Chapter 1

# Basics of Networking

**Learning Outcomes**

After reading this chapter, the reader will be able to:

- Understand the basic principles of computer networking
- List the basic terminologies and technologies
- Relate new concepts of IoT with the basics of networking
- Discuss various network configurations and topologies
- Explain various OSI (open systems interconnections) and TCP/IP (transmission control protocol/Internet protocol) layers and their associated uses
- Describe basics of network addressing

## 1.1 Introduction

In the present era of data- and information-centric operations, everything—right from agriculture to military operations—relies heavily on information. The quality of any particular information is as good as the variety and strength of the data that generates this information. Additionally, the speed at which data is updated to all members of a team (which may be a group of individuals, an organization, or a country) dictates the advantage that the team has over others in generating useful information from the gathered data. Considering the present-day global scale of operations of various organizations or militaries of various countries, the speed and nature of germane information are crucial for maintaining an edge over others in the same area. To sum it up, today's world relies heavily on data and networking, which allows for the instant availability of information from anywhere on the earth at any moment.

Typically, networking refers to the linking of computers and communication network devices (also referred to as hosts), which interconnect through a network

(Internet or Intranet) and are separated by unique device identifiers (Internet protocol, IP addresses and media access control, MAC addresses). These hosts may be connected by a single path or through multiple paths for sending and receiving data. The data transferred between the hosts may be text, images, or videos, which are typically in the form of binary bit streams [1].

Points to ponder

- The data generated from a camera sensor tells us more about a scene compared to the data generated from, say, a proximity sensor, which only detects the presence of people in its sensing range.
- Furthermore, the simultaneous data generated from multiple cameras focusing on the same spot from various angles tell us even more about the scene than a single camera focused at that scene.

As the primary aim of this chapter is to provide the reader with an overview of networking, we have structured the text in such a manner that the general concepts are covered. Additional *Check yourself* suggestions to review various associated technologies are provided along with the topics.

We start our discussion with the different types of networks, followed by an overview of two popularly used layered network models: ISO-OSI (the open systems interconnection developed by the International Organization of Standardization) and TCP/IP (transmission control protocol/Internet protocol) suite. Subsequently, we will touch upon the various types of addressing mechanisms and set up the basic premise of how a message is transmitted between two devices/computers/hosts.

## 1.2 Network Types

Computer networks are classified according to various parameters: 1) Type of connection, 2) physical topology, and 3) reach of the network. These classifications are helpful in deciding the requirements of a network setup and provide insights into the appropriate selection of a network type for the setup.

### 1.2.1 Connection types

Depending on the way a host communicates with other hosts, computer networks are of two types—(Figure 1.1): *Point-to-point* and *Point-to-multipoint*.

(i) **Point-to-point**: Point-to-point connections are used to establish direct connections between two hosts. Day-to-day systems such as a remote control for an air conditioner or television is a point to point connection, where the connection has the whole channel dedicated to it only. These networks were

designed to work over duplex links and are functional for both synchronous as well as asynchronous systems. Regarding computer networks, point to point connections find usage for specific purposes such as in optical networks.

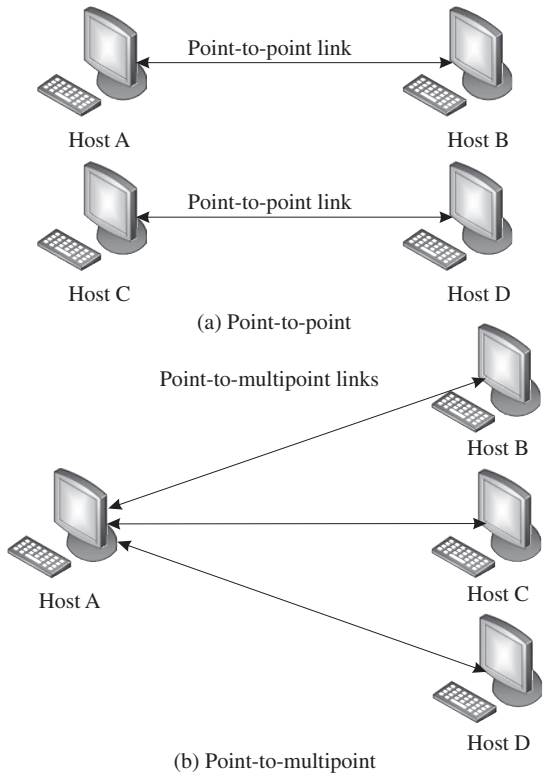



**Figure 1.1** Network types based on connection types

**Point-to-point Requests for Comments (RFCs)**

The following requests for comments (RFCs) are associated with point-to-point communication and its derivatives. **RFC 1332**: point-to-point (PPP) Internet protocol control protocol (IPCP); **RFC 1661**: PPP; **RFC 5072**: IP Version 6 over PPP; **RFC 2516**: PPP over Ethernet; **RFC 1963**: PPP serial data transport protocol; **RFC 1962**: PPP compression control protocol (CCP); **RFC 1990**: PPP multilink protocol (MP); **RFC 2615**: PPP over SONET/SDH (synchronous optical networking/synchronous digital hierarchy).

(ii) **Point-to-multipoint**: In a point-to-multipoint connection, more than two hosts share the same link. This type of configuration is similar to the one-to-many connection type. Point-to-multipoint connections find popular use in wireless networks and IP telephony. The channel is shared between the various hosts,

either spatially or temporally. One common scheme of spatial sharing of the channel is frequency division multiple access (FDMA). Temporal sharing of channels include approaches such as time division multiple access (TDMA). Each of the spectral and temporal sharing approaches has various schemes and protocols for channel sharing in point-to-multipoint networks. Point-to-multipoint connections find popular use in present-day networks, especially while enabling communication between a massive number of connected devices.

