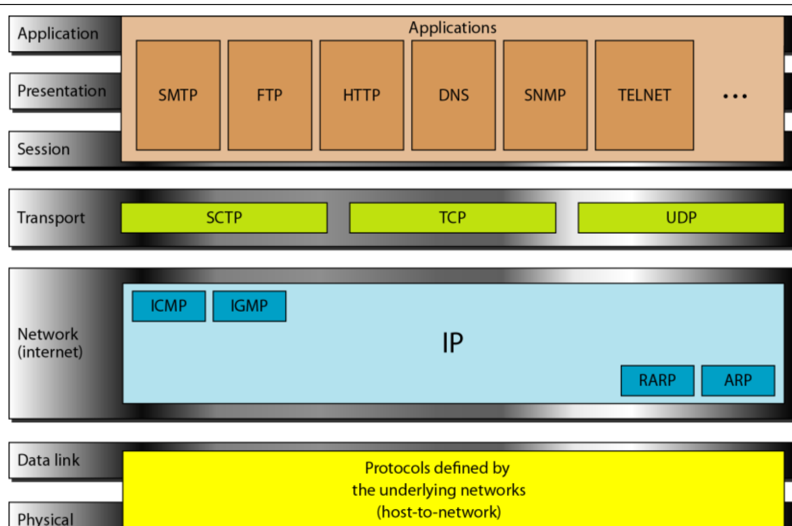
IP Protocols (IPv4 vs IPv6)

1

TCP/IP and OSI model



4

Network Layer

- The network layer is concerned with getting packets from the source all the way to the destination.
- Getting to the destination may require making many hops at intermediate routers along the way.
- The network layer is the lowest layer that deals with end-to-end transmission.
- To achieve its goals, the network layer must know about the topology of the network (i.e., the set of all routers and links) and choose appropriate paths through it, even for large networks.
- It must also take care when choosing routes to avoid overloading some of the communication lines and routers while leaving others idle.

The Network Layer Principles

1. Make sure it works
   - Do not finalize the design or standard until multiple prototypes have successfully communicated with each other.
2. Keep it simple
   - When in doubt, use the simplest solution. Put in modern terms: fight features. If a feature is not absolutely essential, leave it out, especially if the same effect can be achieved by combining other features.
3. Make clear choices
   - If there are several ways of doing the same thing, choose one. Having two or more ways to do the same thing is looking for trouble.

The Network Layer Principles

4. Exploit modularity
   - This principle leads directly to the idea of having protocol stacks, each of whose layers is independent of all the other ones.
5. Expect heterogeneity
   - Different types of hardware, transmission facilities, and applications will occur on any large network. To handle them, the network design must be simple, general, and flexible.
6. Avoid static options and parameters
   - If parameters are unavoidable (e.g., maximum packet size), it is best to have the sender and receiver negotiate a value rather than defining fixed choices.

The Network Layer Principles

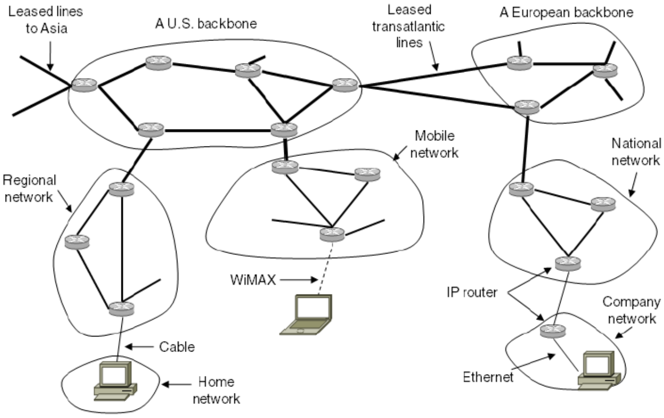
7. Look for good design (need not be perfect)
   - Often, the designers have a good design but it cannot handle some weird special case.
   - Rather than messing up the design, the designers should go with the good design and put the burden of working around it on the people with the strange requirements.
8. Strict sending, tolerant receiving
   - In other words, send only packets that rigorously comply with the standards, but expect incoming packets that may not be fully conformant and try to deal with them.

