Fortinet Networking Fundamentals

Lesson Scripts

1.0

FORTINET®
Training Institute

Fortinet Training Institute - Library

https://training.fortinet.com

Fortinet Product Documentation

https://docs.fortinet.com

Fortinet Knowledge Base

https://kb.fortinet.com

Fortinet Fuse User Community

https://fusecommunity.fortinet.com/home

Fortinet Forums

https://forum.fortinet.com

Fortinet Product Support

https://support.fortinet.com

FortiGuard Labs

https://www.fortiguard.com

Fortinet Training Program Information

https://www.fortinet.com/nse-training

Fortinet | Pearson VUE

https://home.pearsonvue.com/fortinet

Fortinet Training Institute Helpdesk (training questions, comments, feedback)

https://helpdesk.training.fortinet.com/support/home

9/14/2023

TABLE OF CONTENTS

LAN Basics Module

LAN, WLAN, and WAN Lesson

Welcome to the *LAN, WLAN, and WAN* lesson.

After completing this lesson, you will be able to achieve these objectives.

A LAN is a local area network. A WLAN is a wireless local area network. A LAN uses cabling connectivity between attached devices. A WLAN uses radio frequency connectivity between devices.

Both LANs and WLANs operate within a limited geographic area and allow multiple users to have access to shared, high-bandwidth media.

Both are controlled privately under local administration and provide corporations with many benefits.

A LAN is made up of the following components.

LANs provide full-time connectivity to local services and connect physically adjacent devices like switches and APs. In a LAN, the attached devices can both send and receive at the same time, in most implementations.

WLANs also provide connectivity to local services and allow access to those services through a device called an AP. APs connect the radio frequency environment and the wired network environment. In a WLAN, connected devices generally take turns sending and receiving, in most implementations.

A LAN that extends over several buildings at a single site is referred to as a campus area network.

Campus area networks allow network access, resource sharing, and sharing of data storage across the entire site. They are usually locally managed, but may outsource connectivity and management services.

Here are some common examples.

A WAN, or wide area network, connects local networks and campus area networks to other networks and can connect one city with other cities, across the globe. Common examples of WANs are the internet and the telephone network.

In most cases, service providers provide the infrastructure for these connections.

The service providers tend to be cable, cellular, and telephone companies. The data rates available to customers are generally lower than inside a corporate LAN.

While a campus area network may span a city, an enterprise network may include connections that span over a wider, even global, area. An enterprise network provides the benefits of network access, resource sharing, and access to data storage for the entire enterprise.

Enterprise networks may provide security using tunneling and encryption for data transfer. They may allow the adding of resources to existing network capacity, referred to as scalability.

An enterprise network can consist of any or all types of networks, including personal area networks—PANs, local area networks—LANs, wireless local area networks—WLANs, campus area networks—CANs, metropolitan area networks—MANs, storage area networks—SANs, and wide area network—WANs connections. Most aspects of enterprise networks are managed under a single administrative body with contractual services offered by a service provider for MANs, WANs, and, occasionally, CANs.

Enterprise networks are large corporate networks that belong to, for example, an automobile manufacturer or commercial airplane manufacturer. They can easily span several continents and include administrative offices, manufacturing facilities, suppliers, and so on.

You've completed the lesson. You can now achieve these objectives.

LAN Topologies Lesson

Welcome to the *LAN Topologies* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

The topology of a network refers to the way in which its components are connected. That is, it refers to how the computers, printers, and other equipment are connected to the network.

Two different concepts of topology can be distinguished: logical topology and physical topology.

Logical topology refers to the path that signals take from source to destination.

Physical topology refers to the way in which devices are cabled.

In this lesson, you will take a closer look at two physical topologies: bus topology and star topology.

LAN bus network topology was developed by Xerox Corp. and is the basis for how Ethernet works.

The bus topology is one of the simplest of the network topologies to use. Click **Start** to begin.

That’s it. All devices receive the signal and all devices share the same physical bandwidth. Click **Next** to continue.

Bus networks work best with a small number of devices. If more than a few dozen computers are added to a network bus, performance problems will likely result.

A major disadvantage of a physical bus topology is that if a single station fails, it may take down the rest of the network. In addition, if the backbone cable fails, the entire network fails.

In a physical star topology, all stations are connected to the LAN through a central point, usually a device called a hub or a switch.

A principal advantage of the star topology, compared to a physical bus topology, is that if a single station fails, it does not generally take down the rest of the network.

It is also much easier to add a new station to the network. All that is required is to plug the new station in to an available port on the central device.

Multiple hubs or switches can be interconnected to form tree or hierarchical network topologies.

Star topologies are most often implemented with low-cost UTP cabling.

Hubs are no longer seen in industrial and corporate networks, though some hubs may still be in use in small office or home office environments. For the most part, switches have replaced hubs as the central point of star topologies.

You've completed the lesson. You can now achieve these objectives.

Access Methods Lesson

Welcome to the *Access Methods* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

An access method is the term given to the set of rules by which networks arbitrate which station is allowed to transmit to the medium at any given time.

In this way, the LAN protocol prevents data transmissions from crashing into each other on the network. If such a thing occurs, it is called a collision.

Networks need access methods for the same reason streets need traffic lights-to prevent vehicles from hitting each other.

Think of an access method as a form of traffic regulation.

The network cable is the street. Traffic laws (or the access method) regulate the use of the street (or cable), determining who can drive (or send data), where, and when.

On a network, if two or more stations try to send data at exactly the same time, their signals will interfere with each other, ruining the data being transmitted. The access method prevents this.

Although many LAN standards exist, two major ones are most often used today: Ethernet implemented on full-duplex physical connections has taken the place of CSMA/CD and CSMA/CA is most commonly implemented on WLANs.

A switch is a network device that permits LAN devices to connect and communicate using full-duplex connections.

The most common access method in use today is Ethernet full duplex in a dedicated switched environment.

A dedicated switched environment means that a single station is attached to each switch port. In this case, because there are four wires inside the LAN cable, two wires can be used to transmit and two wires can be used to receive simultaneously.

Collisions are not possible.

In wireless networks, full-duplex mode is not possible. If two stations send data simultaneously, then the radio frequency (RF) signals interfere with each other and the data is corrupted.

In wireless networks, the CSMA/CA access method specifies half duplex mode only. Half duplex means that a station can send or receive, but not both at once.

A station that has data to send must first listen to the medium (in this case, the radio frequency) to determine whether another station, is already sending data. If so, the station that needs to send data, the second station, must wait until the original station has finished sending data. Often times, the transmitting station does not hear the transmission of another station across the RF. This results in a collision and causes a delay.

The CSMA/CA protocol specifies: listen before talk.

In the topology depicted on screen, a desktop station is connected by a cable to switch. The switch is also connected to an AP (wireless LAN access point), and a notebook is connected, by radio frequencies, to the same AP.

You've completed the lesson. You can now achieve these objectives.

Network Models Lesson

Welcome to the *Network Models* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

In a peer-to-peer model, all nodes in the network communicate equally. No centralized server is implemented.

It is an inexpensive solution and appropriate for a network with 10 computers or less.

In most cases, no backup facility is available. It is the preferred solution for small office or home office environments where there is no budget for a network administrator or high-performance servers.

Up until the early eighties, processing power and memory were expensive. Therefore, all processing intelligence and application software capability was centered in a computer referred to as a mainframe. Users accessed the applications and capabilities of the mainframe through end devices that had little or no processing intelligence themselves. These were referred to as "dumb terminals".

Because all of the intelligence was in the mainframe, the mainframe would sometimes become bogged down processing all of the transactions for all of the terminals.

Today, client-server architecture distributes intelligence.

Intelligent personal computers (clients) share processing of data for applications and databases running on a server. Each user has the power of a computer on the desktop rather than sharing the power of the mainframe with many other users. The data is centralized, but the computing power is distributed.

A server is a computer that runs service applications to which the users on the network have access. This can include word processing, spreadsheet applications, email, or a database.

A server can also hold files for users, or have an optical drive attached to it, such as a DVD-ROM for mass storage, or a laser printer, to which many users have access.

In a classical server farm, a server is a single piece of hardware dedicated to a single function. Therefore, multiple pieces of hardware are required to perform several functions. One server might be dedicated to directory and security services; a second server might be dedicated to file storage; a third server might be dedicated to managing a database, email, or some other functionality.

As discussed previously, a server is a computer that runs software applications to which the users on the network have access. This can include word processing, spreadsheet applications, email, or a database.

A server may also offer access to computing, processing, and networking resources that are divided among different sets of services that are isolated from each other on the server. This is a concept referred to as virtualization and improves efficient use of server resources.

All computers on the network run an operating system (OS). End-user computers run operating systems appropriate to the tasks typically performed on end-user systems. Servers have special OSs.

The clients, or users on the network, usually run a computer operating system, also just referred to as their OS. The OS moves information around in the computer-from input device to memory to hard disk, and supports a wide variety of applications ranging from photo processing, to word processing, email, web browsers, and more.

Here are some common examples of an OS. They also run a client portion of the network operating system in order to communicate with servers and other components on the network.

The servers run a network operating system (NOS). Network devices such as routers or switches are part of the network infrastructure and also run a NOS. This type of OS allows users to log in, connect, share resources, and manage network traffic. Here are some common examples.

Both computer OSs and the NOSs are required for the network to operate.

Putting everything together, we have peer-to-peer network models, as well as client server network models. Each device in the networking model runs an operating system.

Click each network type to review it.

You've completed the lesson. You can now achieve these objectives.

Ethernet Media Types Lesson

Welcome to the *Ethernet Media Types* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

Ethernet is one of the most widely used networking protocols in the world. The Ethernet protocol specifies an access method for governing access to a shared network environment. It offers a format for the transmission of data across the shared network and specifications for cabling to connect devices to the network.

For Ethernet, the digital information or data to be transmitted on the network is sent in a structured format called a frame. This frame is like an envelope of additional information put around the actual data. When a letter is sent through the postal system, it is first placed inside an envelope. Specific address information must be on the envelope so that the postal service can deliver it to the right person.

Like the envelope, frames also contain address information; both the address of where it is to go and the address of where it came from, so the network can deliver the information to the right station.

There are certain rules that define how frames are structured and handled, just like there are rules that must be followed when sending letters through the postal system. These network rules are referred to as network protocols.

Network protocols specify how large or small frames can be, what information can be contained in the frame, and so on. Frames can vary in size.

In a shared network environment, rules facilitate communication between hosts.

A network technology specifies rules for each of the following.

Ethernet is one of the most common network technologies today. Ethernet, as a LAN technology, includes data transmission speeds between 10 megabits per second and 400 gigabits per second. All the devices on the network share this bandwidth. This is known as shared media.

Ethernet can use a logical bus topology, in which all of the end devices connect to the same physical network and take turns transmitting, or it can be implemented on full duplex connections. In today’s Ethernet networks, full duplex is the access method that is usually used in cabled environments.

Ethernet also allows for a variety of physical topologies, as well as many different types of cabling.

Ethernet supports the following cable types.

A common trend in wiring Ethernet networks was to use shielded twisted-pair (UTP) cable.

1000BASE-T, which uses UTP cable, has been a popular implementation for Ethernet.

"T" stands for twisted pair. This type of cabling has eight wires, two of which are twisted around each other in pairs to reduce electromagnetic interference (EMI). It is based on the IEEE 802.3 standard.

1000BASE-T supports a data rate of 1000 Mbps using baseband transmission. BASE stands for baseband and means that there is one channel on the wire, contrasted with broadband transmission, which has multiple channels on the same wire.

The cable uses RJ-45 connectors, and the network interface card has an RJ-45 socket built in.

Here are some recommendations.

The introduction of Fast Ethernet, running at 100 Mbps, required different cabling specifications. The 100BASE-X standard provides cabling specifications for supporting the 100 Mbps data rate in different areas of a network. It is also an older specification that is now used only to connect end devices to the network.

High-capacity networks require even higher transfer rates than the 100 Mbps offered by the Fast Ethernet standard. The Gigabit Ethernet standard was originally developed for up to 10 Mbps, but it now runs at up to 400 Gbps.

Gigabit Ethernet links are often used from the access layer to the backbone or core network. Also, most server platforms have Gigabit Ethernet interface cards that provide high-speed access.

The standards for Gigabit Ethernet are IEEE 802.3x and 802.3z. Most of Gigabit Ethernet protocols are adapted from fiber channel protocols.

Gigabit Ethernet can use either single mode or multimode fiber-optic cables. According to the type of fiber-optic cable that is used, the extendibility of the medium will differ.

The 10 gigabit Ethernet standard (10GbE or 10GE) encompasses eight transmission technologies for fiber-optic cable, and two for copper cable standards for higher Ethernet speeds, to encompass even more transmission technologies.

10 gigabits, 40 gigabits, 100 gigabits, and 400 gigabits Ethernet are now all implemented in LAN, MAN, and WAN environments.

Here are some current standards.

Ethernet is one of the most rapidly evolving network technologies and dozens of different cabling specifications have been defined for different use cases.

You've completed the lesson. You can now achieve these objectives.

OSI Reference Model Module

Standards Overview Lesson

Welcome to the *Standards Overview* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

Networking standards are documented agreements containing technical specifications or other precise criteria used consistently as rules, guidelines, or definitions of characteristics. Networking standards ensure that materials, products, processes, and services are fit for their purpose.

Before there were networking standards, most manufacturers designed and developed their own systems, hardware, and software as "purpose-built systems".

These systems were often unable to communicate or exchange data with systems built by a different manufacturer or even another generation of the system developed by the same manufacturer!

Networking standards have been developed to ensure that diverse hardware and software in a network can interoperate smoothly.

Networking standards accurately define the interaction and interrelation of the various components of the network architecture and enable multi-vendor interoperability.

Networking standards contribute to making life simpler, and to increasing the reliability and effectiveness of the goods and services we use.

Adherence to networking standards ensures full compatibility among open systems to foster healthy competition among producers, and offer real options to users, since competition is a powerful catalyst for innovation, improved productivity, and cost-cutting.

There are numerous agencies and organizations tasked with developing standards for all aspects of data and networking.

Here are some examples.

You've completed the lesson. You can now achieve these objectives.

The Seven Layer Model Lesson

Welcome to the *Seven Layer Model* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

Prior to standardization of the network industry, many systems had their own method of communication. Networking standards had to be developed and the ISO introduced the OSI model, commonly known as the Seven Layer model.

The Seven Layer model offers many benefits. Here are a few.

The Seven Layer model is a conceptual model used to describe standards of communication functions within a networking system.

The upper three layers, the application layers, focus on the programmatic functions of network systems.

The lower four layers, the data flow layers, focus on the networking or data flow of networks systems.

The transport layer, or Layer 4, is the primary interface between the programmer and the networker.

Here are the upper layers of the Seven Layer model.

Here are the data flow layers of the Seven Layer model.

Host-to-host communication, often referred to as peer-to-peer communication, sends data from a host to the corresponding layer on another host or horizontal communication.

Because connectivity occurs at the physical layer, the data must be programmatically moved down the stack layers, sent across the physical layer, and sent up the stack layers on the receiving device.

This is called vertical communication, and is often referred to as encapsulation down the layers and decapsulation up the layers or stack.

The upper layer, or Applications layers, of the Seven Layer model, generate data. The generated data is handed down to the Transport layer, where it is divided and encapsulated into segments.

Segments are passed down to the Network layer, where they are encapsulated into packets that are addressed with the logical address of the destination.

Packets are then passed down to the Data Link layer, where they are encapsulated in frames that are addressed with the physical address of the next device in the network path.

Finally, the frames are converted to bits, which are most commonly represented as either ones or zeroes, and sent across the media.

On the receiving device, the opposite occurs.

At the Physical layer, the bits are collected into frames, and the frames are passed up to the Data Link layer.

The packets are removed from the frames, and are then passed up to the Network layer.

From the Network layer, the segments are handed up to the correct Transport layer.

Finally, the Transport layer hands the data up to the correct application.

The TCP/IP model is the most common protocol stack. It is what operates on the internet.

Now, take a moment to compare the TCP/IP protocol model against the OSI model.

Application layer

The OSI model Layers 5 through 7 map to a single TCP/IP Application layer. This Application layer is often referred to as Layers 5 through 7.

Transport layer

The OSI model Layer 4 maps to the TCP/IP Transport layer, which is often referred to as Layer 4.

Internet layer

The OSI model Layer 3, or the Network layer, maps to the TCP/IP Internet layer. This is why the world wide web is called the "internet", because it operates at the Internet layer in the TCP/IP model. Note that the Internet layer and its addressing is often referred to as Layer 3 IP addressing, in order to differentiate it from Layer 2 MAC addressing.

Network layer

The OSI model Layers 1 and 2 map to a single TCP/IP Network Access layer or Link layer.

Click **Next** to continue.

As the data moves down the stack, encapsulation occurs, and each layer defines a PDU.

The application data produced by the upper layers can be very large, so the Transport layer often needs to break up the data into smaller segments, so that it will fit into individual packets.

The Transport layer then prepends a transport header, which includes an application or port number.

When the segment is passed down to the Network layer, the packet header is prepended with the logical address of the destination.

The packet is sent down the Data Link layer. The Data Link layer might prepend an LLC header to provide for acknowledged delivery of the frame, if needed, but will prepend a MAC header and an FCS trailer to the frame.

The frame is sent over the media as bits.

Click each button to explore an example of data encapsulation.

Example Part 1

The user types an email message, and the email application generates data.

Example Part 2

The data is passed down to the Transport layer, where the segment/TCP header is prepended.

Example Part 3

The segment is passed down to the Network layer, where the network/IP header is prepended to the packet.

Example Part 4

The packet is passed down to the Data Link layer, and is encapsulated into a frame with a frame header and a frame trailer.

Example Part 5

The frame header and frame trailer become ones and zeros, or bits, at the Physical layer, and the bits are sent over the media.

What happens when the data is decapsulated? Click the buttons from bottom to top, to explore a data decapsulation example.

Example Part 6

The MAC header and FCS trailer are verified and removed. If there is an LLC, it is inspected and removed.

Example Part 7

The packet is then passed up to the appropriate IP layer, either IPv4 or IPv6, where the IP header is inspected and removed.

Example Part 8

The segment is then passed up to the Transport layer, where the TCP header is used to verify flow control, perform error checking sequencing, and evaluate the segment for possible retransmission.

Example Part 9

The data is then sent up to the correct application, resulting in an email.

You've completed the lesson. You can now achieve these objectives.

Physical and Data Link Layers Lesson

Welcome to the *Physical and Data Link Layers* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

The Physical and Data Link layers are very dependent on each other.

In fact, in the TCP/IP Model, these two layers are combined into a single link layer, also called the Network Access layer.

Data Link Layer

At the Data Link layer, data is formatted into a frame. The frame will include a destination address, a channel number, or circuit number, to direct the traffic to the target device. Depending on the standard defining the frame structure, the frame may also include the source address.

The service access point or type code will indicate the service or type of network packet being carried inside the frame.

Each data link standard may include network topology information, whether connectionless or connection oriented. Connection-oriented network topologies may utilize frame sequencing and flow control.

Physical Layer

The Physical layer defines physical media properties, such as media type, connector type, and signaling type.

The media types include, but are not limited to, copper, single-mode or multi-mode fiber, and wireless—both radio and microwave.

The connector types depend on the media types, and can include RJ45 connectors on UTP copper, or various fiber pair connectors.

Signaling type will define the frequencies, clocking, modulations, voltages, and distances of the connection.

Here are just a few of the Data Link and Physical standards.

It is important to note that while many standards are distinctly separated between the two layers, Ethernet II, as a standard, defines both the Data Link and Physical layers.

LAN media may use the LLC to refer to the upper layer and network layer packet types.

The 802.3, 802.11, or MAC refers to the lower hardware level functions.

The LAN physical address of the network interface card (NIC) goes by many names, depending on the system or documentation.

Examples include, the MAC address, burned-in address, NIC address, or Ethernet address.

The physical address is 48 bits in length, and is represented as 12 Hex characters in various formats. Here are some examples.

The first 24 bits (6 Hex) are the vendor code or the organization unique indicator (OUI) of the NIC. This is followed by the 24 bit (6 Hex) serial number of the NIC.

Ethernet frames have a preamble. The preamble allows the receiver to lock in to the timing of the data stream before the actual frame begins.

The preamble is composed of 7 bytes of alternating ones and zeroes. The eighth byte ends in 10101011, indicating the next byte is the beginning of the frame header.

The preamble is followed by the 6-byte destination MAC address (12 Hex characters) and 6-byte source MAC address of the sending NIC.

After the source MAC address comes the 2-byte type/length code, followed by the data content of the frame.

An Ethernet II frame has a 2-byte type code in the type/length field, indicating the type of data carried in the frame, followed immediately by the data type indicated.

The older 802.3 frame will then have a 2-byte length code in the type/length field, indicating that there is an 802.2 LLC header, which defines the packet type.

After the source MAC address comes 2-byte type/length code.

The original 802.3 frame format, using the length indicator and LLC header, is an older version and is rarely used today.

You've completed the lesson. You can now achieve these objectives.

Network Layer Lesson

Welcome to the *Network Layer* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

The Network layer defines, amongst many things, the logical addressing of both the source system and the destination system. The function of the Network layer is to differentiate between the physical MAC address and the logical network address.

The Network layer logical address is generally divided into two components.

The first, or upper most section, defines the network address used to get to a particular network.

The second, or lowermost section, defines a specific node or host on that network.

Consider the postal example from earlier.

Country, state, city, and street would be the network number that the postal service would use to find the street or network. The house number, 7, would be the node address.

Devices called routers are used to forward data between networks.

Routers use the network component to select best path routing to the network, and the very last router uses the node component to forward the packet to the destination system.

As an example, think about a very simple network number, with the address divided as seen here.

In comparison, IPv4 uses a 32-bit address, presented in a format called dotted decimal notation, for example, 10.8.2.48.

In IPv4, a portion of the address indicates the network, and another portion indicates the node. The router could be looking at the network component of the address, in this case 10.8.2, to determine the path to the network.

The router would then use the node or host component, .48, to forward the packet to the destination host.

The routers will store the known network portion of the address, called the network number, in a routing table. The routers then use that information to determine the best path to a particular network.

In this example, a PC application operates at Layer 7. The application will generate data which is encapsulated down through the seven-layer stack.

The routers operate at the Network layer, Layer 3, and forward the packets based on the best path determination.

Ultimately, the last router in the path will forward the packet to the server, which decapsulates the packet right up to the Application layer.

You've completed the lesson. You can now achieve these objectives.

Transport Layer Lesson

Welcome to the *Transport Layer* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

The Transport layer will multiplex many different application data streams by distinguishing different applications with well-known port numbers.

The Transport layer will also segment traffic from the data streams, encapsulate segments by adding a transport header, and hand each segment to the Network layer to encapsulate into a packet for forwarding.

Depending on the needs of the application, the protocol used at the Transport layer may establish an end-to-end connection and guarantee delivery and handling service for the data. This is referred to as a "connection-oriented" or "reliable" connection.

The connection-oriented (or reliable) connections handle sequencing, reordering, retransmission, and flow control of the data. However, as a consequence, these types of connections have more overhead and may be slower.

Depending on the needs of the application, the protocol used at the Transport layer may send the data to the Network layer, without establishing an end-to-end connection or guaranteed delivery and handling for the data. This is referred to as a "connectionless" or "best-effort" connection.

Connectionless connections are very fast, but do not handle error checks, retransmission, or sequencing.

You've completed the lesson. You can now achieve these objectives.

Upper Layers Lesson

Welcome to the *Upper Layers* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

The upper three layers are often a concern to application programmers. The upper layers are usually incorporated into the functionality of the application or the operating system that applications run on.

The upper three layers govern access to system resources, the look and feel of the applications, and how the applications exchange data.

The Session layer coordinates the activities of applications as they interact on different hosts. It also manages access to system resources, such as memory, disk space, and processing.

The Session layer also handles service requests and service reply interactions between the client and servers.

The Presentation layer is also incorporated into the application or operating system. This layer ensures that there is a standard format for the data between the two systems, as well as the look and feel of the user interface.

The Application layer supplies the actual networking services to the computer application. It also supplies the interface through which the user enters data and performs tasks.

For example, the user opens a Thunderbird email application client, which allows them to generate an email. The email is then handled by the Simple Mail Transport Protocol (SMTP) network application.

How the user interacts with the application and how the application client interacts with the network application service, is handled by the Application layer.

You've completed the lesson. You can now achieve these objectives.

TCP/IP Internet Layer Module

Internet Layer Protocols Lesson

Welcome to the *Internet Layer Protocols* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

Several protocols operate on top of Ethernet II with accordant Layer 2 protocol numbers. This figure shows a few examples.

Ethernet indicates these higher layer protocols in a type field. Note that both IPv4 and IPv6 are protocols that operate over Ethernet and are Internet layer protocols.

The Internet layer, or Layer 3 in the OSI model, is responsible for moving data through a set of networks, also known as the internetwork.

The Internet layer uses an addressing scheme in which devices can determine the destination of data as data moves through the networks.

Protocols that have no Internet layer must be switched, cannot be routed, and do not scale well.

Such protocols commonly use only a name (or MAC address) to identify the computer on a network. The problem with this approach is that as the network grows in size, it becomes increasingly difficult to organize all the names and to ensure that two computers are not using the same name.

Protocols that support the Internet layer use a hierarchical addressing scheme that mandates unique addresses across network boundaries. This provides a method for finding a path for data to travel between networks.

MAC addresses, by contrast, use a flat addressing scheme which makes it difficult to locate devices on other networks. Take a moment to explore an example.

A Layer 3 IP packet consists of the data from upper layers plus an IP header, which consists of the following fields.

Network devices need an addressing scheme that allows them to forward data packets through the internetwork.

The internetwork can be understood as a set of networks composed of multiple segments all using the same type of addressing.

There are several Network layer protocols with different addressing schemes that allow devices to forward data through an internetwork.

The internet protocol (IP) is the most popular implementation of a hierarchical network addressing scheme.

IP is the network protocol the internet uses. IP determines the form of the IP packet header. This includes addressing and other control information. IP does not concern itself with user data. IP accepts whatever is passed down from the higher layers.

A field in the IPv4 header indicates which upper-layer protocol sent the data down to the Internet layer and which protocol will receive the data on the target station.

The protocol field in the IPv4 header determines the Layer 4 protocol being carried within an IP packet. Although most IP traffic uses transmission control protocol (TCP) or user datagram protocol (UDP), there are other protocols that can be carried by IP.

Each IP header must identify the destination Layer 4 protocol within the datagram. Transport layer protocols are numbered by IEEE/IANA (Institute of Electrical and Electronics Engineers/Internet Assigned Numbers Authority), similar to port numbers.

Take a moment to explore an example.

Internet Control Message Protocol (ICMP) is one of the core protocols of the internet protocol suite. It is also known as ICMPv4.

It is primarily used by the operating systems of networked computers to send diagnostic or error messages. For example, it indicates that a requested service is not available or that a host or router cannot be reached.

ICMP relies on IPv4 to perform its tasks and is an integral part of IPv4.

ICMP differs in purpose from transport protocols such as TCP and UDP in that it is not typically used to send and receive data between end systems.

ICMP is rarely used directly by user network applications, with some notable exceptions being the ping and trace route utilities.

ICMP is implemented by all TCP/IP hosts.

ICMP messages are carried in IP datagrams and are used to send error and control messages.

The table on the screen indicates the various types of diagnostic and error messages that ICMP supports.

A ping command is often used to test connectivity between IP hosts.

If the ping fails, results could include other ICMP messages. For example, destination unreachable or timeout messages.

The ping is known as an ICMP echo request.

The ping response, which is a successful reply to a ping, is called an ICMP echo reply.

If a router receives a packet that it cannot deliver to its destination, the router sends an ICMP destination unreachable (Type 3) message to the source.

The message might be undeliverable because there is no known route to the destination, or because the host is not responding, or because the path is locked by an administrative filter, such as an access control list.

You've completed the lesson. You can now achieve these objectives.

IP Addressing Lesson

Welcome to the *IP Addressing* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

In the 1980s and early 1990s, IP addresses were assigned to corporations and organizations based on the number of potential host addresses required for the devices on their networks. These addresses were hierarchical in structure and designed to allow different classes of addresses to be assigned, based on the size of the network. Class A network addresses were assigned to the very largest networks, Class B addresses to smaller networks, and Class C addresses to the smallest networks.

Given that there were only 126 Class A addresses available, it quickly became clear that it was impractical and unfair to distribute these addresses to the first 126 corporations or organizations that happened to ask for one because they were relatively large.

Responsibility for distributing IP addresses more fairly was transferred to an organization called the Internet Assigned Numbers Authority, also known as IANA.

IANA distributes contiguous blocks of addresses to regional internet registries (RIRs). The RIRs and their geographical domains of IP address distribution are the following.

These RIRs further distribute contiguous blocks of IP addresses to local country internet registries and/or service providers within their respective regions.

In a TCP/IP environment, network nodes, such as end stations and servers, communicate with other network nodes. This communication depends upon each node using the TCP/IP protocol suite and a unique logical address. This 32-bit long address is known as the IP address.

IP addressing has a hierarchical structure, which is realized in the two parts of an IP address: a network ID and a host ID.

Together, these two parts uniquely identify each device connected to the internet. It is the unique combination of both the host and the network address portions that makes it possible to access any given host on the internet.

IP addresses are like postal mail addresses, which identify a location by providing a country, a city, a street, and a house number. The IP packet is like a letter inside an envelope containing the information a user wants to send, plus, outside the envelope, an address for delivering the letter through the post office.

Routers use the network portion of the IP address, or network ID, in an IP packet to identify the destination network of the packet within an internetwork.

The host portion of the IP address, or host ID, is used to identify the target device on the local network. A local network administrator assigns the host portions of IP addresses according to a predetermined network addressing plan.

The allocation of network IDs is always managed by a central authority.

Network numbers are administered by IANA.

The task of providing internet resource allocations, registration services, and coordination activities, which support the operation of the Internet globally, has been transferred by the IANA to five RIRs.

Public IP addresses are globally unique 32-bit numbers.

These 32 bits are written as four sets of octets, each of which is represented as a decimal number between 0 and 255, and separated by a period.

This representation is called the dotted-decimal notation.

Each bit in each octet has a binary value. These values are allocated starting with the lowest value of 1 on the right side of each octet, and ending with the highest value of 128 on the left side.

Hence, the decimal values for each bit in an octet are, in order: 128, 64, 32, 16, 8, 4, 2, and 1.

To accommodate different size networks and aid in classifying them, IP addresses are divided into classes.

This is known as classful addressing. Each IP address is broken down into a network portion and a host portion.

There are three classes of IP addresses that an organization can receive from the registry responsible for the region in which the network (or the organization's ISP) is located.

They are: Class A, Class B, and Class C.

The IANA now reserves Class A addresses for governments throughout the world, although a few large companies, such as Hewlett Packard, have received one in the past.

Class B addresses serve medium-sized companies.

All other requestors are issued Class C addresses.

Class A

The first bit of a Class A address is always 0. An example of a Class A IP address is 124.95.44.15. The number 124 identifies the network number assigned by the IANA; the internal administrators of the network assign the remaining 24 bits. An easy way to recognize whether a device is part of a network of a certain class is to look at the first octet of its IP address. The first octet in an assigned Class A network will range from 1–126. The IP addresses that begin with 0 and 127 are reserved for special purposes and are not assigned to corporate or industrial networks. Class A IP addresses use only the first 8 bits to identify the network portion of the address; the remaining three octets can be used for the host portion. A network that uses a Class A IP address can assign up to 224 minus 2, or 16,777,214, possible IP addresses to devices attached to the network.

Class B

The first two bits of a Class B address are always 10. An example of a Class B IP address is 151.10.13.28. The first two octets identify the network number assigned by the IANA; the local administrators of the network assign the remaining 16 bits to hosts. A device is part of a Class B network if its IP address has values ranging from 128 to 191 in its first octet. Class B IP addresses use the first 16 bits to identify the network portion of the address; the two remaining octets of the IP address can be used for the host portion. A network that uses a Class B IP address can assign up to 216 minus 2, or 65,534, possible IP addresses to devices attached to the network.

Class C

The first three bits of a Class C address are always 110. An example of a Class C IP address is 201.110.213.28. The first three octets identify the network number assigned by the IANA; the local administrators of the network assign the remaining 8 bits to hosts. A device is part of a Class C network if its IP address has values ranging from 192 to 223 in its first octet. Class C IP addresses use the first 24 bits to identify the network part of the address; only the last octet of a Class C IP address is used for the host portion of the address. A network that uses a Class C IP address can assign up to 28 minus 2, or 254, possible IP addresses to devices attached to the network.

There are also Class D IP addresses, which are reserved for multicasting, and Class E addresses, which are reserved by the Internet Engineering Task Force, or IETF, for it's own research.

Class E

The first octet range for class E addresses is 11110000 to 11111111 or 240 to 255. The Internet Engineering Task Force (IETF) reserves these addresses for its own research.

Class D

The first octet range for class D addresses is 11100000 to 11101111 or 223 to 239. The IP addresses in this range are reserved for use when exchanging multicast data. Multicasting is when a specified group of hosts are all receiving the same data at one time.

The most significant bit pattern determines the class of the address, as well as how many bits are used for the network portion of the address. The bit pattern also directly impacts the number of available assignable host addresses there are for each address class.

Here are some examples.

You've completed the lesson. You can now achieve these objectives.

Reserved, Public, and Private IP Addresses Lesson

Welcome to the *Reserved, Public, and Private IP Addresses* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

Certain addresses are reserved and cannot be assigned to devices on a network.

These reserved addresses include network and broadcast addresses.

Network address

An IP address that has binary 0s in all host bit positions is reserved for the network address, which is used to identify a network. A network address should have binary 0s in all host octets.

The IP address of a network is not used for any device attached to a network. A router uses the network address when searching its routing table for destination network locations.

Directed broadcast address

To send data to all devices in a certain subnet, a directed broadcast address is used. This IP address sets the network and subnet portions to the value of the target subnet, and the remaining host bits are set to binary 1s.

Local broadcast address

To send data to all devices on the local segment, the local broadcast address is used.

Network ID

The network ID of an IP address is important because most hosts on a network can directly communicate only with devices in the same network. If a host needs to communicate with devices that have interfaces assigned to another network ID, there must be a network device that can transfer data between different networks. This device is usually a router.

A network ID enables a router to forward a packet to the appropriate destination network.

Hosts for classes of IP addresses

Each network class allows a fixed number of hosts. In a Class A network, the first octet is assigned to the network, leaving the last three octets (24 bits) to be assigned to hosts.

The Internet Assigned Numbers Authority (IANA) manages the supply of IP addresses to ensure that duplication of publicly used addresses does not occur. Such duplication would cause instability in the internet and compromise the internet’s ability to deliver packets to networks using the duplicated addresses.

Internet service providers (ISPs) distribute IP addresses or blocks of addresses.

The growth of the internet in the 1990s led to a shortage of globally unique IP addresses.

As a result, the Internet Engineering Task Force (IETF) began work on defining a longer address format. This new format allowed the assignment of more unique addresses for use in the global internet.

IPv6 is the result of that work and uses 128-bit addresses.

To approach the address shortage issue in the short term, the IETF has defined three blocks of addresses that can only be used inside of private networks to address devices.

Addresses from these private address blocks cannot be routed in the public internet. Therefore, these addresses must be changed in any packets that are sent out of the private network. A function called address aliasing, or address translation, is used to swap private addresses for globally unique addresses, in packets that are leaving the private network.

Public IP addresses are unique because they are global and standardized. All machines connected directly to the internet agree to adhere to this system. To communicate directly over the global internet, internet hosts require a globally unique IP address.

Private hosts not connected to the internet can use any valid address, as long as it is unique within the private network.

Because many private networks exist alongside public networks, grabbing just any address is strongly discouraged. RFC 1918, from the IETF, sets aside three blocks of IP addresses for private internal use.

Addresses in this range are not routed on the internet backbone. Internet routers immediately discard packets with private destination addresses.

Connecting a network that uses private addresses to the internet requires the translation of the private addresses to a public address or addresses. This translation process is referred to as network address translation (NAT). NAT is often performed by a router.

You've completed the lesson. You can now achieve these objectives.

IPv6 Lesson

Welcome to the *Internet Protocol Version 6 (IPv6)* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

In the 1980s and early 1990s, IPv4 offered an addressing scheme which, although scalable for its original purpose, resulted in a less than optimal allocation of addresses. The architects of TCP/IP could not have predicted that their protocol would eventually sustain a global network of information, commerce, and entertainment.

Class A and B addresses make up 75 percent of the IPv4 address space. However, only a handful of organizations can be assigned Class A or B network numbers. While Class A and Class B addresses can provision large numbers of assignable addresses, the number of addresses provisioned is often more than the needs of the networks to which they are assigned.

Class C addresses account for only 12.5 percent of the possible 4 billion IP addresses, yet they are far more numerous than Class A and Class B addresses. Class C addresses do not support more than 254 hosts, and this does not meet the needs of medium size to large organizations.

In 1992, the IETF identified two concerns regarding the exhaustion of the remaining unassigned IPv4 addresses.

In the late 1990s, a more scalable version of IP was defined. This is IPv6.

IPv6 has a 128-bit binary address structure which provides 3.34 x 1038 IP addresses.

IPv6 has already been implemented in multiple networks and will replace IPv4 eventually.

Here is the IPv6 datagram.

Click each region of the datagram to learn more.

Here are the extension headers.

An IPv6 address consists of 128 binary digits (1s and 0s).

The addresses are commonly written in hexadecimal digits, consisting of eight fields of four hex digits each. Colons separate these fields. An example is shown here in the figure.

The hexadecimal digits A, B, C, D, E, and F in IPv6 addresses are not case sensitive.

Use the following guidelines for IPv6 address string notations.

Take a moment to explore some examples.

The IPv6 address consists of a network portion and a host portion. The first 64 bits constitute the network portion, and the second 64 bits constitute the host portion.

The network portion contains several fields that facilitate network aggregation and routing.

These include fields that identify the source registry and the service provider to which the address has been assigned.

The host portion is an interface identifier (IID) and is equivalent to an Ethernet MAC address.

The 48-bit MAC address can be used to create a 64-bit extended unique identifier (EUI), and is incorporated as the host portion of an IPv6 address.

The IEEE EUI-64 address represents a new standard for network interface addressing. EUI-64 addresses are often assigned to a network adapter to provide a unique host address for the interface. This is one of several ways that a unique address can be assigned to a station on the network.

Addresses can also be assigned by an IPv6 DHCP server. Unlike addressing in IPv4, stations may have multiple addresses assigned to a network interface for different purposes. Each station should have a unicast address assigned.

Because of the many ways that addresses can be assigned in IPv6, it is important to have a method for differentiating whether the address is for use on the local private network only, or if it is globally unique and can be used to connect directly to the global internet.

There are codings included in the address structure itself to indicate the use and type of the address.

These codings are known as universal/local versus individual/group indicators.

Take a moment to explore each one.

Here are examples of the different types of addresses that can be assigned for use to the network interface of a station.

There are also different ranges of addresses assigned for special use.

Like IPv4 addresses, IPv6 addresses for use on the public internet are defined by IANA and other address registries for special purposes and are also assigned to different groups for use.

For example, some IPv6 addresses have already been assigned to service providers and their customers.

Here is a list of some of the ranges of IPv6 addresses that have been defined for special purposes, such as tunneling IPv4 packets across IPv6 networks.

Similar to the addressing for IPv4, IPv6 also has some addresses that are reserved for special use cases.

ICMP has also been re-invented to support IPv6 and is now designated ICMPv6. The IPv6 next header number 58 (0x3A) specifies ICMPv6.

The diagnostic functions were carried over and enhanced for IPv6 functionality.

Some of the diagnostic functions that were carried over include error messages as well as messages used for troubleshooting. The enhanced functionality includes information exchanged for the purpose of managing addressing and packet size.

A new function is the Neighbor Discovery Protocol (NDP). As broadcast is no longer supported in IPv6, a new neighbor discovery mechanism was needed to replace ARP. NDP fulfills this function.

IPv6 also offers new ways to assign addresses which were not possible in IPv4. NDP is part of this new functionality as well.

In IPv4, a request for a neighbor's physical address uses the ARP request and is sent to the local broadcast address.

In IPv6, the NDP fulfills this function using a multicast address instead. The neighbor reply is unicast, because the ARP replies in IPv4.

The NDP then messages neighbor solicitation (NS) and neighbor advertisement (NA), which are both carried by ICMPv6 and are individually identified by different ICMPv6 type numbers.

Type 135 applies for the NS message and type 136 applies for the NA message.

IPv6 stations maintain a cache of learned neighbor addresses, similar to the IPv4 ARP cache, which is called the IPv6 Neighbor Table.

The IPv6 NDP defines another ICMPv6 packet type whereby devices can learn about nearby routers or default gateways.

IPv6 routers send periodic announcements of their presence; the message is called a router advertisement (RA), and is sent to a reserved IPv6 multicast address.

By listening to these multicast announcements, all hosts on the segment can learn the IP address of the local router or default gateway.

One advantage of RA messages is that IPv6 does not require a DHCP server for automatic addressing.

IPv6 clients can learn their network address prefixes directly from local routers by simply listening to the RAs.

RAs contain the IPv6 prefix for the local segment and the address of the DNS server (RFC 8106).

The client no longer needs the IP address of the default gateway, as in IPv4, because IPv6 uses the link local address for this purpose. DHCP options are not supported by the RAs.

A client does not need to wait for a periodic RA; it may send a router solicitation (RS) message at any time.

If a router on the same physical segment is active and so configured, then it replies with an RA.

If no such router is present, then the client may send a multicast DHCPv6 discovery message.

Wireshark often captures these DHCP requests on an active link because most current operating systems have the IPv6 stack running by default.

Here is an example of Wireshark IPv6 in action. In this case, a client DHCP packet decode is being shown.

In addition to static addressing and configuration using DHCP, like IPv4, IPv6 also offers several other methods of assigning addresses.

IP has gone through significant changes since its inception as a set of protocols originally defined for military network use.

The protocol was never intended for use in the global networking environment that it now supports.

But like many other protocols, IP has been adapted for use, as needed. These adaptations include changes in address assignment, address format, as well as the addition of new mechanisms for managing addressing and address translation. These updates have enabled IP to become one of the most widely used set protocols in the world.

Today, there are working groups and task force members focused on improvements for the future of IP networking. The latest whitepapers and RFCs can be found at IANA.org where ongoing work documentation is published.

You've completed the lesson. You can now achieve these objectives.

IP Addressing, Assignment, and Resolution Lesson

Welcome to the *IP Addressing Assignment and Resolution* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

In order for devices in the same segment to communicate, the sending device needs both the Internet Protocol (IP) address and the MAC address of the destination device.

When a destination IP address is known, but not the corresponding MAC address, the known IP address must be resolved to the unknown MAC address.

There are a variety of ways that devices can determine the MAC addresses they need to add to the encapsulated data.

For example, the Transmission Control Protocol/Internet Protocol (TCP/IP) suite includes a protocol called the address resolution protocol (ARP), which does exactly this.

ARP enables one computer to learn the MAC address of another computer in order for the other computer's MAC address to be associated with the other computer's IP address.

Computers keep tables that contain MAC addresses and IP addresses of other devices which have been learned using ARP or have been statically configured. These tables are called address resolution protocol, also known as ARP tables or ARP cache.

ARP tables are maintained automatically on each of the devices. Only on rare occasions must an ARP table entry be made manually.

Each computer on a network maintains its own ARP table. Whenever a network device wants to send data across a network, it uses information provided by its ARP table.

When a source determines the IP address for a destination, the source consults its ARP table in order to locate the MAC address for the destination.

If the source locates an entry in its table, such as the destination IP address and destination MAC address, it binds, or associates, the IP address to the MAC address and uses it to encapsulate the data.

If the source host is unable to locate a MAC address for the destination in its own ARP table, it sends an ARP request. An ARP request enables it to discover the destination MAC address.

A host builds an ARP request packet and sends it as a broadcast to all devices on the network.

To ensure that all devices see the ARP request, the destination MAC address is the local broadcast MAC address. This means all bits of the destination MAC address are set to binary ones, or hexadecimal F.

Thus, a MAC broadcast address would have the following form.

Because ARP request packets have the broadcast destination address, all devices on the local network receive the packets and are required to process them and pass them up to the ARP protocol running on the device for further examination.

If the IP address of a receiving device matches the requested IP address in the ARP request, that device responds by sending its MAC address back to the requesting station. This is known as an ARP reply.

After determining the addressing scheme for a network, a method must be chosen for assigning addresses to hosts. There are essentially two methods available for assigning IP addresses: static addressing and dynamic addressing. Regardless of which addressing scheme one uses, it is important to remember that no two interfaces can have the same IP address.

If you assign IP addresses statically, you must go to each individual device and configure it with an IP address. This method requires you to keep meticulous records, because problems can occur on the network if duplicate IP addresses exist.

Some operating systems, such as Windows 8 and Windows 10, send an ARP request, also called a gratuitous ARP, to check for duplicate IP addresses when they attempt to initialize TCP/IP. If they discover a duplicate IP address, the operating systems will not initialize TCP/IP and will generate an error message.

Record keeping is also important because not all operating systems identify duplicate IP addresses.

There are a few different methods that can be used to assign IP addresses dynamically. The most commonly implemented method is dynamic host configuration protocol (DHCP).

DHCP helps a host obtain an IP address quickly and automatically. All that is required is a DHCP server with a defined range of IP addresses for distribution to clients. As hosts come online they contact the DHCP server and request an address. The DHCP server selects an address from a pool of available addresses and allocates it to the host.

DHCP can transmit other network relevant configuration parameters in a single message, including the IP address, subnet mask, default gateway address, DNS address, as well as other OS-specific parameters such as domain name or workgroup.

DHCP can make TCP/IP network administration much more efficient by dynamically assigning IP addresses to hosts. This practically eliminates the need to configure host addresses manually.

A DHCP client can even move to a different subnet and obtain a new IP address without any need for manual reconfiguration.

The DHCP process requires two elements: DHCP servers which assign IP addresses and DHCP clients which request IP addresses.

A single DHCP server can supply addresses for more than one network. To support DHCP on an internetwork, routers must be configured with DHCP-relay forwarding.

DHCP clients initially communicate using broadcast frames, which normally do not pass routers. However, routers configured as so-called relay agents assist clients to reach servers on different network segments by acting as a kind of go-between or proxy between the client and the server.

The DHCP server maintains pools of IP addresses, called scopes. When a DHCP client first starts up, it requests and is usually granted a lease to use an address from an appropriate scope for a specific period of time.

If the DHCP server is on a different subnet, the DHCP relay agent tags each of the requests for an IP address with the DHCP relay interface address. This informs the DHCP server to assign an address to the client within the agent's subnet and the client's subnet.

The concept of leasing is important because it allows a DHCP server to distribute a smaller number of addresses than there are clients in total, assuming not all of the clients are online at the same time. When a client goes offline, its address is returned to the pool to be used by some other client.

Clients must periodically renew their leases. This renewal ensures that hosts receive current configuration parameters on a routine basis.

Click each button to explore how DHCP messages are sent using the discover, offer, request, acknowledge method; commonly abbreviated as DORA.

DNS, sometimes called domain name service, is the telephone book of the internet.

A successful telephone book search provides a telephone number for a known name. Likewise, a successful DNS resolution provides an IP address for a known name. A reverse DNS can resolve a name, if the IP address is known.

Both DNS and reverse DNS can be used for resolving names and addresses both locally, such as computer names, and on the internet, such as URLs.

There is no central DNS database which maintains a list of all names and IP addresses. DNS is distributed over thousands of servers all over the world. DNS servers worldwide maintain a distributed database which is periodically updated and synchronized.

Take a moment to explore an example of DNS service in action.

The DNS naming convention consists of three tiers. Here is an example.

Note each tier is separated by a dot from the others.

The right-most portion is the top-level domain (TLD). There are several types of TLDs, including country-coded domains; organizational domains; and the original infrastructure domain, the Address and Routing Parameter Area domain, denoted as .arpa. These are defined by standard organizations.

Second-level domains (SLDs) are assigned to persons or corporations or other entities, such as museums and schools, when they register for their domain names.

The combination of a TLD and an SLD enables a person or organization to set up a server and offer services under the name so assigned.

An example would be hosting a website under the name www.fortinet.com or a file server under the name ftp.fortinet.com.

A further layer of functionality can also be offered, called third-level domains or subdomains. In this case, smaller and smaller units can be addressed—for example, by appending a hostname or computer name to the URL chain.

A URL which is complete is said to be a fully-qualified domain name (FQDN). An example of an FQDN is www.fortinet.com.

You've completed the lesson. You can now achieve these objectives.

IP Subnetting and Calculation Lesson

Welcome to the *IP Subnetting and Calculation* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

The outside world sees an organization as a single network, with no detailed knowledge of the internal structure required.

Without subnets, both routing and the utilization of network addressing space would be highly inefficient. Routing tables become very large and their size leads routers to experience delays when performing path determination. Furthermore, in a flat switched network, broadcasts and multicasts are flooded by switches to all network segments.

These factors represent a significant burden on end station processors, which must process every inbound broadcast packet.

Empirical experience has shown that when the broadcast level in a network exceeds 25 percent, CPUs tend to struggle with the load because they are almost permanently distracted by their own network cards.

Network administrators typically need to divide networks, especially large ones, into smaller networks.

These smaller divisions are called sub-networks and provide addressing flexibility, routing efficiency, and broadcast traffic control.

Most often, sub-networks are simply referred to as subnets.

Similar to the host number portion of Class A, Class B, and Class C addresses, subnet addresses are assigned locally, usually by the network administrator.

Also, like other IP addresses, each subnet address is unique within the private network.

For each of the first three classes of IP addresses, a default mask applies. Together with the IP address, the mask indicates which part of the IP address signifies the network and which part of the IP address signifies the host.

Each position in the mask which carries a binary value of 1 corresponds to the network portion in the IP address. Each position in the mask which carries a binary value of 0 corresponds to the host portion the IP address.

These masks are sometimes called natural masks, or classful masks, because they apply to the full class of each address.

Take a moment to explore a few examples.

The primary reasons for using subnets is to improve routing efficiency and reduce the size of a broadcast domain.

Broadcasts are sent to all hosts on a network or sub-network and are flooded by switches, bridges, and wireless LAN access points.

When broadcast traffic begins to consume too much of the available resources, network administrators may choose to reduce the size of the broadcast domain by implementing subnetting. In this case, some of the switches are replaced by routers, which prevent broadcasts from being flooded to all network segments.

For example, a Class B address assigned by the service provider can be broken up into many subnets. With subnets, the network address utilization is more efficient.

Consider the questions on screen. Click each view the answer.

Question 1: What are the routing bits (network portion) in the address 131.108.0.0?

Answer 1: 131.108—that is the 16-bit Class B network number as determined by the class B default mask, 255.255.0.0.

Question 2: What are the other two octets (16 bits) of the address 131.108.0.0 used for?

Answer 2: As far as the internet knows, that is a 16-bit host field, because that is what a Class B address is: a 16-bit network number and a 16-bit host number.

Question 3: What part of the address 131.108.0.0 is the subnet field?

Answer 3: The answer depends on when you decide to create subnets. You divide the original host field (16 bits in the case of Class B) into two parts: subnet field and host field. This is sometimes referred to as "borrowing" some of the original host bits to create a subnet field. The network administrator can decide how many host bits he wants to re-designate as subnet bits. The minimum number of bits that can be borrowed is two, regardless of whether you are working with a Class A, B, or C network address. The maximum number of bits that can be borrowed depends on the address class.

A router which is directly connected to one or more subnets will know the appropriate subnet mask for each subnet because these are part of the configuration file of the router itself.

A subnet mask is not a default mask, but one which redefines some host bits as subnet bits. It is therefore essential for a network administrator to recognize address classes immediately without having to calculate or learn the default masks by heart.

A network administrator should be able to tell at a single glance whether an IP address and a mask are sufficient to recognize whether the default mask or a subnet mask is being employed.

The subnet mask, also referred to as the extended network prefix, determines which part of an IP address is the subnetwork field and which part is the host field.

An IP address without a mask is meaningless, and conversely, a mask without an IP address is meaningless.

To determine the subnet mask for a particular subnetwork IP address, complete the following steps.

Subnet bits are taken from the high-order bits of the host field. Only contiguous-bit subnet masks are valid. Contiguous-bit means that, counting from the left, only a row of all 1s with no 0s in between, are valid masks.

Take a moment to explore an example.

To put it another way, a valid subnet mask consists of a consecutive row of binary 1s immediately followed by a consecutive row of binary 0s. To help remember, imagine an invisible line, wall, or border between the row of 1s and the row of 0s.

Whenever the subnet portion of the address is extended into the host portion, in other words, when you move your invisible line to the right, you borrow bits from the host portion. In terms of the subnet mask, this means the more bits you set to binary 1s, the greater the decimal number your subnet mask will become.

The invisible line, or rather its imagined position within the two consecutive rows making up the 32-bit subnet mask, can be written as an equivalent of the dotted-decimal notation of a subnet mask (RFC 4632).

This CIDR notation, or prefix notation, is commonly called the slash-notation, because the slash (“/”) character is utilized.

The term operations in mathematics refers to rules which define how one number combines with other numbers. Decimal number operations include addition, subtraction, multiplication, and division.

There are related, but different, operations for working with binary numbers. The basic Boolean operations are AND, OR, and NOT.

The lowest numbered address in an IP network is the network address. The network number plus 0 in the entire host field.

This also applies to a subnet. The lowest numbered address is the address of the subnet, but it is not always 0.

In order to route a data packet, a router must first determine the destination network/subnet address by performing a logical AND. This is actually multiplication using the destination host IP address and the subnet mask.

The result is the network/subnet address.

If no subnetting is implemented, then the mask is the default mask of the respective address class. This is sometimes also called a natural mask.

A router extracts the IP destination address from an inbound packet. The mask is then retrieved from the routers configuration file.

The router performs a logical AND operation to obtain the network number that the IP destination address of the inbound packet belongs to. During the logical AND operation, the host portion of the destination address is multiplied by the 0s in the mask and ignored.

In this case, routing decisions are based on the network number only.

Each time you borrow one bit from a host field, there is one fewer bit remaining in the host field which can be used for numbering hosts.

Each time you borrow a bit from the host field, the number of addresses which can be assigned to hosts decreases by a power of two. In other words, the number of addresses gets cut in half.

In this example, with 8 bits of subnetting, the new network (subnet) number is 172.16.2.0.

Consider a Class C network address. If no subnetting is implemented, all 8 bits in the last octet are used for the host field. Therefore, there are 256 possible addresses available, 254 of which can be assigned to hosts after subtracting 2 for the network address and the broadcast address.

Now, imagine that this Class C network is divided into subnets. If you borrow 2 bits from the default 8-bit host field, the host field decreases in size to 6 bits.

If you write out all of the possible combinations of 0s and 1s that could occur in the remaining 6 bits, you would discover the following result.

In the same Class C network, if you borrow 3 bits, the size of the host field decreases to 5 bits and the total number of hosts which could be assigned in each subnet would be reduced to 30 (32 minus 2).

The number of possible host addresses that can be assigned to a subnet is related to the number of subnets which have been created.

In this example, with 4 bits of subnetting, the number of actual host bits is reduced to 4 bits.

The new network (subnet) number is 192.64.2.160 as the dotted decimal representation, which is incorporating the subnet bits that have been reassigned from the host address field.

In this case, using the remaining 4 bits for assigning host addresses, there would be 16 useable subnets created and there would be only 14 (16 minus 2) usable host addresses.

To calculate actual subnet addresses, complete the process using an existing host address and subnet mask.

Step 1

Take in dotted decimal notation—the last field of the subnet provided.

Step 2

Subtract the total number of binary combinations for an 8-bit byte.

Step 3

Take the network ID portion of the address and starting with the value of 0, add multiples of 64 to find each applicable subnet address for that network ID and mask combination:

1st subnet: 192.168.2.0

2nd subnet: 192.168.2.64

3rd subnet: 192.168.2.128

4th subnet: 192.168.2.192

Step 4

To find the range of hosts, take the subnet ID and add 1. