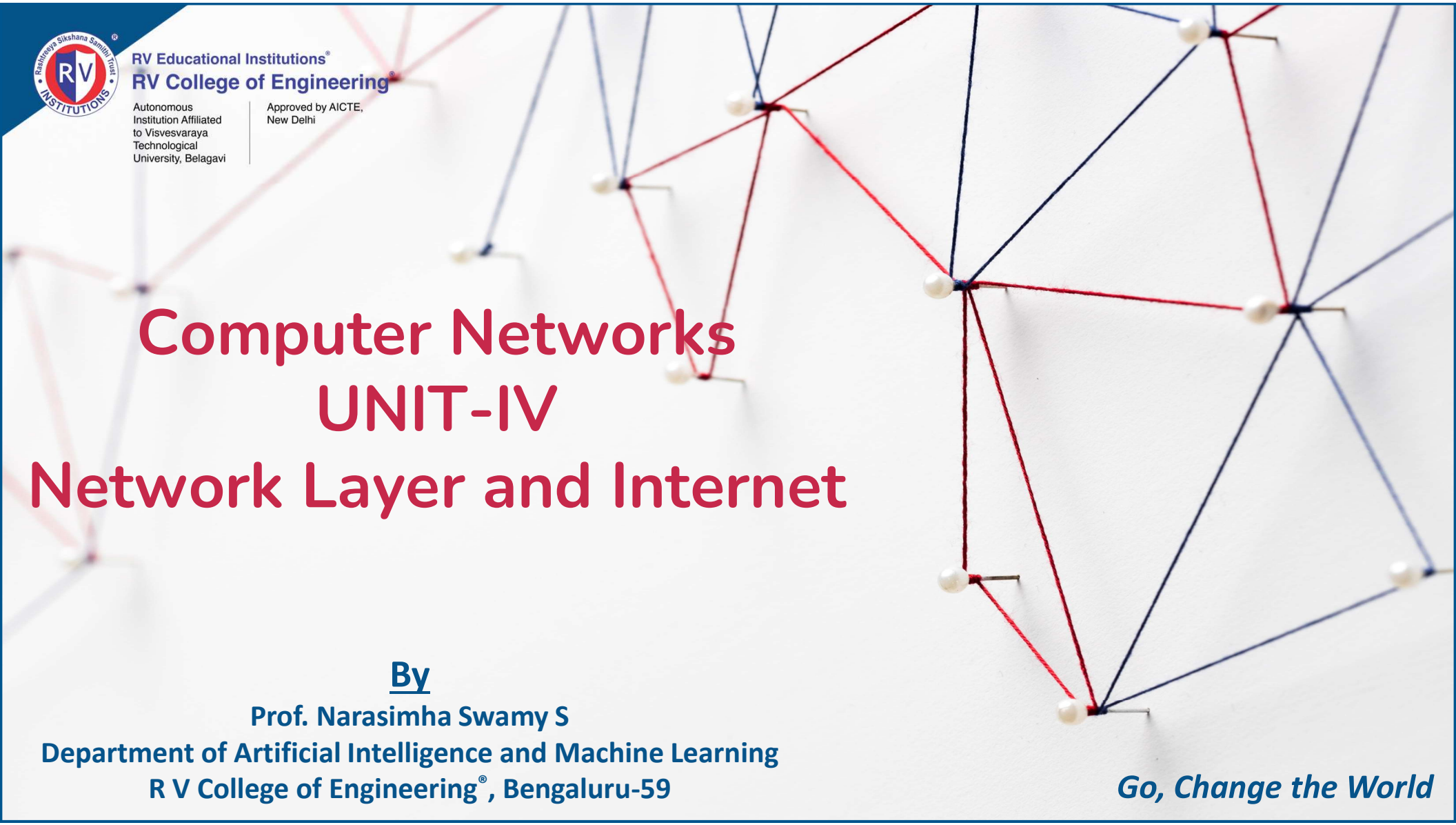
RV Educational Institutions®
RV College of Engineering®
Autonomous Institution Affiliated to Visvesvaraya Technological University, Belagavi
Approved by AICTE, New Delhi



Computer Networks
UNIT-IV
Network Layer and Internet

By
Prof. Narasimha Swamy S
Department of Artificial Intelligence and Machine Learning
R V College of Engineering®, Bengaluru-59

Go, Change the World

Outline

→ Internetworking
- > How networks differ
- > How networks can be connected
- > Connectionless Internetworking
- > Tunnelling Internetwork Routing, Fragmentation

→ The Network Layer in the Internet
- > The IP Protocol
- > IP Addresses
- > Internet Control Protocols
- > OSPF- Interior Gateway Routing Protocol
- > BGP- Exterior Gateway Routing Protocol
- > IPv6

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 PANs, LANs, MANs, and WANs
- Realistically, we have Ethernet, Internet over cable, the fixed and mobile telephone networks, 802.11, 802.16, and more
- Numerous protocols are in widespread use across these networks in every layer
- In the following sections, we will take a careful look at the issues that arise when two or more networks are connected to form an internetwork, or more simply an internet
- It would be much simpler to join networks together if everyone used a single networking technology, and it is often the case that there is a dominant kind of network, such as Ethernet

How Networks Differ

- Some of the differences that can be exposed to the network layer.
- It is papering over these differences that makes internetworking more difficult than operating within a single network

Item	Some Possibilities
Service offered	Connectionless versus connection oriented
Addressing	Different sizes, flat or hierarchical
Broadcasting	Present or absent (also multicast)
Packet size	Every network has its own maximum
Ordering	Ordered and unordered delivery
Quality of service	Present or absent; many different kinds
Reliability	Different levels of loss
Security	Privacy rules, encryption, etc.
Parameters	Different timeouts, flow specifications, etc.
Accounting	By connect time, packet, byte, or not at all

- When packets sent by a source on one network must transit one or more foreign networks before reaching the destination network, many problems can occur at the interfaces between networks

How Networks Differ (Contd.)

- What do we do if the source is on an Ethernet network and the destination is on a WiMAX network?
- Assuming we can even specify a WiMAX destination from an Ethernet network, packets would cross from a connectionless network to a connection-oriented one
- This may require that a new connection be set up on short notice, which injects a delay, and much overhead
- How do we multicast a packet to a group with some members on a network that does not support multicast?
- The differing max packet sizes used by different networks can be a major nuisance, too.
  - How do you pass an 8000-byte packet through a network whose maximum size is 1500 bytes
- If packets on a connection-oriented network transit a connectionless network, they may arrive in a different order than they were sent. That is something the sender likely did not expect, and it might come as an (unpleasant) surprise to the receiver as well

How Networks Differ (Contd.)

- The clearest example is quality of service
  - If one network has strong QoS and the other offers best effort service, it will be impossible to make bandwidth and delay guarantees for real-time traffic end to end
- Most of the time Security mechanisms are problematic, but at least encryption for confidentiality and data integrity can be layered on top of networks that do not already include it.

How Networks Can Be Connected

- There are two basic choices for connecting different networks
  - We can build devices that translate or convert packets from each kind of network into packets for each other network
  - Solve the problem by adding a layer of indirection and building a common layer on top of the different networks
- Early on, Cerf and Kahn (1974) argued for a common layer to hide the differences of existing networks. This approach has been tremendously successful, and the layer they proposed was eventually separated into the TCP and IP protocols
- Almost four decades later, IP is the foundation of the modern Internet. For this accomplishment, Cerf and Kahn were awarded the Turing Award, informally known as the Nobel Prize of computer science
- IP provides a universal packet format that all routers recognize and that can be passed through almost every network. IP has extended its reach from computer networks to take over the telephone network
- It also runs on sensor networks and other tiny devices that were once presumed too resource-constrained to support it.

How Networks Can Be Connected (Contd.)

- We have discussed several different devices that connect networks, including repeaters, hubs, switches, bridges, routers, and gateways
  - Repeaters and hubs just move bits from one wire to another. They are mostly analog devices and do not understand anything about higher layer protocols
  - Bridges and switches operate at the link layer. They can be used to build networks, but only with minor protocol translation in the process, for example, between 10, 100 and 1000 Mbps Ethernet switches
  - Our focus in this section is interconnection devices that operate at the network layer, namely the routers
  - We will leave gateways, which are higher-layer interconnection devices, until later

How Networks Can Be Connected (Contd.)

- Let us first explore at a high level how interconnection with a common network layer can be used to interconnect dissimilar networks
- An internet comprised of 802.11, MPLS, and Ethernet networks is shown in Fig. 5-39(a)
- Suppose that the source machine on the 802.11 network wants to send a packet to the destination machine on the Ethernet network
- Since these technologies are different, and they are further separated by another kind of network (MPLS), some added processing is needed at the boundaries between the networks
- Because different networks may, in general, have different forms of addressing, the packet carries a network layer address that can identify any host across the three networks
- The first boundary the packet reaches is when it transitions from an 802.11 network to an MPLS network. 802.11 provides a connectionless service, but MPLS provides a connection-oriented service
- This means that a virtual circuit must be set up to cross that network. Once the packet has traveled along the virtual circuit, it will reach the Ethernet network

How Networks Can Be Connected (Contd.)

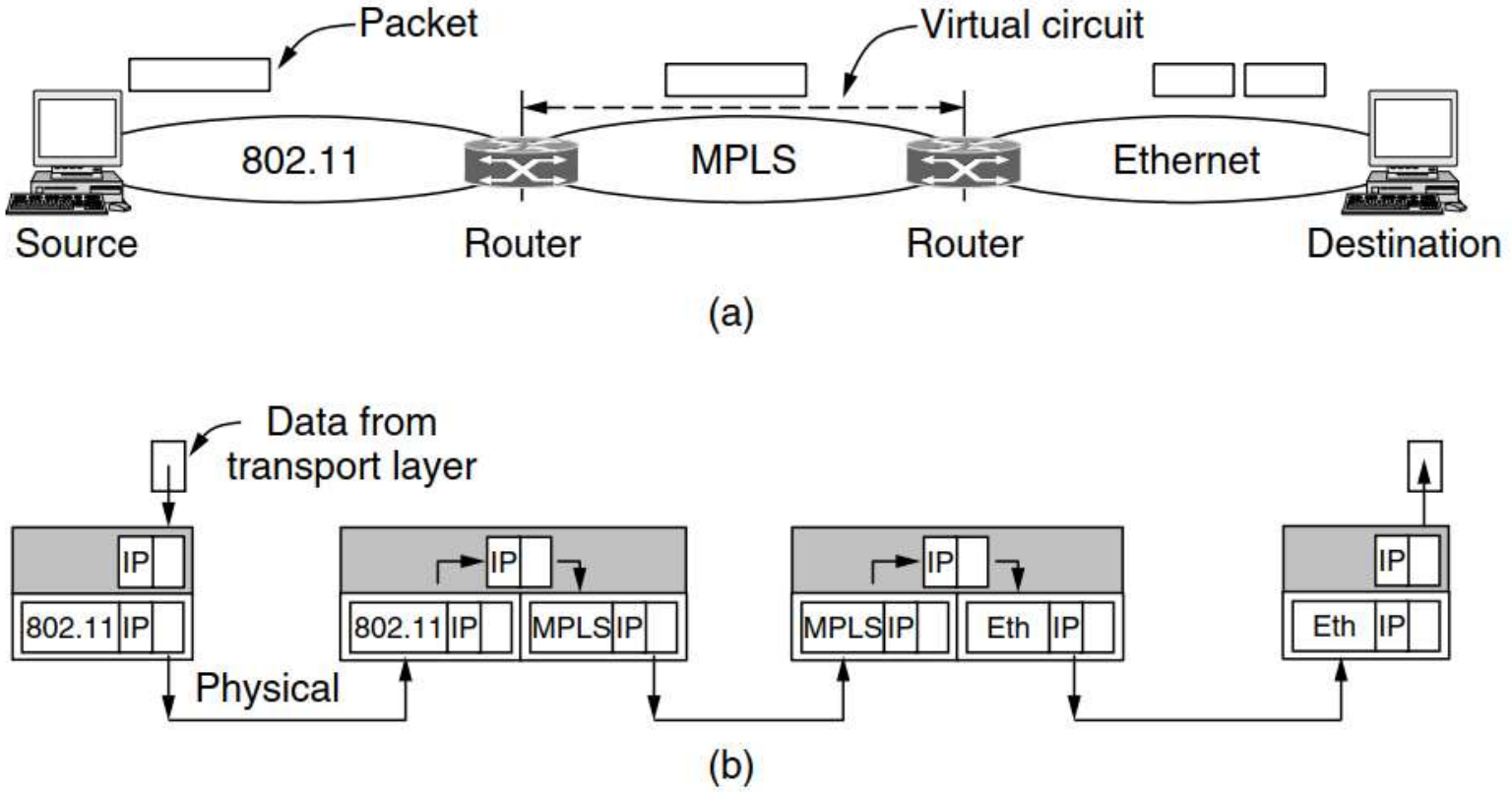


Figure 5-39. (a) A packet crossing different networks. (b) Network and link layer protocol processing.

