

5.5 Internetworking

5.5.1 How Networks Differ

5.5.2 How Networks Can Be Connected

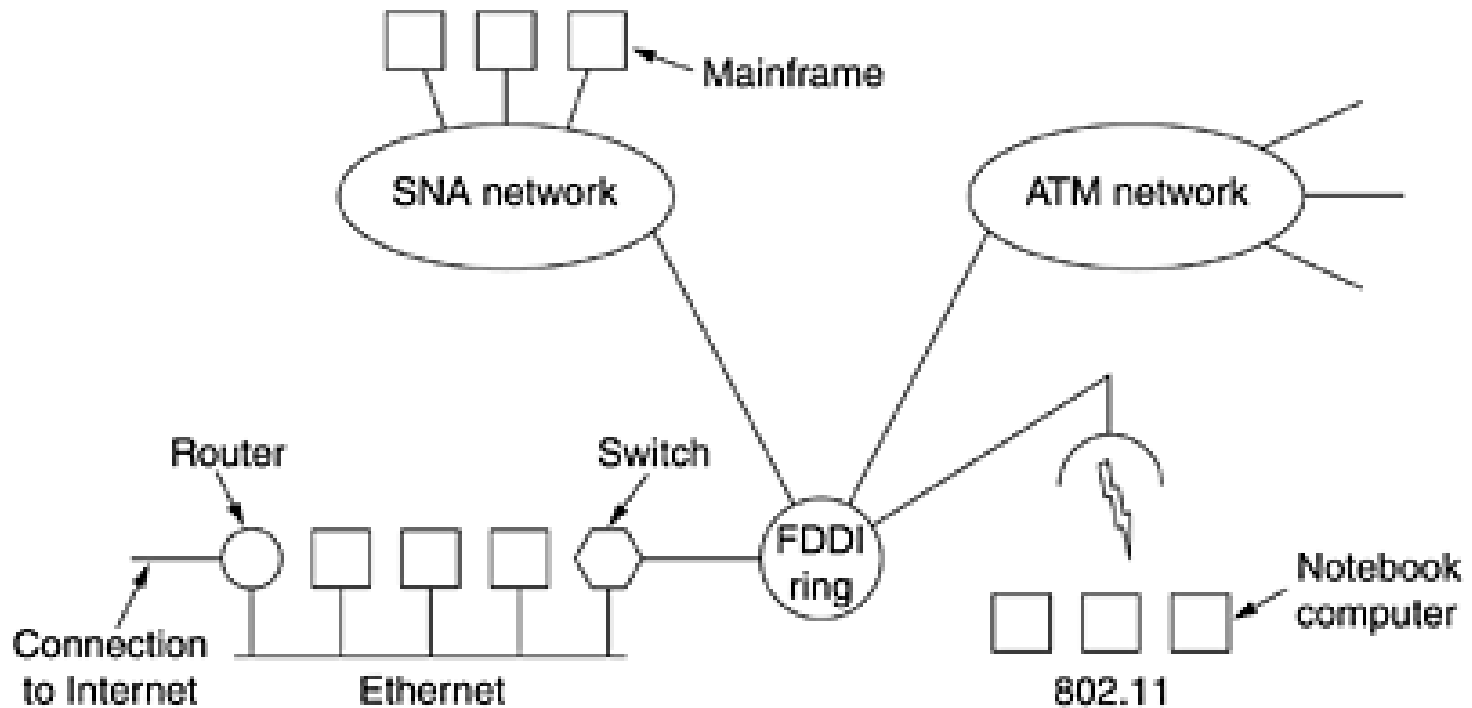
5.5.7 Fragmentation

Internetworking

- Until now, we have implicitly assumed that there is a single homogeneous network, with each machine using the same protocol in each layer.
- Unfortunately, this assumption is wildly optimistic. Many different networks exist, including LANs, MANs, and WANs.
- Numerous protocols are in widespread use in every layer.
- Issues that arise when two or more networks are connected to form an **internet**.
- Having different networks invariably means having different protocols.

- We believe that a variety of different networks (and thus protocols) will always be around.
- **Internetworking:** interconnecting multiple computer networks, such that any pair of hosts in the connected networks can exchange messages irrespective of their networking technology.
- The resulting system of interconnected networks are called an internetwork, or simply an internet.
- The purpose of interconnecting all these networks is to allow users on any of them to communicate with users on all the other ones and also to allow users on any of them to access data on any of them.

Figure 5-42. A collection of interconnected networks.



-an FDDI (Fiber Distributed Data Interface) optical backbone is used to connect an Ethernet, an 802.11 wireless LAN, and the corporate data center's SNA mainframe network.

How Networks Differ

Figure 5-43. Some of the many ways networks can differ.

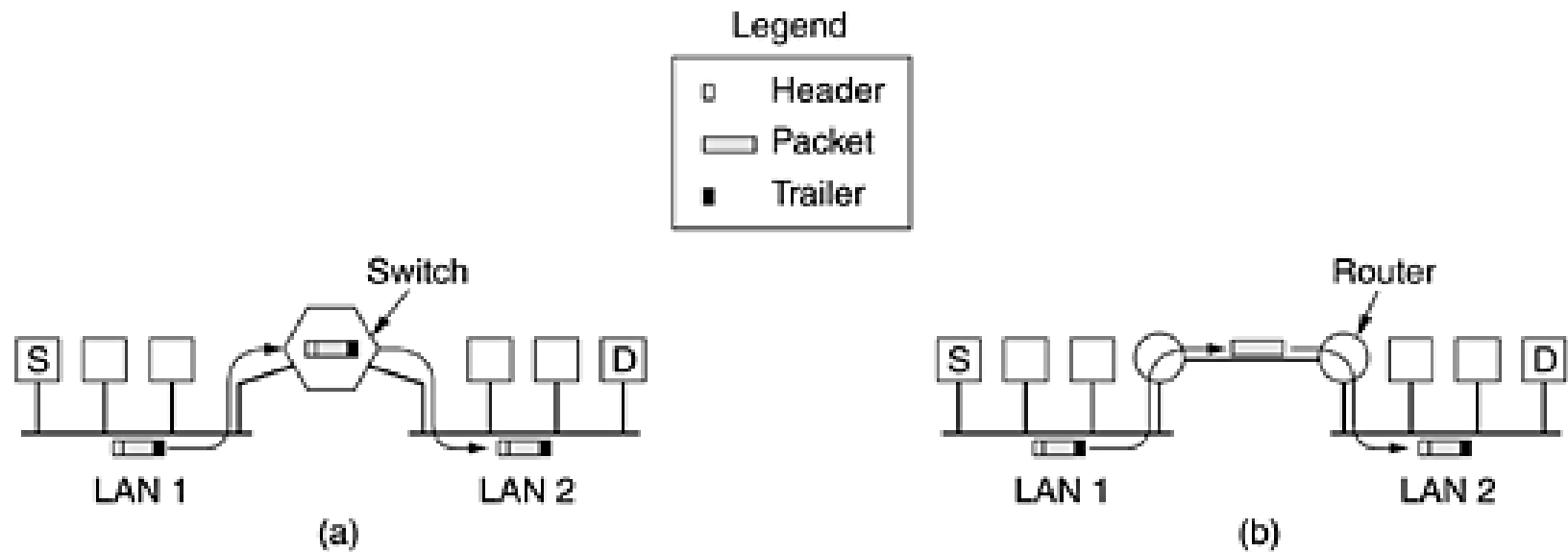
Item	Some Possibilities
Service offered	Connection oriented versus connectionless
Protocols	IP, IPX, SNA, ATM, MPLS, AppleTalk, etc.
Addressing	Flat (802) versus hierarchical (IP)
Multicasting	Present or absent (also broadcasting)
Packet size	Every network has its own maximum
Quality of service	Present or absent; many different kinds
Error handling	Reliable, ordered, and unordered delivery
Flow control	Sliding window, rate control, other, or none
Congestion control	Leaky bucket, token bucket, RED, choke packets, etc.
Security	Privacy rules, encryption, etc.
Parameters	Different timeouts, flow specifications, etc.
Accounting	By connect time, by packet, by byte, or not at all

