Internet Protocol

Reference
Sections 18.1, 18.2 and 18.4 Stallings
Study Guide 10

Learning Objectives

- describe the basic functions of networking protocols;
- understand the principles of internetworking;
- use and understand terms related to internetworking; and
- explain the network addressing and packet structures of the Internet Protocol.

Protocol Functions

- Small set of functions that form basis of all protocols
  - Encapsulation
  - Fragmentation and reassembly
  - Connection control
  - Ordered delivery
  - Flow control
  - Error control
  - Addressing
  - Multiplexing
  - Transmission services

Basic Protocol Functions

- Not all protocols have all functions; this would involve a significant duplication of effort
- However, there are many instances where the same type of function are present in protocols at different levels
- Protocol functions can be categorised as follows:
  - Encapsulation
    - For all protocols, data are transferred in blocks, called protocol data units (PDUs)

Basic Protocol Functions Contd.

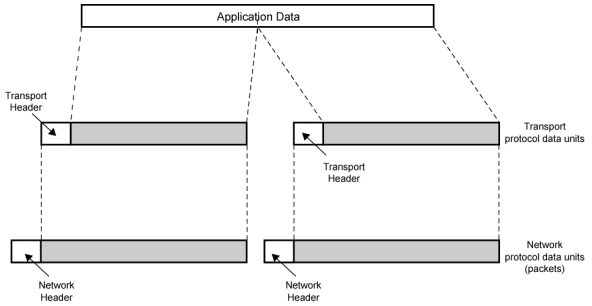
- Each PDU contains not only data but also control information
- The addition of control information to data is referred to as encapsulation

- Fragmentation and Reassembly
  - At the application level, we refer to a logical unit of data transfer as a message
  - Whether the application entity sends data in messages or in a continuous stream, lower level protocols may need to break the data up into blocks of some smaller bounded size
    - This process is called fragmentation

Basic Protocol Functions Contd.

- Some typical reasons for fragmentation are:
  - The communication network may only accept blocks of data up to a certain size
    - An ATM network is limited to blocks of 53 octets and Ethernet imposes a maximum size of 1526 octets
  - Error control may be more efficient with a smaller PDU size
    - Fewer bits need to be retransmitted when suffers from an error
  - More equitable access to shared transmission facilities, with shorter delay, can be provided
  - A smaller PDU size mean the receiving entities can allocate smaller buffers

Basic Protocol Functions Contd.



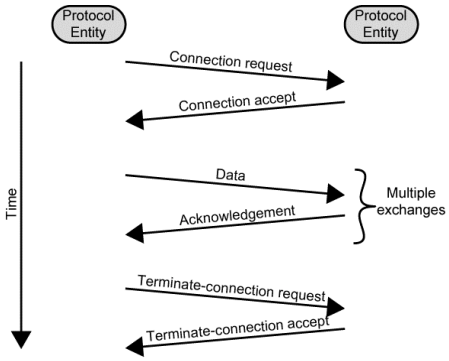
Basic Protocol Functions Contd.

- The arguments against fragmentation – to make larger PDUs- are:
  - Each PDU contains a certain amount of control information; smaller the blocks, the greater the percentage of overhead
  - PDU arrival may generate an interrupt that must be serviced; smaller blocks results in more interrupts
  - More time is spent processing smaller, more numerous, PDUs
- The counterpart of fragmentation is reassembly
  - The segmented data must be reassembled to messages appropriate o the application level, eventually
  - If the PDUs arrive out of order, the task is complicated

Basic Protocol Functions Contd.

- Connection Control
  - Connection-oriented data transfer is to be preferred (even required) if stations anticipate a lengthy exchange of data and/or certain details of their protocol must be worked out dynamically
  - In the above situations, a logical association, or connection, is established between the entities
  - The three phases occur with respect to a connection are:
    - Connection establishment
    - Data transfer
    - Connection termination

Basic Protocol Functions Contd.



Basic Protocol Functions Contd.