How Networks Can Be Connected (Contd.)

- At this boundary, the packet may be too large to be carried, since 802.11 can work with larger frames than Ethernet
- To handle this problem, the packet is divided into fragments, and each fragment is sent separately
- When the fragments reach the destination, they are reassembled. Then the packet has completed its journey
- The protocol processing for this journey is shown in Fig. 5-39(b)
- The source accepts data from the transport layer and generates a packet with the common network layer header, which is IP in this example
- The network header contains the ultimate destination address, which is used to determine that the packet should be sent via the first router
- So the packet is encapsulated in an 802.11 frame whose destination is the first router and transmitted
- At the router, the packet is removed from the frame's data field and the 802.11 frame header is discarded. The router now examines the IP address in the packet and looks up this address in its routing table.

How Networks Differ (Contd.)

- Based on this address, it decides to send the packet to the second router next. For this part of the path, an MPLS virtual circuit must be established to the second router and the packet must be encapsulated with MPLS headers that travel this circuit. At the far end, the MPLS header is discarded and the network address is again consulted to find the next network layer hop
- It is the destination itself. Since the packet is too long to be sent over Ethernet, it is split into two portions. Each of these portions is put into the data field of an Ethernet frame and sent to the Ethernet address of the destination. At the destination, the Ethernet header is stripped from each of the frames, and the contents are reassembled. The packet has finally reached its destination
- Observe that there is an essential difference between the routed case and the switched (or bridged) case.
  - With a router, the packet is extracted from the frame and the network address in the packet is used for deciding where to send it
  - With a switch (or bridge), the entire frame is transported on the basis of its MAC address. Switches do not have to understand the network layer protocol being used to switch packets.

How Networks Can Be Connected (Contd.)

- Unfortunately, internetworking is not as easy as we have made it sound. In fact, when bridges were introduced, it was intended that they would join different types of networks, or at least different types of LANs
- They were to do this by translating frames from one LAN into frames from another LAN
- However, this did not work well, for the same reason that internetworking is difficult: the differences in the features of LANs, such as different maximum packet sizes and LANs with and without priority classes, are hard to mask
- Today, bridges are predominantly used to connect the same kind of network at the link layer, and routers connect different networks at the network layer

Tunneling

- Handling the general case of making two different networks interwork is exceedingly difficult
- However, there is a common special case that is manageable even for different network protocols. This case is where the source and destination hosts are on the same type of network, but there is a different network in between
- As an example, think of an international bank with an IPv6 network in Paris, an IPv6 network in London and connectivity between the offices via the IPv4 Internet. This situation is shown in Fig. 5-40.

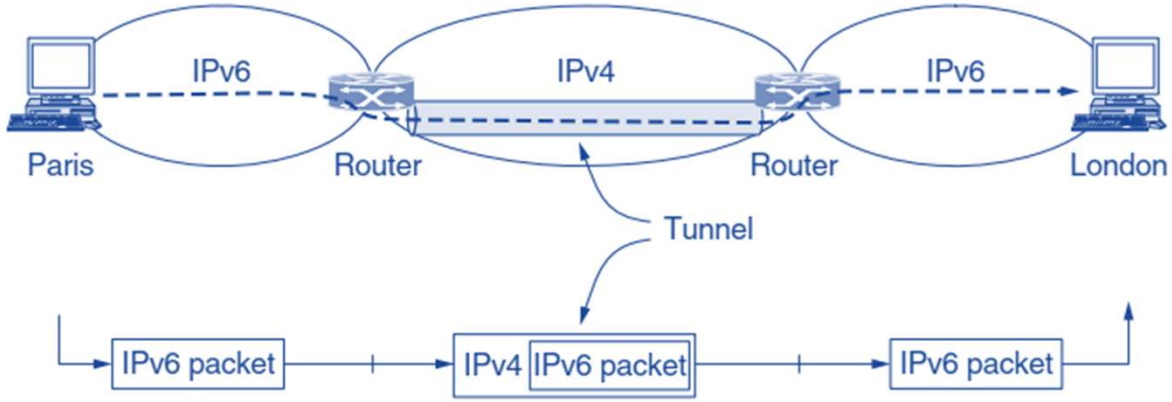


Figure 5-40. Tunneling a packet from Paris to London.

Tunneling (Contd.)

- The solution to this problem is a technique called **Tunneling**
- To send an IP packet to a host in the London office, a host in the Paris office constructs the packet containing an IPv6 address in London, and sends it to the multiprotocol router that connects the Paris IPv6 network to the IPv4 Internet
- When this router gets the IPv6 packet, it encapsulates the packet with an IPv4 header addressed to the IPv4 side of the multiprotocol router that connects to the London IPv6 network. That is, the router puts a (IPv6) packet inside a (IPv4) packet
- When this wrapped packet arrives, the London router removes the original IPv6 packet and sends it onward to the destination host
- Only the multi-protocol routers have to understand both IPv4 and IPv6 packets. In effect, the entire trip from one multiprotocol router to the other is like a hop over a single link
- An analogy may make tunneling clearer. Consider a person driving her car from Paris to London

Tunneling (Contd.)

- Within France, the car moves under its own power, but when it hits the English Channel, it is loaded onto a high-speed train and transported to England through the Chunnel (cars are not permitted to drive through the Chunnel). Effectively, the car is being carried as freight, as depicted in Fig. 5-41.
- At the far end, the car is let loose on the English roads and once again continues to move under its own power. Tunneling of packets through a foreign network works the same way

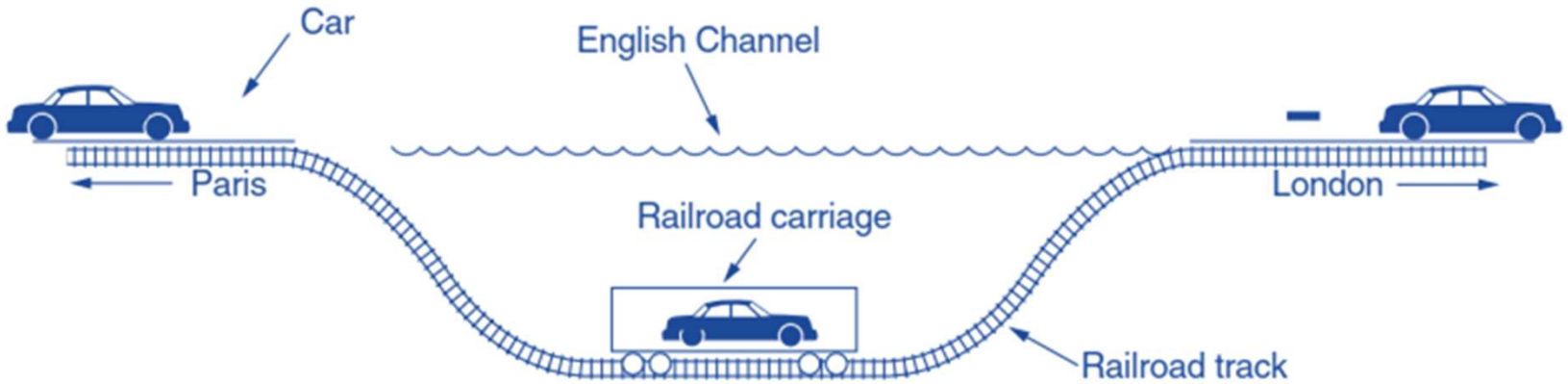


Figure 5-41. Tunneling a car from France to England.

Tunneling (Contd.)

- Tunneling is widely used to connect isolated hosts and networks using other networks.
- The network that results is called an overlay since it has effectively been overlaid on the base network.

Internetwork Routing

- Routing through an internet poses the same basic problem as routing within a single network, but with some added complications
- To start, the networks may internally use different routing algorithms. For example, one network may use link state routing and another distance vector routing
- Since link state algorithms need to know the topology but distance vector algorithms do not, this difference alone would make it unclear how to find the shortest paths across the internet
- Networks run by different operators lead to bigger problems. First, the operators may have different ideas about what is a good path through the network.
  - One operator may want the route with the least delay, while another may want the most inexpensive route. This will lead the operators to use different quantities to set the shortest-path costs (e.g., milliseconds of delay vs. monetary cost).
  - The weights will not be comparable across networks, so shortest paths on the internet will not be well defined
- Worse yet, one operator may not want another operator to even know the details of the paths in its network, perhaps because the weights and paths may reflect sensitive information (such as the monetary cost) that represents a competitive business advantage

Internetwork Routing (Contd.)

- Finally, the internet may be much larger than any of the networks that comprise it. It may therefore require routing algorithms that scale well by using a hierarchy
- All of these considerations lead to a two-level routing algorithm
- Within each network, an intradomain or interior gateway protocol is used for routing. (“Gateway” is an older term for “router.”)
- Across the networks that make up the internet, an interdomain or exterior gateway protocol is used
- The networks may all use different intradomain protocols, but they must use the same interdomain protocol
- In the Internet, the interdomain routing protocol is called BGP (Border Gateway Protocol).
- There is one more important term to introduce. Since each network is operated independently of all the others, it is often referred to as an AS (Autonomous System)

Internetwork Routing (Contd.)

- In the Internet, a large determining factor is the business arrangements between ISPs
- Each ISP may charge or receive money from the other ISPs for carrying traffic. Another factor is that if internetwork routing requires crossing international boundaries, various laws may suddenly come into play, such as Sweden's strict privacy laws about exporting personal data about Swedish citizens from Sweden.
- All of these nontechnical factors are wrapped up in the concept of a routing policy that governs the way autonomous networks select the routes that they use

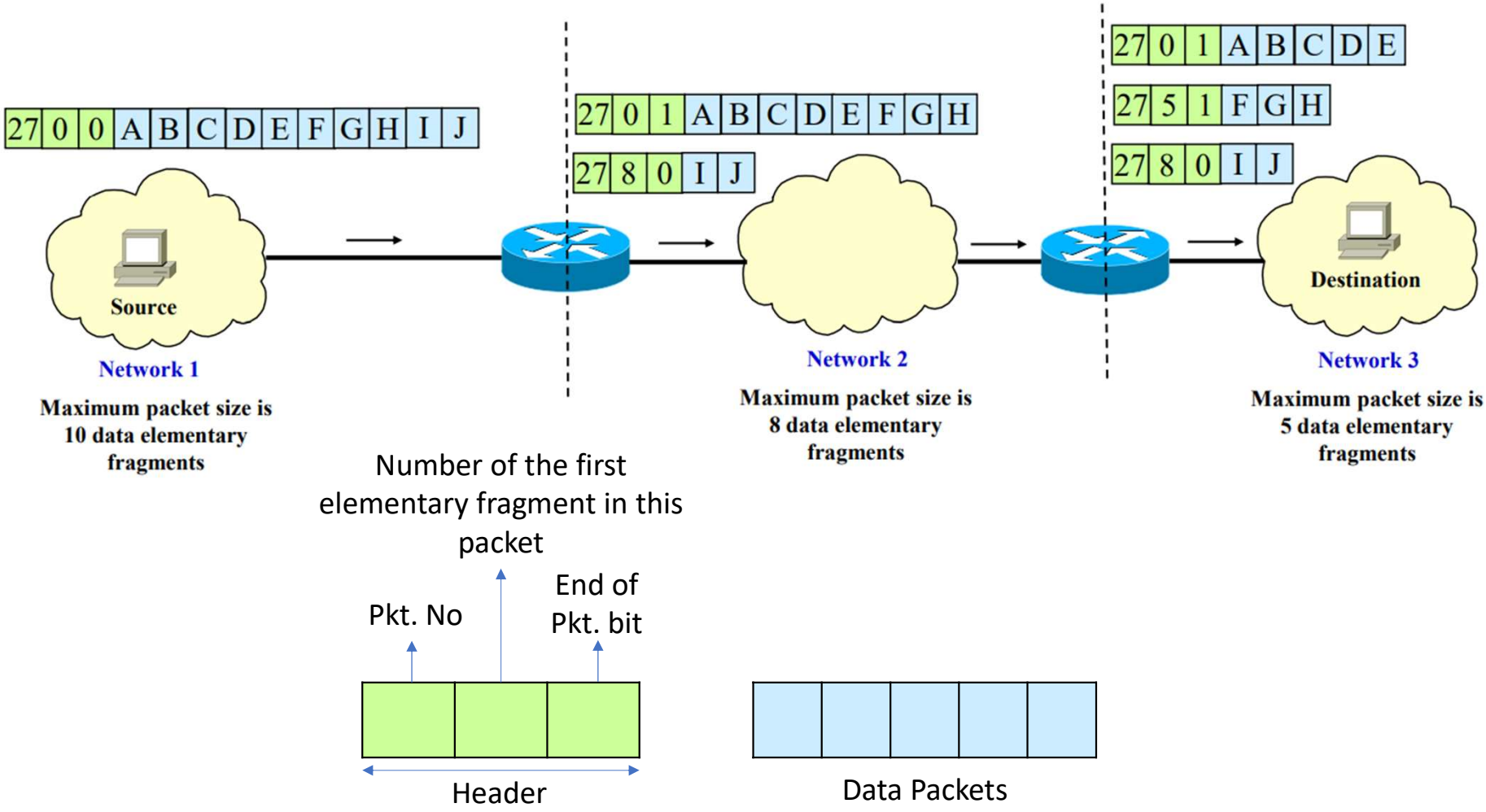
Packet Fragmentation

- **Fragmentation** is done by the network layer when the maximum size of datagram is greater than maximum size of data that can be held in a frame i.e., its Maximum Transmission Unit (MTU)
- Each network or link 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
- The result of all these factors is that the network designers are not free to choose any old maximum packet size they wish. Maximum payloads for some common technologies are 1500 bytes for Ethernet and 2272 bytes for 802.11. IP is more generous, allows for packets as big as 65,515 bytes