How Networks Can Be Connected

- Networks can be interconnected by different devices.
- In the **physical layer**, networks can be connected by **repeaters or hubs**, which just move the bits from one network to an identical network.
- These are mostly analog devices and do not understand anything about digital protocols (they just regenerate signals).
- One layer up we find **bridges** and **switches**, which operate at the **data link layer**. They can accept frames, examine the MAC addresses, and forward the frames to a different network.

- In the **network layer**, we have **routers** that can connect two networks. If two networks have dissimilar network layers, the router may be able to translate between the packet formats.
- A router that can handle multiple protocols is called a **multiprotocol router**.
- In the **transport layer** we find **transport gateways**, which can interface between two transport connections.
- Finally, in the **application layer**, **application gateways** translate message semantics.

Figure 5-44. (a) Two Ethernets connected by a switch. (b) Two Ethernets connected by routers.



- The source machine, **S**, wants to send a packet to the destination machine, **D**. These machines are on different Ethernets, connected by a switch.
- **S** encapsulates the packet in a frame and sends it on its way. The frame arrives at the switch, which then determines that the frame has to go to LAN 2 by looking at its MAC address.
- The switch just removes the frame from LAN 1 and deposits it on LAN 2.

- Now let us consider the same situation but with the two Ethernets connected by a pair of routers instead of a switch.
- The routers are connected by a point-to-point line, possibly a leased line thousands of kilometers long.
- Now the frame is picked up by the router and the packet removed from the frame's data field. The router examines the address in the packet (e.g., an IP address) and looks up this address in its routing table.
- Based on this address, it decides to send the packet to the remote router, potentially encapsulated in a different kind of frame, depending on the line protocol.
- At the far end, the packet is put into the data field of an Ethernet frame and deposited onto LAN 2.

Difference between the switched (or bridged) case and the routed case

- With a switch (or bridge), the entire frame is transported on the basis of its MAC address.
- With a router, the packet is extracted from the frame and the address in the packet is used for deciding where to send it.
- Switches do not have to understand the network layer protocol being used to switch packets. Routers do.

Fragmentation

- Each network imposes some maximum size on its packets. These limits have various causes, among them:
 1. Hardware (e.g., the size of an Ethernet frame).
 2. Operating system (e.g., all buffers are 512 bytes).
 3. Protocols (e.g., the number of bits in the packet length field).
 4. Compliance with some (inter)national standard.
 5. Desire to reduce error-induced retransmissions to some level.
 6. Desire to prevent one packet from occupying the channel too long.
- Maximum payloads range from 48 bytes (ATM cells) to 65,515 bytes (IP packets)

Issues

- When a large packet wants to travel through a network whose maximum packet size is too small.

Solution

- Allow gateways to break up packets into **fragments**, sending each fragment as a separate internet packet.
- Packet-switching networks, too, have trouble putting the fragments back together again.

- Reaching at the destination, how to recombine the fragments there are two strategies

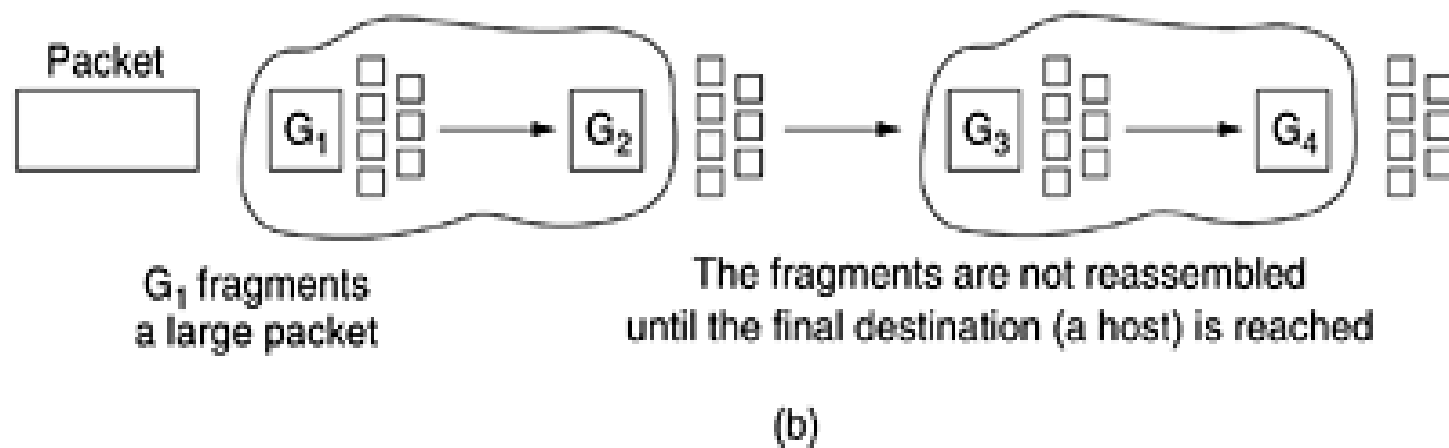
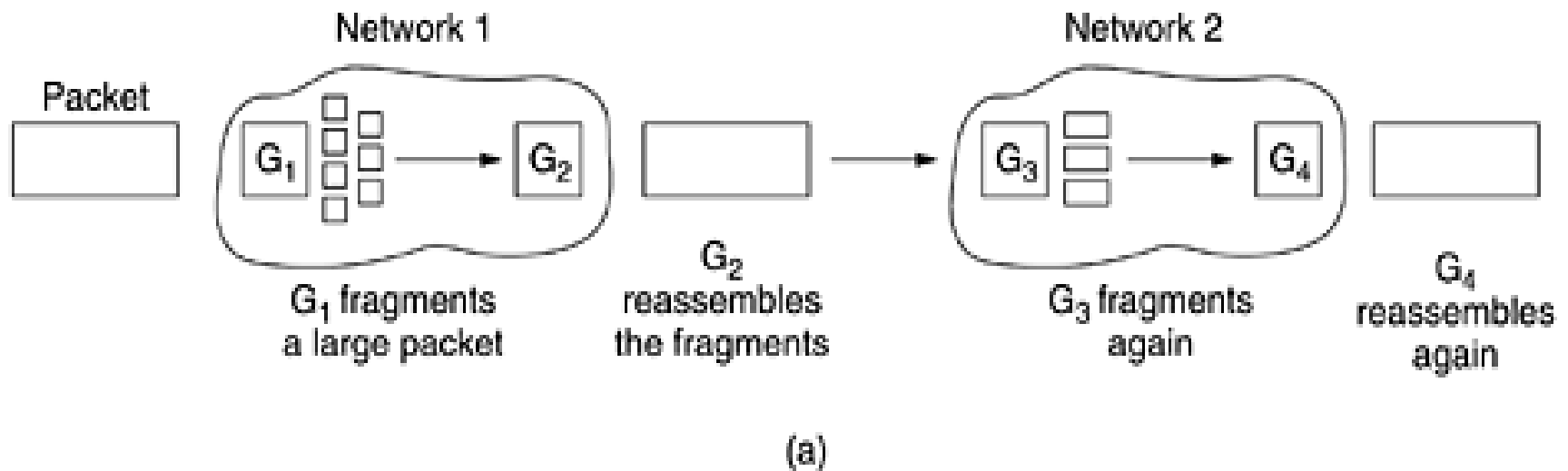
(1) Transparent Fragmentation