- A Key characteristic of many connection-oriented data transfer protocols is that sequencing is used
  - Each side sequentially numbers the PDUs that it sends to the other side
  - As each side remembers that it is engaged in a logical connection, it can keep track of both outgoing numbers, which it generates, and incoming numbers, which are generated by the other side
    - One can essentially define a connection-oriented data transfer as one in which both sides number PDUs and keep track of both incoming and outgoing numbers
  - Sequencing supports 3 main functions:
    - Ordered delivery, flow control, and error control

Basic Protocol Functions Contd.

- Ordered Delivery
  - If two communicating entities are in different hosts connected by a network, there is a risk that PDUs will not arrive in the order in which they were sent
    - This is because they may traverse different paths through the network
  - In connection-oriented protocols, it is generally required that PDU order be maintained
  - If each PDU is given a unique number, and numbers are assigned sequentially, then it is a logically simple task for the receiving entity to reorder PDUs on the basis of the sequence number

Basic Protocol Functions Contd.

- A problem with the above scheme is that, with a finite sequence number field, sequence numbers repeat (modulo some maximum number)
- The maximum sequence number must be greater than the maximum number of PDUs that could be outstanding at any time

- Flow Control
  - A function performed by a receiving entity to limit the amount or rate of data that is sent by a transmitting entity
  - Flow control is a good example of a function that must be implemented in several protocols
    - The network will need to exercise flow control over X via network access protocol, to enforce network traffic control

Basic Protocol Functions Contd.

- At the same time, Y's network access module has only limited buffer space and needs to exercise flow control over X's network access module via the transport protocol
- Finally, even though Y's network access module can control its data flow, Y's application may be vulnerable to overflow
  - The application may hung up waiting for disk access
  - Thus flow control is also needed over the application-oriented protocol

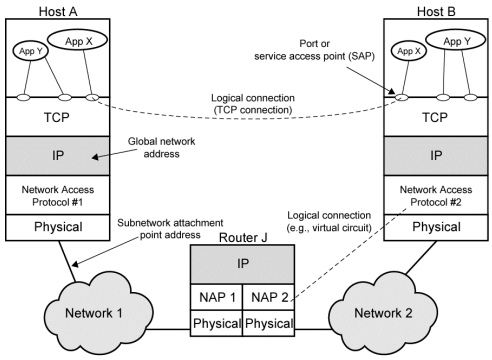
- Error Control
  - Error control techniques are needed to guard against loss or damage of data and control information

Basic Protocol Functions Contd.

- As with flow control, error control is a function that must be performed at various layers of protocol
  - The network access protocol should include error control to assure that data are successfully exchanged between station and network
  - However, a packet of data may be lost inside the network, and the transport protocol should be able to recover from this loss

- Addressing
  - The concept of addressing in a communications architecture is a complex one and covers the following issues:
    - Addressing level
      - Refers to the level in the communications architecture at which an entity is named

Basic Protocol Functions Contd.



Basic Protocol Functions Contd.

- Typically, a unique address is associated with each end system (e.g., workstation or server) and each intermediate system (e.g., router)
- Such an address, in general, is a network-level address
- In TCP/IP architecture, this is known as an IP address, or simply an internet address
- In OSI architecture, this is referred to as a network service access point (NSAP)
- The network-level address is used to route a PDU through a network or networks to a system indicated by a network-level address in the PDU
- Once data arrive at a destination system, they must be routed to some process or application in a system

Basic Protocol Functions Contd.

- Typically, a system will support multiple applications and an application may support multiple users
- Each application and, perhaps, each concurrent user of an application, is assigned a unique identifier, referred to as a port in the TCP/IP architecture and as a service access point (SAP) in the OSI architecture

- Addressing Scope
  - The internet address or NSAP address referred to previously is a global address
  - A global address identifies a unique system (global nonambiguity)
  - It is possible at any global address to identify any other global address, in any system, by means of the global address of the other system (global applicability)

Basic Protocol Functions Contd.

