Internetwork Protocols

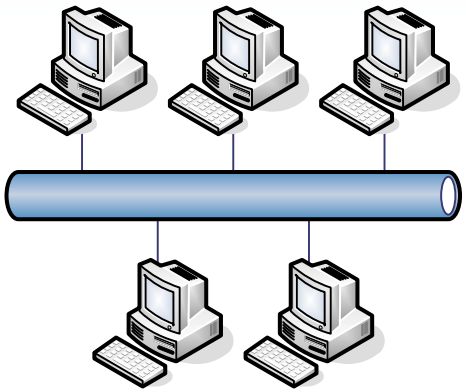
Dr Steve Gordon
ICT, SIIT

Contents

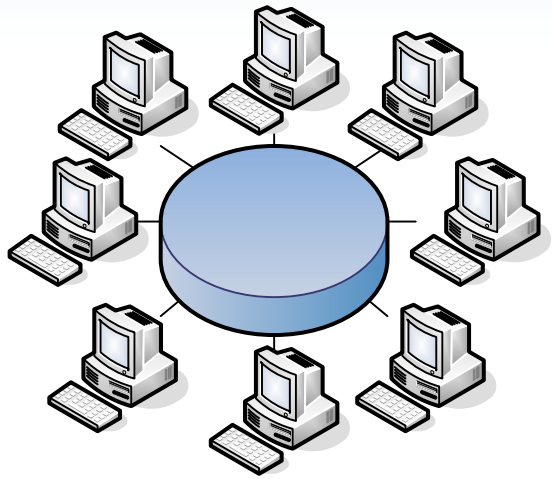
- Motivation for Internetworking
- Basic Functions of Internetworking Protocols
- Internetworking with the Internet Protocol
- IP Addresses
- Other Network Layer Functions

Motivation for Internetworking

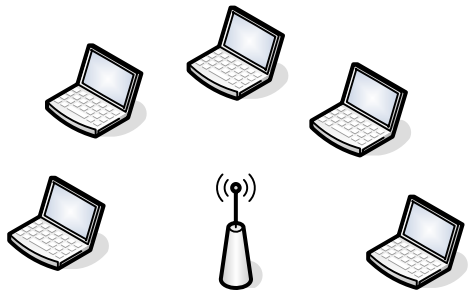
Local Area Networks



Ethernet



Token Ring



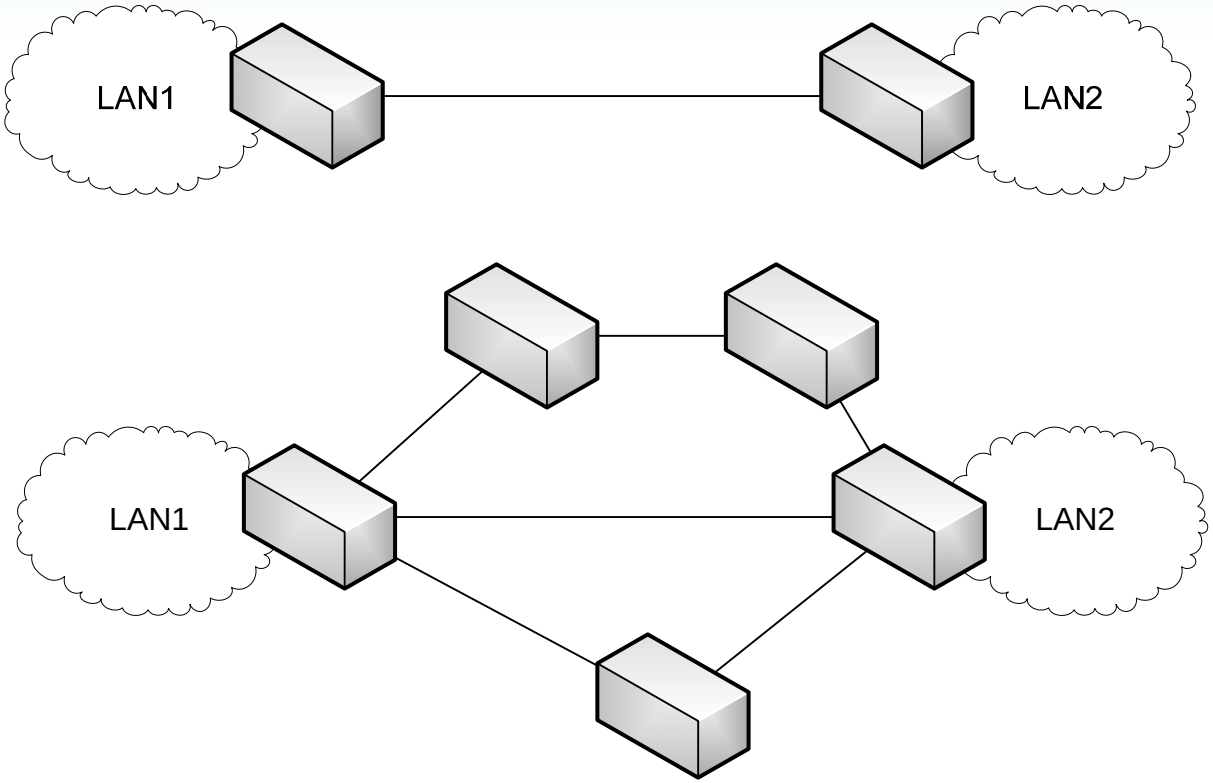
Wireless LAN

Different types of LANs: different topologies, different technologies, different purposes

Many LANs operate at layers 1 and 2 (Physical and Data Link Layer) using switches and hubs

(There are techniques for connecting 2 or more LANs (e.g. bridging) between it cannot be used for all networks and becomes complex)

Wide Area Networks



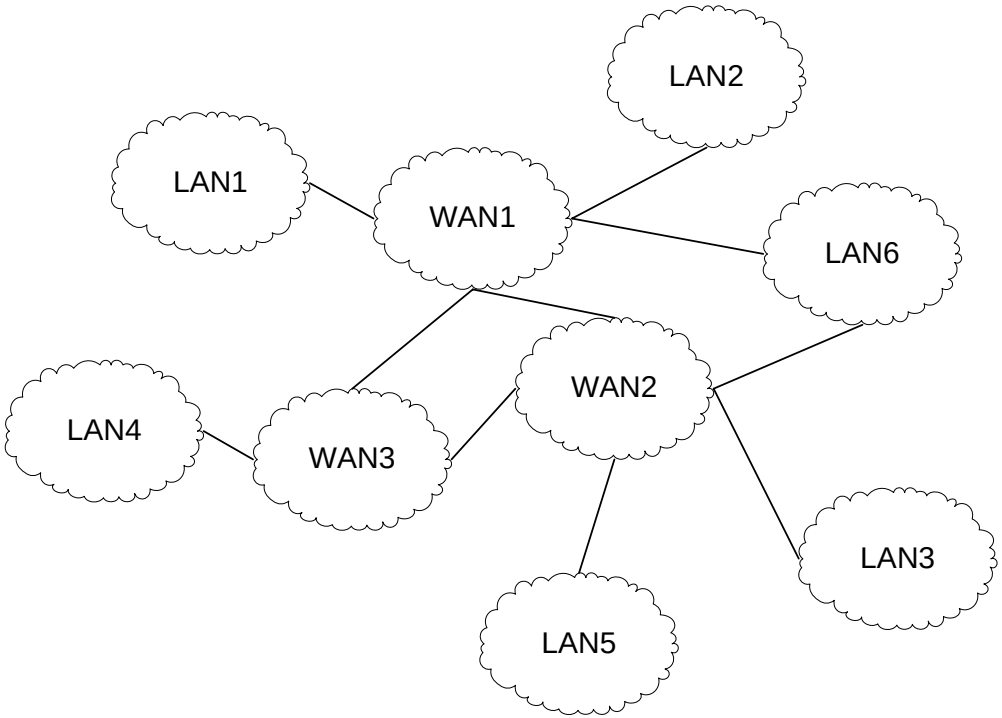
WANs can interconnect LANs over a larger distance. The WAN can either be a point-to-point link (e.g. ADSL, PDH) or a network (e.g. ATM, SDH, telephone) using packet or circuit switching.

