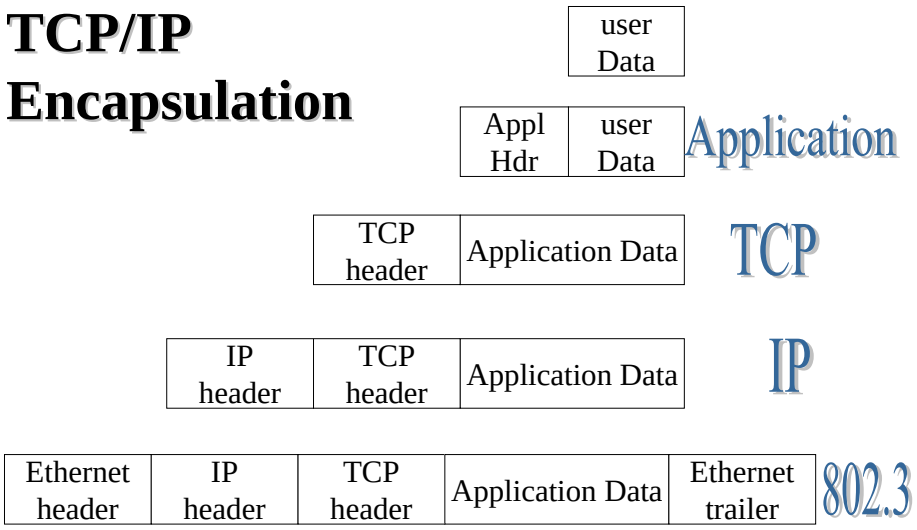
IP Address Notation

- A decimal dot notation is used to break down the IP address.
- Example
  - 10001000 11001110 00001011 00000110
  - gives the address 136.206.11.6 aka boole !
  - Note that this is a Class B address (first zero in second position) and the subnet is defined with 14 bits, the host address with 16 bits.

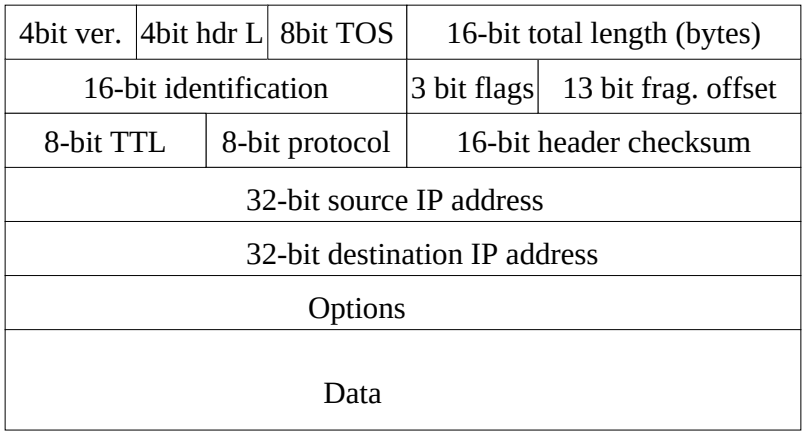
IP Allocations

- A central authority has responsibility for allocation of IP addresses. They are the network Information center, or NIC.

TCP/IP Encapsulation



IP Packet Header



IP Header Description

- *Version*: Currently V 4.
- *Header Length*:Specifies length of header as some fields are optional.
- *Type of Service*:This is the same as the QOS mentioned previously.
- *Total length*:Specifies the length of the datagram.

- *Identification*: Used to identify a set of datagrams which were formed from a single user message, but which got fragmented while traversing possibly several networks.
- *D bit*:Indicates that routers should not fragment a datagram i.e. Don't fragment bit.
- *M bit*:Indicates that there are more fragments to follow in later datagrams.
- *Fragment offset*:Where this fragments fits into the original fragmented datagram.

- *Time to live*: Datagram loses a life (or some time to live) on each hop across the internet. Datagram destroyed when time\lives run out. Prevents Datagrams from wandering endlessly.
- *Header Checksum*: Checks header only.
- IP addresses (Source, Destination): As described previously.

IP Routing

- Central function of IP is routing along with fragmentation and re-assembly of data across an internet.
- Routing information organised in a hierarchy. With hosts and gateways involved.
- ARP address resolution protocol maps IP to Ethernet addresses, an Interior Gateway Protocol (IGP)

- Exterior Gateway Protocol (EGP) knows about other routers on the internet and can route from network to network.
- Distance Vector and Link State routing are most popular, Link State is superior.
- Subnet addressing may be performed on a group of related networks (owned by one organisation).
- More on Routing later…

Special IP Addresses

- Some addresses are reserved for special use.
- IP address composed of all 0 means this host.
- Network part all 0, Host part not, host on this network.
- All 1s broadcast on LAN
- Host part 127 is Loopback, useful for debugging.