Check yourself

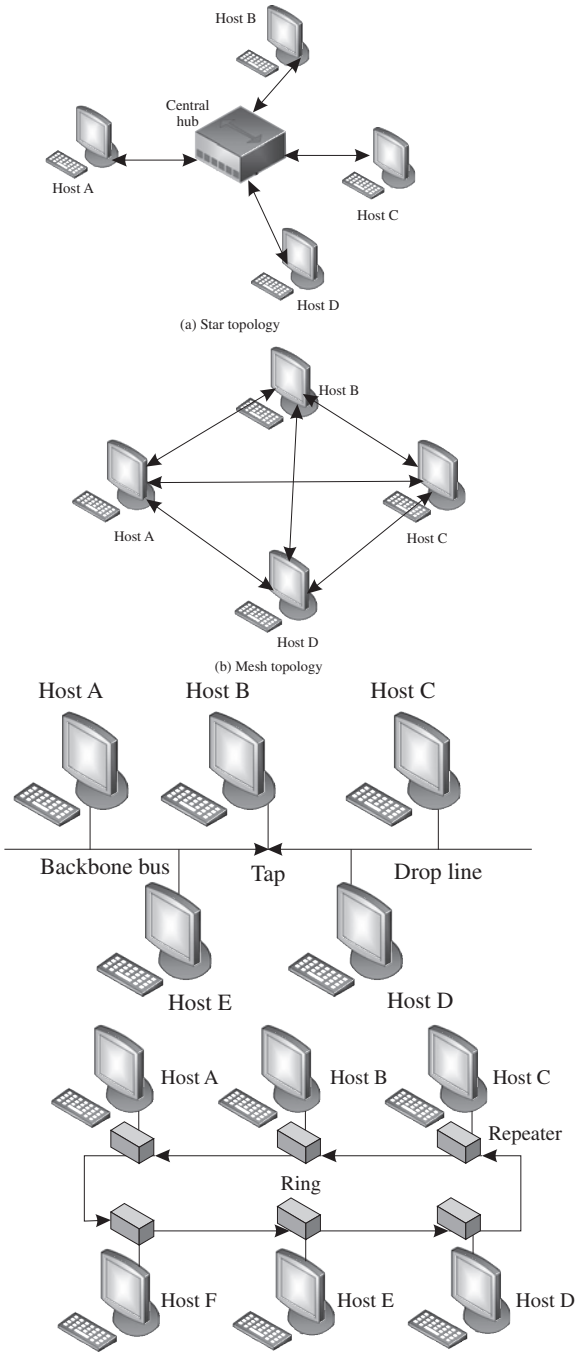
Space division multiplexing, Frequency division multiplexing, Time division multiplexing, Polarization division multiplexing, Orbital angular momentum multiplexing, Code division multiplexing

### 1.2.2 Physical topology

Depending on the physical manner in which communication paths between the hosts are connected, computer networks can have the following four broad topologies—(Figure 1.2): *Star*, *Mesh*, *Bus*, and *Ring*.

(i) **Star**: In a star topology, every host has a point-to-point link to a central controller or hub. The hosts cannot communicate with one another directly; they can only do so through the central hub. The hub acts as the network traffic exchange. For large-scale systems, the hub, essentially, has to be a powerful server to handle all the simultaneous traffic flowing through it. However, as there are fewer links (only one link per host), this topology is cheaper and easier to set up. The main advantages of the star topology are easy installation and the ease of fault identification within the network. If the central hub remains uncompromised, link failures between a host and the hub do not have a big effect on the network, except for the host that is affected. However, the main disadvantage of this topology is the danger of a single point of failure. If the hub fails, the whole network fails.

(ii) **Mesh**: In a mesh topology, every host is connected to every other host using a dedicated link (in a point-to-point manner). This implies that for $n$ hosts in a mesh, there are a total of $n(n-1)/2$ dedicated full duplex links between the hosts. This massive number of links makes the mesh topology expensive. However, it offers certain specific advantages over other topologies. The first significant advantage is the robustness and resilience of the system. Even if a link is down or broken, the network is still fully functional as there remain other pathways for the traffic to flow through. The second advantage is the security and privacy of the traffic as the data is only seen by the intended recipients and not by all members of the network. The third advantage is the reduced data load on a




(c) Bus topology

**Figure 1.2** Network types based on physical topologies

single host, as every host in this network takes care of its traffic load. However, owing to the complexities in forming physical connections between devices and the cost of establishing these links, mesh networks are used very selectively, such as in backbone networks.

(iii) **Bus**: A bus topology follows the point-to-multipoint connection. A backbone cable or bus serves as the primary traffic pathway between the hosts. The hosts are connected to the main bus employing drop lines or taps. The main advantage of this topology is the ease of installation. However, there is a restriction on the length of the bus and the number of hosts that can be simultaneously connected to the bus due to signal loss over the extended bus. The bus topology has a simple cabling procedure in which a single bus (backbone cable) can be used for an organization. Multiple drop lines and taps can be used to connect various hosts to the bus, making installation very easy and cheap. However, the main drawback of this topology is the difficulty in fault localization within the network.