The device that interconnects the WAN to LAN must support both technologies (e.g. an interface for IEEE 802.3 Ethernet and an interface for PDH)

WANs typically operate at Layers 1 and 2

Many Different LANs/WANs

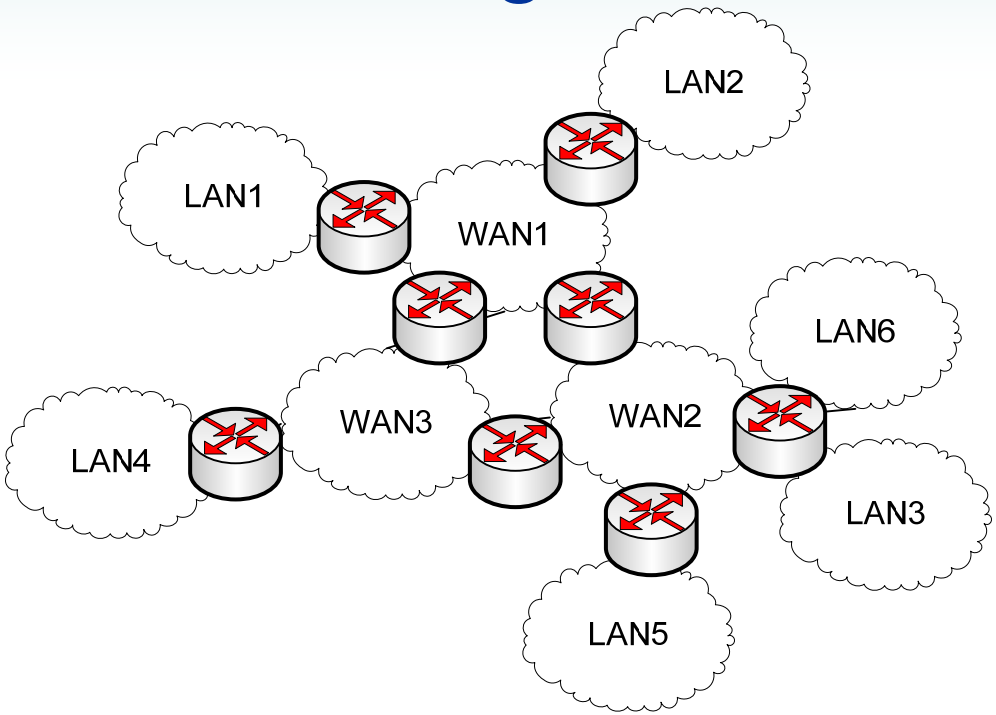
Organisations have different requirements of their network, and therefore may choose different technologies for their LANs/WANs



The aim is to allow any computer to communicate with any other computer, independent of what LAN/WAN they are connected to.

Internetworking involves connecting the many different types of LANs/WANs together to achieve this aim.

Internetworking with Routers



Internetworking is performed using **routers**.

Routers connect two or more LANs or WANs together. Routers are packet switches that operate at **Layer 3** (the Networking layer)

Terminology

- Terminology in data communications is sometimes ambiguous (e.g. one word means two different things, types of nodes have two different names)
- We will try to use the following terminology:
  - **Routers**: nodes that connect networks (LANs/WANs) together. Operate at Layer 3 (Networking)
  - Subnetworks (or **subnets**): individual networks (LANs and WANs)
  - **Internetworking**: connect two or more subnets together using routers
  - An internetwork or an **internet**: the resulting network from internetworking
  - The **Internet**: an internet that uses the Internet Protocol (IP) and used today to connect networks/computers across the globe
  - **Routing**: process of discovering a path from source to destination through a network
  - **Forwarding**: process of sending data along a path through a network
  - **Packet Switch**: a generic device that performs switching in a Packet Switching network. May operate at layers 2 or 3. A packet switch at layer 3 is called a router
  - **Circuit Switch**: a generic device that performs circuit switching in a Circuit Switching network
  - **Ethernet switch**: an IEEE 802.3 switch (either Ethernet, Fast Ethernet or Gigabit Ethernet). Operates at layer 2 (Data Link Layer)

Basic Functions of Internetworking Protocols

Basic Functions of Internetworking

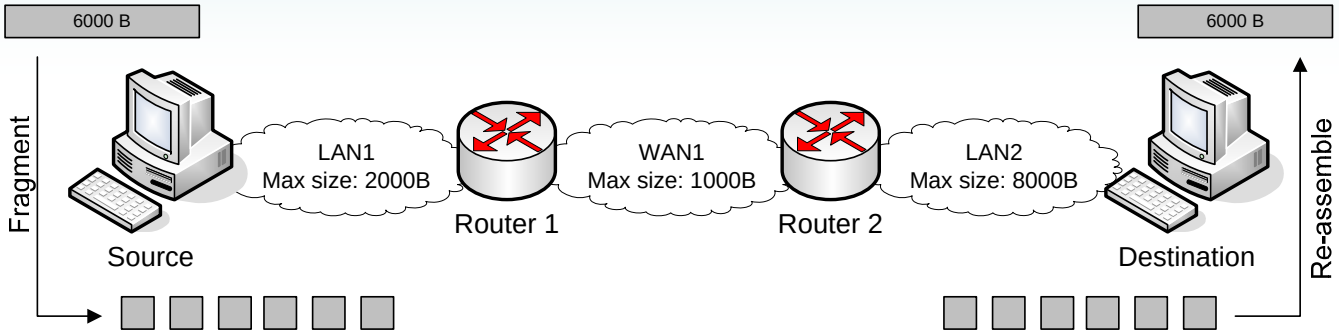
- Various design options for an internetworking protocol
  - In an internet, there are many functions that need to be considered
  - (We have already seen most of these functions in other protocols, e.g. at data link layer)
- The following are some of the basic functions an internetworking protocol may implement
  - Fragmentation and re-assembly
  - Connection control
  - Ordered delivery
  - Flow control
  - Error control
  - Addressing
  - Multiplexing
  - Priority
  - Quality of service
  - Security
- Not all internetworking protocols implement all functions

Fragmentation and Reassembly

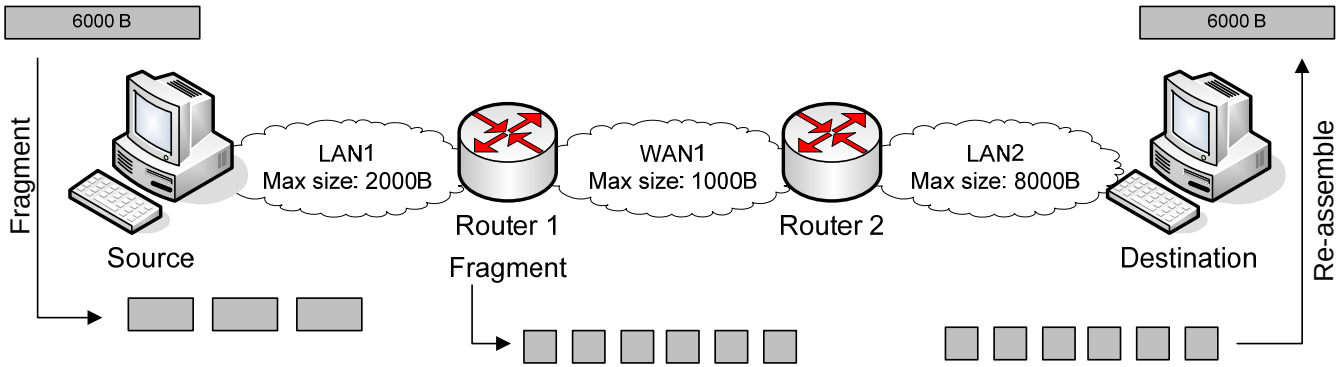
- An LAN or WAN (or internetworking) protocol may limit the size of a packet
  - Therefore, if the amount of data to send is larger than what can be carried in a maximum sized packet, require functions to fragment the data before sending and re-assemble when receiving
- Why is there a limit on the size of the packet?
  - LAN/WAN technologies make trade-offs in packet size considering efficiency, errors, fairness, processing time, buffer sizes, …
  - Different LAN/WANs may have different maximum frame sizes
    - E.g. IEEE 802.3 has maximum size of 1518 bytes; X.25 limits to 128 bytes of data; ATM uses 53 byte frames; …
- Different options for fragmenting a packet
  - Fragment at the source node only: if the source node knows the minimum maximum frame size along the path to destination, then can fragment the data into packets
  - Fragment at the source and routers: if the sending device (source or router) knows about the maximum frame size on the next network, then can fragment packets if necessary
  - Re-assembly can potentially occur only at the destination, or at each device (routers and destination)

Fragmentation and Reassembly

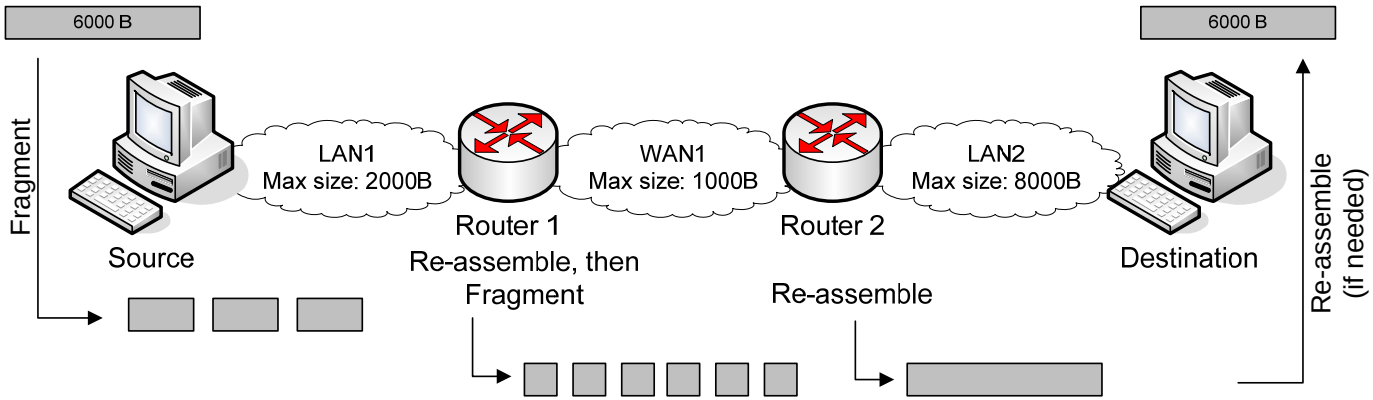
Fragment and re-assemble only at source and destination