CIDR

- Classless Inter Domain Routing -
- Give the IP address space some breathing room!
- Basic idea: allocate the remaining IP addresses in variable-size blocks without regard to classes
  - original name: supernetting
- A site needing 2000 addresses receives a block of 2408 addresses i.e., 8 contiguous class C networks. If need 8000 hosts, then allocate a block of 8192 addresses, i.e., 32 contiguous class C networks.

CIDR Example 1

- A site receives 16 class C addresses a block of class C addresses 192.4.16.0 through 192.4.31.0
- The top 20 bits of all the addresses in this range are the same (11000000 00000100 0001….)
- A 20-bit network number has been created (something between a class B and a class C)
- Written192.4.16.0/20 (20= number of bits in network prefix)
- Representing a network address like this is similar to using a network mask
- Modern routing protocols(like BGP4) understand that network numbers maybe of any length

CIDR Example 2

- If a organization needs 2000 hosts, then allocate it a block of 2048 addresses, i.e.,
  - 8 contiguous class C networks.
- If need 8000 hosts,
  - then allocate a block of 8192 addresses, i.e.,
  - 32 contiguous class C networks.

CIDR Example 3

Suppose an organization is allocated four contiguous class C networks:

205.100.0.0 205.100.1.0 205.100.2.0 205.100.3.0

Question: how to treat these four contiguous networks as one from outside?

Answer: Network mask which will mask out one common prefix for these four networks.

Question: what is the network mask for these four networks?

205.100.0.0 --- 11001101 . 01100100 . 00000000 . 00000000
205.100.1.0 --- 11001101 . 01100100 . 00000001 . 00000000
205.100.2.0 --- 11001101 . 01100100 . 00000010 . 00000000
205.100.3.0 --- 11001101 . 01100100 . 00000011 . 00000000

The common prefix: 11001101 . 01100100 . 000000

Therefore, *network mask:* 11111111 . 11111111 . 111111100 . 00000000,

i.e., 255.255.252.0

In routing table, instead of putting all four networks entries, just put one entry: 205.100.0.0/22, where 22 indicates the *network mask* is 22 bits.

CIDR is also called supernetting because it "supernets" multiple networks into one.

CIDR Challenge 1

- it is possible that both supernet 205.100.0.0/22 and 205.100.0.0/20 appear in the routing table.
  - Therefore, the IP address 205.100.1.1 will match both of them.
  - Solution: *longest prefix match*. 205.100.1.1 will match 205.100.0.0/22.
- A prefix of *arbitrary* length, along with the *network mask* of the same length, indicates a network number.

CIDR Challenge 2

- The multiple contiguous networks can not begin at a random class C network address but must begin at certain boundary.
  - E.g., 16 contiguous networks (i.e., 4096 addresses) can not begin at 194.24.8.0. Instead, they must lie on a 4096-byte boundary. Such as begin from 194.24.16.0 through 194.24.31.0.
- Dropping classes makes forwarding more complicated
- IPv4 days numbered but CIDR buys a little more time
- Larger address space of IPv6 is real solution, but transition to IPv6 takes time.
- All of our machines are IPv6 capable, but none use it!

Variable Length Subnet Masks

- Only works with routing protocols which support CIDR
- Different masks on each router interface. Small number of bits for routers so they have few hosts, few routers. Keep big numbers for LANs
- Match required number of hosts to appropriate mask on each interface.
- Requires careful design so that blocks do not overlap
- Routes may be summarised, providing a hierarchy.

Subnetting – Why?

- Reduces Network traffic
  - Routers create smaller broadcast domains, more smaller domains limits the span of a broadcast.
- Optimizes NW performance
  - Less traffic, things run faster.
- Simplifies management
  - Easier to do fault analysis on a smaller self-contained NW than with a single huge NW
- Facilitates spanning of large geographical distances
  - Single large NW over large distance incurs big overhead of resources. Smaller NWs which keep much traffic local will incur less overhead over the long haul.

Creating Subnets

- Address space
  - [network#, host#]
  - [network#, subnet#, host#]
- Subnet *mask* used to find the host part of IP address and distinguish it from the NW part.