The Network Layer Principles

9. Think about scalability
   - If the system is to handle millions of hosts and billions of users effectively, no centralized databases of any kind are tolerable and load must be spread as evenly as possible over the available resources.
10. Consider performance and cost
    - If a network has poor performance or outrageous costs, nobody will use it.

10

The Network Layer in the Internet



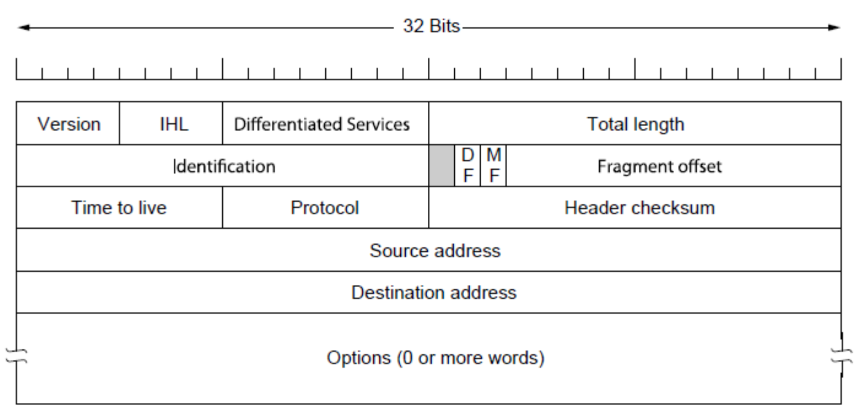
The Internet is an interconnected collection of many networks.

12

The IP Version 4 Protocol

- An appropriate place to start our study of the network layer in the Internet is with the format of the IP datagrams themselves.
- An IPv4 datagram consists of a header part and a body or payload part.
- The header has a 20-byte fixed part and a variable-length optional part. The header format is shown in next slide.
- The bits are transmitted from left to right and top to bottom, with the high-order bit of the Version field going first.
- This is a ''big-endian'' network byte order. On little-endian machines, such as Intel x86 computers, a software conversion is required on both transmission and reception.)

IPv4 datagram



The IPv4 (Internet Protocol) header.

IP datagram format (cont.)

- **Vers** (4 bits): version of IP protocol (IPv4=4)
- **IHL (**4 bits): Header length in 32 bit words, without options (usual case) = 20
- **Type of Service (Differentiated Services)-** (8 bits): little used in past, now being used for QoS
- **Total length** (16 bits): length of datagram in bytes, includes header and data
- **Time to live – TTL** (8bits): specifies how long datagram is allowed to remain in internet
  - Routers decrement by 1
  - When TTL = 0 router discards datagram
  - Prevents infinite loops
- **Protocol** (8 bits): specifies the format of the data area
  - Protocol numbers administered by central authority to guarantee agreement, e.g. TCP=6, UDP=17 …

16

IP Datagram format (cont.)

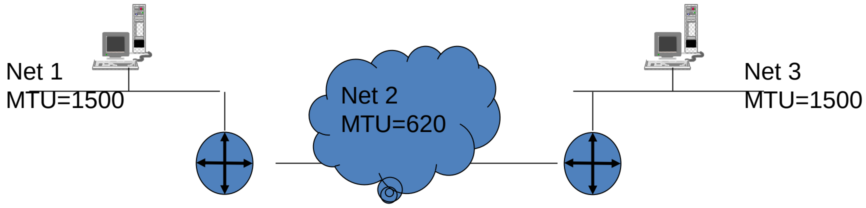
- **Source & destination IP address** (32 bits each): contain IP address of sender and intended recipient
- **Options** (variable length): Mainly used to record a route, or timestamps, or specify routing

17

17

IP Fragmentation

- How do we send a datagram of say 1400 bytes through a link that has a *Maximum Transfer Unit (MTU)* of say 620 bytes?
- Answer the datagram is broken into fragments