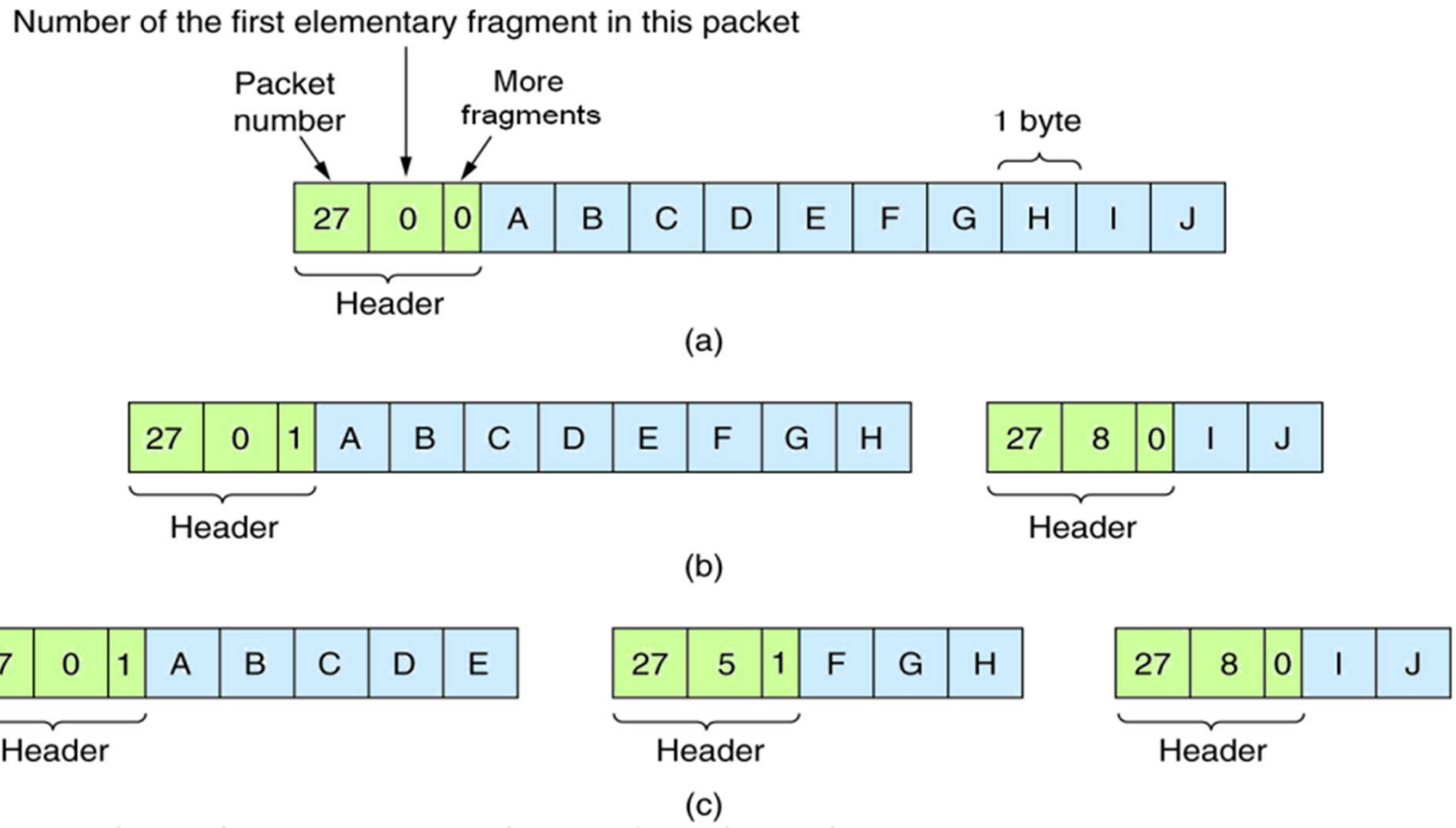
Packet Fragmentation

- A source does not usually know the path a packet will take through the network to a destination, so it certainly does not know how small packets must be to get there. This packet size is called the Path MTU (Path Maximum Transmission Unit).
- Even if the source did know the path MTU, packets are routed independently in a connectionless network such as the Internet. This routing means that paths may suddenly change, which can unexpectedly change the path MTU
- The solution to the problem is to allow routers to break up packets into fragments, sending each fragment as a separate network layer packet.
- Two opposing strategies exist for recombining the fragments back into the original packet.
  - The first strategy is to make fragmentation caused by a "small packet" network transparent to any subsequent networks through which the packet must pass on its way to the ultimate destination. This option is shown in Fig. 5-42(a).

Packet Fragmentation (Contd.)



Packet Fragmentation (Contd.)



Packet Fragmentation (Contd.)

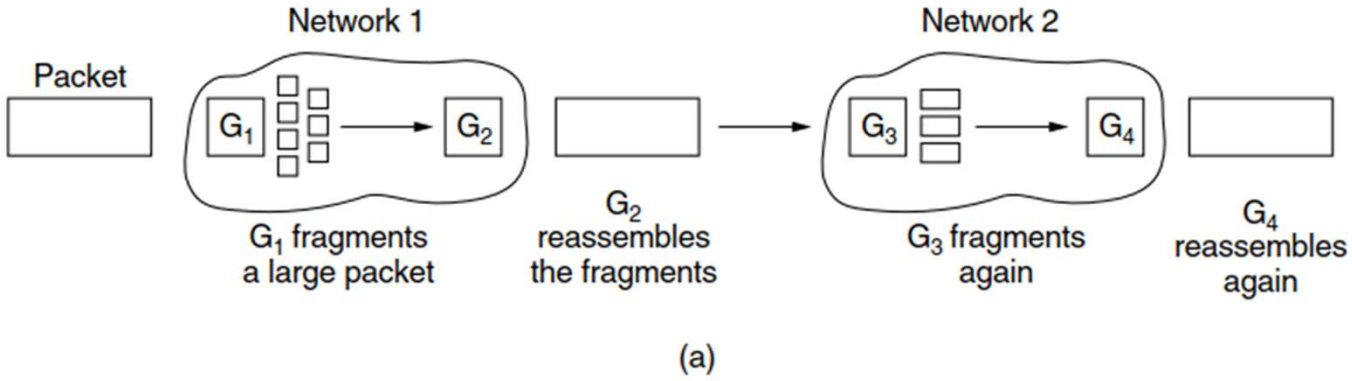


Figure 5-42. (a) Transparent fragmentation.

- In this approach, when an oversized packet arrives at G1, the router breaks it up into fragments. Each fragment is addressed to the same exit router, G2, where the pieces are recombined. In this way, passage through the small-packet network is made transparent. Subsequent networks are not even aware that fragmentation has occurred
- Transparent fragmentation is straightforward but has some problems. For one thing, the exit router must know when it has received all the pieces, so either a count field or an "end of packet" bit must be provided. Also, because all packets must exit via the same router so that they can be reassembled, the routes are constrained. By not allowing some fragments to follow one route to the ultimate destination and other fragments a disjoint route, some performance may be lost

Packet Fragmentation (Contd.)

- The other fragmentation strategy is to refrain from recombining fragments at any intermediate routers. Once a packet has been fragmented, each fragment is treated as though it were an original packet. The routers pass the fragments, as shown in Fig. 5-42(b), and reassembly is performed only at the destination host.

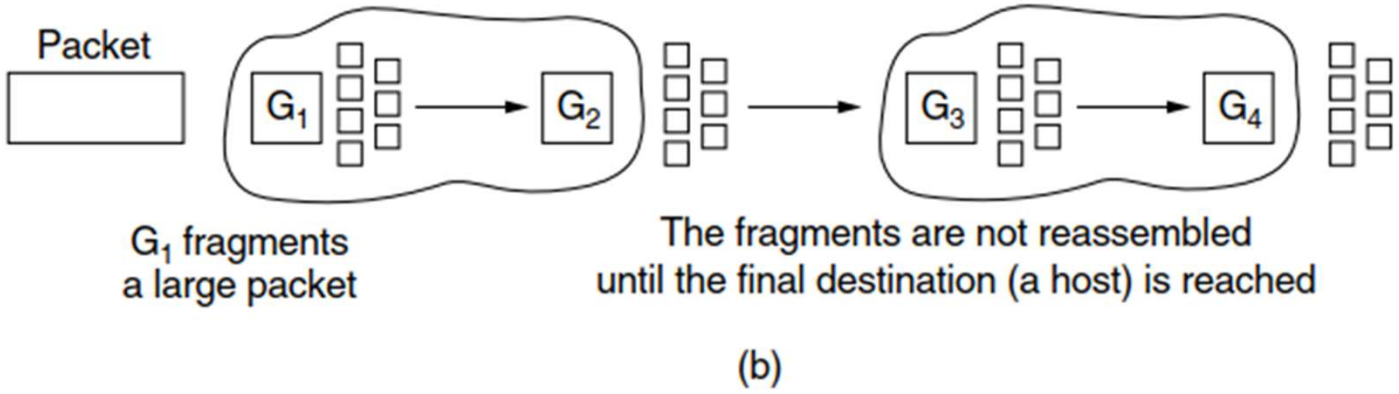


Figure 5-42. (b) Nontransparent fragmentation.

- The main advantage of non-transparent fragmentation is that it requires routers to do less work. IP works this way. A complete design requires that the fragments be numbered in such a way that the original data stream can be reconstructed.

The Network Layer in the Internet

- It is now time to discuss the network layer of the Internet in detail
- But before getting into specifics, it is worth taking a look at the principles that drove its design in the past and made it the success that it is today
- These principles are enumerated and discussed in RFC 1958, which is well worth reading (and should be mandatory for all protocol designers—with a final exam at the end)
- This RFC draws heavily on ideas put forth by Clark (1988) and Saltzer et al. (1984)

The Network Layer in the Internet (Contd.)

- Make sure it works
  Do not finalize the design or standard until multiple prototypes have successfully communicated with each other. All too often, designers first write a 1000-page standard, get it approved, then discover it is deeply flawed and does not work. Then they write version 1.1 of the standard. This is not the way to go
- Keep it simple
  When in doubt, use the simplest solution. William of Occam stated this principle (Occam's razor) in the 14th century. Put in modern terms: fight features. If a feature is not absolutely essential, leave it out, especially if the same effect can be achieved by combining other features
- Make clear choices
  If there are several ways of doing the same thing, choose one. Having two or more ways to do the same thing is looking for trouble. Standards often have multiple options or modes or parameters because several powerful parties insist that their way is best. Designers should strongly resist this tendency. Just say no.

The Network Layer in the Internet (Contd.)