Fragment at source and routers; re-assemble at destination



Fragment at source and routers; re-assemble at routers and destination

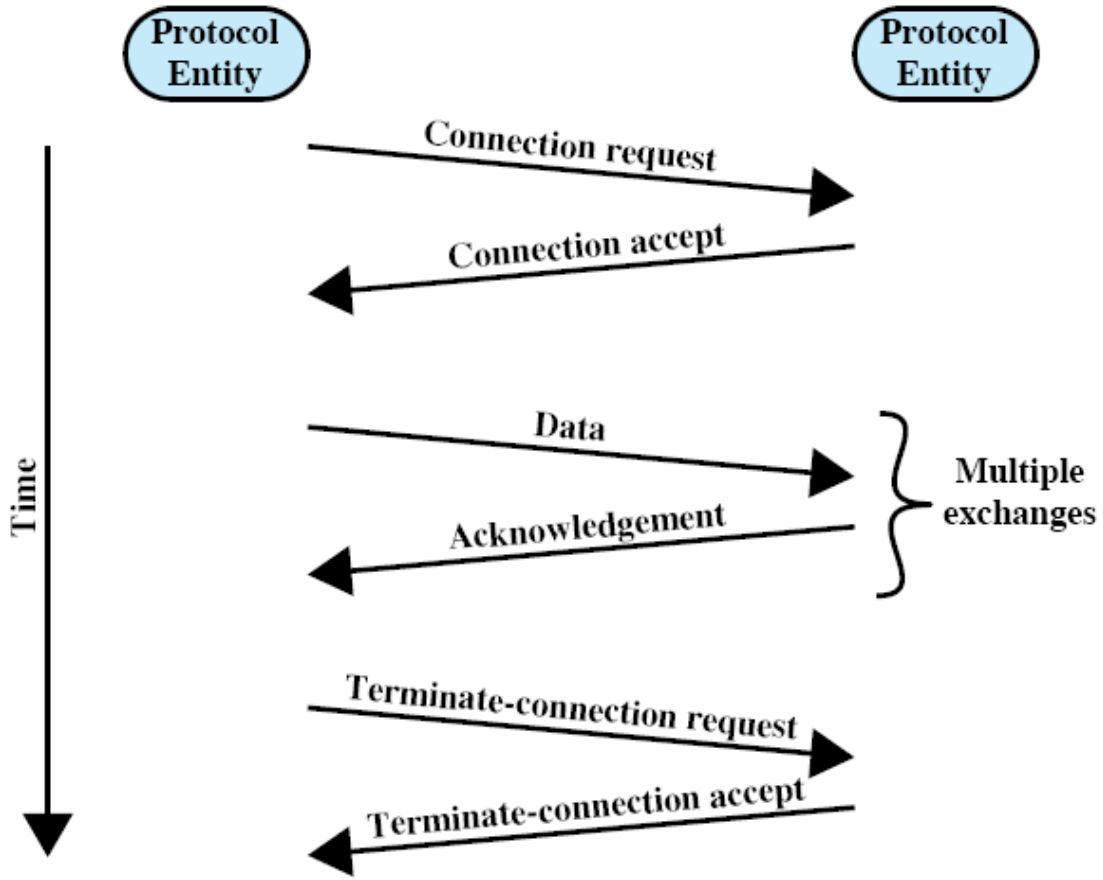


Connection-oriented or Connection-less

- Two modes of operation with respect to connections between source and destination
  - Connection-oriented: a logical connection is created between source and destination for data transfer
    - Packets are associated with each other
    - E.g. a connection is established to transfer a 10MB file from source to destination, or to participate in a video conference call
    - Involves: connection establishment (setup); data transfer; connection termination (teardown, close)
    - A connection will normally have some parameter values associated with it (e.g. maximum packet size, encryption algorithms and keys, flow control algorithms)
  - Connection-less: no connection between source and destination
    - Packets are treated independently by the internetworking protocol

- (Note: these *concepts* are similar to (virtual) circuit vs datagram packet switching; however they are not identical; for example, later we will see that a connection-oriented protocol can be used in a datagram packet switching network)

Example: Connection-Oriented Phases



Ordered Delivery

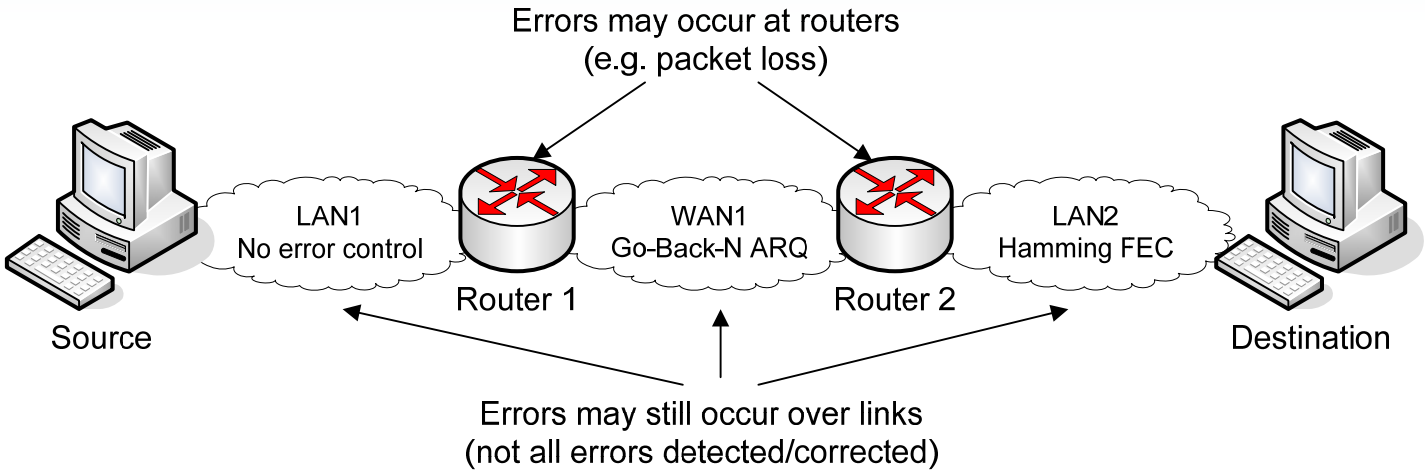
- Usually a sequence of packets should be delivered to the receiver in order
  - E.g. a sender application has a file that is sent as 10 packets; the receiving application must receive the 10 packets in order to create the original file
- But networks can sometimes deliver packets out of order
  - E.g. in a datagram packet switching network, one packet may take one path, and the second packet may take a different (shorter) path arriving first
- Ordering is often achieved using sequence numbers; each packet is given a sequence number
  - The receiver then knows which order to reassemble the PDUs
  - The sequence number is carried in the header
    - Note that there is a finite number of bits available in the header, hence the sequence number is limited
    - "Wrap" the sequence number (e.g. for a 7 bit sequence number):
      - » 0, 1, 2 , 3, …., 126, 127, 0, 1, 2, …

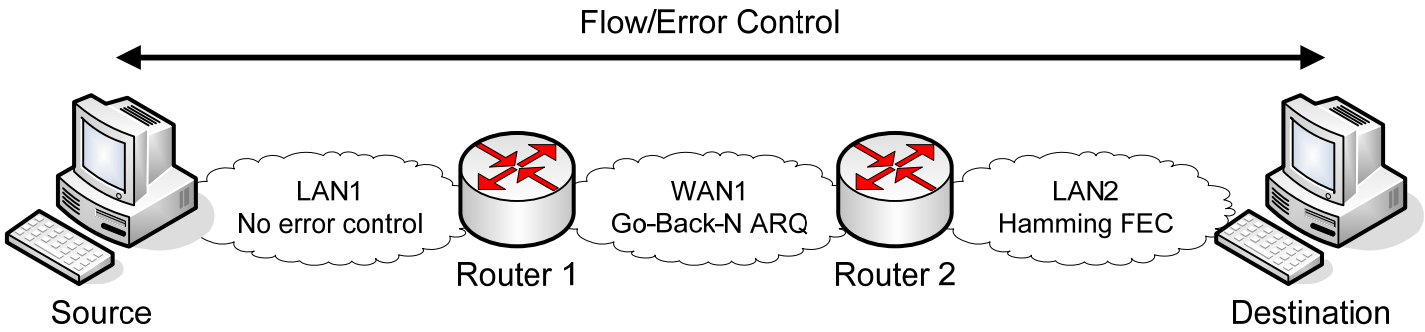
Flow and Error Control

- Flow Control: limit the rate at which source sends so that destination is not overflowed
- Error Control: guard against the loss or damage of data

- We have seen examples of flow and error control across links (Stop-and-Wait, Go-Back-N, FEC)
- Why may an internetworking protocol also need flow/error control?
  - Not all LAN/WAN protocols use flow/error control, and not the same algorithms (same may be better than others)
  - Even with error control (e.g. Hamming FEC), errors are still possible
  - Errors such as packet loss, may occur in routers
    - If a router is busy and its buffer is full of packets, newly arriving packets may be discarded

Flow and Error Control





Addressing

- LAN/WAN technologies may use different addressing schemes for devices
  - IEEE 802 is very common, but not the only scheme
- An internetworking protocol should define an addressing scheme so that all nodes can communicate, independent of the LAN/WAN addressing scheme
- Addresses should be globally unique