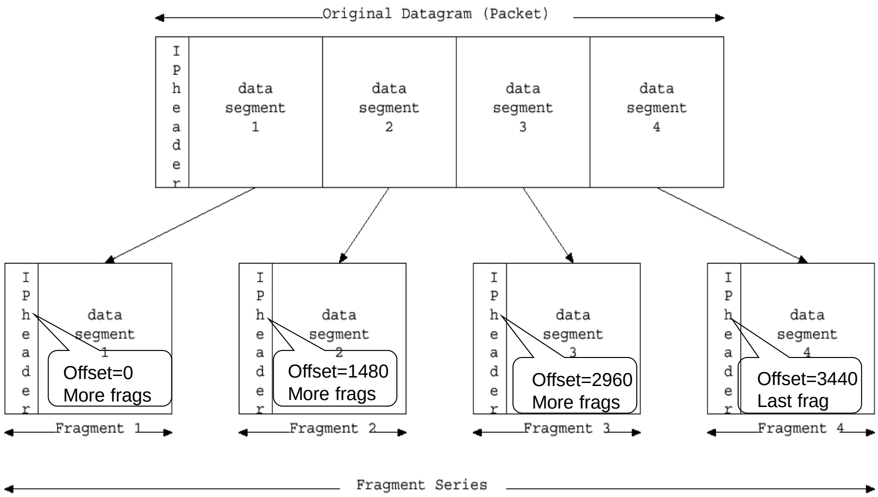


- Router fragments 1400 byte datagrams
  - Into 600 bytes, 600 bytes, 200bytes (note 20 bytes for IP header)
  - Routers do NOT reassemble, up to end host

Fragmentation Control

- **Identification**: copied into fragment, allows destination to know which fragments belong to which datagram
- **Fragment Offset** (13 bits): specifies the offset in the original datagram of the data being carried in the fragment
  - Measured in units of 8 bytes starting at 0
- **Flags** (3 bits): control fragmentation
  - Reserved (0-th bit)
  - Don't Fragment – DF (1st bit):
    - useful for simple (computer bootstrap) application that can't handle
    - also used for MTU discovery (see later)
    - if need to fragment and can't router discards & sends error to source
  - More Fragments (least sig bit): tells receiver it has got last fragment
- TCP traffic is hardly ever fragmented (due to use of MTU discovery). About 0.5% - 0.1% of TCP packets are fragmented .

Fragment series composition



NB. If data segment contains its own header that is not replicated

20

The IPv4 Datagram Options

Option	Description
Security	Specifies how secret the datagram is
Strict source routing	Gives the complete path to be followed
Loose source routing	Gives a list of routers not to be missed
Record route	Makes each router append its IP address
Timestamp	Makes each router append its address and timestamp

Some of the IP options.

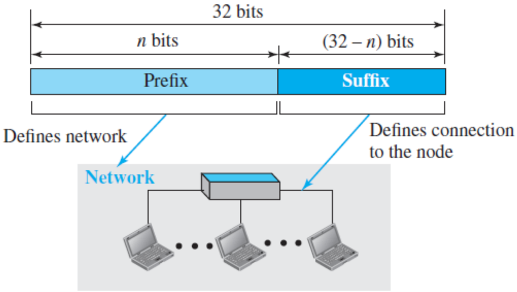
21

IP Addressing

- IP addresses refers to interface rather than host.
- Consists of network and host portions
  - Enables routers to keep 1 entry/network instead of 1/host
- There are two versions of IP addresses:

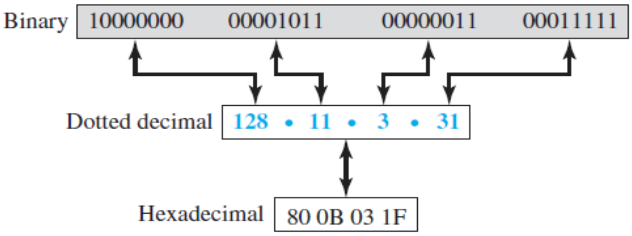
  1) IPv4 (IP version 4) : 32-bits in length

  2) IPv6 (IP version 6) : 128-bits in length



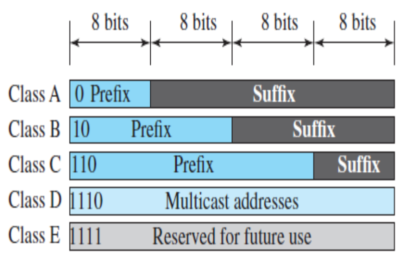
IPv4 Address Notation

- An IPv4 address is a 32-bit address that uniquely and universally defines the connection of a device to the Internet.
- There are two prevalent notations to show an IPv4 address: **binary notation** and **dotted decimal notation**. Hexadecimal notation is also used.