- Exploit modularity
  This principle leads directly to the idea of having protocol stacks, each of whose layers is independent of all the other ones. In this way, if circumstances require one module or layer to be changed, the other ones will not be affected
- Expect heterogeneity
  Different types of hardware, transmission facilities, and applications will occur on any large network. To handle them, the network design must be simple, general, and flexible.
- Avoid static options and parameters
  If parameters are unavoidable (e.g., maximum packet size), it is best to have the sender and receiver negotiate a value rather than defining fixed choices.
- Look for a good design; it need not be perfect
  Often, the designers have a good design but it cannot handle some weird special case. Rather than messing up the design, the designers should go with the good design and put the burden of working around it on the people with the strange requirements.

The Network Layer in the Internet (Contd.)

- Be strict when sending and tolerant when receiving

  In other words, send only packets that rigorously comply with the standards, but expect incoming packets that may not be fully conformant and try to deal with them.

- Think about scalability

  If the system is to handle millions of hosts and billions of users effectively, no centralized databases of any kind are tolerable and load must be spread as evenly as possible over the available resources.

- Consider performance and cost

  If a network has poor performance or outrageous costs, nobody will use it

The Network Layer in the Internet (Contd.)

- Be strict when sending and tolerant when receiving

  In other words, send only packets that rigorously comply with the standards, but expect incoming packets that may not be fully conformant and try to deal with them.

- Think about scalability

  If the system is to handle millions of hosts and billions of users effectively, no centralized databases of any kind are tolerable and load must be spread as evenly as possible over the available resources.

- Consider performance and cost

  If a network has poor performance or outrageous costs, nobody will use it

The Network Layer in the Internet (Contd.)

- In the network layer, the Internet can be viewed as a collection of networks or ASes (Autonomous Systems) that are interconnected
- There is no real structure, but several major backbones exist. These are constructed from high-bandwidth lines and fast routers.
- Attached to the backbones are ISPs (Internet Service Providers) that provide Internet access to homes and businesses, data centers and colocation facilities full of server machines, and regional (mid-level) networks.
- The data centers serve much of the content that is sent over the Internet. Attached to the regional networks are more ISPs, LANs at many universities and companies, and other edge networks
- A sketch of this quasi hierarchical organization is given in Fig. 5-45.

The Network Layer in the Internet (Contd.)

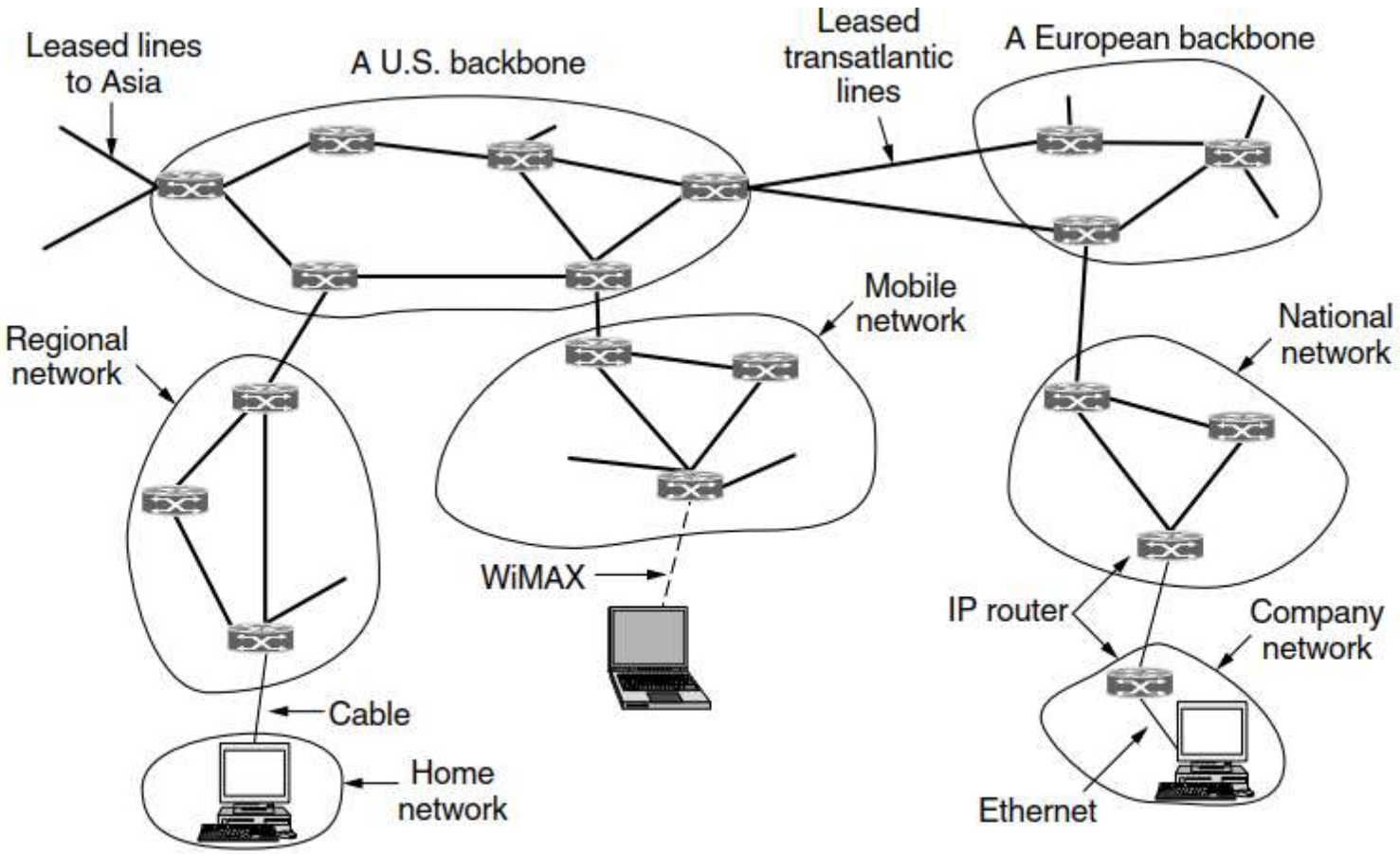


Figure 5-45. The Internet is an interconnected collection of many networks.

The Network Layer in the Internet (Contd.)

- Communication in the Internet works as follows.
  - The transport layer takes data streams and breaks them up so that they may be sent as IP packets
  - In theory, packets can be up to 64 KB each, but in practice they are usually not more than 1500 bytes (so they fit in one Ethernet frame).
  - IP routers forward each packet through the Internet, along a path from one router to the next, until the destination is reached
  - At the destination, the network layer hands the data to the transport layer, which gives it to the receiving process. 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.

The Network Layer in the Internet (Contd.)

The IP Version 4 Protocol

An IPv4 datagram consists of a header part and a body or payload part. The header has a 20-byte fixed part and a variable-length optional part

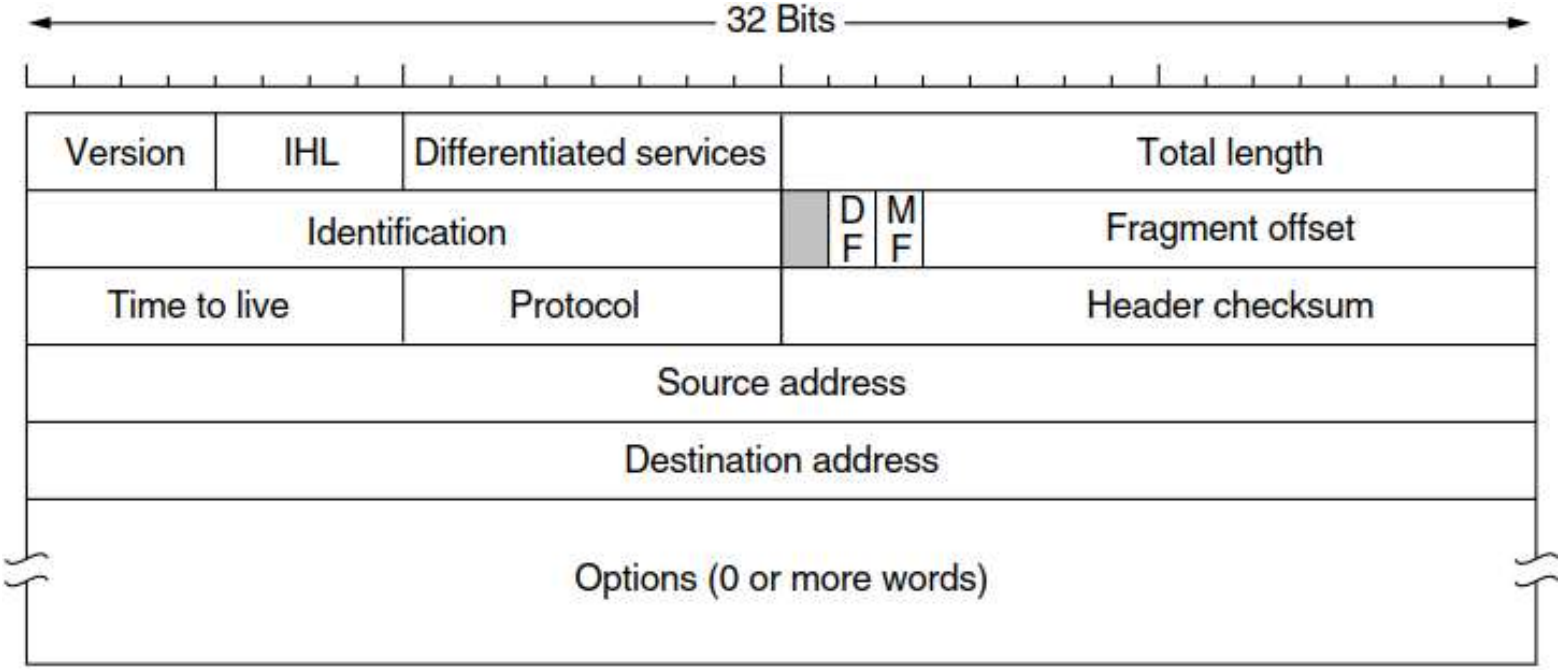
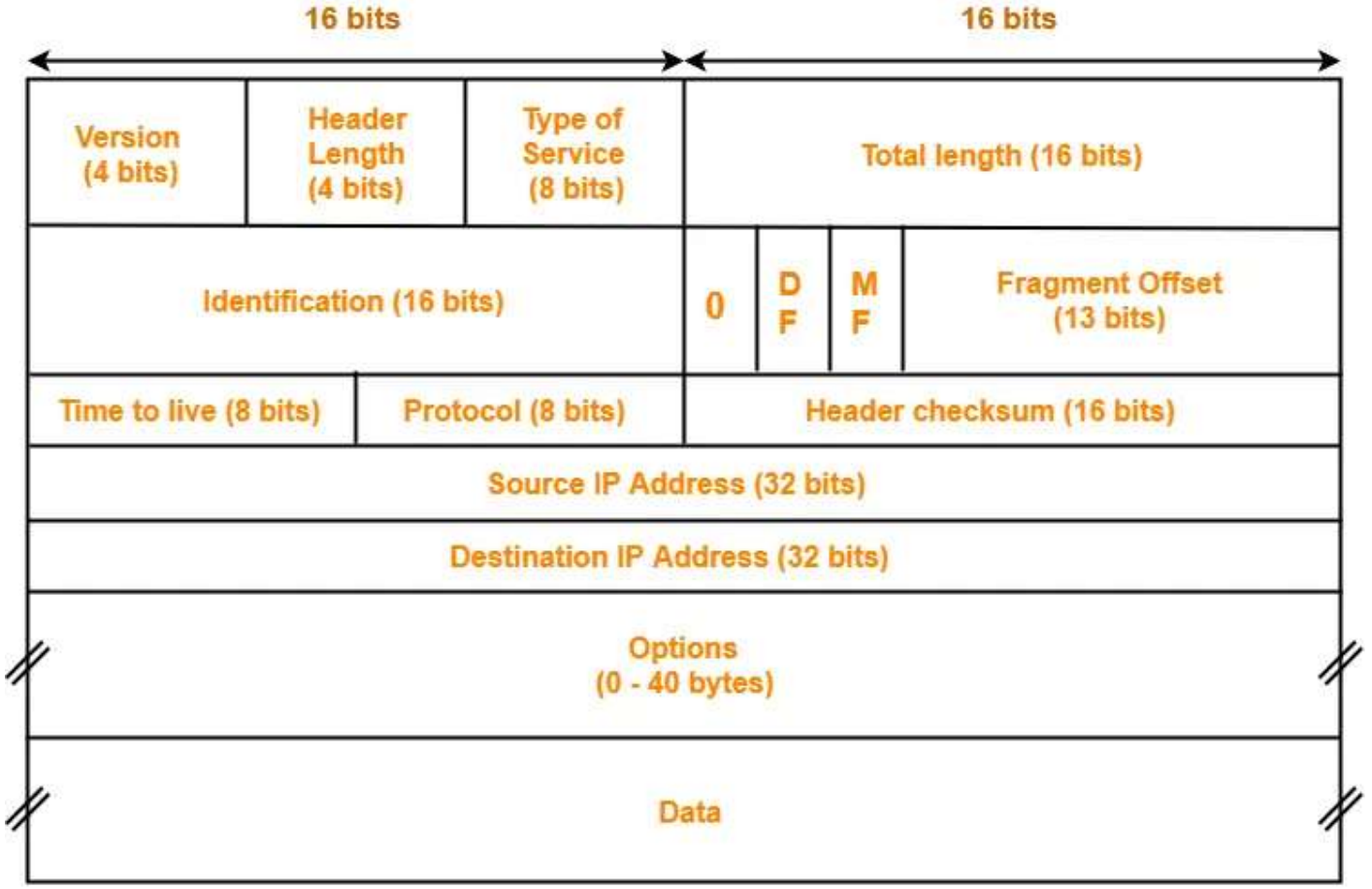


Figure 5-46. The IPv4 (Internet Protocol) header.