(2) Non-Transparent fragmentation

Transparent Fragmentation

- Here the fragmentation is transparent to other network. EX- The packets has to be fragmentated because of its oversize before entering to the network.
- Make fragments caused by a “ small packet” network transparent to any subsequent network.
- When an oversized packet arrives at a gateway, the gateway breaks it up into fragments
- Each fragment is addressed to the same exit gateway, where the pieces are recombined.

Figure 5-50. (a) Transparent fragmentation. (b) Nontransparent fragmentation.



Problems

- The exit gateway must know when it has received all the pieces, so either a count field or an “end of packet” bit must be provided
- All packet must exit via the same gate way
- Overhead required to repeatedly reassemble and then refragment a large packet passing through a series of small packet networks.
- Ex- ATM networks use the transparent fragmentation

Non-Transparent fragmentation

- All the networks are aware of fragments, so the recombination occurs only at the destination host.
- Once a packet has been fragmented, each fragment is treated as though it was an original packet.

Problems-

- Requires every host to be able to do reassembly
- When a large packet is fragmented, the total overhead increases because each fragment must have a header
- Ex: IP

- Fragment consists of three parts
 - (1) Original Packet Number
 - (2) The fragment number
 - (3) Fragment is the last fragment of the packet or not.

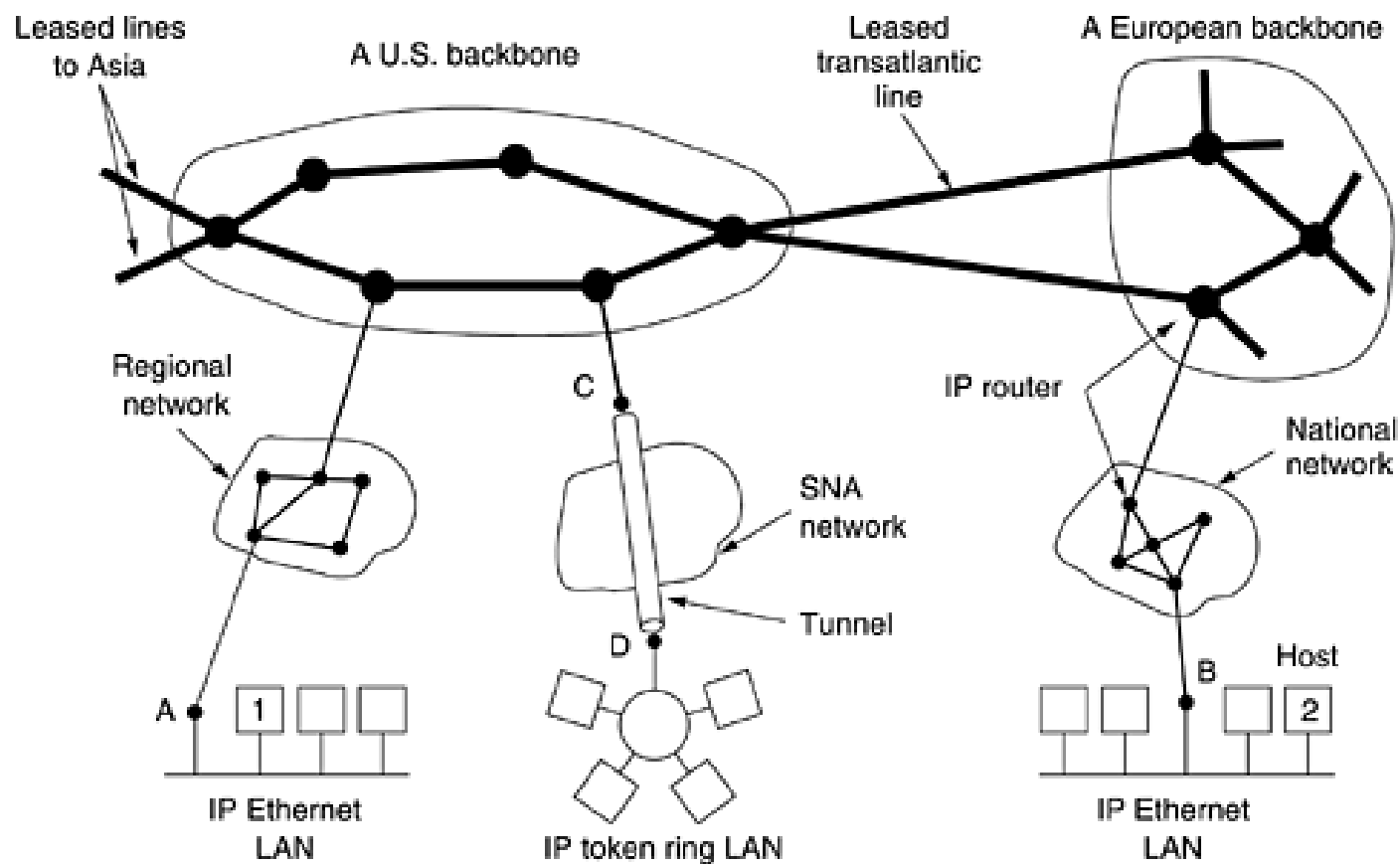
The Network Layer in the Internet

- The IP Protocol
- IP Addresses

The Network Layer in the Internet

- Top 10 principles (from most important to least important) protocol designers should consider in network layer:
 1. Make sure it works.
 2. Keep it simple.
 3. Make clear choices.
 4. Exploit modularity.
 5. Expect heterogeneity.
 6. Avoid static options and parameters.
 7. Look for a good design; it need not be perfect.
 8. Be strict when sending and tolerant when receiving.
 9. Think about scalability.
 10. Consider performance and cost.

Figure 5-52. The Internet is an interconnected collection of many networks.