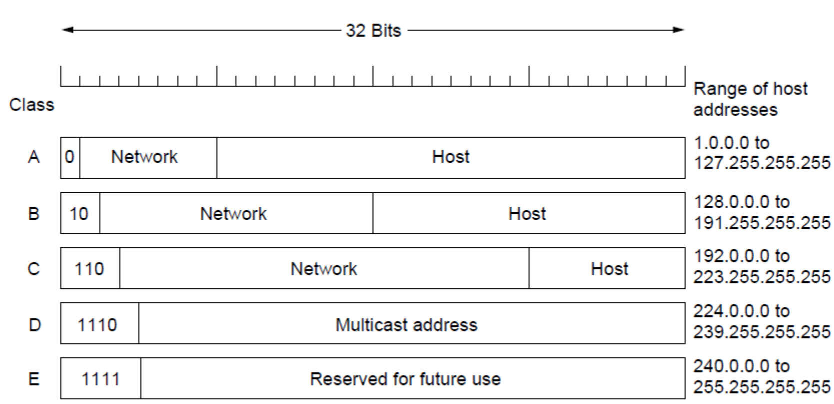
Classful Addressing

- In classful addressing, the address space is divided into five classes: A, B, C, D and E. Each class occupies some part of the address space.
  - Class A, B, C for unicast
  - Class D for multicast
  - Class E reserved



Class	Prefixes	First byte
A	$n = 8$ bits	0 to 127
B	$n = 16$ bits	128 to 191
C	$n = 24$ bits	192 to 223
D	Not applicable	224 to 239
E	Not applicable	240 to 255

Classful Addressing



IP address formats

Internet Class-based addresses

- Class A: large number of hosts, few networks
  - 0nnnnnnn hhhhhhhh hhhhhhhh hhhhhhhh
    - 7 network bits (0 and 127 reserved, so 126 networks), 24 host bits (> 16M hosts/net)
    - Initial byte 1-127 (decimal)
- Class B: medium number of hosts and networks
  - 10nnnnnn nnnnnnnn hhhhhhhh hhhhhhhh
    - 16,384 class B networks, 65,534 hosts/network
    - Initial byte 128-191 (decimal)
- Class C: large number of small networks
  - 110nnnnn nnnnnnnn nnnnnnnn hhhhhhhh
    - 2,097,152 networks, 254 hosts/network
    - Initial byte 192-223 (decimal)
- Class D: 224-239 (decimal) Multicast [RFC1112]
- Class E: 240-255 (decimal) Reserved

26

Subnet Mask

- A mask or default mask is a 32-bit number made of contiguous 1s followed by contiguous 0s.
- The default mask can help us to find the net-id and the host-id for the sub network.
- This concept does not apply to classes D and E.
- For example, the mask for a class A address has eight 1s, which means the first 8 bits of any address in class A define the net-id; the next 24 bits define the host-id.
- **Default Mask for classful addressing:**

CLASS	BINARY	DOTTED-DECIMAL
A	11111111 00000000 00000000 00000000	255.0.0.0
B	11111111 11111111 00000000 00000000	255.255.0.0
C	11111111 11111111 11111111 00000000	255.255.255.0

27

Subnet Mask Conversions

Prefix Length	Subnet Mask	Prefix Length	Subnet Mask
/1	128.0.0.0	/17	255.255.128.0
/2	192.0.0.0	/18	255.255.192.0
/3	224.0.0.0	/19	255.255.224.0
/4	240.0.0.0	/20	255.255.240.0
/5	248.0.0.0	/21	255.255.248.0
/6	252.0.0.0	/22	255.255.252.0
/7	254.0.0.0	/23	255.255.254.0
/8	255.0.0.0	/24	255.255.255.0
/9	255.128.0.0	/25	255.255.255.128
/10	255.192.0.0	/26	255.255.255.192
/11	255.224.0.0	/27	255.255.255.224
/12	255.240.0.0	/28	255.255.255.240
/13	255.248.0.0	/29	255.255.255.248
/14	255.252.0.0	/30	255.255.255.252
/15	255.254.0.0	/31	255.255.255.254
/16	255.255.0.0	/32	255.255.255.255

Decimal Octet	Binary Number
128	1000 0000
192	1100 0000
224	1110 0000
240	1111 0000
248	1111 1000
252	1111 1100
254	1111 1110
255	1111 1111