- Because a global address is unique and globally applicable, it enables an internet to route data from any system attached to any network to any other system attached to any other network
- Each network must maintain a unique address for each device interface on the network
- Examples are MAC address on an IEE 802 network and an ATM host address
- This address enables the network to route data units (e.g., MAC frames, ATM cells) through the network and deliver them to the intended attached system
- Such an address is referred to as a network attachment point address

Basic Protocol Functions Contd.

- The issue of addressing scope is generally only relevant for network-level addresses
- A port or SAP above the network level is unique within a given system but need not be globally unique

- Connection identifiers
  - The concept of connection identifiers comes into play when we consider connection-oriented data transfer (e.g., virtual circuit) rather than connectionless data transfer
  - For connectionless data transfer, a global identifier is used with each data transmission
  - For connection-oriented transfer, it is sometimes desirable to use only a connection identifier during data transfer phase

Basic Protocol Functions Contd.

- Addressing mode
  - Most commonly, an address refers to a single system or port; in this case it is referred to a s an individual or unicast address
  - It is also possible for an address to refer to more than one entity or port; such an address identifies multiple simultaneous recipients for data
  - An address for multiple recipients may be broadcast, intended for all entities within a domain, or multicast, intended for a specific subset of entities

- Multiplexing
  - One form of multiplexing is supported by means of multiple connections into a single system

Basic Protocol Functions Contd.

- For example, there can be multiple data link connections terminating in a single end system
  - We can say that these data link connections are multiplexed over the single physical interface between the end system and the network

- Multiplexing can also be accomplished via port names, which also permit multiple simultaneous connections
  - For example, there can be multiple TCP connections terminating in a given system, each connection supporting a different pair of ports
- Multiplexing is used in another context as well, namely, mapping of connections from one level to another

Basic Protocol Functions Contd.

- In a network, for each process to process connection established at the higher level, a data link connection could be created at the network access level
  - This is one-to-one relationship, but need not be so.
- Multiplexing can be used in on of two directions
  - Upward multiplexing, or inward multiplexing, occurs when multiple higher-level connections are multiplexed on, or share, a single lower-level connection
  - Downward multiplexing, or splitting, means that a single higher-level connection is built on top of multiple lower-level connections, the traffic on the higher connection being divided among the various lower connections

Basic Protocol Functions Contd.

- Transmission Service
  - A protocol may provide a variety of additional services to the entities that use it
  - Three common examples are:
    - Priority
      - Certain messages, such as control messages, may need to get through to the destination entity with minimum delay
      - Thus, priority could be assigned on a message basis, or on a connection basis
    - Quality of service
      - Certain classes of data may require a minimum throughput or a maximum delay threshold

Basic Protocol Functions Contd.

- Security
  - Security mechanisms, restricting access, may be invoked

- All of these services depends on the underlying transmission system and any intervening lower-level entities

Principles of Internetworking

- Packet-switching and packet-broadcasting networks grew out of a need to allow the computer user to have access to resources beyond that available in a single system
  - Resources of a single network are often inadequate to meet user's needs
- As the networks that might be interest exhibit so many differences, it is impractical to consider merging them into a single network
  - Rather, what is needed is the ability to interconnect various networks so that any 2 stations on any of the constituent networks can communicate

Principles of Internetworking Contd.

- An interconnected set of networks, from a user's point of view, may appear simply a large network
  - However, if each of the constituent networks retain its identity and special mechanisms are for communicating across multiple networks, then the entire configuration is often referred to as an internet
- Each constituent network in an internet supports communication among the devices attached to the network
  - These devices are referred to as end systems (ESs)

Principles of Internetworking Contd.

- In addition, networks are connected by devices referred to in the ISO documents as intermediate nodes (ISs)
  - ISs provide a communications path and perform the necessary relaying and routing functions so that data can be exchanged between devices attached to different networks in the internet
  - Two types of ISs of particular interest are bridges and routers
    - A bridge operates at layer 2 of the OSI 7 layer architecture and acts as a relay of frames between similar networks
    - A router operates at layer 3 of the OSI architecture and routes packets between potentially different networks

Principles of Internetworking Contd.