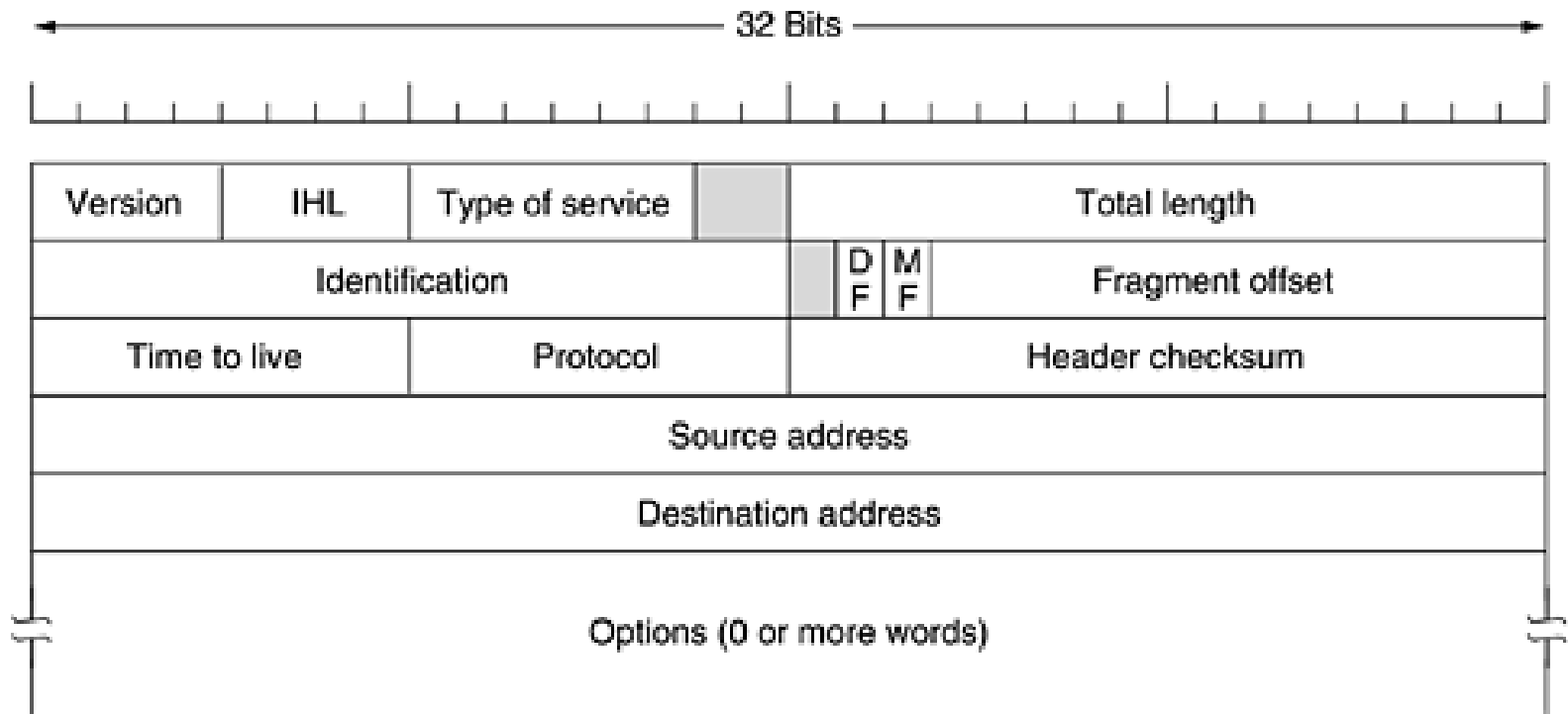


- The glue that holds the whole Internet together is the network layer protocol, **IP (Internet Protocol)**.
- Unlike most older network layer protocols, it was designed from the beginning with internetworking in mind.
- A good way to think of the network layer is this.
- Its job is to provide a best-efforts (i.e., not guaranteed) way to transport datagrams from source to destination, without regard to whether these machines are on the same network or whether there are other networks in between them.

- Communication in the Internet works as follows.
- The transport layer takes data streams and breaks them up into datagrams.
- Each datagram is transmitted through the Internet, possibly being fragmented into smaller units as it goes.
- When all the pieces finally get to the destination machine, they are reassembled by the network layer into the original datagram.
- This datagram is then handed to the transport layer, which inserts it into the receiving process' input stream.
- As can be seen from Fig. 5-52, a packet originating at host 1 has to traverse six networks to get to host 2. In practice, it is often much more than six.

The IP Protocol

Figure 5-53. The IPv4 (Internet Protocol) header



- An IP datagram consists of a header part and a text part.
- The header has a 20-byte fixed part and a variable length optional part.
- **Version:** The Version field keeps track of which version of the protocol the datagram belongs to. Ex: IPv4 and IPv6.
- **IHL:** Since the header length is not constant, a field in the header, *IHL*, is provided to tell how long the header is.
- **Type of service:** distinguish between different classes of service. Various combinations of reliability and speed are possible.
- **Total Length:** The Total length includes everything in the datagram—both header and data. The maximum length is 65,535 bytes. At present, this upper limit is tolerable, but with future gigabit networks, larger datagrams may be needed.
- **Identification:** The Identification field is needed to allow the destination host to determine which datagram a newly arrived fragment belongs to. All the fragments of a datagram contain the same Identification value.

- **DF:** DF stands for Don't Fragment. It is an order to the routers not to fragment the datagram because the destination is incapable of putting the pieces back together again.
- **MF:** MF stands for More Fragments. All fragments except the last one have this bit set. It is needed to know when all fragments of a datagram have arrived.
- **Fragment offset:** The Fragment offset tells where in the current datagram this fragment belongs.
- **Time to live:** The Time to live field is a counter used to limit packet lifetimes.
- It is supposed to count time in seconds, allowing a maximum lifetime of 255 sec. It must be decremented on each hop and is supposed to be decremented multiple times when queued for a long time in a router. In practice, it just counts hops. When it hits zero, the packet is discarded and a warning packet is sent back to the source host.

- **Protocol:** When the network layer has assembled a complete datagram, it needs to know what to do with it.
- The Protocol field tells it which transport process to give it to. TCP is one possibility, but so are UDP and some others.
- **Header checksum:** The Header checksum verifies the header only. Such a checksum is useful for detecting errors generated by bad memory words inside a router.
- **Source address and Destination address** indicate the network number and host number.

- **Option:** The Options field was designed to provide an escape to allow subsequent versions of the protocol to include information not present in the original design.
- The options are variable length. Each begins with a 1-byte code identifying the option. Some options are followed by a 1-byte option length field, and then one or more data bytes.

Figure 5-54. Some of the IP 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