(iv) **Ring**: A ring topology works on the principle of a point-to-point connection. Here, each host is configured to have a dedicated point-to-point connection with its two immediate neighboring hosts on either side of it through repeaters at each host. The repetition of this system forms a ring. The repeaters at each host capture the incoming signal intended for other hosts, regenerates the bit stream, and passes it onto the next repeater. Fault identification and set up of the ring topology is quite simple and straightforward. However, the main disadvantage of this system is the high probability of a single point of failure. If even one repeater fails, the whole network goes down.

Table 1.1 compares the various network topologies.

**Table 1.1** Network topology comparison

| Topology | Feature | Advantage | Disadvantage |
|---|---|---|---|
| Star | Point-to-point | Cheap; ease of installation; ease of fault identification | Single point of failure; traffic visible to network entities |
| Mesh | Point-to-point | Resilient against single point of failures; scalable; traffic privacy and security ensured | Costly; complex connections |
| Bus | Point-to-multipoint | Ease of installation; cheap | Length of backbone cable limited; number of hosts limited; hard to localize faults |
| Ring | Point-to-point | Ease of installation; cheap; ease of fault identification | Prone to single point of failure |

### 1.2.3 Network reachability

Computer networks are divided into four broad categories based on network reachability: *personal area networks*, *local area networks*, *wide area networks*, and *metropolitan area networks*.

(i) **Personal Area Networks (PAN)**: PANs, as the name suggests, are mostly restricted to individual usage. A good example of PANs may be connected wireless headphones, wireless speakers, laptops, smartphones, wireless keyboards, wireless mouse, and printers within a house. Generally, PANs are wireless networks, which make use of low-range and low-power technologies such as Bluetooth. The reachability of PANs lies in the range of a few centimeters to a few meters.

(ii) **Local Area Networks (LAN)**: A LAN is a collection of hosts linked to a single network through wired or wireless connections. However, LANs are restricted to buildings, organizations, or campuses. Typically, a few leased lines connected to the Internet provide web access to the whole organization or a campus; the lines are further redistributed to multiple hosts within the LAN enabling hosts. The hosts are much more in number than the actual direct lines to the Internet to access the web from within the organization. This also allows the organization to define various access control policies for web access within its hierarchy. Typically, the present-day data access rates within the LANs range from 100 Mbps to 1000 Mbps, with very high fault-tolerance levels. Commonly used network components in a LAN are servers, hubs, routers, switches, terminals, and computers.

(iii) **Metropolitan Area Networks (MAN)**: The reachability of a MAN lies between that of a LAN and a WAN. Typically, MANs connect various organizations or buildings within a given geographic location or city. An excellent example of a MAN is an Internet service provider (ISP) supplying Internet connectivity to various organizations within a city. As MANs are costly, they may not be owned by individuals or even single organizations. Typical networking devices/components in MANs are modems and cables. MANs tend to have moderate fault tolerance levels.

(iv) **Wide Area Networks (WAN)**: WANs typically connect diverse geographic locations. However, they are restricted within the boundaries of a state or country. The data rate of WANs is in the order of a fraction of LAN's data rate. Typically, WANs connecting two LANs or MANs may use public switched telephone networks (PSTNs) or satellite-based links. Due to the long transmission ranges, WANs tend to have more errors and noise during transmission and are very costly to maintain. The fault tolerance of WANs are also generally low.

Check yourself

ARPANET, BITNET, Cellular network, CYCLADES, FidoNet, Telex, World Wide Web

## 1.3 Layered Network Models

The intercommunication between hosts in any computer network, be it a large-scale or a small-scale one, is built upon the premise of various task-specific layers. Two of the most commonly accepted and used traditional layered network models are the open systems interconnection developed by the International Organization of Standardization (ISO-OSI) reference model and the Internet protocol suite.

### 1.3.1 OSI Model

The ISO-OSI model is a conceptual framework that partitions any networked communication device into seven layers of abstraction, each performing distinct tasks based on the underlying technology and internal structure of the hosts. These seven layers, from bottom-up, are as follows: 1) *Physical layer*, 2) *Data link layer*, 3) *Network layer*, 4) *Transport layer*, 5) *Session layer*, 6) *Presentation layer*, and 7) *Application layer*. The major highlights of each of these layers are explained in this section.

Points to ponder

The OSI or open system interconnect model for networked devices was standardized by the International Standards Organization (ISO). It is a conceptual framework that divides any networked communication system into seven layers, each performing specific tasks toward communicating with other systems [5], [1]. The OSI is a reference model and is maintained by the ISO under the identity of ISO/IEC 7498-1.

(i) **Physical Layer**: This is a media layer and is also referred to as layer 1 of the OSI model. The physical layer is responsible for taking care of the electrical and mechanical operations of the host at the actual physical level. These operations include or deal with issues relating to signal generation, signal transfer, voltages, the layout of cables, physical port layout, line impedances, and signal loss. This layer is responsible for the topological layout of the network (star, mesh, bus, or ring), communication mode (simplex, duplex, full duplex), and bit rate control operations. The protocol data unit associated with this layer is referred to as a *symbol*.

(ii) **Data Link Layer**: This is a media layer and layer 2 of the OSI model. The data link layer is mainly concerned with the establishment and termination of the connection between two hosts, and the detection and correction of errors during communication between two or more connected hosts. IEEE 802 divides the OSI layer 2 further into two sub-layers [2]: Medium access control (MAC) and logical link control (LLC). MAC is responsible for access control and permissions for connecting networked devices; whereas LLC is mainly tasked with error checking, flow control, and frame synchronization. The protocol data unit associated with this layer is referred to as a *frame*.