- The overall requirements for an internetworking facility are:
  - Provide a link between networks
    - At minimum, a physical and link control connection is needed
  - Provide for routing and delivery of data between processes on different networks
  - Provide an accounting service that keeps track of the use of various networks and routers and maintains status information
  - Provide the services just listed in such a way as not to require modifications to the networking architecture of any of the constituent networks

Principles of Internetworking Contd.

- This means that the internetworking facility must accommodate a number of differences among networks:
  - Different addressing schemes
    - The networks may use different endpoint names and address and directory maintenance schemes
    - Some form of global network addressing must be provided, as well as a directory service
  - Different maximum packet size
    - Packets from one network may have to be broken up into smaller pieces for another; this process is referred to as fragmentation
  - Different network access mechanisms
    - The network access mechanism between station and network may be different for stations on different networks

Principles of Internetworking Contd.

- Different timeouts
  - Typically, a connection-oriented transport service will await an acknowledgment until a timeout expires, at which it will retransmit its block of data
  - In general, longer times are required for successful delivery across multiple networks
  - Internetwork timing procedures must allow successful transmission that avoids unnecessary retransmissions
- Error recovery
  - Network procedures may provide anything from no error recovery up to reliable end-to-end (within the network) service

Principles of Internetworking Contd.

  - The internetwork service should not depend on nor be interfered with by nature of the individual network's error recovery capability
- Status reporting
  - Different networks report status and performance differently
  - It must be possible for the internetworking facility to provide such information on internetworking activity to interested and authorised processes
- Routing techniques
  - Internetwork routing may depend on fault detection and congestion control techniques peculiar to each network
  - The internetworking facility must be able to coordinate these to route data adaptively between stations on different networks

Principles of Internetworking Contd.

- User access control
  - Each network will have its own user access control technique
  - These must be invoked by the internetwork facility as needed
  - Further, a separate internetwork access control technique may be required
- Connection, connectionless
  - Individual networks may provide connection-oriented or connectionless service
  - It may be desirable for the internetwork service not to depend on the nature of the connection service of the individual network

Principles of Internetworking Contd.

- A key characteristic of an internet architecture is whether the mode of operation is connection oriented or connectionless
  - Connection-oriented operation
    - It is assumed that each network provides a connection-oriented form of service
      - That is, it is possible to establish a logical network connection between any two end systems attached to the same network
    - ISs are used to connect 2 or more networks
      - Each IS appears as an ES to each of the network to which it is attached

Principles of Internetworking Contd.

- When ES A wishes to exchange data with ES B, a logical connection is set up between them
  - This connection consists of the concatenation of a sequence of logical connections across networks
- The individual network logical connections are spliced together by ISs
  - Any traffic arriving at an IS on one logical connection is retransmitted on a second logical connection and vice versa
- A connection oriented IS performs the following key functions
  - Relaying
    - Data units arriving from one network via the network layer protocol are relayed (retransmitted) on another network

Principles of Internetworking Contd.

  - Routing
    - When an end-to-end logical connection consisting of a sequence logical connections, is to be set up, each IS in the sequence must make a routing decision that determines the next hop in the sequence
- Thus, at layer 3, a relaying operation is performed
  - It is assumed that all of the end systems share common protocols at layer 4 and above for successful end-to-end communication
- Connectionless Operation
  - Connectionless-mode operation corresponds to the datagram mechanism of a packet-switching network
  - Each network protocol data unit is treated independently and routed from source ES to destination ES through a series of routers and networks

Principles of Internetworking Contd.