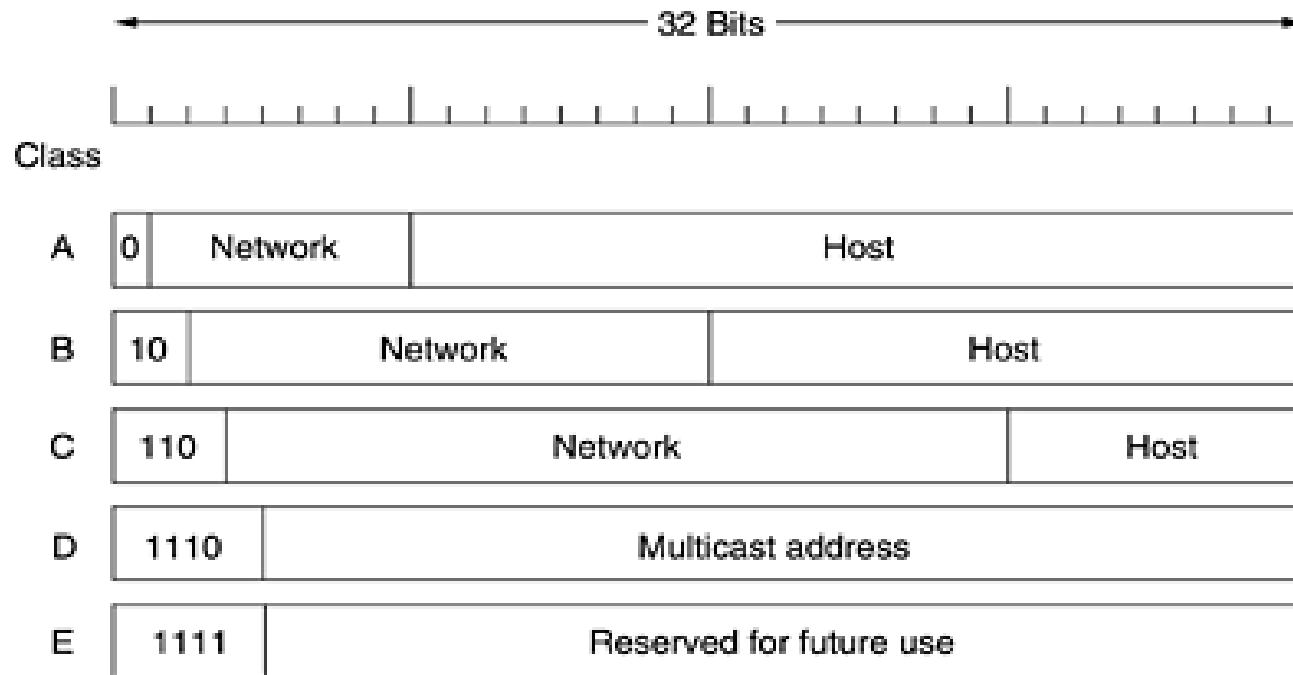
IP Address

- Every host and router on the Internet has an IP address, which encodes its network number and host number.
- The combination is unique: in principle, no two machines on the Internet have the same IP address.
- In the TCP/IP protocol, the unique identifier for a computer is called its IP address.
- There are two standards for IP addresses: **IP Version 4 (IPv4)** and **IP Version 6 (IPv6)**.

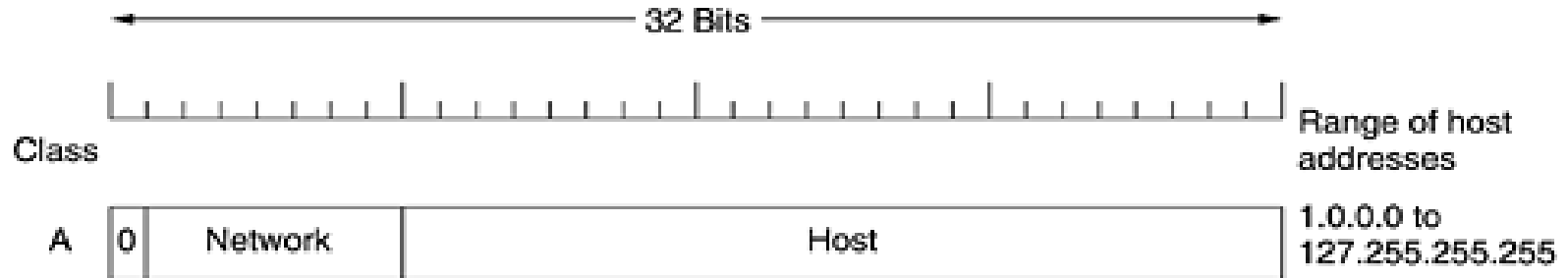
Addresses - IPv4

- All IP addresses are 32 bits long and are used in the *Source address* and *Destination address* fields of IP packets.
- It is important to note that an IP address does not actually refer to a host. It really refers to a network interface, so if a host is on two networks, it must have two IP addresses.
- However, in practice, most hosts are on one network and thus have one IP address.
- The **32** bits of an IPv4 address are broken into **4 octets**, or 8 bit fields (0-255 value in decimal notation).