Class	Format	Default subnet mask
A	nw.node.node.node	255.0.0.0
B	nw.nw.node.node	255.255.0.0
C	nw.nw.nw.node	255.255.255.0

Subnetting Class C Addresses

- Only 8 bits for hosts, $2^8 = 256$ hosts

Binary (host)	Decimal	CIDR
00000000	0	/24
10000000	128	/25
11000000	192	/26
11100000	224	/27
11110000	240	/28
11111000	248	/29
11111100	252	/30

- /31, /32 unused as we must have at least 2 host bits

Subnetting - simply

- 5 Questions
  - How many subnets does my network mask produce
  - How many valid hosts are there per subnet
  - What are the valid subnets
  - What is the broadcast address on each subnet
  - What are the valid hosts in each subnet
- Remember
  - Cannot use the subnet number as a host
  - Must keep a host number for broadcast on subnet

5 Answers (I)

- Subnet 192.168.10.0 using a subnet mask of 255.255.255.192 i.e. /26 [1111 1111 . 1111 1111 . 1111 1111 . 1100 0000]
- How many subnets on this network?
  - 2^x where *x* is number of subnet bits, ones
    - **11**00 0000 gives 2^2 subnets, i.e. 4
- How many hosts per subnet?
  - $(2^y - 2)$ where y is number of unmasked bits, zeros
    - 11**00 0000** gives us $(2^6 - 2)$ hosts, 62 per subnet
    - Subtract 2 for the subnet address and the broadcast address
- What are the valid subnets?
  - (256 – subnet mask) = block size or increment number
    - (256 – 192)= 64 block size
    - Start counting at 0 in blocks of 64 until you reach the mask, these are your subnets, 0, 64, 128, 192

5 Answers (II)

- What is the broadcast address of each subnet?
  - We know our subnets are 0, 64, 128,and 192. Broadcast is always the number just before the next subnet
    - 0 subnet has broadcast address of 63 because next subnet is 64
    - 64 subnet has broadcast address of 127 because next subnet is 128
    - Broadcast address of last subnet is always 255
- What are the valid hosts?
  - The host numbers between the subnets, except the all 1s and all 0s
  - For subnet 64, 127 is broadcast address, then 65 to 126 is valid host range, the numbers between subnet address and broadcast address. See table…

255.255.255.192

- Perform calculations in the following order of steps

	Steps				
Subnets	1	0	64	128	192
1st Host	4	.1	.65	.129	.193
Last Host	3	.62	.126	.190	.254
Broadcast	2	.63	.127	.191	.255

Practice this for Class C Addresses

- Answer the 5 questions for the following addresses, all Class C addresses
  - 255.255.255.224/27
    - Subnet 192.168.10.0 using subnet mask 255.255.255.224
  - 255.255.255.240/28 (this is 16 bits, careful!)
    - Subnet 192.168.10.0 using subnet mask 255.255.255.240
  - 255.255.255.248/29
    - Subnet 192.168.10.0 using subnet mask 255.255.255.248
  - Keep going…
  - Why would I use a mask that provides only 2 hosts?

Another Worked example

- Ok, I'll do this one also: 192.168.10.0 network address, Subnet mask 255.255.255.**252** i.e. /30
  - 1111 1111 . 1111 1111 . 1111 1111. **1111 11**00
  - Subnets: 64 ($2^6 = 64$)
  - Hosts: 2 on each subnet ($2^y - 2$ where y is 2 bits)
  - Valid Subnets 0, 4, 8, 12 … 244, 248, 252
  - Broadcast address for each subnet, see table below

Subnet	0	4	8	12	…	240	244	248	252
First Host	1	5	9	13	…	241	245	249	253
Last Host	2	6	10	14	…	242	246	250	254
Broadcast	3	7	11	15	..	243	247	251	255

Subnetting Class B Addresses

- Notice the pattern below, we saw it before in Class C addresses
- Worth memorising!
- More addresses available here with 16 bits for host part.
- Can use 14 bits for subnets
- Must leave 2 bits minimum for host part

255.255.0.0	/16		
255.255.128.0	/17	255.255.255.0	/24
255.255.192.0	/18	255.255.255.128	/25
255.255.224.0	/19	255.255.255.192	/26
255.255.240.0	/20	255.255.255.224	/27
255.255.248.0	/21	255.255.255.240	/28
255.255.252.0	/22	255.255.255.248	/29
255.255.254.0	/23	255.255.255.252	/30

Example 1 – Class B Address

- Network Address: 172.16.0.0, Subnet mask:255.255.192.0 (/18)
  - 1111 1111 . 1111 1111 . **11**00 0000 . 0000 0000
- Subnets: 2^2 =4, Hosts: (2^{14} – 2) = 16,382
  - 6 bits from 3rd byte, 8 from 4th
- Valid subnets = 256 – 192 = 64 (block)
  - 0, 64, 128, 192

Subnet	0.0	64.0	128.0	192.0
1st Host	0.1	64.1	128.1	192.1
Last Host	63.254	127.254	191.254	255.254
Broadcast	63.255	127.255	191.255	255.255