Other Services

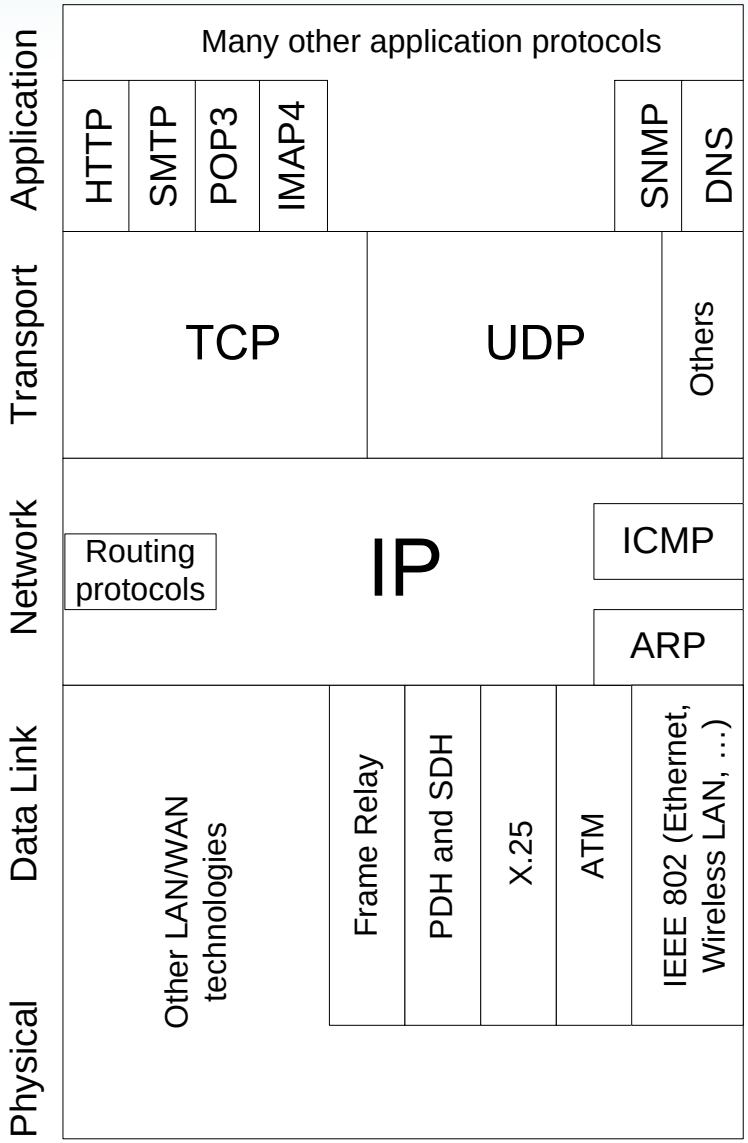
- Multiplexing
  - Allow multiple types of data to be sent over an internet
- Priority
  - Support giving priority to some data over other data
- Quality of service
  - Support guarantees of performance for data
- Security
  - Ensure data is transferred in secure manner (confidential, authentication, …)

Internetworking with the Internet Protocol (IP)

The Internet Protocol (IP)

- IP is one internetworking protocol
  - Initially developed by US Department of Defence and published as RFC791 in 1981 (45 pages)
  - It is now an Internet Standard produced by IETF, and various enhancements have been published and standardised
- IP is the internetworking protocol used in the Internet
  - There are (were) other internetworking protocols
    - IPX (Novell), X.25 (OSI), CLNP (OSI), SCCP (ITU)
- IP Features:
  - Connection-less, network layer, internetworking protocol using datagram packet switching
    - No connection control, no error control, no flow control, no ordered delivery, no priority, no quality of service, no security
      - These functions are left to other protocols/layers, e.g. TCP in Transport Layer, as well as extensions of IP, e.g. IPsec provides security
    - What does it do? Delivery of IP datagrams, multiplexing, addressing, fragmentation/re-assembly
  - Simple

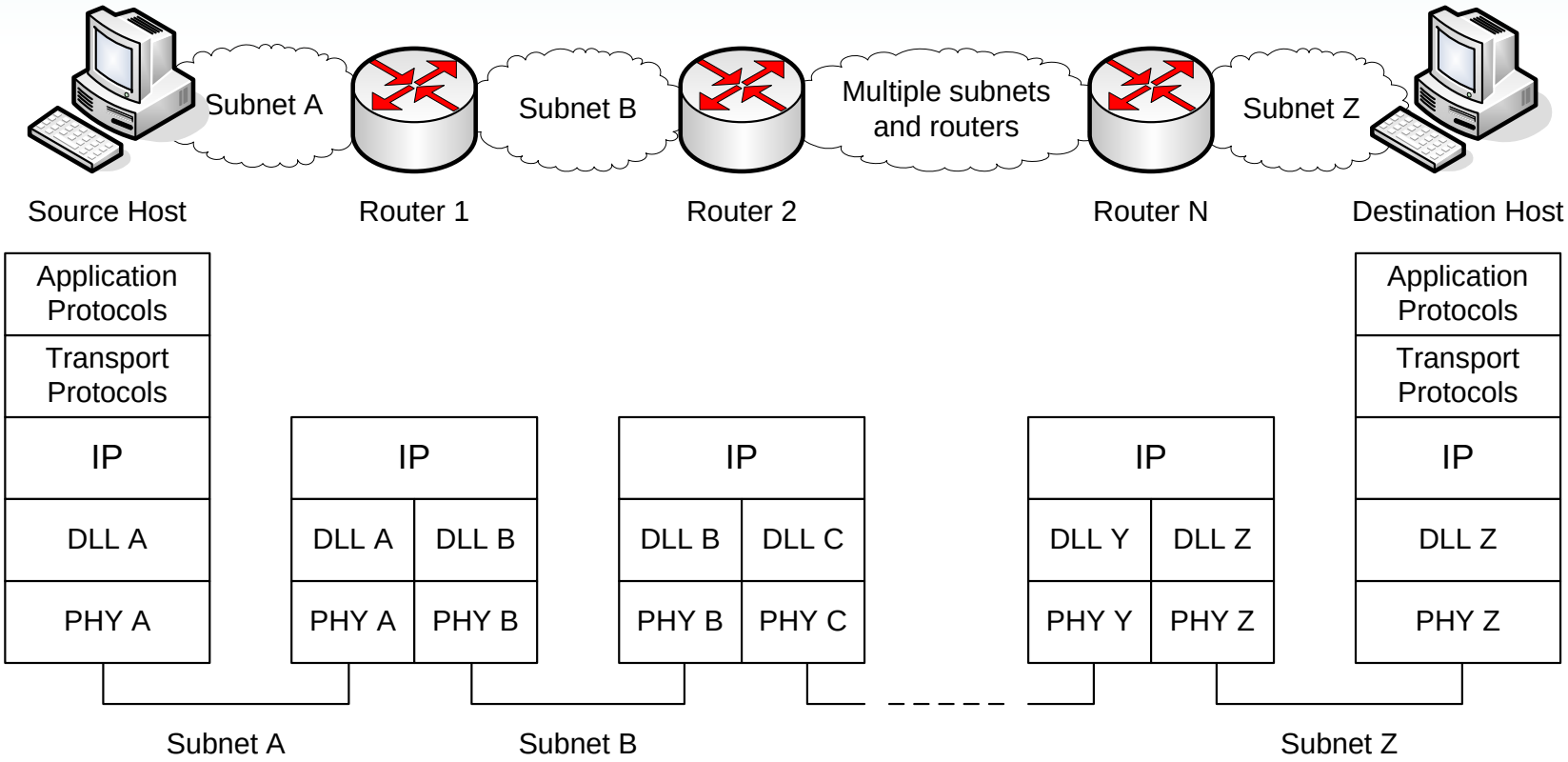
IP in Internet 5 Layer Model



IP Hosts and Routers

- Hosts are the end-devices (stations)
  - Assume hosts have single interface (only attached to one LAN/WAN)
    - In practice, hosts can have multiple interfaces
  - Hosts do not forward datagrams
    - A host is either source or destination; if a host receives a datagram and the host is not the destination, then the host will discard the datagram
- Routers are the datagram packet switches
  - Routers have two or more interfaces (since they connect LANs/WANs together)
  - Routers forward datagrams
  - Routers can act as a source or destination of datagrams (however this is mainly for management purposes)
- IP routing is the process of discovering the best path between source and destination
  - Adaptive routing protocols execute on routers/hosts to find the path; the paths are stored in routing tables on routers and hosts
- IP forwarding is the process of delivering an IP datagram from source to destination