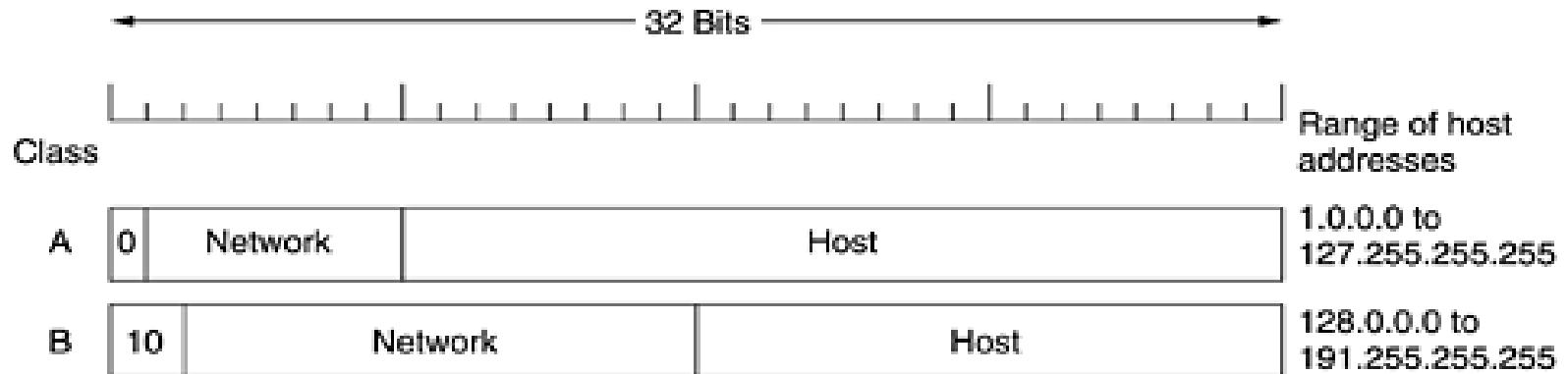
IP address formats




IP address formats



IP address formats



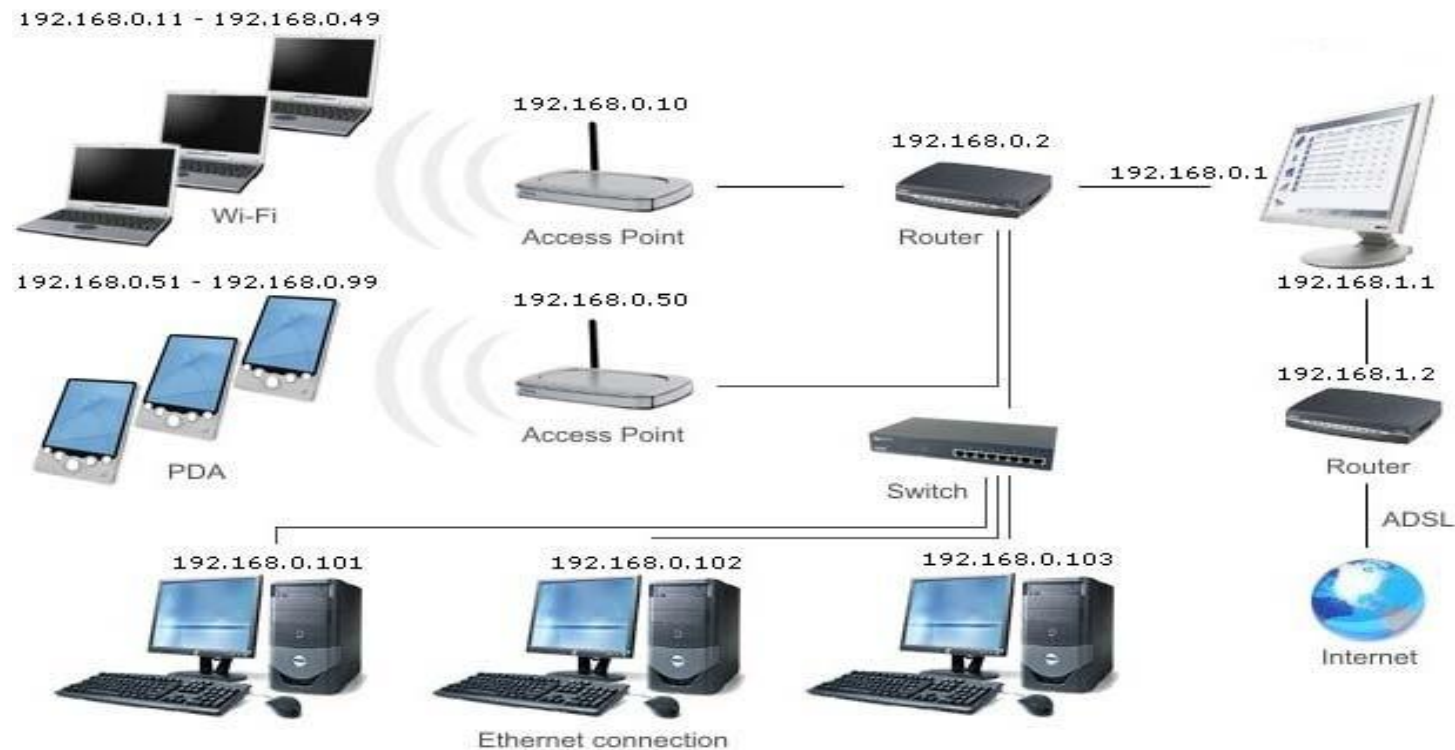
IP address formats

	← 32 Bits →			
				Range of host addresses
Class				
A	0	Network	Host	1.0.0.0 to 127.255.255.255
B	10	Network	Host	128.0.0.0 to 191.255.255.255
C	110	Network	Host	192.0.0.0 to 223.255.255.255
D	1110	Multicast address		224.0.0.0 to 239.255.255.255
E	1111	Reserved for future use		240.0.0.0 to 255.255.255.255

- For networks of different size, The first one (for large networks) to three (for small networks) octets can be used to identify the **network**, while the rest of the octets can be used to identify the **node** on the network.
- The class A formats allow for up to 128 networks with 16 million hosts each,
- The class B formats allow 16,384 networks with up to 64K hosts, and
- The class C formats allow 2 million networks (e.g., LANs) with up to 256 hosts each (although a few of these are special).
- Also supported is multicast, in which a datagram is directed to multiple hosts. Addresses beginning with 1111 are reserved for future use.

- Over 500,000 networks are now connected to the Internet, and the number grows every year. Network numbers are managed by a nonprofit corporation called **ICANN (Internet Corporation for Assigned Names and Numbers)** to avoid conflicts.
- In turn, ICANN has delegated parts of the address space to various regional authorities, which then dole out IP addresses to ISPs and other companies.
- Network addresses, which are 32-bit numbers, are usually written in **dotted decimal notation**. In this format, each of the 4 bytes is written in decimal, from 0 to 255.
- For example, the 32-bit hexadecimal address C0290614 is written as 192.41.6.20.
- The lowest IP address is 0.0.0.0 and the highest is 255.255.255.255.

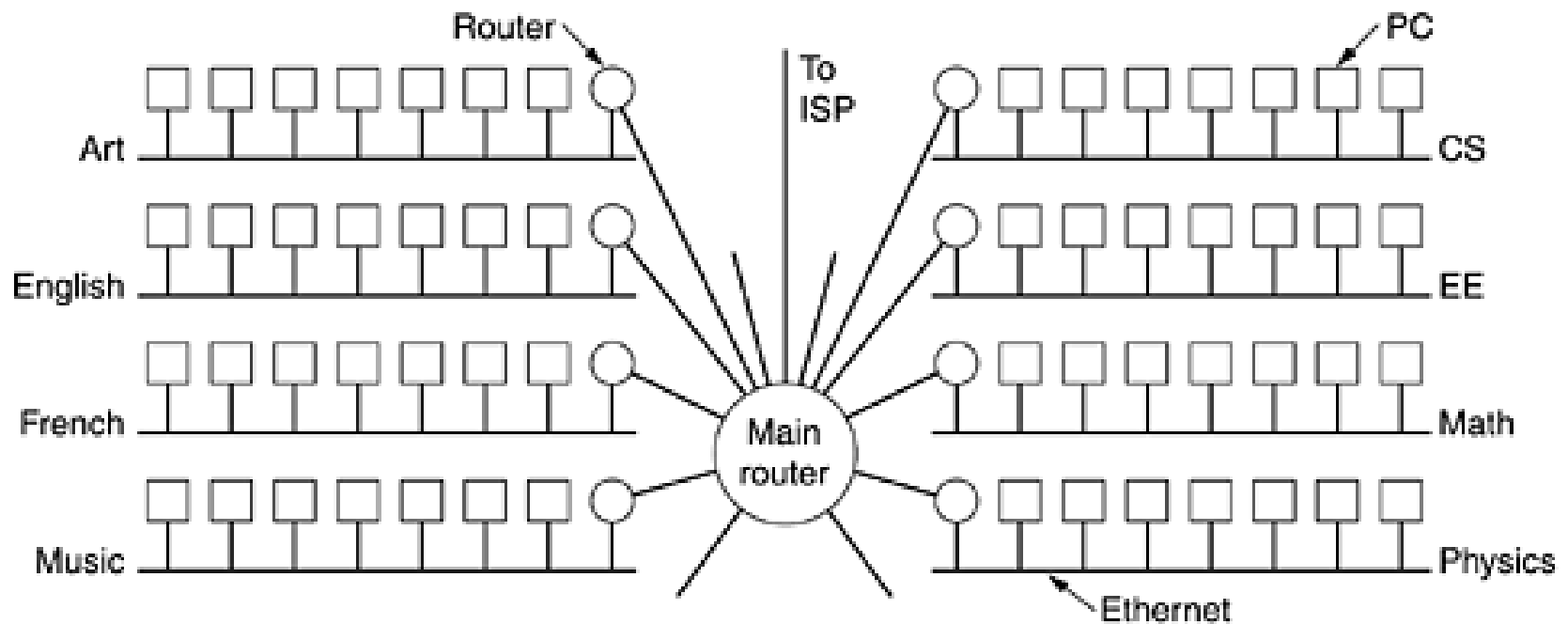
Local Area Network Addresses - IPv4



Subnets

- All the hosts in a network must have the same network number. This property of IP addressing can cause problems as networks grow.
- The solution is to allow a network to be split into several parts for internal use but still act like a single network to the outside world.
- A typical campus network with a main router connected to an ISP or regional network and numerous Ethernets spread around campus in different departments.
- Each of the Ethernets has its own router connected to the main router.
- In the Internet literature, the parts of the network (in this case, Ethernets) are called **subnets**.