- For each data unit transmitted by A, A makes a decision as to which router should receive the data unit
- The data unit hops across the internet from one router to the next until it reaches the destination network
  - At each router a routing decision is made (independently for each data unit) concerning the next hop
    - Thus, different data units may travel different routes between source and destination ES
- All Ess and routers share a common network-layer protocol known generally as the internet protocol
- An Internet Protocol (IP) was initially developed for the DARPA internet project and published as RFC 791 and has become an Internet Standard

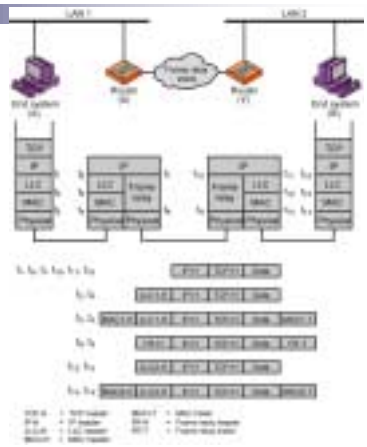
Connectionless Internetworking

- In this section we refer specifically to the Internet Standard IP, but it applies any connectionless Internet Protocol, such as IPv6
- IP provides connectionless, or datagram, service between end systems
- The advantages of this approach are:
  - Connectionless internet facility is flexible
    - It can deal with a variety of networks, some of which are themselves connectionless
    - In essence, IP requires very little from the constituent networks

Connectionless Internetworking Contd.

- A connectionless internet service can be made highly robust
  - This is basically the same argument made for a datagram network service versus a virtual circuit service
- A connectionless internet service is best for connectionless transport protocols, as it does not impose unnecessary overhead

Connectionless Internetworking Contd.



Connectionless Internetworking Contd.

- The figure in the previous slide depicts a typical example using IP, in which two LANs are interconnected by a frame relay WAN
- End System A has a datagram to transmit to end system B
  - The datagram includes the internet address of B
- The IP module in A recognises that the destination B is on another network
  - So the first step is to send the data to a router, in this case router X

Connectionless Internetworking Contd.

- To send data to router X, IP passes the datagram down to the next lower layer (in this case LLC) with instruction to send it to router X
- LLC in turn passes this information down to MAC layer, which inserts the MAC-level address of router X into the MAC header
- When the packet reaches router X, the router removes MAC and LLC fields and analyse the IP header to determine the ultimate destination of the data – in this case B

Connectionless Internetworking Contd.

- The router must now make a routing decision; there are 3 possibilities
  - The destination station B is connected directly to one of the networks to which the router is attached
    - If so, the router sends the datagram directly to the destination
  - To reach the destination, one or more additional routers must be traversed
    - If so, a routing decision must be made: to which router the datagram must be sent?

Connectionless Internetworking Contd.

    - In both above cases, the IP module in the router sends the datagram down to the next lower layer with the destination network address
  - The router does not know the destination address
    - Router returns an error message to the source of the datagram
- In the above example, the data must pass through router Y before reaching the destination
  - So the router X constructs a new frame by appending a frame relay header and trailer to the IP data unit
  - The frame relay header indicates a logical connection to router Y

Connectionless Internetworking Contd.

- When the frame arrives at router Y, the frame header and the trailer are stripped off
  - The router determines that this IP data unit is destined for B, which is connected directly to a network to which this router is attached
  - The router therefore creates a frame with layer-2 destination address of B and sends it out onto LAN 2
- The data finally arrive at B, where the LAN and IP headers can be stripped off

Connectionless Internetworking Contd.

- At each router, before the data can be forwarded, the router may need to fragment the data unit
  - This is done to accommodate a smaller maximum packet size limitation on the outgoing network
- The data units split into two or more fragments, each of which becomes an independent IP data unit
- Each new data unit is wrapped in a lower-layer packet and queued for transmission

Connectionless Internetworking Contd.

- The process described above continues through as many routers as it takes for the data unit to reach its destination
- As with routers, the destination end systems recovers the IP data unit from its network wrapping
- If fragmentation has occurred, the IP module in the destination end system buffers the incoming data until the entire original data field can be reassembled

Connectionless Internetworking Contd.