The Network Layer in the Internet (Contd.)

The IP Version 4 Protocol



The Network Layer in the Internet (Contd.)

The IP Version 4 Protocol

- **Version -4 bits (0-3 bits)**
  - The Version field keeps track of which version of the protocol the datagram belongs to.
  - Version 4 dominates the Internet today, and that is where we have started our discussion
  - By including the version at the start of each datagram, it becomes possible to have a transition between versions over a long period of time. In fact, IPv6, the next version of IP, was defined more than a decade ago, yet is only just beginning to be deployed.
- **Internet Header Length (IHL) – 4 bits (4-7 bits)**
  - Since the header length is not constant, a field in the header, IHL, is provided to tell how long the header is, in 32-bit words
  - The minimum value is 5, which applies when no options are present. The maximum value of this 4-bit field is 15, which limits the header to 60 bytes, and thus the Options field to 40 bytes
  - For some options, such as one that records the route a packet has taken, 40 bytes is far too small, making those options useless.

The Network Layer in the Internet (Contd.)

The IP Version 4 Protocol

- **Differentiated Services- 8 bits (8-15 bits)**
  - Originally, it was called the Type of service field. It was and still is intended to distinguish between different classes of service
  - Various combinations of reliability and speed are possible. For digitized voice, fast delivery beats accurate delivery
  - For file transfer, error-free transmission is more important than fast transmission. The Type of service field provided 3 bits to signal priority and 3 bits to signal whether a host cared more about delay, throughput, or reliability.
  - Now, the top 6 bits are used to mark the packet with its service class; we described the expedited and assured services earlier in this chapter.
  - The bottom 2 bits are used to carry explicit congestion notification information, such as whether the packet has experienced congestion
- **Total length 8 bits (16-31 bits)**
  - 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 networks, larger datagrams may be needed

The Network Layer in the Internet (Contd.)

The IP Version 4 Protocol

- **Identification Field-16 bits**
  - The Identification field is needed to allow the destination host to determine which packet a newly arrived fragment belongs to. All the fragments of a packet contain the same Identification value
- **Unused Field/ bit - 1 bit**
  - Next comes an unused bit, which is surprising, as available real estate in the IP header is extremely scarce
  - This would greatly simplify security, as packets with the "evil" bit set would be known to have been sent by attackers and could just be discarded. Unfortunately, network security is not this simple
- **Don't Fragment (DF) - 1 bit**
  - This is a 1-bit fields related to fragmentation. Originally, it was intended to support hosts incapable of putting the pieces back together again.
  - Now it is used as part of the process to discover the path MTU, which is the largest packet that can travel along a path without being fragmented

The Network Layer in the Internet (Contd.)

The IP Version 4 Protocol

- **More Fragments (MF) - 1 bit**
  - 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 – 13 bits**
  - The Fragment offset tells where in the current packet this fragment belongs
  - This offset tells the exact position of the fragment in the original IP Packet
  - Working together, the Identification, MF, and Fragment offset fields are used to implement fragmentation
- **Time-to-Leave – 8 bits**
  - This field is a counter used to limit packet lifetimes. It was originally 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 a packet is queued for a long time in a router. In practice, it just counts hops.

The Network Layer in the Internet (Contd.)

The IP Version 4 Protocol

- When it hits zero, the packet is discarded and a warning packet is sent back to the source host

- **Protocol field – 8 bits**
  - TCP is one possibility, but so are UDP and some others. The numbering of protocols is global across the entire Internet.
- **Header checksum – 16 bits**
  - This field is used to keep checksum value of entire header which is then used to check if the packet is received error-free
- **Source address and Destination address – 32 bits**
  - The Source address and Destination address indicate the IP address of the source and destination network interfaces
- **Options field – 40 bytes**
  - It was designed to provide an escape to allow subsequent versions of the protocol to include information not present in the original design, to permit experimenters to try out new ideas, and to avoid allocating header bits to information that is rarely needed

The Network Layer in the Internet (Contd.)

The IP Version 4 Protocol

- The options are of 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. The Options field is padded out to a multiple of 4 bytes. Originally, the five options listed in Fig. 5-47 were defined

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

Figure 5-47. Some of the IP options.

- The Security option tells how secret the information is. In theory, a military router might use this field to specify not to route packets through certain countries the military considers to be ‘‘bad guys.’’ In practice, all routers ignore it, so its only practical function is to help spies find the good stuff more easily.

The Network Layer in the Internet (Contd.)

The IP Version 4 Protocol

- The Strict source routing option gives the complete path from source to destination as a sequence of IP addresses. The datagram is required to follow that exact route. It is most useful for system managers who need to send emergency packets when the routing tables have been corrupted, or for making timing measurements.
- The Loose source routing option requires the packet to traverse the list of routers specified, in the order specified, but it is allowed to pass through other routers on the way. Normally, this option will provide only a few routers, to force a particular path.
- The Record route option tells each router along the path to append its IP address to the Options field. This allows system managers to track down bugs in the routing algorithms ("Why are packets from Houston to Dallas visiting Tokyo first?"). When the ARPANET was first set up, no packet ever passed through more than nine routers, so 40 bytes of options was plenty. As mentioned above, now it is too small.
- Finally, the Timestamp option is like the Record route option, except that in addition to recording its 32-bit IP address, each router also records a 32-bit time stamp. This option, too, is mostly useful for network measurement.

The Network Layer in the Internet (Contd.)

IP Version 6

- It has worked extremely well, as demonstrated by the exponential growth of the Internet
- IPv6 (IP version 6) is a replacement for IPv4 and it uses 128-bit addresses; a shortage of these addresses is not likely any time in the foreseeable future. However, IPv6 has proved very difficult to deploy
- It is a different network layer protocol that does not really interwork with IPv4, despite many similarities
- Its major goals were
  1. Support billions of hosts, even with inefficient address allocation
  2. Reduce the size of the routing tables
  3. Simplify the protocol, to allow routers to process packets faster
  4. Provide better security (authentication and privacy)
  5. Pay more attention to the type of service, particularly for real-time data
  6. Aid multicasting by allowing scopes to be specified
  7. Make it possible for a host to roam without changing its address
  8. Allow the protocol to evolve in the future
  9. Permit the old and new protocols to coexist for years

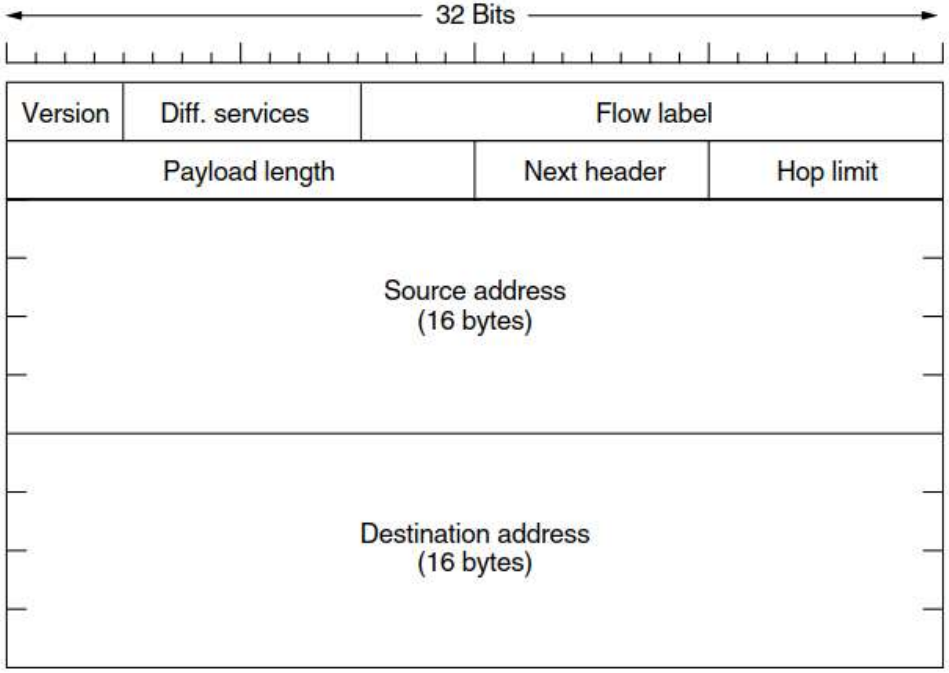
The Network Layer in the Internet (Contd.)

IP Version 6 Advantages

- First and foremost, IPv6 has longer addresses than IPv4. They are 128 bits long, which solves the problem that IPv6 set out to solve: providing an effectively unlimited supply of Internet addresses
- The second major improvement of IPv6 is the simplification of the header. It contains only seven fields (versus 13 in IPv4). This change allows routers to process packets faster and thus improves throughput and delay
- The third major improvement is better support for options. This change was essential with the new header because fields that previously were required are now optional (because they are not used so often). In addition, the way options are represented is different, making it simple for routers to skip over options not intended for them.

The Network Layer in the Internet (Contd.)

IP Version 6 Header



Version: It is a 4-bit field. It represents version IPV6 in binary 0110 (6)

Differentiated Services: This field (originally called Traffic class) is used to distinguish the class of service for packets with different real-time delivery

Flow label: It is 20-bits field. Sequential flow of the packets is maintained belonging to a datagram. This field avoid to re-ordering of data packets because all data travel in a single path. It is designed for streaming/real-time media

Payload Length. It is 16-bits field. It tells the routers about the size of payload which belongs to a particular packet. With 16 bits, up to 65535 bytes can be indicated of data

Next Header: it is 8-bits field. It tells the type of Extension Header which are additionally used with base header to send more data or information's. Some extension headers are given below

The Network Layer in the Internet (Contd.)

IP Version 6 Header

Extension header	Description
Hop-by-hop options	Miscellaneous information for routers
Destination options	Additional information for the destination
Routing	Loose list of routers to visit
Fragmentation	Management of datagram fragments
Authentication	Verification of the sender's identity
Encrypted security payload	Information about the encrypted contents

Figure 5-57. IPv6 extension headers.

Hop Limit: it is 8-bits field. This field is used to stop packet to loop in the network infinitely. This is same as TTL in IPv4. The Hop-Limit field value is decremented by 1 as it passes a link (i.e. router). When the value of Hop-limit field reaches 0 the packet is discarded

Source Address(128-bits): This field indicates the address of originator of the packet.

Destination Address(128-bits): This field provides the address of intended recipient of the packet.

The Network Layer in the Internet (Contd.)

IP Version 6 Header

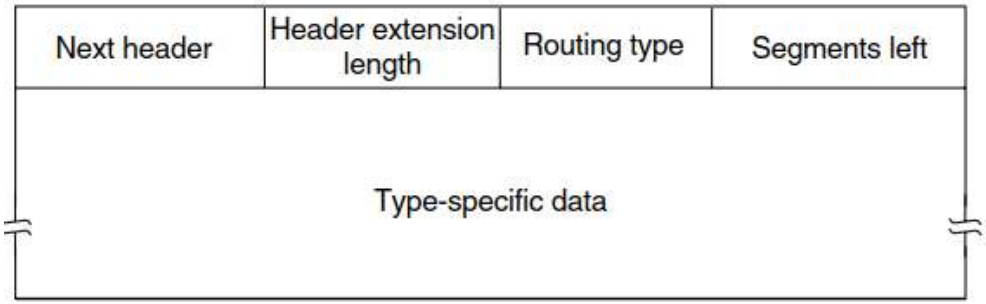


Figure 5-59. The extension header for routing.

The Network Layer in the Internet (Contd.)