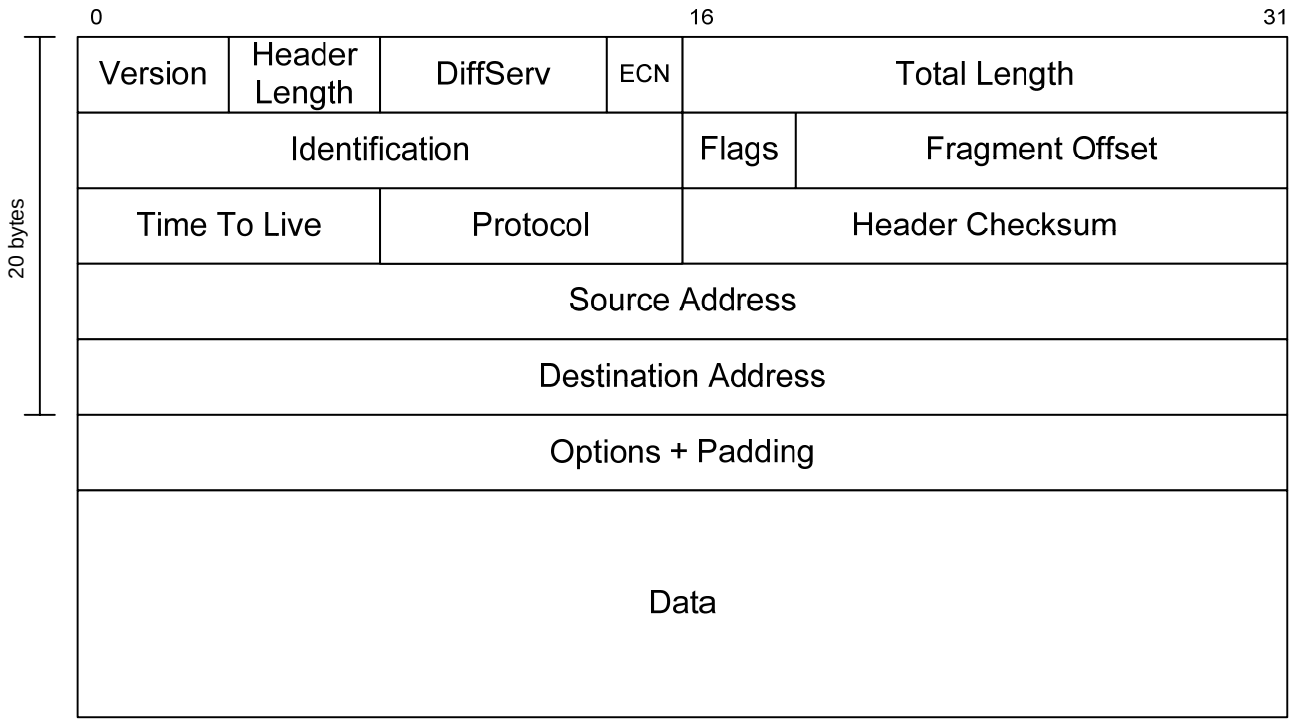
IP Hosts and Routers



- IP is implemented at Layer 3 (Networking layer) in Hosts and Routers
  - Typically as software in a host or router operating system
- There may be 0 or more Routers between a source Host and destination Host

IP Datagram

- IP datagram consists of a variable length header and variable length of data
  - Header has 20 bytes for required fields; then optional fields bringing maximum size to 60 bytes
  - Data length is variable (but must be integer multiple of 8 bits in length); maximum size of datagram (that is, header + data) is 65,535 bytes



IP Datagram Fields

- **Version** [4 bits]: version number of IP; current value is 4 (IPv4)
- **Header Length** [4 bits]: length of header, measured in 4 byte words; minimum value is 5 (20 bytes); maximum is 15 (60 bytes)
- **DiffServ** [6 bits]: Used for quality of service control. DiffServ and ECN used to be called Type of Service field.
- **ECN** [2 bits]: Used for notifying nodes about congestion
- **Total Length** [16 bits]: total length of the datagram, including header, measured in bytes. Max 65535 bytes in datagram
- **Identification**: sequence number for datagram
- **Flags**: 2 bits are used for Fragmentation and Re-assembly, the third bit is not used
  - Don't Fragment bit: if set to 1, then the datagram will not be fragmented (it will be discarded if fragmentation is needed)
  - More Fragments bit: if datagram is fragmented, then set to 1 on all fragments except the last fragment
- **Fragment Offset** [13 bits]: Indicates where this fragment belongs in the original datagram, measured in blocks of 8 bytes
- **Time To Live** [8 bits]: how long datagram should remain in internet. In practice used as a hop counter (a router decrements every time it is forwarded)
- **Protocol** [8 bits]: indicates the next higher layer protocol with a code (e.g. TCP = 6; UDP = 17; ICMP = 1)
- **Header Checksum** [16 bits]: error-detecting code applied to header only (to check for errors in the header); recomputed at each router
- **Source Address** [32 bits]: IP address of source host
- **Destination Address** [32 bits]: IP address of destination host
- **Options**: variable length fields to include options
- **Padding**: used to ensure datagram is multiple of 4 bytes in length
- **Data**: variable length of the data

IP Routing

- No routing protocol is specified for IP
  - Any of the available routing protocols can be used depending on the network topology and requirements of network administrator
    - RIP, EIGRP, OSPF, BGP, …
  - Each routing protocol creates and updates a routing table, which stores information to determine the path from source to destination
- IP uses the information in the routing tables to forward datagrams
- In order to make routing tables manageable, three strategies are used in the Internet:
  - Storing Next-Hop Routes
  - Network-specific Routing
  - Default Routes

- (We will see detailed examples later, after looking at IP addresses)

Routing Tables based on Next Hop

- The routing table stores the next-hop that the packet is to be sent to in order to reach the destination

IP uses this approach

a. Routing tables based on route

Destination	Route
Host B	R1, R2, host B

Routing table for host A

Destination	Route
Host B	R2, host B

Routing table for R1

Destination	Route
Host B	Host B

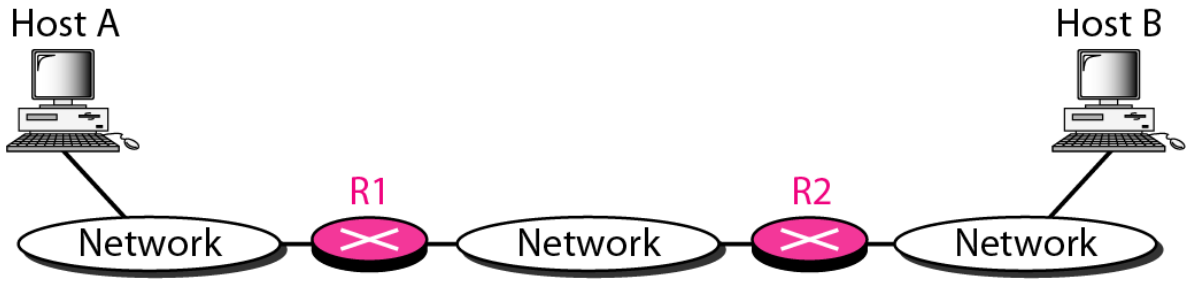
Routing table for R2

b. Routing tables based on next hop

Destination	Next hop
Host B	R1

Destination	Next hop
Host B	R2

Destination	Next hop
Host B	---



Network-Specific Routing

- Routing table stores paths to networks, rather than to individual hosts (its only the last router that deals with individual hosts)
  - How is a network identified?
    - IP addressing scheme

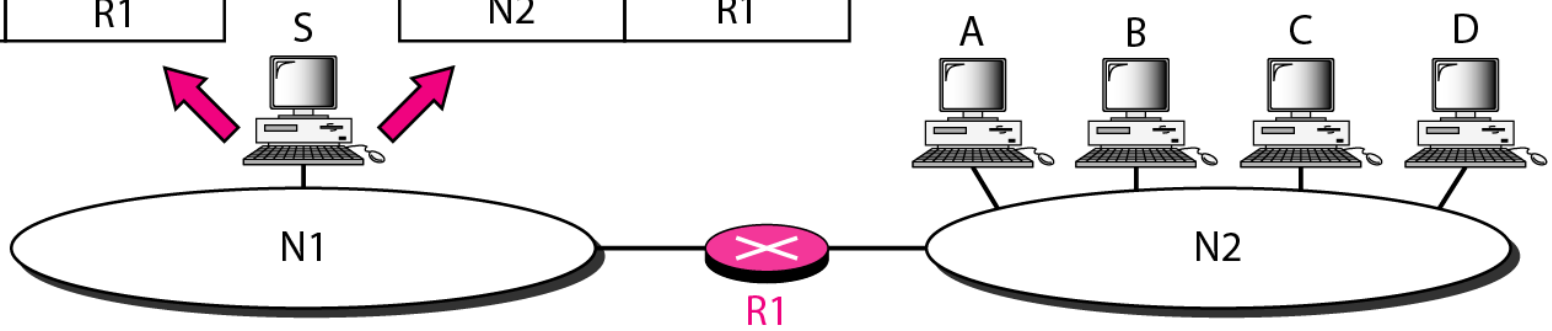
Routing table for host S based on host-specific method