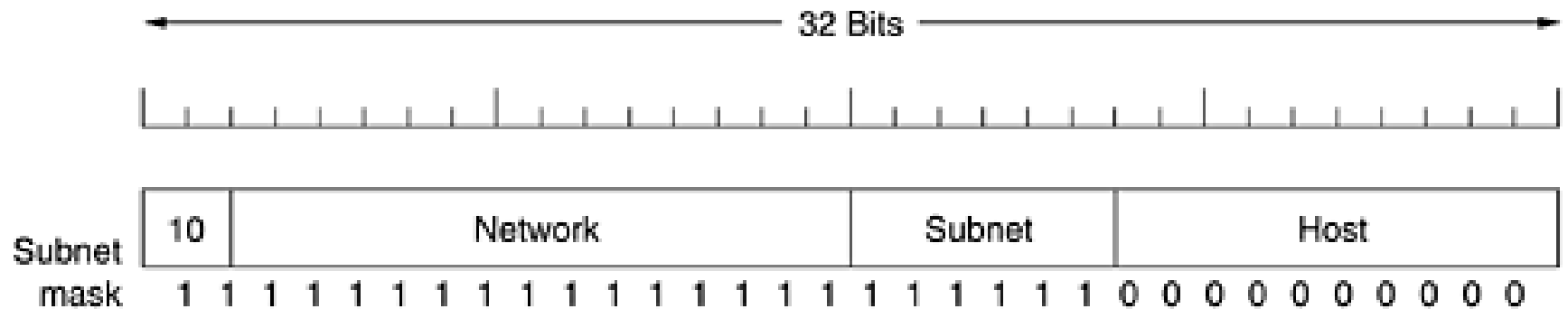
Figure 5-57. A campus network consisting of LANs for various departments.



- When a packet comes into the main router, how does it know which subnet (Ethernet) to give it to?
- One way would be to have a table with large number of entries in the main router telling which router to use for each host on campus.
- This idea would work, but it would require a very large table in the main router and a lot of manual maintenance as hosts were added, moved, or taken out of service.

- Instead, a different scheme was invented. Basically, instead of having a single class B address with 16 bits for the network number and 16 bits for the host number, some bits are taken away from the host number to create a subnet number.
- For example, if the university has 35 departments, it could use a 6-bit subnet number and a 10-bit host number, allowing for up to 64 Ethernets, each with a maximum of 1024 hosts.
- To implement subnetting, the main router needs a **subnet mask** that indicates the split between network + subnet number and host.

Figure 5-58. A class B network subnetted into 64 subnets.



- Subnet masks are also written in dotted decimal notation, with the addition of a slash followed by the number of bits in the network + subnet part.
- For the example of Fig. 5-58, the subnet mask can be written as 255.255.252.0. An alternative notation is /22 to indicate that the subnet mask is 22 bits long.

Internet Control Protocols

- In addition to IP, which is used for data transfer, the Internet has several control protocols used in the network layer.

ICMP: Internet Control Message Protocol

ARP: The Address Resolution Protocol

RARP: Reverse Address Resolution Protocol

BOOTP: bootstrap protocol

DHCP: Dynamic Host Configuration Protocol

ICMP: Internet Control Message Protocol

- The operation of the Internet is monitored closely by the routers. When something unexpected occurs, the event is reported by the **ICMP** (Internet Control Message Protocol).
- About a dozen types of ICMP messages are defined.
- Each ICMP message type is encapsulated in an IP packet.

Figure 5-61. The principal ICMP message types.

Message type	Description
Destination unreachable	Packet could not be delivered
Time exceeded	Time to live field hit 0
Parameter problem	Invalid header field
Source quench	Choke packet
Redirect	Teach a router about geography
Echo	Ask a machine if it is alive
Echo reply	Yes, I am alive
Timestamp request	Same as Echo request, but with timestamp
Timestamp reply	Same as Echo reply, but with timestamp

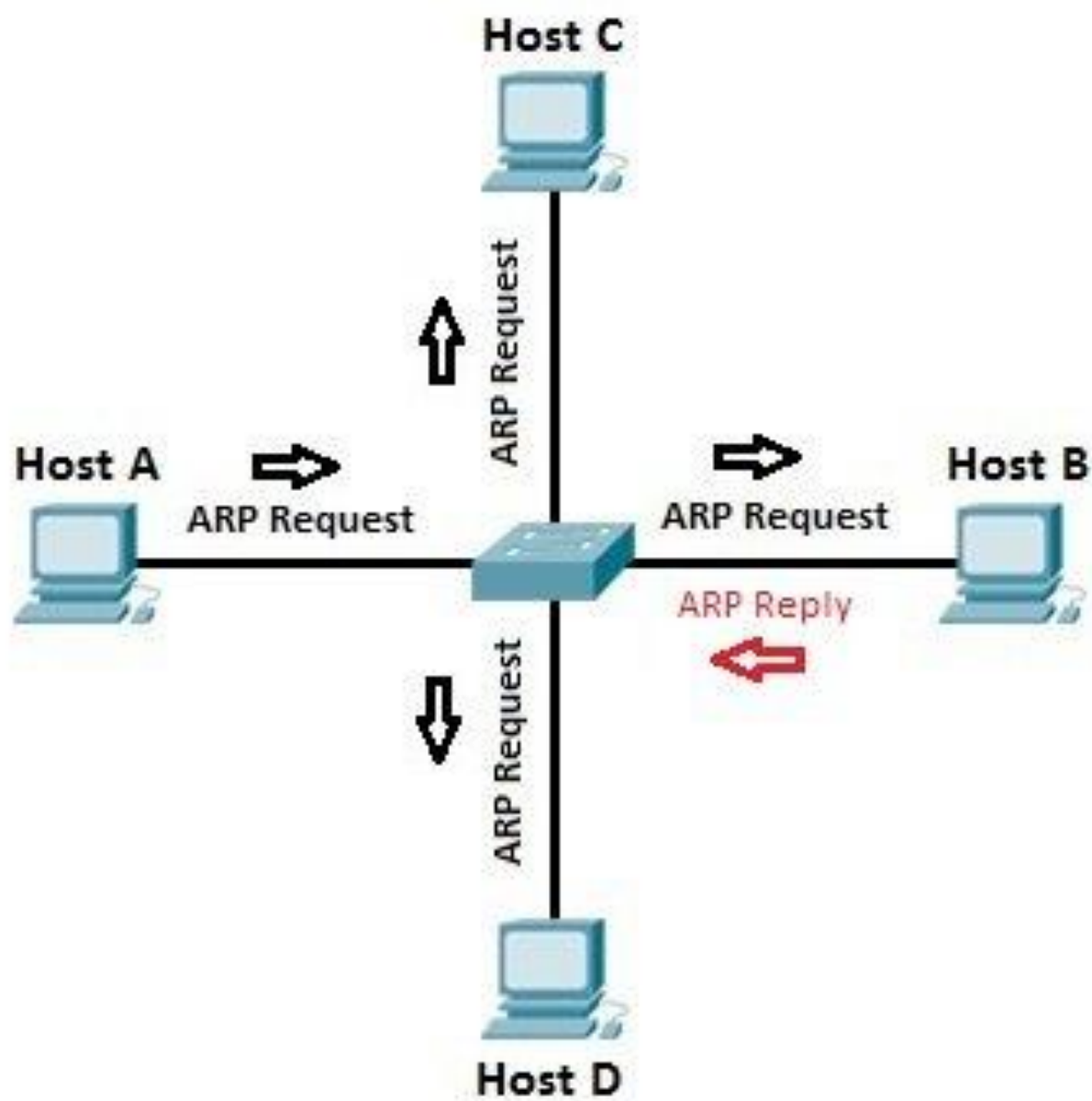
- The **DESTINATION UNREACHABLE** message is used when the subnet or a router cannot locate the destination or when a packet cannot be delivered because a "small-packet" network stands in the way.
- The **TIME EXCEEDED** message is sent when a packet is dropped because its counter has reached zero.
- The **PARAMETER PROBLEM** message indicates that an illegal value has been detected in a header field. This problem indicates a bug in the sending host's IP software or possibly in the software of a router transited.
- The **SOURCE QUENCH** message was formerly used to throttle hosts that were sending too many packets. When a host received this message, it was expected to slow down.