IPv4 vs. IPv6

IPv4	IPv6
Deployed 1981	Deployed 1998
32-bit IP address	128-bit IP address
4.3 billion addresses Addresses must be reused and masked	7.9×10^{28} addresses Every device can have a unique address
Numeric dot-decimal notation 192.168.5.18	Alphanumeric hexadecimal notation 50b2:6400:0000:0000:6c3a:b17d:0000:10a9 (Simplified - 50b2:6400::6c3a:b17d:0:10a9)

Internet Control Protocols

- In addition to IP, which is used for data transfer, the Internet has several companion control protocols that are used in the network layer. They include ICMP, ARP, and DHCP
- In this section, we will look at each of these in turn, describing the versions that correspond to IPv4 because they are the protocols that are in common use
- ICMP and DHCP have similar versions for IPv6; the equivalent of ARP is called NDP (Neighbor Discovery Protocol) for IPv6.

Internet Control Protocols (Contd.)

IMCP—The Internet Control Message Protocol

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

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 and echo reply	Check if a machine is alive
Timestamp request/reply	Same as Echo, but with timestamp
Router advertisement/solicitation	Find a nearby router

- The most important ones are listed in Fig. 5-60.

Figure 5-60. The principal ICMP message types.

Internet Control Protocols (Contd.)

IMCP—The Internet Control Message Protocol

- The DESTINATION UNREACHABLE message is used when the router cannot locate the destination or when a packet with the DF bit 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 TtL (Time to live) 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 long ago 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 incorrectly. It is used by the router to tell the sending host to update to a better route
- The ECHO and ECHO REPLY messages are sent by hosts to see if a given destination is reachable and currently alive. Upon receiving the ECHO message, the destination is expected to send back an ECHO REPLY message. These messages are used in the ping utility that checks if a host is up and on the Internet.

Internet Control Protocols (Contd.)

IMCP—The Internet Control Message Protocol

- 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
- The ROUTER ADVERTISEMENT and ROUTER SOLICITATION messages are used to let hosts find nearby routers. A host needs to learn the IP address of at least one router to be able to send packets off the local network

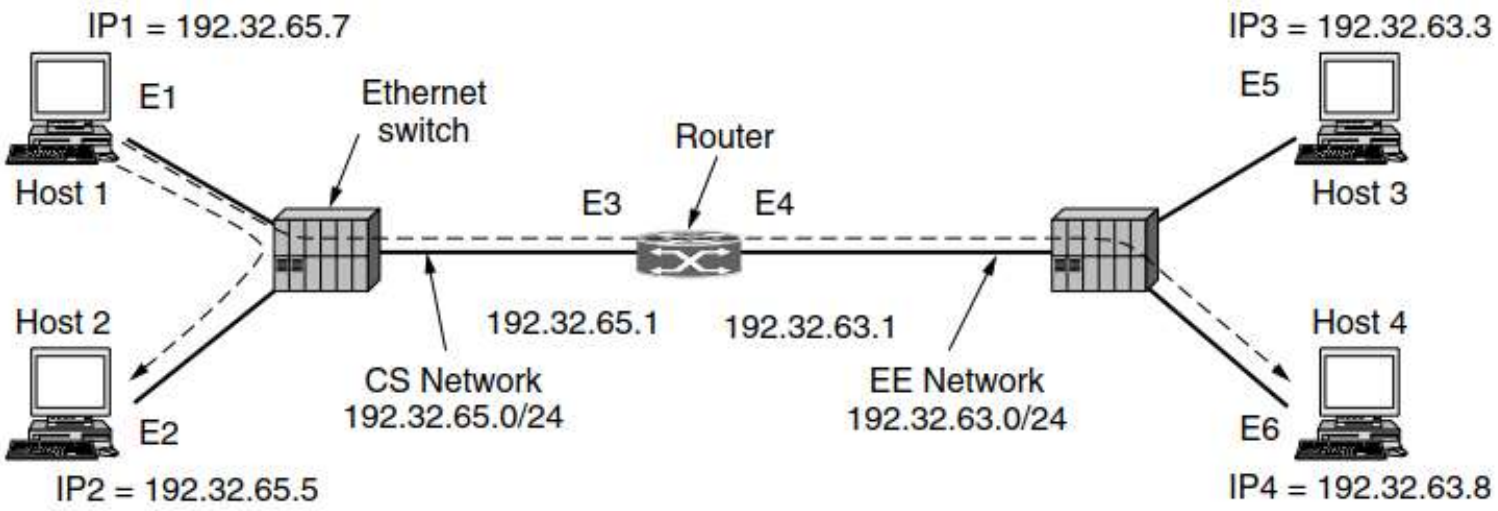
Internet Control Protocols (Contd.)

ARP—The Address Resolution Protocol

- Although every machine on the Internet has IP addresses, these addresses are not sufficient for sending packets
- Data link layer NICs (Network Interface Cards) such as Ethernet cards do not understand Internet addresses
- In the case of Ethernet, every NIC ever manufactured comes equipped with a unique 48-bit Ethernet address
- The question now arises, how do IP addresses get mapped onto data link layer addresses, such as Ethernet?
- To explain how this works, let us use the example of Fig. 5-61, in which a small university with two /24 networks is illustrated. One network (CS) is a switched Ethernet in the Computer Science Dept. It has the prefix 192.32.65.0/24. The other LAN (EE), also switched Ethernet, is in Electrical Engineering and has the prefix 192.32.63.0/24. The two LANs are connected by an IP router. Each machine on an Ethernet and each interface on the router has a unique Ethernet address, labeled E1 through E6, and a unique IP address on the CS or EE network.

Internet Control Protocols (Contd.)

ARP—The Address Resolution Protocol



Frame	Source IP	Source Eth.	Destination IP	Destination Eth.
Host 1 to 2, on CS net	IP1	E1	IP2	E2
Host 1 to 4, on CS net	IP1	E1	IP4	E3
Host 1 to 4, on EE net	IP1	E4	IP4	E6

Figure 5-61. Two switched Ethernet LANs joined by a router.

Internet Control Protocols (Contd.)

ARP—The Address Resolution Protocol

- Algorithm/ Steps

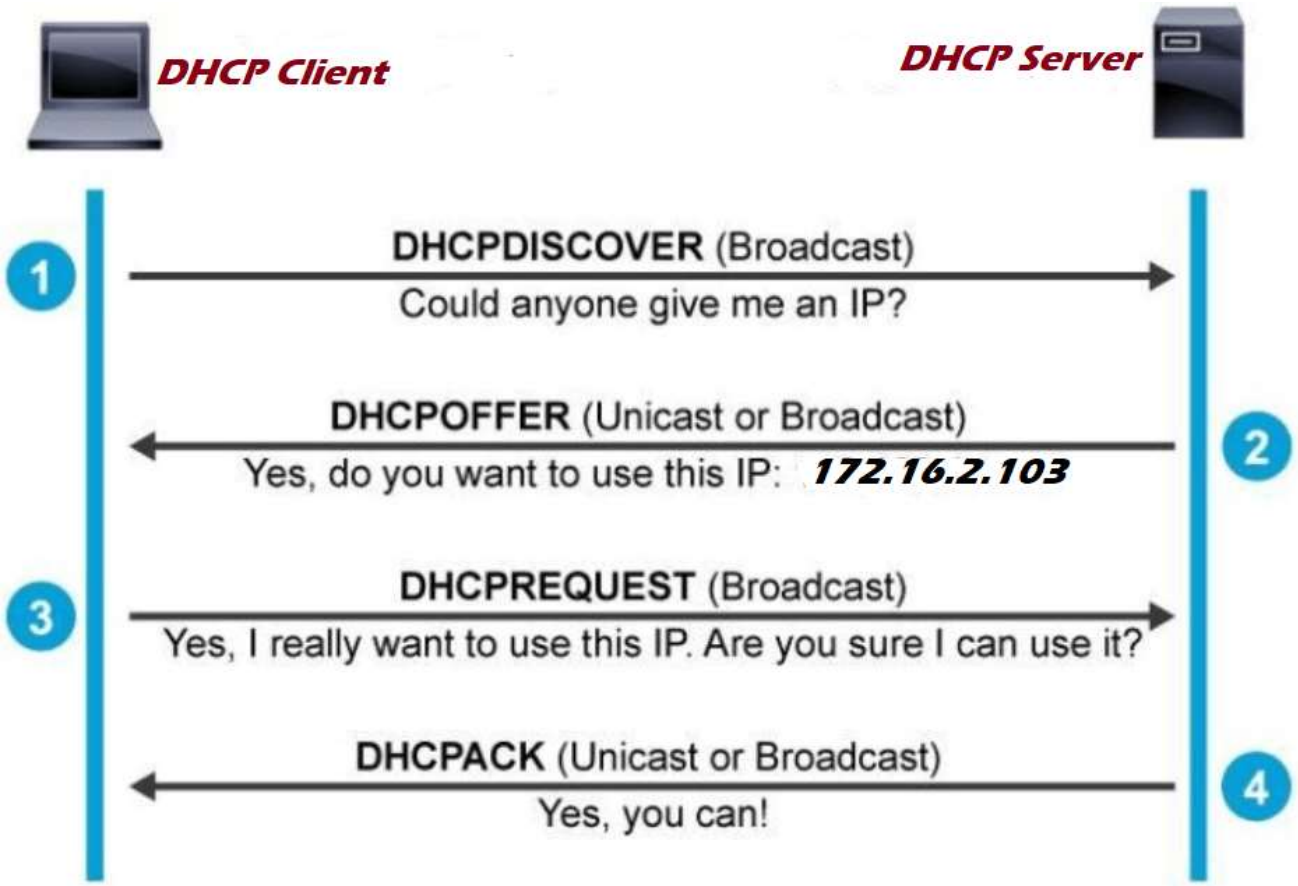
Internet Control Protocols (Contd.)

DHCP—The Dynamic Host Configuration Protocol

- DHCP (Dynamic Host Configuration Protocol) is a network management protocol used to dynamically assign an IP address to any device, or node, on a network so it can communicate using IP
- Every network must have a DHCP server that is responsible for configuration
- When a computer is started, it has a built-in link layer address embedded in the NIC, but no IP address
- Much like ARP, the computer broadcasts a request for an IP address on its network. It does this by using a DHCP DISCOVER packet
- This packet must reach the DHCP server. If that server is not directly attached to the network, the router will be configured to receive DHCP broadcasts and relay them to the DHCP server, wherever it is located
- When the server receives the request, it allocates a free IP address and sends it to the host in a DHCP OFFER packet (which again may be relayed via the router)
- To be able to do this work even when hosts do not have IP addresses, the server identifies a host using its Ethernet address (which is carried in the DHCP DISCOVER packet)

Internet Control Protocols (Contd.)