Example 2 – Class B Address

- Network Address: 172.16.0.0, Subnet mask:255.255.240.0 (/20)
  - 1111 1111 . 1111 1111 . 1111 0000 . 0000 0000
- Subnets: 2^4 =16, Hosts: (2^{12} – 2) = 4094
- Valid subnets = 256 – 240 = 16 (block)
  - 0, 16, 32, 48… 224, 240
- Depicted below are 1st four and last two subnets.
- Continue a few more as an exercise

Subnet	0.0	16.0	32.0	48.0	…	224.0	240.0
1st Host	0.1	16.1	32.1	48.1	…	224.1	240.1
Last Host	15.254	31.254	47.254	53.254	…	239.254	255.254
Broadcast	15.255	31.255	47.255	53.255	…	239.255	255.255

Example 2 – Class B Address

- Network Address: 172.16.0.0, Subnet mask:255.255.255.128 (/25) (this is a tricky one)
  - 1111 1111 . 1111 1111 . **1111 1111 . 1**000 0000
- Subnets: 2^9 =512, Hosts: (2^7 – 2) = 126
- Valid subnets (the tricky bit) = 256 – 255 = 1
  - 0, 1, 2, 3, 4, etc for the 3rd byte & 1 bit for 4th byte
  - You get two subnets for each third byte bit value, e.g. when 3rd byte shows subnet 3, the two subnets would be 3.0 and 3.128
- This is a really useful subnet size, 512 subnets with 126 hosts each

Subnet	0.0	0.128	1.0	1.128	2.0	2.128	3.0	3.128	…	255.0	255.128
1st Host	0.1	0.129	1.1	1.129	2.1	2.129	3.1	3.129	…	255.1	255.129
Last Host	0.126	0.254	1.126	1.254	2.126	2.254	3.126	3.254	…	255.126	255.254
Broadcast	0.127	0.255	1.127	1.255	2.127	2.255	3.127	3.255	...	255.127	255.255

Transmission Control Protocol

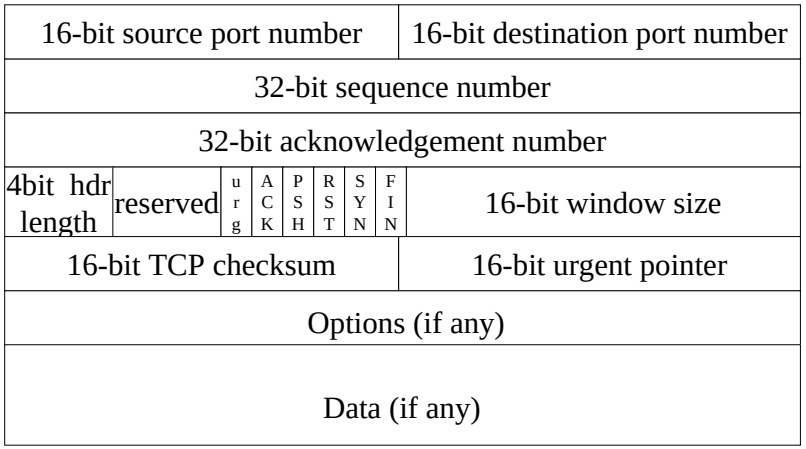
OSI Transport Layer

TCP Services

- Provides connection-oriented, reliable, byte stream service.
- Segments passed to IP for routing, timer attached for each segment.
- Sliding window protocol utilised with go-back-n or selective-repeat for retransmission.
- All TCP segments acknowledged.

- TCP segments may arrive out of order, sliding window will sort order.
- TCP segments may be duplicated, duplicated are discarded.
- TCP provides flow control, no process\host will be swamped, helps avoid congestion.
- TCP utilised by many internet applications such as Telnet, Rlogin, FTP, E-mail, WWW Browsers.

TCP Segment Header



TCP Header Description

- *Source Port* and *Destination Port* identify transport end-points of connection.
- *Sequence Number* and *Acknowledgement Number* perform usual functions, Ack numbers next byte expected.
- *TCP Header Length* indicates number of 32 bit words in header. Length varies because of options.
- Not used. No bug fixes required !

- Six one bit flags…
  - URGent pointer in use, used for indicating interrupts and offset from seq no. to urgent data.
  - ACK bit used to indicate piggybacked acknowledgement.
  - PSH requests that receiver does not buffer but to deliver.
  - RST is reset connection, means problems !
  - SYN used in conjunction with ACK to request connection.
  - FIN release connection

- *Window size* used for variable-sized sliding window. Size of zero indicates a choke packet.
- Checksum checks header.
- Options field for things like specification of maximum TCP payload. Negotiated at startup lowest bid wins.
- A *selective repeat* instead of *go-back-n* sliding window protocol may be specified as an option.

TCP Addressing

- TCP uses notion of Port Number to access transport endpoint on a single host. Many Ports may be in use simultaneously.
- Combination of IP address and port number uniquely identifies a port for process running on a particular machine.
- Process may even have several ports open.