- The service offered by IP is an unreliable one
  - That is, IP does not guarantee that all data will be delivered or that the data that are delivered will arrive in the proper order
    - It is the responsibility of the next higher layer (e.g., TCP) to recover from any errors that occur
  - This approach provides a great deal of flexibility
- As the sequence of delivery is not guaranteed, successive data units can follow different paths through the internet
  - This allows the protocol to react to both congestion and failure in the internet by changing routes

Internet Protocol

- In this section, we will look at version 4 of IP, officially defined in RFC 791
- Although it is intended that IPv4 will eventually be replaced by IPv6, it is currently the standard IP used in TCP/IP networks
- As with any protocol standard, IP is specified in two parts:
  - The interface with higher layer (e.g., TCP), specifying the services that IP provides
  - The actual protocol format and mechanisms

Internet Protocol Contd.

- The services to be provided across adjacent protocol layers (e.g., IP and TCP) are expressed in terms pf primitives and parameters
  - A primitive specifies the function to be performed
    - The actual form of a primitive is implementation dependent
    - An example is a subroutine call
  - Parameters are used to pass data and control information
- IP provides two service primitives at the interface to the interface to the next higher layer

Internet Protocol Contd.

  - The send primitive is used to request transmission of a data unit
  - The delivery primitive is used by IP to notify a user of the arrival of data unit
- The parameters associated with the two primitives are as follows:
  - Source address
  - Destination address
  - Protocol
    - Recipient protocol entity (such as TCP)

Internet Protocol Contd.

  - Type of service indicators
    - Used to specify the treatment of the data unit in its transmission through component networks
  - Identification
    - Used in combination with the source and destination addresses and user protocol to identify the data unit uniquely
    - This parameter is required for reassembly and error reporting
  - Don't fragment identifier
  - Time to live

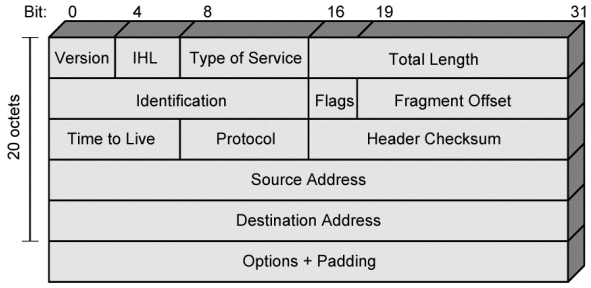
Internet Protocol Contd.

  - Data length
  - Option data
  - Data
- The identification, don't fragment identifier, and time to live parameters are present in the Send primitive but not in the Deliver primitive
  - These 3 parameters provide instructions to IP that are not of concern to the recipient IP user

Internet Protocol Contd.

- The options parameter allows for future extensibility and inclusion of parameters that are usually not invoked
  - The currently defined options are
    - Security
      - Allow a security label to be attached to a datagram
    - Source routing
      - A sequenced list of router addresses that specifies the route to be followed
    - Route recording
    - Stream identification
    - Timestamping

Internet Protocol Contd.



Internet Protocol Contd.

- The protocol between IP entities is best described with reference to IP datagram format, shown in the previous slide
- The fields are:
  - Version
    - Indicates version number, to allow evolution of the protocol; the value is 4
  - Internet Header Length (IHL)
    - The length of header in 32-bit words
    - The minimum value is 5, for minimum header length of 20 octets

Internet Protocol Contd.

- Type of Service
  - Specifies reliability, precedence, delay, and throughput parameters
  - This field is rarely used
- Total length
  - Total datagram length, in octets
- Identification
  - A sequence number that, together with the source address, destination address, and user protocol, is intended to identify a datagram uniquely
  - Thus this number should be unique for the datagram's source address, destination address, and user protocol for the time during which the datagram will remain in the internet

Internet Protocol Contd.

- Flags
  - Only 2 bits are currently used
    - The more bit is used for fragmentation and reassembly
    - The Don't fragment bit prohibits fragmentation when set
- Fragment Offset
  - Indicates where in the original datagram this fragment belongs, measured in 64-bit units
  - This implies that fragments other than the last fragment must contain data field that is a multiple of 64 bits in length
- Time to Live
  - Specifies how long, in seconds, a datagram is allowed to remain in the internet

Internet Protocol Contd.