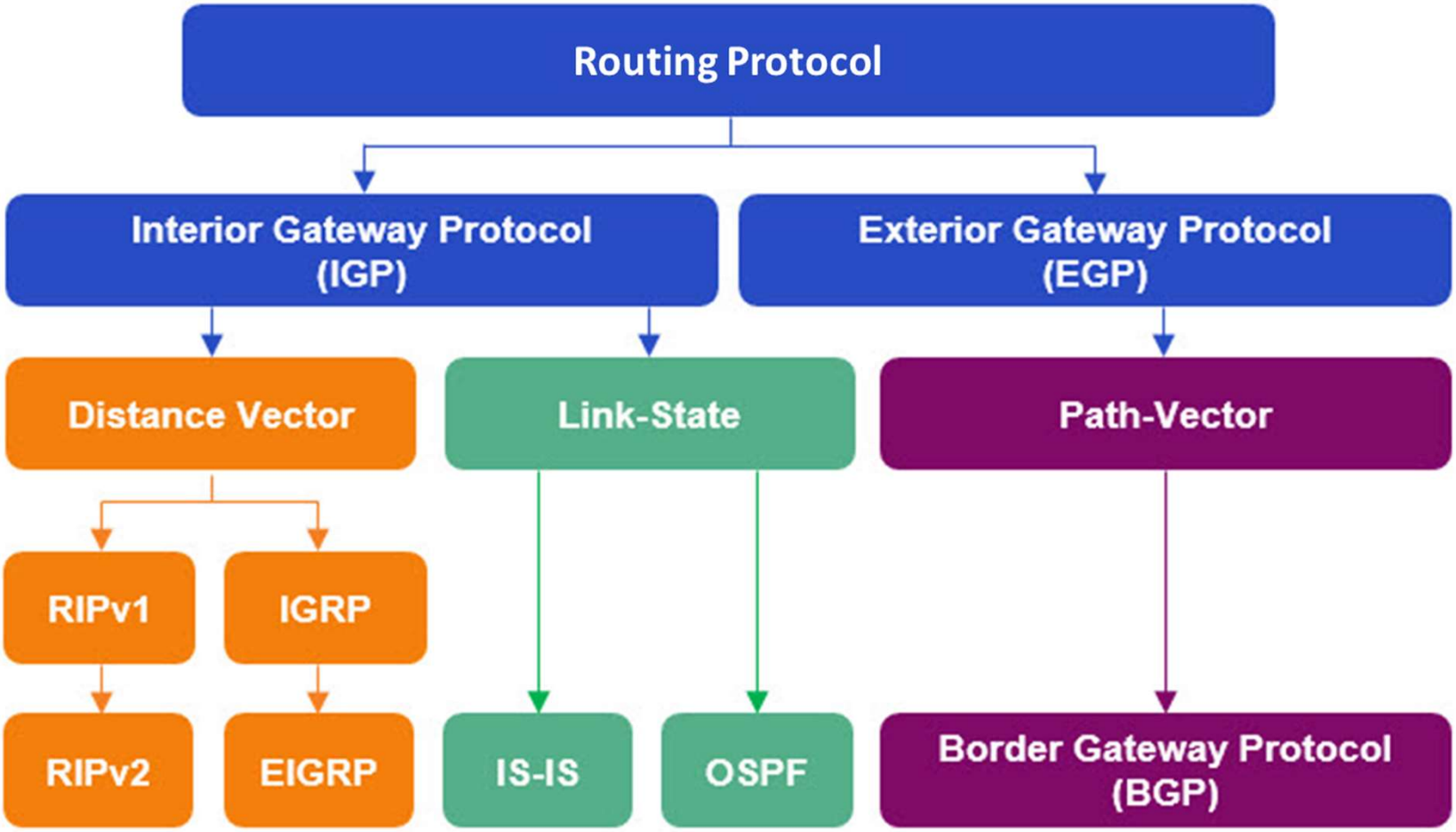
DHCP—The Dynamic Host Configuration Protocol



Internet Control Protocols (Contd.)

- It is time to move on to the next topic: routing in the Internet
- As we mentioned earlier, the Internet is made up of a large number of independent networks or ASes (Autonomous Systems) that are operated by different organizations, usually a company, university, or ISP
- Inside of its own network, an organization can use its own algorithm for internal routing, or intradomain routing, as it is more commonly known
- Early intradomain routing protocols used a distance vector design, based on the distributed Bellman-Ford algorithm inherited from the ARPANET
- RIP (Routing Information Protocol) is the main example that is used to this day. It works well in small systems, but less well as networks get larger. It also suffers from the count-to-infinity problem and generally slow convergence.

Internet Control Protocols (Contd.)



Internet Control Protocols (Contd.)

OSPF—An Interior Gateway Routing Protocol

- The ARPANET switched over to a link state protocol in May 1979 because of these problems, and in 1988 IETF began work on a link state protocol for intradomain routing
- That protocol, called OSPF (Open Shortest Path First), became a standard in 1990
- It drew on a protocol called IS-IS (Intermediate-System to Intermediate-System), which became an ISO standard.
- Given the long experience with other routing protocols, the group designing OSPF had a long list of requirements that had to be met
  1. The algorithm had to be published in the open literature, hence the "O" in OSPF
  2. The new protocol had to support a variety of distance metrics, including physical distance, delay, and so on
  3. It had to be a dynamic algorithm, one that adapted to changes in the topology automatically and quickly
  4. it had to support routing based on type of service.

Internet Control Protocols (Contd.)

OSPF—An Interior Gateway Routing Protocol

5. The new protocol had to be able to route real-time traffic one way and other traffic a different way. At the time, IP had a Type of service field, but no existing routing protocol used it. This field was included in OSPF but still nobody used it, and it was eventually removed. Perhaps this requirement was ahead of its time, as it preceded IETF's work on differentiated services, which has rejuvenated classes of service
6. OSPF had to do load balancing, splitting the load over multiple lines. Most previous protocols sent all packets over a single best route, even if there were two routes that were equally good. The other route was not used at all. In many cases, splitting the load over multiple routes gives better performance
7. Some degree of security was required to prevent fun-loving students from spoofing routers by sending them false routing information
8. Finally, provision was needed for dealing with routers that were connected to the Internet via a tunnel. Previous protocols did not handle this well.

Internet Control Protocols (Contd.)

OSPF—An Interior Gateway Routing Protocol

- OSPF supports both point-to-point links and broadcast networks (e.g., most LANs)
- Actually, it is able to support networks with multiple routers, each of which can communicate directly with the others (called multiaccess networks) even if they do not have broadcast capability. Earlier protocols did not handle this case well

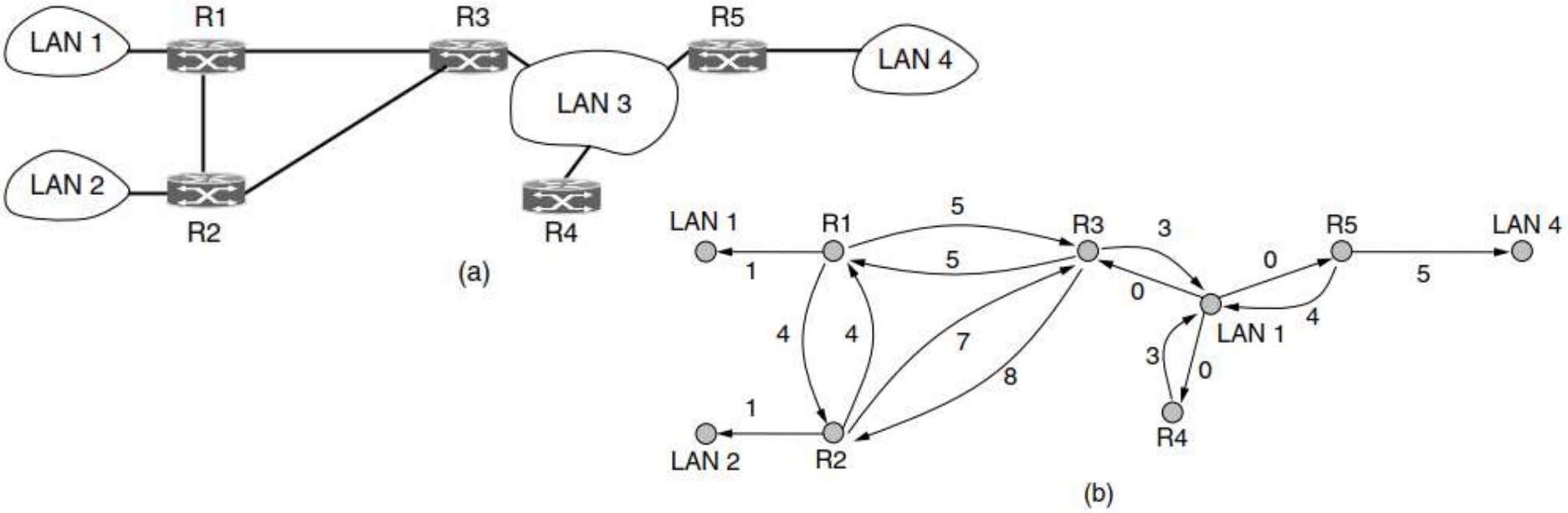


Figure 5-64. (a) An autonomous system. (b) A graph representation of (a).

Internet Control Protocols (Contd.)

OSPF—An Interior Gateway Routing Protocol

- An example of an autonomous system network is given in Fig. 5-64(a). Hosts are omitted because they do not generally play a role in OSPF, while routers and networks (which may contain hosts) do
- Most of the routers in Fig. 5-64(a) are connected to other routers by point-to-point links, and to networks to reach the hosts on those networks
- However, routers R3, R4, and R5 are connected by a broadcast LAN such as switched Ethernet

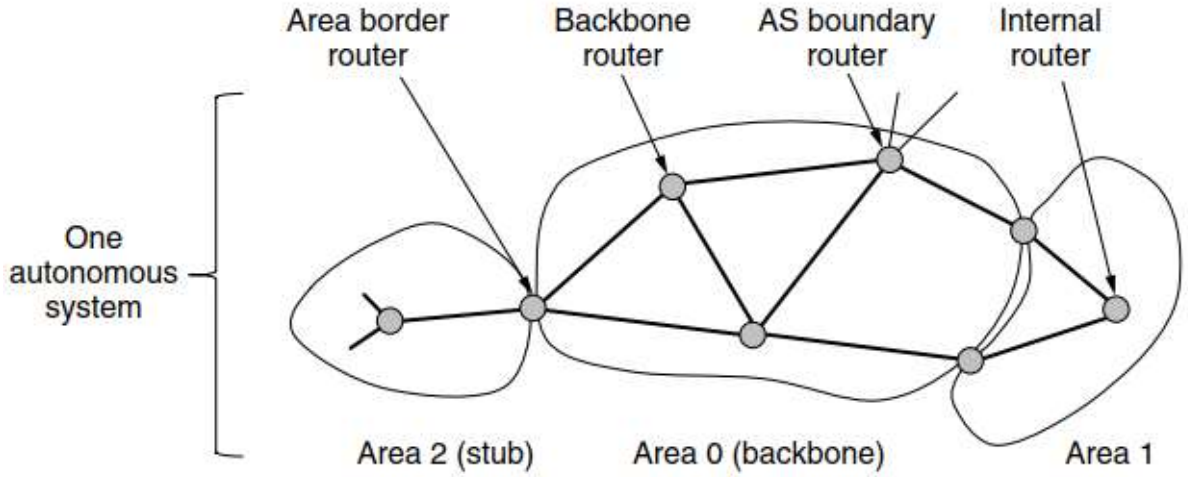


Figure 5-65. The relation between ASes, backbones, and areas in OSPF.

Internet Control Protocols (Contd.)

OSPF—An Interior Gateway Routing Protocol

- **Working**
  - When a router boots, it sends HELLO messages on all of its point-to-point lines and multicasts them on LANs to the group consisting of all the other routers. From the responses, each router learns who its neighbors are. Routers on the same LAN are all neighbors.
  - OSPF works by exchanging information between adjacent routers, which is not the same as between neighboring routers. In particular, it is inefficient to have every router on a LAN talk to every other router on the LAN
  - To avoid this situation, one router is elected as the designated router. It is said to be adjacent to all the other routers on its LAN, and exchanges information with them
  - Neighboring routers that are not adjacent do not exchange information with each other. A backup designated router is always kept up to date to ease the transition should the primary designated router crash and need to be replaced immediately
  - During normal operation, each router periodically floods LINK STATE UPDATE messages to each of its adjacent routers. These messages gives its state and provide the costs used in the topological database
  - The flooding messages are acknowledged, to make them reliable. Each message has a sequence number, so a router can see whether an incoming LINK STATE UPDATE is older or newer than what it currently has. Routers also send these messages when a link goes up or down or its cost changes

Internet Control Protocols (Contd.)

OSPF—An Interior Gateway Routing Protocol

- **Working**
  - DATABASE DESCRIPTION messages give the sequence numbers of all the link state entries currently held by the sender. By comparing its own values with those of the sender, the receiver can determine who has the most recent values. These messages are used when a link is brought up.
  - Either partner can request link state information from the other one by using LINK STATE REQUEST messages. The result of this algorithm is that each pair of adjacent routers checks to see who has the most recent data, and new information is spread throughout the area this way
  - All these messages are sent directly in IP packets. The five kinds of messages are summarized in Fig. 5-66.

Message type	Description
Hello	Used to discover who the neighbors are
Link state update	Provides the sender's costs to its neighbors
Link state ack	Acknowledges link state update
Database description	Announces which updates the sender has
Link state request	Requests information from the partner

Figure 5-66. The five types of OSPF messages.

Internet Control Protocols (Contd.)

BGP—The Exterior Gateway Routing Protocol

- Within a single AS, OSPF and IS-IS are the protocols that are commonly used. Between ASes, a different protocol, called BGP (Border Gateway Protocol), is used
- Aim of intradomain protocol is move packet as efficiently as possible from the source to the destination. It does not have to worry about politics
- In contrast, interdomain routing protocols have to worry about politics a great deal (Metz, 2001). For example, a corporate AS might want the ability to send packets to any Internet site and receive packets from any Internet site
- However, it might be unwilling to carry transit packets originating in a foreign AS and ending in a different foreign AS, even if its own AS is on the shortest path between the two foreign ASes (“That’s their problem, not ours”)
- On the other hand, it might be willing to carry transit traffic for its neighbors, or even for specific other ASes that paid it for this service
- Exterior gateway protocols in general, and BGP in particular, have been designed to allow many kinds of routing policies to be enforced in the interAS traffic

Internet Control Protocols (Contd.)