- The **REDIRECT** message is used when a router notices that a packet seems to be routed wrong. It is used by the router to tell the sending host about the probable error.
- The **ECHO** and **ECHO REPLY** messages are used to see if a given destination is reachable and alive. Upon receiving the ECHO message, the destination is expected to send an ECHO REPLY message back.
- The **TIMESTAMP REQUEST** and **TIMESTAMP REPLY** messages are similar, except that the arrival time of the message and the departure time of the reply are recorded in the reply. This facility is used to measure network performance.

ARP: The Address Resolution Protocol

- Although every machine on the Internet has IP address, these cannot actually be used for sending packets because the data link layer hardware does not understand Internet addresses.
- Whenever a machine needs to communicate with another machine on a local area network(LAN), it needs the MAC address for that machine.
- MAC address is the physical address of the machine.
- ARP used to resolve IP addresses to MAC addresses.
- ARP finds MAC address of a machine from its known IP address.

- Computer A wants to communicate with computer B. Now computer already knows the IP address for computer B. But in order to communicate with computer B, it needs its MAC address.
- Now an IP address is used to locate a device on network and MAC address is what identifies the actual device.
- Computer A will send out a **broadcast message** out on the network asking every device which computer has the specific IP address and will ask for their MAC address.
- Then the computer has the matching IP address will then respond back and tell computer A its MAC address.
- Then once its receives the MAC address, the communication can takes place between A and B.
- The protocol used for asking this question and getting the reply is called **ARP** (Address Resolution Protocol). Almost every machine on the Internet runs it.

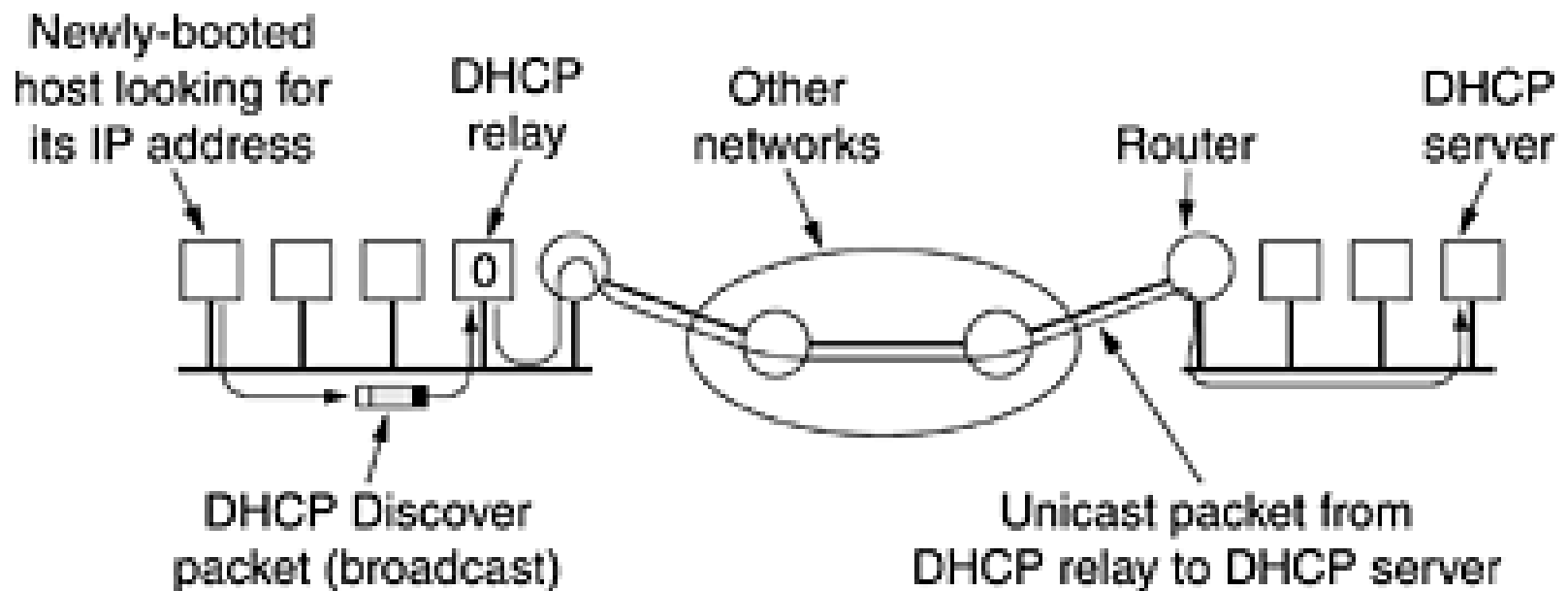


DHCP: Dynamic Host Configuration Protocol

- Every computer or device on a network has an IP address for communication purposes.
- There are two ways that a computer can be assigned an IP address: **Static IP** or **Dynamic IP**.
- **Static IP** is where a user assigns a computer or device with an IP address manually.
- Assigning IP address manually for a large networks that has a lot of computers is difficult. And also make sure that all IP addresses are unique (To avoid IP conflict).

- There is a better and easier way to assign a computer an IP address and this is called a dynamic IP.
- A dynamic IP is where the computer gets an IP address automatically from a DHCP server.
- A DHCP server automatically assigns a computer an IP address.
- DHCP is based on the idea of a special server that assigns IP addresses to hosts asking for one. This server need not be on the same LAN as the requesting host.
- Since the DHCP server may not be reachable by broadcasting, a **DHCP relay agent** is needed on each LAN.

Figure 5-63. Operation of DHCP



- To find its IP address, a newly-booted machine broadcasts a DHCP DISCOVER packet.
- The DHCP relay agent on its LAN intercepts all DHCP broadcasts.
- When it finds a DHCP DISCOVER packet, it sends the packet as a unicast packet to the DHCP server, possibly on a distant network.
- The only piece of information the relay agent needs is the IP address of the DHCP server.

- An issue that arises with automatic assignment of IP addresses from a pool is how long an IP address should be allocated.
- If a host leaves the network and does not return its IP address to the DHCP server, that address will be permanently lost. After a period of time, many addresses may be lost.
- To prevent that from happening, IP address assignment may be for a fixed period of time, a technique called **leasing**.
- Just before the lease expires, the host must ask the DHCP for a renewal. If it fails to make a request or the request is denied, the host may no longer use the IP address it was given earlier.