Destination	Next hop
A	R1
B	R1
C	R1
D	R1

IP uses this approach

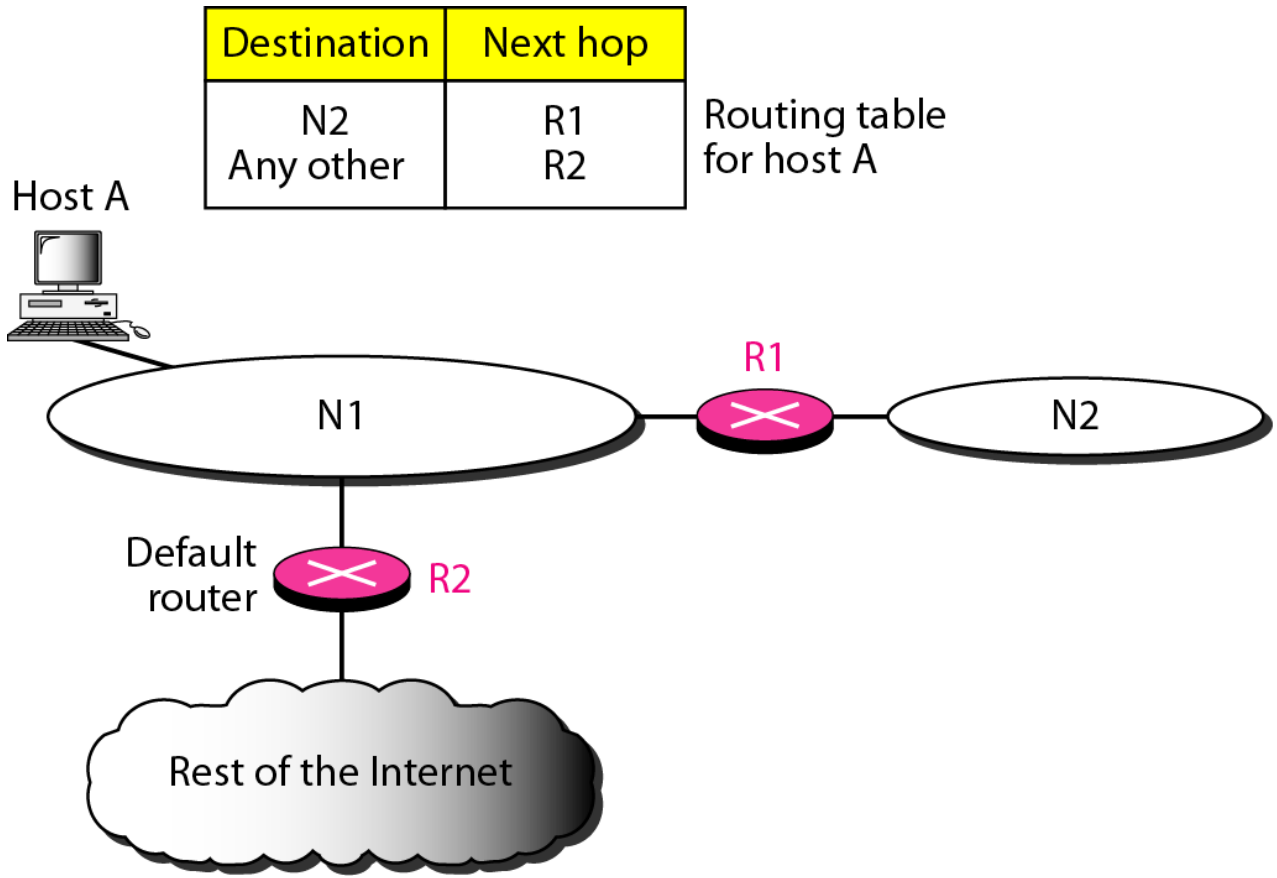
Routing table for host S based on network-specific method

Destination	Next hop
N2	R1



Default Routes

- Instead of routers storing routes to all possible networks, they maintain default routes for “all other networks”

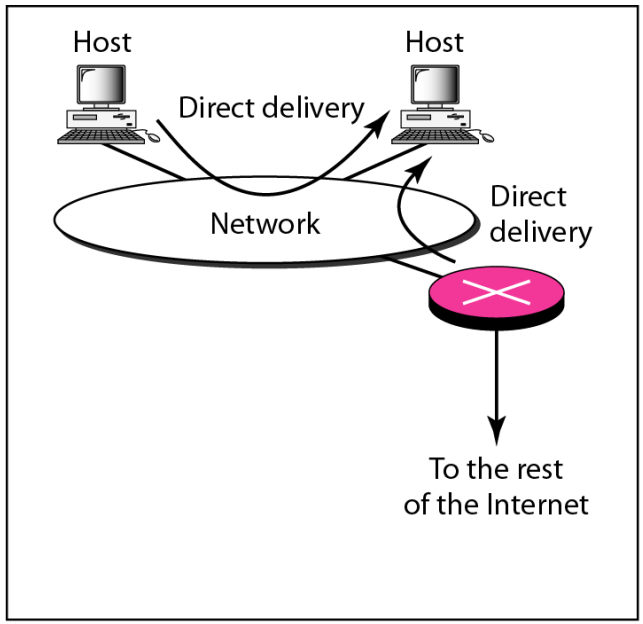


IP Forwarding

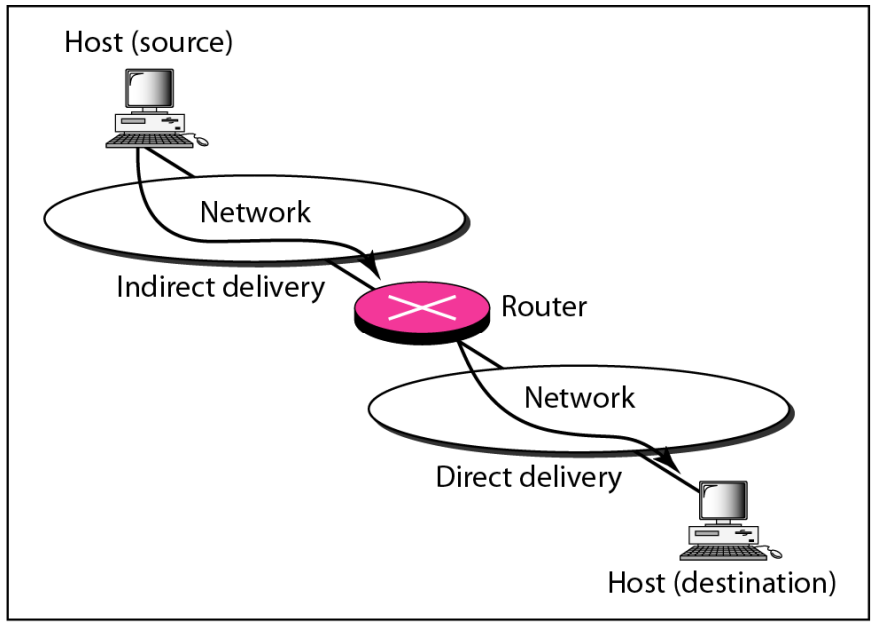
- **Forwarding**: when a node has a packet to send, the node uses the destination IP address and the routing table to determine where to send the packet, and then sends (or forwards) the packet
- Datagram delivery can be divided into two forms:
  - Direct and Indirect Delivery

Direct vs Indirect Delivery

- Sending a packet to the destination involves two different delivery methods:
  - **Direct delivery**: if the final destination is on the same physical network as the "deliverer", send the packet direct to the destination
  - **Indirect delivery**: if the final destination is NOT on the same physical network as the "deliverer", send the packet indirectly to the destination, via a router (which will become the "deliverer")
  - How does the "deliverer" know if the destination is on the same physical network?
    - IP addressing scheme