BGP—The Exterior Gateway Routing Protocol

- Typical policies involve political, security, or economic considerations. A few examples of possible routing constraints are:
  1. Do not carry commercial traffic on the educational network
  2. Never send traffic from the Pentagon on a route through Iraq
  3. Use TeliaSonera instead of Verizon because it is cheaper
  4. Don't use AT&T in Australia because performance is poor
  5. Traffic starting or ending at Apple should not transit Google

Fig a. Stub Connection

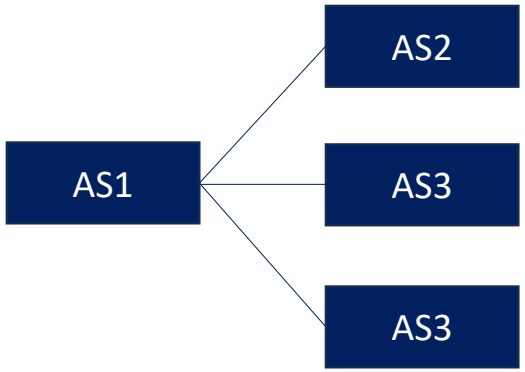


Fig b. Multi Homed AS

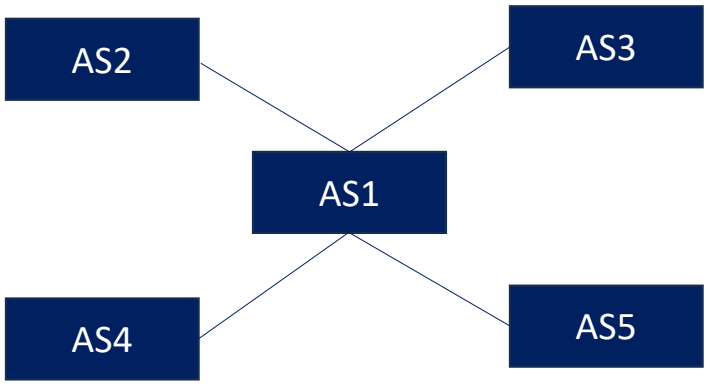


Fig c. Transit Service

Internet Control Protocols (Contd.)

BGP—The Exterior Gateway Routing Protocol

- We can see an example of transit service in Fig. 5-67.
- There are four Ases that are connected. The connection is often made with a link at IXPs (Internet eXchange Points), facilities to which many ISPs have a link for the purpose of connecting with other ISPs

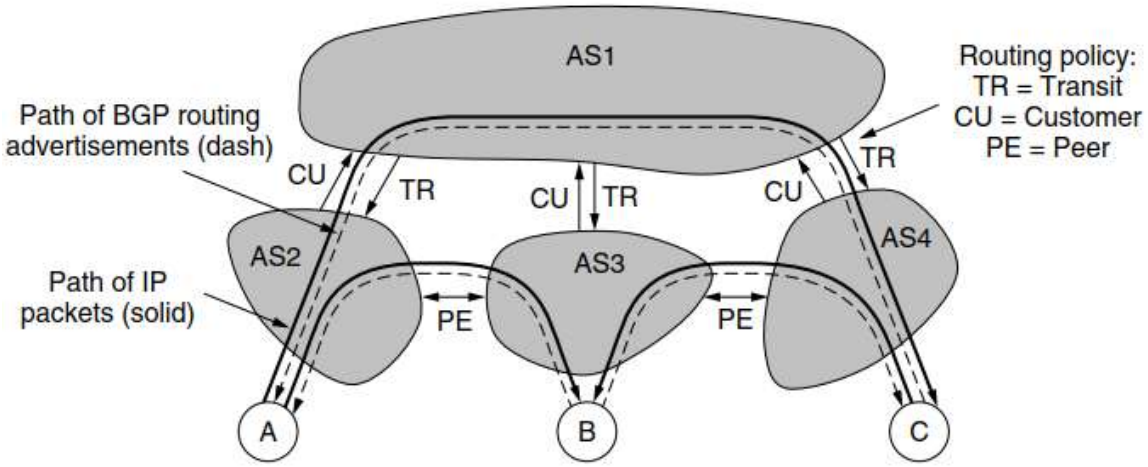
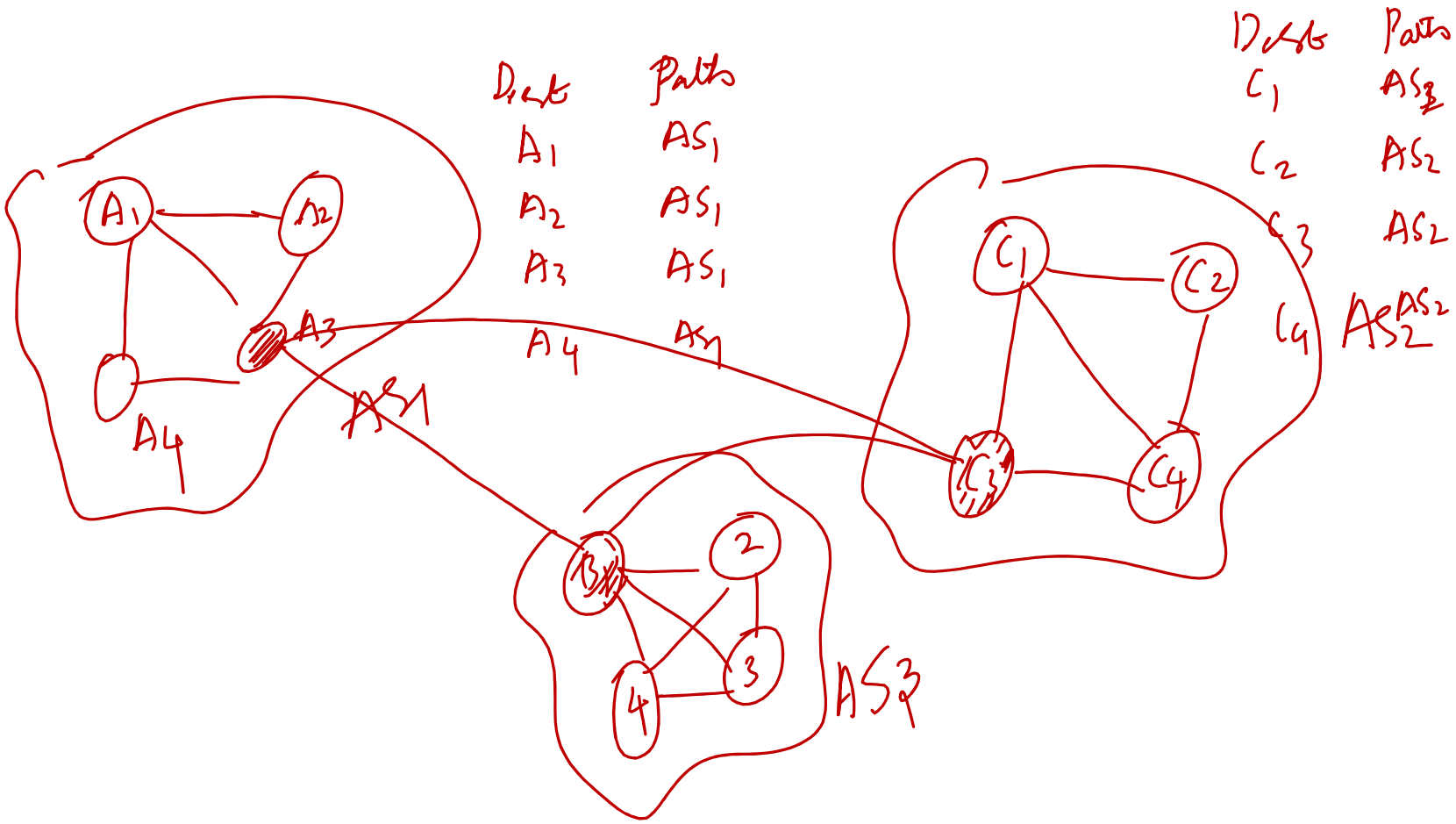


Figure 5-67. Routing policies between four autonomous systems.

- AS2, AS3, and AS4 are customers of AS1. They buy transit service from it.
- Thus, when source A sends to destination C, the packets travel from AS2 to AS1 and finally to AS4
- The routing advertisements travel in the opposite direction to the packets
- The AS4 advertises C as a destination to its transit provider, AS1, to let sources reach C via AS1
- Later, AS1 advertises a route to C to its other customers, including AS2, to let the customers know that they can send traffic to C via AS1


A1
A2
A3
A4
AS1
B1
2
3
4
AS3
C1
C2
C3
C4
AS2
Dest
Path
A1 AS1
A2 AS1
A3 AS1
A4 AS1
C1 AS2
C2 AS2
C3 AS2
C4 AS2 AS2

Internet Control Protocols (Contd.)

BGP—The Exterior Gateway Routing Protocol

- An example of how BGP routes are advertised is shown in Fig. 5-68.
- There are three ASes and the middle one is providing transit to the left and right ISPs

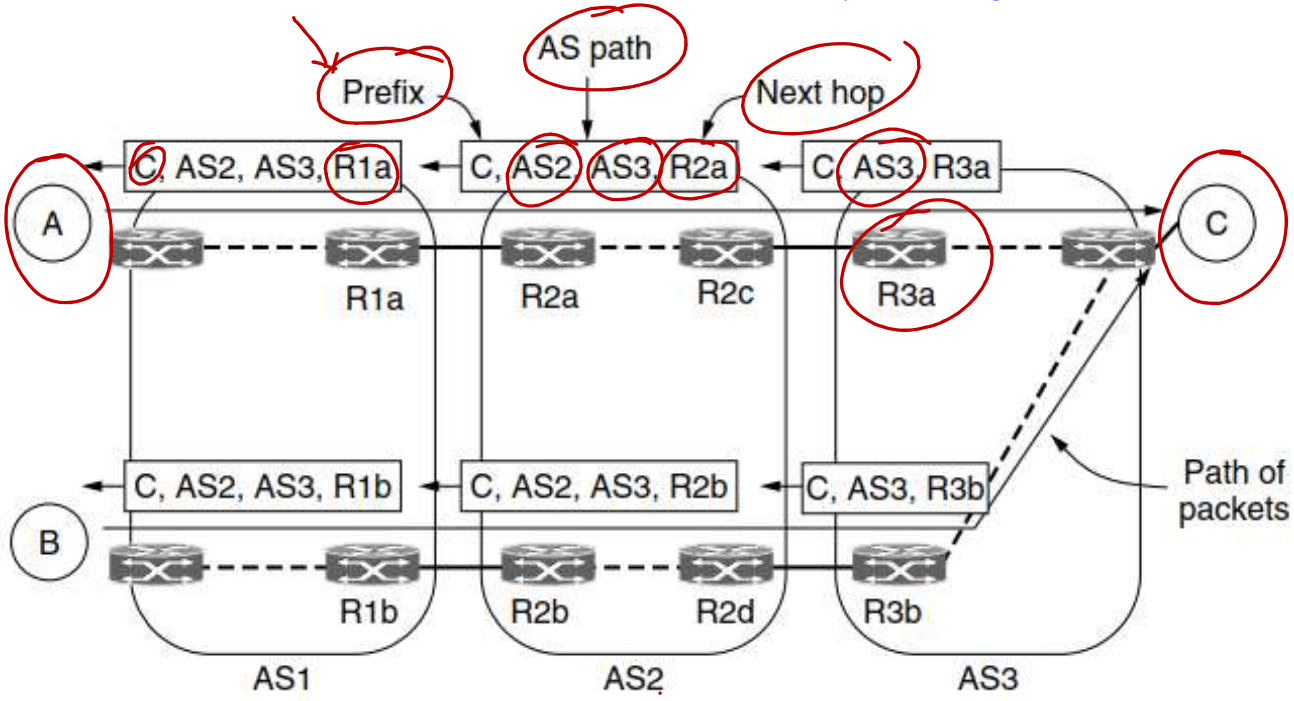


Figure 5-68. Propagation of BGP route advertisements.

- A route advertisement to prefix C starts in AS3.
- When it is propagated across the link to R2c at the top of the figure, it has the AS path of simply AS3 and the next hop router of R3a. At the bottom, it has the same AS path but a different next hop because it came across a different link. This advertisement continues to propagate and crosses the boundary into AS1. At router R1a, at the top of the figure, the AS path is AS2, AS3 and the next hop is R2a.

Thank You