(iii) **Network Layer**: This layer is a media layer and layer 3 of the OSI model. It provides a means of routing data to various hosts connected to different networks through logical paths called virtual circuits. These logical paths may pass through other intermediate hosts (nodes) before reaching the actual destination host. The primary tasks of this layer include addressing, sequencing of packets, congestion control, error handling, and Internetworking. The protocol data unit associated with this layer is referred to as a *packet*.

(iv) **Transport Layer**: This is layer 4 of the OSI model and is a host layer. The transport layer is tasked with end-to-end error recovery and flow control to achieve a transparent transfer of data between hosts. This layer is responsible for keeping track of acknowledgments during variable-length data transfer between hosts. In case of loss of data, or when no acknowledgment is received, the transport layer ensures that the particular erroneous data segment is re-sent to the receiving host. The protocol data unit associated with this layer is referred to as a *segment* or *datagram*.

(v) **Session Layer**: This is the OSI model's layer 5 and is a host layer. It is responsible for establishing, controlling, and terminating of communication between networked hosts. The session layer sees full utilization during operations such as remote procedure calls and remote sessions. The protocol data unit associated with this layer is referred to as *data*.

(vi) **Presentation Layer**: This layer is a host layer and layer 6 of the OSI model. It is mainly responsible for data format conversions and encryption tasks such that the syntactic compatibility of the data is maintained across the network, for which it is also referred to as the *syntax layer*. The protocol data unit associated with this layer is referred to as *data*.

(vii) **Application Layer**: This is layer 6 of the OSI model and is a host layer. It is directly accessible by an end-user through software APIs (application program interfaces) and terminals. Applications such as file transfers, FTP (file transfer protocol), e-mails, and other such operations are initiated from this layer. The application layer deals with user authentication, identification of communication hosts, quality of service, and privacy. The protocol data unit associated with this layer is referred to as *data*.

A networked communication between two hosts following the OSI model is shown in Figure 1.3. Table 1.2 summarizes the OSI layers and their features, where PDU stands for protocol data unit.

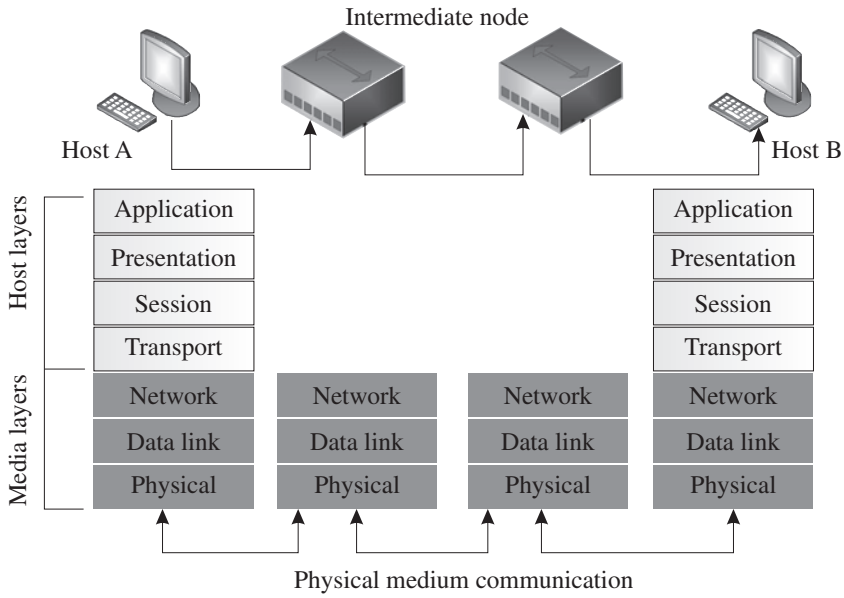



**Figure 1.3** Networked communication between two hosts following the OSI model

**Check yourself**

Ethernet, FDDI, B8ZS, V.35, V.24, RJ45, PPP, FDDI, ATM, IEEE 802.5/ 802.2, IEEE 802.3/802.2, HDLC, Frame Relay, AppleTalk DDP, IP, IPX, NFS, NetBios names, RPC, SQL, ASCII, EBCDIC, TIFF, GIF, PICT, JPEG, MPEG, MIDI, NFS, SNMP, Telnet, HTTP, FTP

**Table 1.2** Summary of the OSI layers and their features

| Layer | Name | Location | PDU | Function | Examples |
|---|---|---|---|---|---|
| 1 | Physical | Media | Symbol | Communication over physical medium | Ethernet, FDDI, B8ZS, V.35, V.24, RJ45 |
| 2 | Data link | Media | Frame | Reliability of communication over physical medium | IEEE 802.5/ 802.2, IEEE 802.3/802.2, PPP, HDLC, Frame Relay, ATM, FDDI |
| 3 | Network | Media | Packet | Structuring of data and routing between multiple nodes | DDP, IP, AppleTalk, IPX |
| 4 | Transport | Host | Segment | Reliability of communication over networks or between hosts | SPX, TCP, UDP |
| 5 | Session | Host | Data | Establishment, management, and termination of remote sessions | NetBios names, NFS, RPC, SQL |
| 6 | Presentation | Host | Data | Syntactic conversion of data and encryption | Encryption, ASCII, MIDI, PICT, JPEG, EBCDIC, TIFF, GIF, MPEG |
| 7 | Application | Host | Data | User identification, authentication, privacy, and quality of service | SNMP, Telnet, WWW browsers, HTTP, NFS, FTP |

### 1.3.2 Internet protocol suite

The Internet protocol suite is yet another conceptual framework that provides levels of abstraction for ease of understanding and development of communication and networked systems on the Internet. However, the Internet protocol suite predates the OSI model and provides only four levels of abstraction: 1) Link layer, 2) Internet layer, 3) transport layer, and 4) application layer. This collection of protocols is commonly referred to as the TCP/IP protocol suite as the foundation technologies of this suite are transmission control protocol (TCP) and Internet protocol (IP) [3], [4], [6]. The TCP/IP protocol suite comprises the following four layers:

**Points to ponder**

The development of the TCP/IP protocol suite is originally attributed to DARPA, which is part of the United States Department of Defence. The Internet protocol suite or the TCP/IP protocol suite is sometimes also referred to as the Department of Defence (DoD) model.