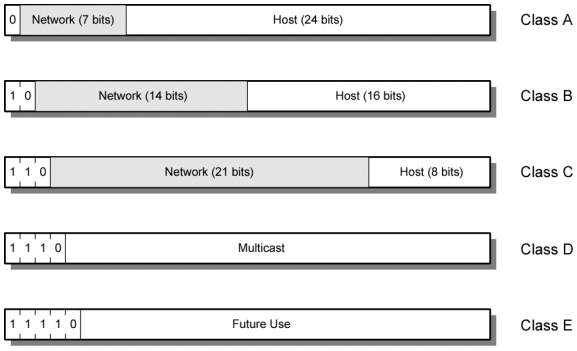
- Protocol
  - Indicates the next higher level protocol that is to receive the data field at the destination
- Header Checksum
  - An error-detecting code applied to the header only
  - Because some header fields may change during transit, this is reverified and recalculated at each router
- Source Address
- Destination Address
- Options

Internet Protocol Contd.

- Padding
  - Used to ensure that the datagram header is a multiple of 32 bits in length
- Data
  - Must be an integer multiple of 8 bits in length
  - The maximum length of that datagram is 65,535 octets

- The source and destination address fields in the IP header each contain a 32-bit global internet address, generally consisting of a network identifier and a host identifier
- The address is coded to allow a variable allocation of bits to specify network and host, as shown in the next slide

Internet Protocol Contd.



Internet Protocol Contd.

- This encoding provides flexibility in assigning addresses to hosts and allows a mix of network sizes on an internet
- The 3 principal network classes are best suited to the following conditions:
  - Class A
    - Few networks, each with many hosts
  - Class B
    - Medium number of networks, each with a medium number of hosts
  - Class C
    - Many networks, each with a few hosts

Internet Protocol Contd.

- A mixture of classes is appropriate for an internetwork consisting of a few large networks, many small networks, plus some medium-sized networks
- IP addresses are usually written in what is called dotted decimal notation, with a decimal number representing each of the octets of 32-bit address
  - For example, the IP address 11000000 11100100 00010001 00111001 is written as 192.228.17.57
- All class A network addresses begin with a binary 0

Internet Protocol Contd.

- Network addresses with a first octet of 0 (00000000) and 127 (01111111) are reserved
  - So there are 126 potential Class A network numbers, which have a first decimal number in the range 1 to 126
- Class B network addresses begin with a binary 10
  - So the range of first decimal numbers in a class B address is 128 to 191 (binary 10000000 to 10111111)
  - The 2nd octet is also part of the Class B address
    - So there are 2^{14} = 16,384 Class B addresses

Internet Protocol Contd.

- For Class addresses, the first decimal number ranges from 192 to 223 (11000000 to 11011111)
- The total number of Class C addresses is 2^{21}= 2,097,152

Type of Service

- Sending IP user may request types of service parameter
  - One or more services
    - Precedence
    - Reliability
    - Delay
    - Throughput
- It can be used for routing priorities
- Passed down to Network access layer

Type of Service…

- Precedence – relative importance
  - 8 levels (higher number better preference)
- Reliability
  - Normal or high : High - datagram should not be lost
- Delay
  - Normal or low : low – minimum delay should be experienced
- Throughput
  - Normal or high
- If precedence selected and network supports this feature
  - Mapping to network level for this hop

Subnets and Subnet Masks

- Allow arbitrary complexity of internetworked LANs within organization
- Insulate overall internet from growth of network numbers and routing complexity
- Site looks to rest of internet like single network
- Each LAN assigned subnet number
- Host portion of address partitioned into subnet number and host number
- Local routers route within subnetted network
- Subnet mask indicates which bits are subnet number and which are host number
  - 255.255.255.0 Subnet mask

Routing Using Subnets (255.255.255.224)