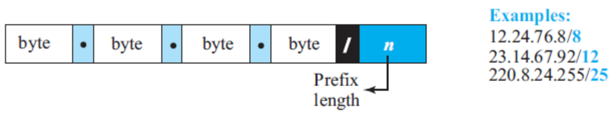
28

Classless Addressing(CIDR)

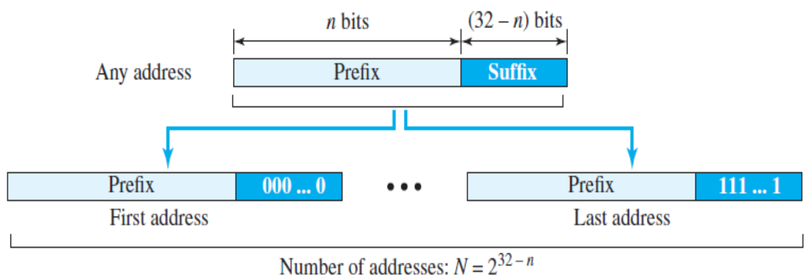
- In this scheme, there are no classes, but the addresses are still granted in blocks.
- In classless addressing, when an entity, needs to be connected to the Internet, it is granted a block (range) of addresses.
- The size of the block (the number of addresses) varies based on the nature and size of the entity.
- To **overcome** address depletion and give more organizations access to the Internet, classless addressing was designed and implemented.

29

Slash Notation in CIDR



Classless Address Information



Address depletion

- In 1991, Internet Architecture Board (IAB) identified 3 dangers
  - Running out of class B addresses
  - Increase in nets has resulted in routing table explosion
  - Increase in net/hosts exhausting 32 bit address space
- Four strategies to address
  - Creative address space allocation {RFC 2050}
  - Private addresses {RFC 1918}, Network Address Translation (NAT) {RFC 1631}
  - Classless Inter Domain Routing (CIDR) {RFC 1519}
  - IP version 6 (IPv6) {RFC 1883}

Creative IP address allocation

- Class A addresses 64 – 127 reserved
  - Handle on individual basis
- Class B only assigned given a demonstrated need
- Class C
  - divided up into 8 blocks allocated to regional authorities
  - 208-223 remains unassigned and unallocated
- Three main registries handle assignments
  - APNIC (Asia pacific Network Information Center) – Asia & Pacific www.apnic.net
  - ARIN (American Registry for Internet Numbers) – N. & S. America, Caribbean & sub-Saharan Africa www.arin.net
  - RIPE NCC (Réseaux IP Européens Network Coordination Centre) – Europe and surrounding areas www.ripe.net

32

Private IP Addresses

- IP addresses that are not globally unique, but used exclusively in an organization
- Three ranges:
  - 10.0.0.0 - 10.255.255.255 a single class A net
  - 172.16.0.0 - 172.31.255.255 16 contiguous class Bs
  - 192.168.0.0 – 192.168.255.255 256 contiguous class Cs
- Connectivity provided by Network Address Translator (NAT)
  - translates outgoing private IP address to Internet IP address, and a return Internet IP address to a private address
  - Only for TCP/UDP packets

33

Class InterDomain Routing (CIDR)

- Many organization have > 256 computers but few have more than several thousand
- Instead of giving class B (16384 nets) give sufficient contiguous class C addresses to satisfy needs
  - < 256 addresses assign 1 class C
  - …
  - < 8192 addresses assign 32 contiguous Class C nets

34

CIDR & Supernetting

- Since assigned contiguously, class C CIDR has same most significant bits & so only needs one routing table entry
- CIDR block represented by a prefix and prefix length
  - **Prefix** = single address representing block of nets, e.g
    - 192.32.136.0 = 11000000 00100000 10001000 00000000 while
    - 192.32.143.0 = 11000000 00100000 10001111 00000000

    21 bit prefix (2048 host addresses)
  - **Prefix length** indicates number of routing bits, e.g.

    192.32.136.0/21 means 21 bits used for routing
    - CIDR collects all nets in range 192.32.136.0 through 143.0 into a single router entry – **reduces router table entries**
- Removes address classes A, B & C boundaries
- For more details see RFC 1519

35