(i) **Link Layer**: The first and base layer of the TCP/IP protocol suite is also known as the network interface layer. This layer is synonymous with the collective physical and data link layer of the OSI model. It enables the transmission of TCP/IP packets over the physical medium. According to its design principles, the link layer is independent of the medium in use, frame format, and network access, enabling it to be used with a wide range of technologies such as the Ethernet, wireless LAN, and the asynchronous transfer mode (ATM).

(ii) **Internet Layer**: Layer 2 of the TCP/IP protocol suite is somewhat synonymous to the network layer of the OSI model. It is responsible for addressing, address translation, data packaging, data disassembly and assembly, routing, and packet delivery tracking operations. Some core protocols associated with this layer are address resolution protocol (ARP), Internet protocol (IP), Internet control message protocol (ICMP), and Internet group management protocol (IGMP). Traditionally, this layer was built upon IPv4, which is gradually shifting to IPv6, enabling the accommodation of a much more significant number of addresses and security measures.

(iii) **Transport Layer**: Layer 3 of the TCP/IP protocol suite is functionally synonymous with the transport layer of the OSI model. This layer is tasked with the functions of error control, flow control, congestion control, segmentation, and addressing in an end-to-end manner; it is also independent of the underlying network. Transmission control protocol (TCP) and user datagram protocol (UDP) are the core protocols upon which this layer is built, which in turn enables it to have the choice of providing connection-oriented or connectionless services between two or more hosts or networked devices.

(iv) **Application Layer**: The functionalities of the application layer, layer 4, of the TCP/IP protocol suite are synonymous with the collective functionalities of the OSI model's session, presentation, and application layers. This layer enables an end-user to access the services of the underlying layers and defines the protocols for the transfer of data. Hypertext transfer protocol (HTTP), file transfer protocol (FTP), simple mail transfer protocol (SMTP), domain name system (DNS), routing information protocol (RIP), and simple network management protocol (SNMP) are some of the core protocols associated with this layer.

A networked communication between two hosts following the TCP/IP model is shown in Figure 1.4

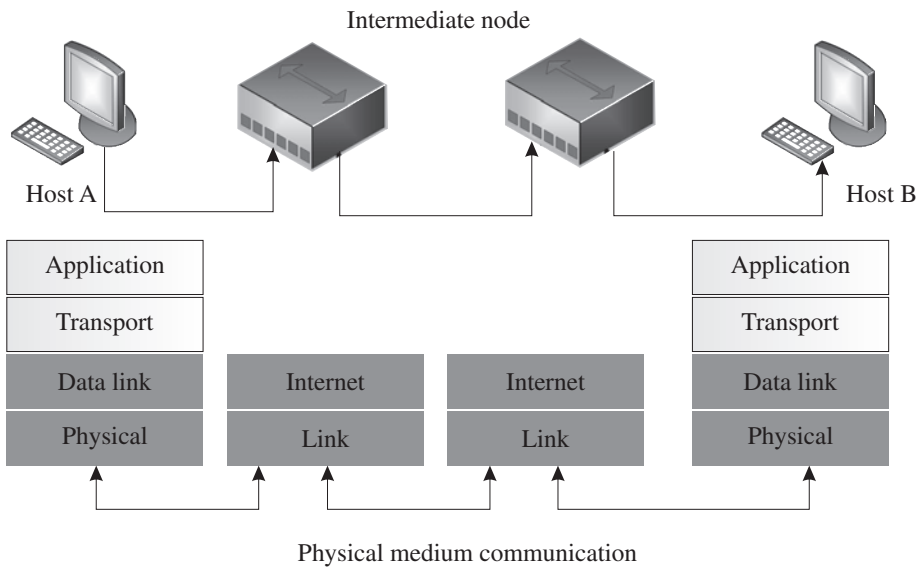



**Figure 1.4** Networked communication between two hosts following the TCP/IP suite

## 1.4 Addressing

Addressing in networked devices plays a crucial role in ensuring the delivery of packets to the designated/intended receivers. The addressing scheme is synonymous with postal addresses used in real-life scenarios. Addressing mechanisms can be divided into two parts: one focusing on data link layer addressing, while the other focuses on network layer addressing.

### 1.4.1 Data link layer addressing

Data link layer addressing deals with media access control (MAC) addresses of devices, which work at the MAC sub-layer of the data link layer.

Points to ponder

Data link layer addressing handles the host/device network interface of physical addresses. These physical addresses are also known as MAC addresses. MAC addresses are unique 48-bit hardware addresses provided by the device manufacturers.

MAC addresses are 48-bits long; the first 24 bits are organizational identifiers, while the last 24 bits are network interface controller identifiers. These addresses are unique globally. Data link layer addressing is broadly divided into three types: *1) Unicast, 2) Multicast, and 3) Broadcast.*

(i) **Unicast**: This addressing is meant for one-to-one communication. The data flow from a transmitting host is restricted to only one receiving host in the link.

(ii) **Multicast**: This addressing is meant for one-to-many communication within a single link. The data flow from a transmitting host is intended for multiple hosts within the same link. It is to be noted that more than one host can transmit data streams, which are designed for multiple receiving hosts in the link.

(iii) **Broadcast**: This addressing is meant for one-to-all communication within a link. The data from a transmitting host is received by all other hosts connected to that link.

### 1.4.2 Network layer addressing

Network layer addressing is also termed as IP-based addressing or logical addressing. IPv4 addressing uses 32-bits long addresses, whereas IPv6 uses addresses that are 128 bits long. These addresses can identify the source or destination addresses from the address itself. The mapping of a device/host's logical address to its hardware address is done through a mechanism called address resolution protocol (ARP). During transmission of a packet from a host, the IPv4 sends an IPv4 packet, the next-hop address, and the next-hop interface to the ARP.

Points to ponder

Network layer addressing deals with 32-bit (in case of IPv4) or 128-bit (in case of IPv6) logical addresses assigned to networked devices. These addresses are not hard-coded or provided by the manufacturers.

Direct delivery is performed by the ARP if the destination address for delivery matches the next-hop address. In contrast, if the addresses do not match, the ARP performs an indirect delivery by forwarding the packet to a router or an intermediate node. The resolution of the mapping of a packet's next-hop address to its MAC

address is made using broadcasting ARP requests. The returning ARP reply frame to the sender contains the MAC address corresponding to the packet's next-hop address.

In the context of addressing, we will discuss the structure of IPv4 and IPv6 packets, which will provide a much clearer understanding of the workings of these two protocols.

(i) **IPv4**: The IPv4 header packet shown in Figure 1.5 has 13 distinct fields, the functions of which are enumerated as follows.

- **VER**: It is 4 bits long and represents the version of IP. In the example given in Figure 1.5, it is 4 bits (binary: 0100).

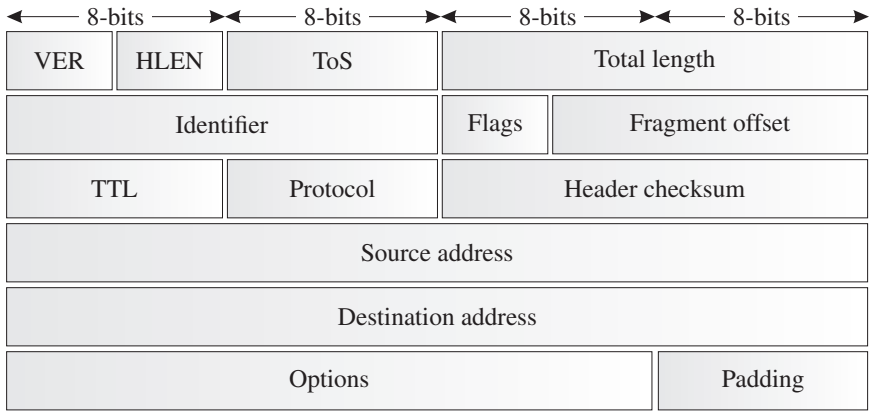



**Figure 1.5** An IPv4 packet header structure

- **HLEN**: It is 4 bits long and denotes the length of the IPv4 packet header.
- **ToS**: It is 8 bits long. The first six most significant bits represent the differentiated services code point (DSCP) to be provided to this packet (by the routers). Explicit congestion notification (ECN), which gives information about the congestion witnessed in the network, is handled by the last 2 bits.
- **TOTAL LENGTH**: It is 16 bits long and identifies the length of the entire IPv4 packet, including the header and the payload.
- **IDENTIFIER**: It is 16 bits long and used to identify the original packets in case of packet fragmentation along the network.
- **FLAGS**: It is a 3-bit field with the most significant bit always set to 0. FLAGS indicates whether a packet can be fragmented or not in case the packet is too big for the network resources.
- **FRAGMENT OFFSET**: It identifies the exact offset or fragment position of the original IP packet and is 13 bits long.
- **TTL**: It is 8 bits long and prevents a packet from looping infinitely in the network. As it completes a link, its value is decremented by one.

- **PROTOCOL**: It is 8 bits long. This field identifies the protocol of the packet as user datagram protocol, UDP (17), transmission control protocol, TCP (6), or Internet control message protocol, ICMP (1). The identification is made at the network layer of the destination host.
- **HEADER CHECKSUM**: It is 16 bits long and used for identifying whether a packet is error-free or not.
- **SOURCE ADDRESS**: It indicates the origin address of the packet and is 32 bits long.
- **DESTINATION ADDRESS**: It indicates the destination address of the packet and is 32 bits long.
- **OPTIONS and PADDING**: It is an optional field, which may carry values for security, time stamps, route records, and others.

(ii) **IPv6**: The IPv6 header packet shown in Figure 1.6 has eight distinct fields, the functions of which are enumerated as follows.

- **VER**: It is 4 bits long and represents the version of IP. In the example given in Figure 1.6, it is 6 (binary: 0110).

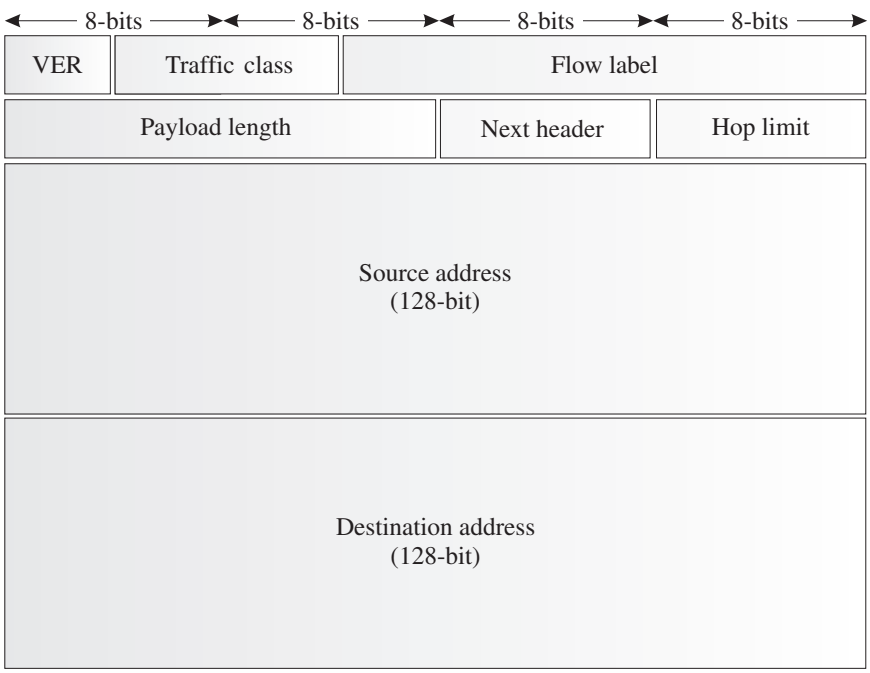



**Figure 1.6** An IPv6 packet header structure

- **TRAFFIC CLASS**: It is 8 bits long. The first six most significant bits represent the type of service to be provided to this packet (by the routers); explicit congestion notification (ECN) is handled by the last 2 bits.

- **FLOW LABEL**: It is 20 bits long and designed for streaming media or real-time data. The FLOW LABEL allows for information flow ordering; it also avoids packet resequencing.
- **PAYLOAD LENGTH**: It is 16 bits long and provides a router with information about a packet's payload length or the amount of data contained in the packet's payload.
- **NEXT HEADER**: It is 8 bits long and informs the router about the type of extension header the packet is carrying. Some of the extension headers and their corresponding values are as follows: Hop-by-hop options header (0), routing header (43), fragment header (44), destination options header (60), authentication header (51), and encapsulating security payload header (50). In case an extension header is absent, it represents the upper layer protocol data units (PDUs).
- **HOP LIMIT**: It is 8 bits long and prevents a packet from looping infinitely in the network. As it completes a link, the limit's value is decremented by one.
- **SOURCE ADDRESS**: It is 128 bits long and indicates the origin address of the packet.
- **DESTINATION ADDRESS**: It is 128 bits long and indicates the destination address of the packet.

Check yourself

Classful addressing, Classless addressing, CIDR notation, Subnetting, NAT, DHCP, RARP

## 1.5 TCP/IP Transport layer

The transport layer is the third layer in the TCP/IP protocol suite and is an important connectivity entity as it acts as the interlocutor for the clients and the servers in a client–server paradigm. This layer forms the core of the TCP/IP protocol suite as it provides logical mechanisms for data exchanges between two or more points over the Internet. As mentioned in the previous sections, the transport layer engages in networking functionalities such as process-to-process communication, encapsulation and decapsulation of data, multiplexing and demultiplexing of virtual pathways, flow control, error control, and congestion control.

From a broader perspective, the transport layer provides two types of services: 1) connectionless and 2) connection-oriented. These service types at the transport layer determine the degree of interdependence between transmitted packets. The application layer for both service types first divides a message into smaller chunks,

which are then acceptable by the transport layer for further transmission. These chunks are sequentially forwarded from the application layer to the transport layer. Upon receiving these chunks, the transport layer encapsulates these into packets for transmission. Generally, the packets in a connectionless transport layer service are independent of one another; whereas the packets in a connection-oriented service are dependent on one another. Figure 1.7 shows these two service types side-by-side for the sake of comparison of their working.

### 1.5.1 Connectionless service

Owing to its working, the connectionless service at the transport layer treats each incoming chunk from the application layer as independent units. After these chunks have been packetized, the packets are transmitted over the network with basic information of the source and destination addresses and ports. Even if the packets at the receiving end arrive out of order, they are submitted to the receiving host's application layer as it is, without any sequence maintenance. Additionally, no dedicated connection is established between the client and the server processes, as shown in Figure 1.7(a). The client transport layer has four message chunks M0, M1, M2, and M3, which are transmitted in that sequence. However, M2 arrives at the server at a time much later than its subsequent packet M3. As this is a connectionless service, the sequence is not maintained, and the packets are forwarded to the server's application layer as it is (out of sequence). Voice-over-IP (VoIP) is a popular usage of this service type. The most famous protocol associated with this service type is the user datagram protocol (UDP).

**Points to ponder**

Datagrams are packet-switched network transfer units that are generally used for connectionless services. There is no guarantee of the time of arrival, sequence, and surety of delivery of datagrams.

### 1.5.2 Connection-oriented service

The connection-oriented service, in contrast to the connectionless service, has a high dependency on the sequence of packets. Before the transmission of data to the server from the client (refer to Figure 1.7(b)), the client and server establish a connection employing handshaking using SYN and ACK frames. Once the data transmission is complete, the connection is terminated. In case another message has to be transmitted, the connection establishment process is again followed. This service type ensures that the packets arriving at the client's transport layer from its application layer are delivered in the exact sequence as in the server's application layer, in turn ensuring the quality of service (QoS) for the connection. However, ensuring QoS makes this type of service quite slow in comparison to connectionless services. Illustrating its working

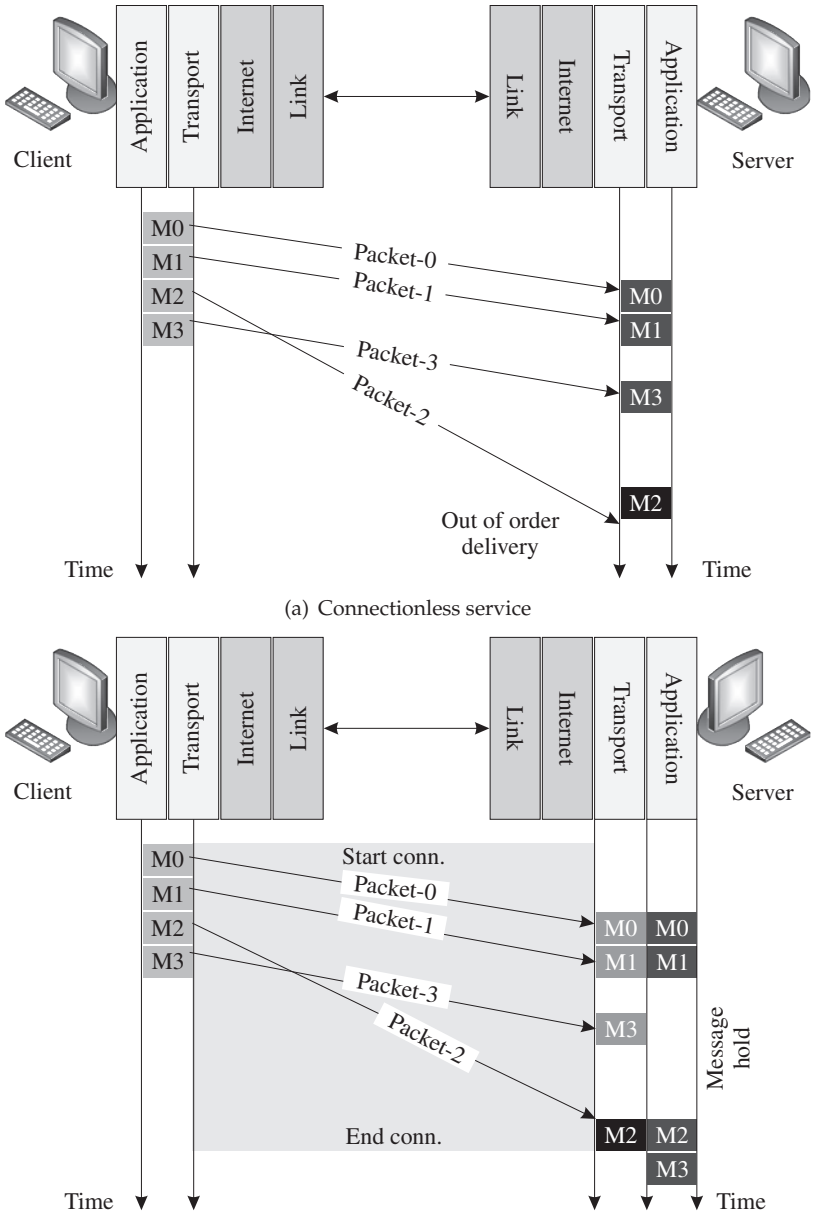



**Figure 1.7** Transport layer service types during client–server data transfer

using Figure 1.7(b), the client transport layer has four message chunks M0, M1, M2, and M3 (from the client's application layer), which are packetized and transmitted in that sequence once a connection is established between the client and the server. Even if the M2 packet arrives out of sequence at the server's transport layer, the subsequent

packets are held back until M2 is received. Upon receiving M2, M2 and the held back M3 packet are forwarded to the server's application layer in the same sequence that it was transmitted from the client's transport layer. Application layer protocols such as HTTP (hyper text transfer protocol) and HTTPS (hyper text transfer protocol secure) rely on connection-oriented services for their operation. The popular transport layer protocol, transmission control protocol (TCP), is a means of achieving connection-oriented service. The features of TCP and UDP are compared in Table 1.3.

**Table 1.3** Comparison of the features of TCP and UDP

| Feature | UDP | TCP |
|---|---|---|
| Name | User datagram protocol | Transmission control protocol |
| Type of service | Connectionless | Connection-oriented |
| Reliability | Low | High |
| Time-criticality | High | Low |
| Packet sequencing | No sequencing required | High level of sequencing involved |
| Speed of transfer | High | Relatively low |
| Error checking | Present, but it simply discards errorenous packets | Present; Errorenous packets are re-transmitted from the source |
| Error recovery | Absent | Present |
| Acknowledgment | Absent | Present; Done by means of ACK frames |
| Handshake | None | Done by SYN, SYN-ACK, ACK frames |
| Weight | Lightweight protocol | Heavyweight protocol |
| Usage | SNMP, TFTP, RIP, VoIP, DNS, DHCP | HTTP, HTTPs, FTP, SMTP, Telnet |

Check yourself

Client–server architecture, Connection-oriented service, Connection-less service

## Summary

This chapter covered the very basics of networking, which would prove handy in the following chapters covering the Internet of Things and its various associated paradigms. We discussed different network types based on connection types, topologies, and network reachability. We then outlined the two popular layered network models: the ISO-OSI model and the TCP/IP protocol suite. Subsequently,