a. Direct delivery



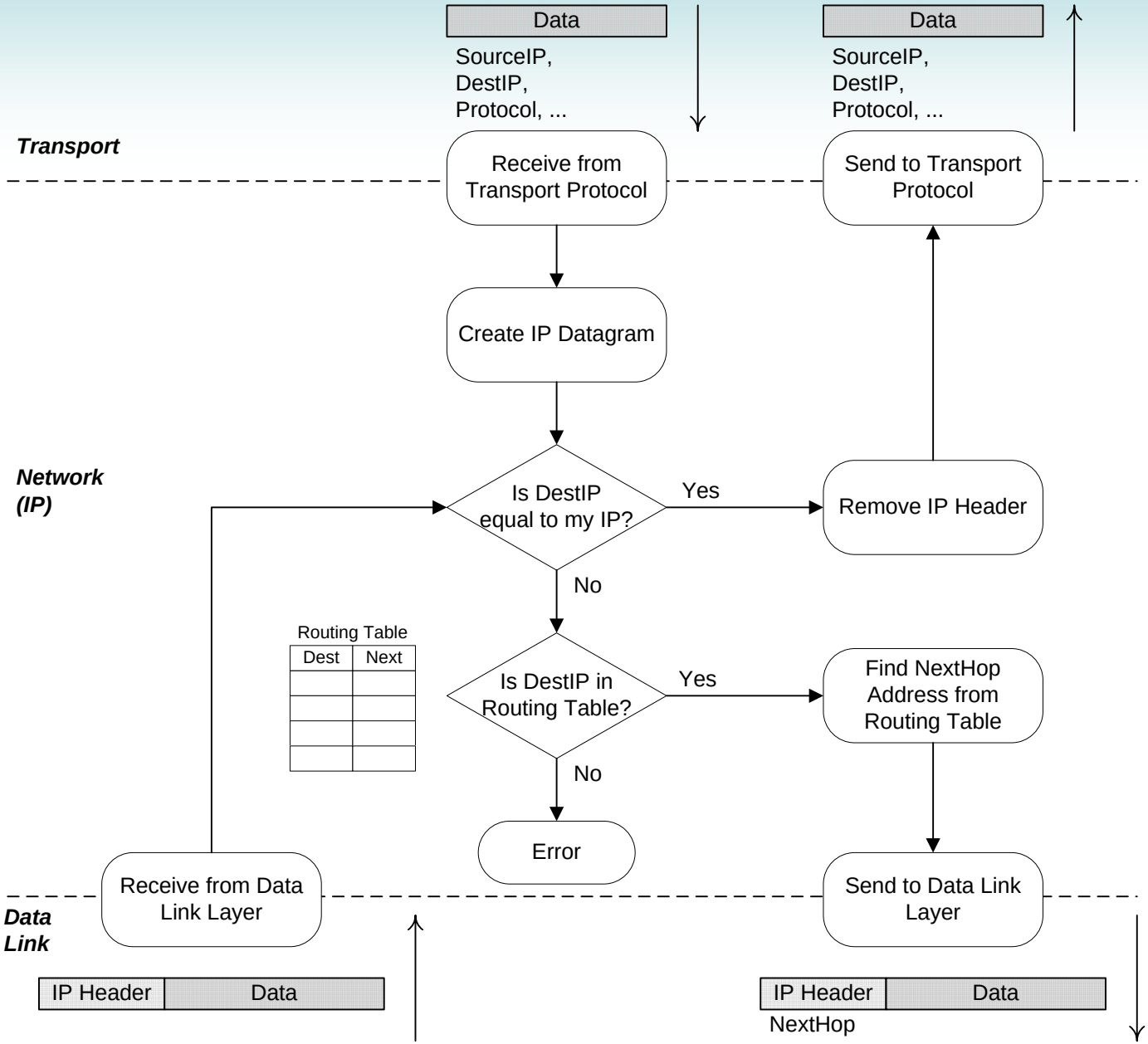
b. Indirect and direct delivery

IP Datagram Processing

- Data received from Transport Protocol
  1. Transport protocol must also indicate destination IP (other information optional)
  2. IP Datagram created
  3. Forwarding Algorithm
- Datagram received from Data Link Protocol
  1. Forwarding Algorithm

- Forwarding Algorithm
  1. If Destination IP address of datagram is same as nodes IP address
     - Remove the header and send Data to Transport Protocol (Source, Destination address is also sent to Transport Protocol)
  2. Else if Destination IP address is in Routing Table
     - Extract the corresponding IP address of the Next Hop
     - Send IP Datagram and Next Hop address to Data Link Protocol

- Data Link Protocol receives IP Datagram and Next Hop address
  1. Determine the corresponding physical address for the Next Hop IP address
  2. Create frame
  3. Transmit frame to the physical address


Data
SourceIP,
DestIP,
Protocol, ...
Data
SourceIP,
DestIP,
Protocol, ...
Transport
Receive from Transport Protocol
Send to Transport Protocol
Create IP Datagram
Network (IP)
Is DestIP equal to my IP?
Yes
Remove IP Header
No
Routing Table
Dest
Next
Is DestIP in Routing Table?
Yes
Find NextHop Address from Routing Table
No
Error
Receive from Data Link Layer
Send to Data Link Layer
Data Link
IP Header
Data
IP Header
Data
NextHop

IP Fragmentation and Re-assembly

- An IP Source host will typically choose a datagram size that will fit into a frame of the attached subnetwork
  - Example: A Source Host is attached to an Ethernet network. The transport protocol will break the application data into blocks such that an IP datagram has maximum size of 1500 bytes (Ethernet's maximum data size)
    - Therefore, the IP datagram is not fragmented
- IP fragmentation may occur at routers
  - If a router receives a datagram that is larger than the maximum frame size the next subnetwork can support, the router will fragment the datagram
  - Example: a router receives a 1500 byte datagram. It has to be sent on a subnetwork that has maximum data size of 1000 bytes. The router will fragment the datagram into two fragments
    - Each fragment has the same IP header as the original datagram, except for the fragmentation-related fields
- IP re-assembly occurs only at the destination host

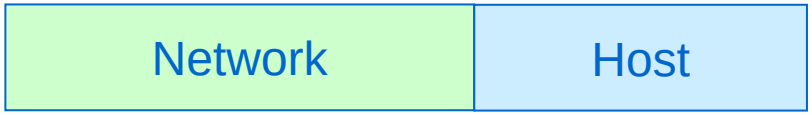
IP Time To Live

- Time To Live (TTL) field in IP datagram is used as a hop limit
  - Source host sets TTL to initial value (default: 255)
  - Each router that forwards the datagram will decrement the value
  - If the TTL is 0, a router will discard the datagram (and possibly send a special ICMP error message back to source)
- This prevents IP datagrams remaining in the Internet “forever”
  - Even if there are errors in routing tables that create loops

IP Addresses

IP Addresses

- IP addresses (used to identify hosts and routers) are 32-bit addresses generally consisting of network portion identifier and host portion identifier



- 32-bits gives 4.2×10^9 possible values
- But IP addresses have different structures (and these have changed over time)
  - First, the set of addresses was separated into five different classes (**Classful addressing**)
  - Then in 1980's, for organisations to have multiple IP networks (e.g. SIIT Bangkadi is one network, SIIT Rangsit is another), Subnet Addressing was introduced
  - Then in 1990's, **Classless Addressing** was allowed so can fully utilise the address space
- We will look at Classful Addressing (original scheme) and Classless Addressing (used today)

Why network and host portion?

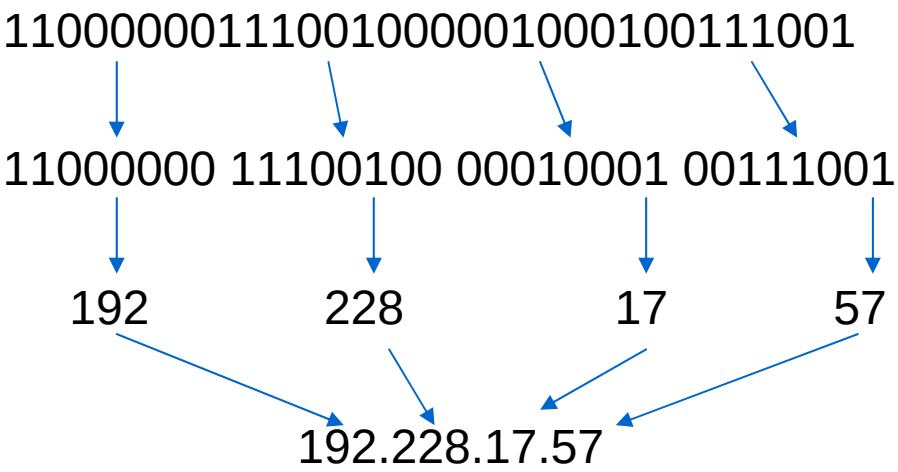
- Network portion: identifies a subnetwork in an internet
- Host portion: identifies a host in a subnetwork
- Splitting the IP address into two parts allows for hierarchical addressing. This makes routing in the global Internet possible!
- Example (also see slide on “Network Specific Addressing”):
  - A reminder: routing protocols provide information to routers about how to reach destinations
    - E.g. “if the destination is D, then send to the next router R”
  - If we did not split the IP address into two ports then routers must know routes to hosts (that is, host-specific routing)
    - “if the destination is host H, then send to the next router R”
      - But on a single subnet, there may be 100’s or 1000’s of hosts
      - Worst case: Routers must store routes to every host on Internet (100,000,000+)
  - But with hierarchical addressing, routers only need to know routes to subnetworks (network-specific routing)
    - “if the destination is subnetwork N, then send to the next router R”
    - Routers only need to know about hosts on their own subnetwork
      - Worst case: routers must know routes to every subnetwork on Internet (100,000)

Representing IP Addresses

- Writing (and remembering) 32 bits is difficult

  11000000111001000001000100111001

- IP addresses are usually written in dotted decimal notation
  - Decimal number represents the bytes of the 32 bit address
  - Decimal numbers are separated by dots

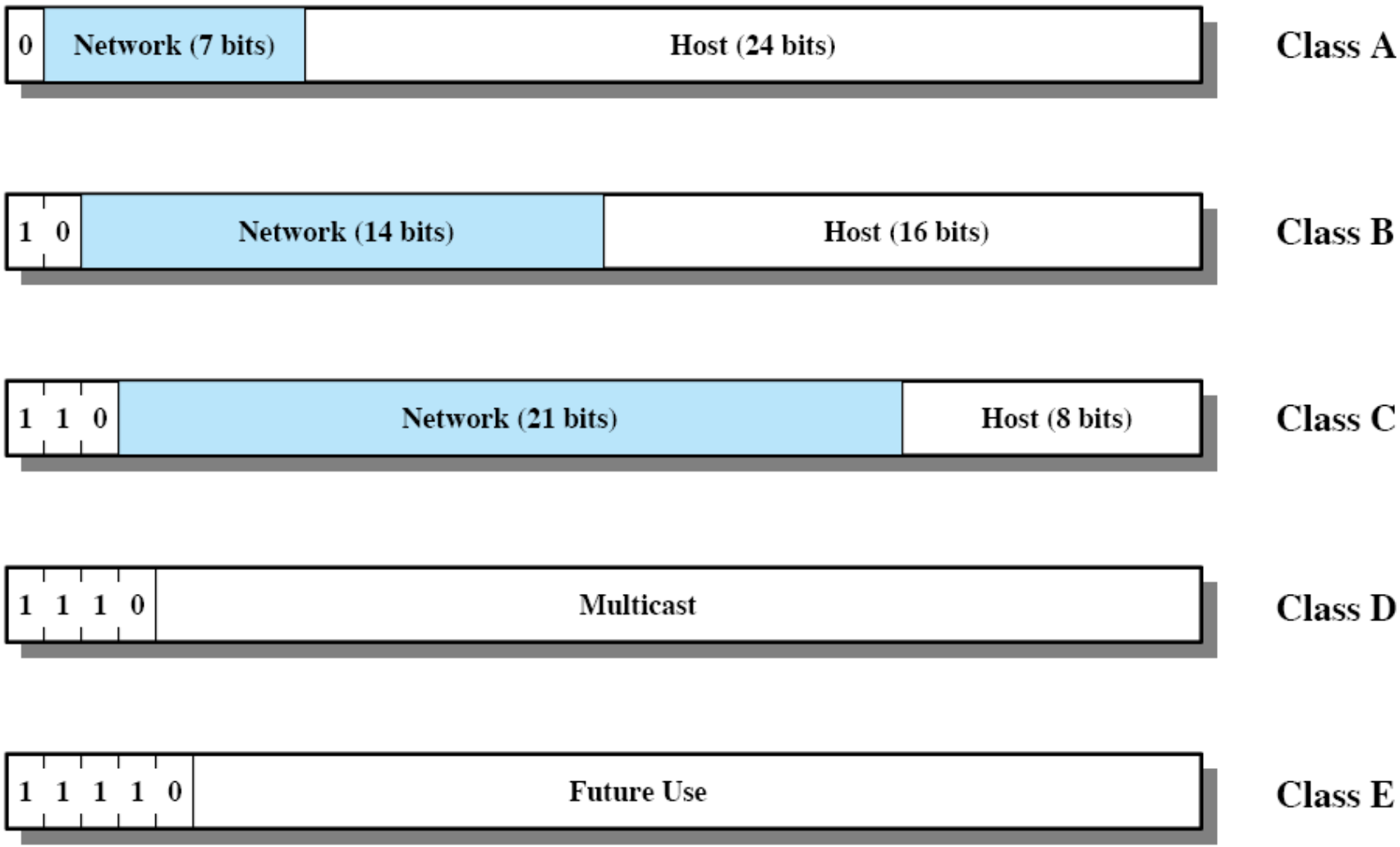


This approach can be used for any type of IP address (for example, classful And classless addresses in the following slides, as well as for subnet masks)

Classful IP Addressing

- The split between network and host portion is pre-defined into one of five different classes
  - Class A: suitable if few networks, many hosts
    - First bit is 0; network portion is 7 bits, host portion 24 bits
    - Note: network addresses with first byte 00000000 (0) or 01111111 (127) are reserved
    - Maximum of 126 networks, each with 16 million hosts
  - Class B: medium number of networks, medium number of hosts
    - First 2 bits are 10; network portion 14 bits, host portion 16 bits
    - Maximum of 16384 networks, each with 65534 hosts
  - Class C: many networks, each with a few hosts
    - First 3 bits are 110; network portion 21 bits, host portion 8 bits
    - Maximum of 2 million networks, each with 254 hosts
  - Class D: use for multicast addressing
    - First 4 bits are 1110
  - Class E: reserved for future use
    - First 5 bits are 11110

IP Address Classes

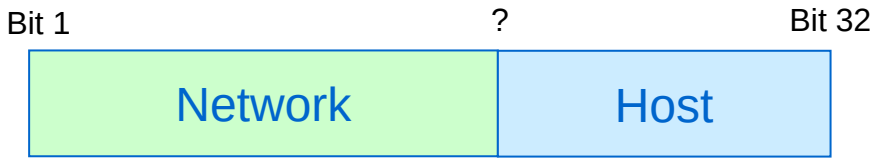


Obtaining an IP Address

- The Internet Assigned Numbers Authority (IANA) manages the assignment of IP addresses
  - If your organisation wants an IP address (or set of addresses) you obtain it through a local or national registry
  - The national registry obtains IP addresses from regional centre, e.g. Asia Pacific Network Information Centre (APNIC)
  - APNIC is assigned address spaces from IANA
- Example:
  - My new company, Steve's Super Solutions, needs IP addresses for its new IP network with around 20 to 30 hosts (and 1 router)
  - Most suitable address type is Class C because not likely to have more than 256 hosts in the network
  - My company is assigned the network address 197.100.7.0, this means I can have 254 computers on my network with addresses 197.100.7.1, 197.100.7.2, 197.100.7.3, …, 197.100.7.254
    - The broadcast address (to send to all computers) is 197.100.7.255
    - It is up to me to properly assign valid IP addresses to my computers
- Problem with Classful Addresses:
  - If every organisation in the world wants a unique network address for each of its IP networks (remember some organisations have many IP networks), then not enough IP addresses!
  - Hence subnet addressing and then classless addressing was developed

Classless IP Addressing

- IP address has 32 bits: Where is the split between network and host portion?



- In classless addressing, an additional item, called an address mask or subnet mask indentifies where the split is
  - The mask is 32 bits: a bit 1 indicates the corresponding bit in the IP address is the network portion; a bit 0 indicates the corresponding bit in the IP address is the host portion

IP address, 130.17.41.129: 10000010 00010001 00101001 10000001
Subnet mask, 255.255.252.0: 11111111 11111111 11111100 00000000

Network portion | Host portion

Network, 130.17.40.0: 10000010 00010001 00101000 00000000

  - The mask can be given in dotted decimal form or a shortened form, which counts the number of 1 bits
    - The above example can be written as /22, and the IP address as 130.17.41.129/22

Special Cases for IP Addresses

- There are special case addresses that cannot be used to identify a particular host:
  - Network Address
    - The bits of the Host portion are 0
    - Used to identify the network, e.g. for routers to send to a network
    - E.g. host 130.17.41.129/22 is on the network 130.17.40.0/22
  - Broadcast Address (Directed)
    - The bits of the Host portion are 1
    - Used as a destination for broadcast directed to a specific network
    - E.g. host 130.17.41.129/22 sends to 116.42.2.255/24, then all hosts on network 116.42.2.0/24 will receive the datagram
  - Loopback Address
    - The first 8 bits of Network portion are 01111111 (decimal: 127)
    - Used as a destination address when a host sends to itself
    - E.g. host 130.17.41.129/22 sends to 127.0.0.1, then the datagram will not be sent on the network, but instead to itself (130.17.41.129)
  - Local Broadcast Address
    - All 32 bits are 1 (255.255.255.255)
    - Used as a destination for broadcast to the local network
    - E.g. host 130.17.41.129/22 sends to 255.255.255.255, then all hosts on network 130.17.40.0/22 will receive the datagram
  - Startup Source Address
    - All 32 bits are 0 (0.0.0.0)
    - Used as a source address by a host if the host doesn't know its own IP address
    - E.g. host sends an address to a known server (or local broadcast address) asking for its own IP address; 0.0.0.0 is used as the source

Other Network Layer Functions

Internet Control Message Protocol

- IP is used for transferring datagrams between hosts
- ICMP is used for sending feedback about problems between routers and hosts
  - Error reporting examples:
    - Destination unreachable: if a router cannot send an IP datagram to the destination host (e.g. no route because a router or link is down, or the destination doesn't exist), then it returns an ICMP Destination Unreachable message to the source host
    - Echo and Echo Reply: used to test if two computers can communicate. Source sends an ICMP Echo message to destination; if destination receives ICMP Echo, it will respond with ICMP Echo Reply
      - Used by `ping` and other network tools
- ICMP uses IP to send messages

Address Resolution Protocol

- IP (network layer) uses IP addresses to identify each interface. These are logical addresses
- Each subnetwork of an internet uses its own (data link layer) addressing mechanism, for example:
  - Ethernet uses IEEE 48 bit addresses
  - Frame Relay and ATM use path/channel identifiers
  - HDLC uses 8 bit addresses

  These are physical addresses.
- A mapping must be made from an IP address to a physical address
  - When IP sends Data to the Data Link layer protocol, it also sends a Next Hop IP address
  - The Address Resolution Protocol (ARP) is used in each subnetwork to perform the mapping from the Next Hop IP address to the physical address of that node
  - ARP creates a table in each host in the local network that has a list of:
    - IP address and corresponding physical address

IP Address	Physical Address
10.10.1.1	00-50-ba-be-34-cc
10.10.1.101	00-50-ba-87-61-eb
10.10.1.133	00-50-ba-be-34-d2

Other Features

- IPv6
  - The current version is IPv4, uses 32 bit addresses; estimates that there will not be enough addresses in 5 to 15 years
  - IPv6 has been designed to improve/replace IPv4
    - 128 bit addresses (enough for 1028 addresses per person!)
    - Still not in widespread use; no strong motivation for ISPs to change from IPv4 to IPv6
- Multicasting
  - Unicast is one-to-one communication
  - Broadcast is one-to-all communication
  - Multicast is one-to-many communication
    - Uses different parts of IP addresses and requires different routing
    - Useful for multimedia communication, e.g. video conferences, TV and audio streaming, collaborative applications
- Quality of Service Control
  - IP provides a “best effort” service: a datagram is sent with no guarantee of arriving at the destination within some time, or at some speed; also no priority is given to datagrams – everyone’s data is treated the same
  - Many multimedia applications (video, voice) work better if they have guaranteed bandwidth/delay, or at least priority over other datagrams
  - QoS includes mechanisms controlling priority of transfer, and guaranteeing certain level of service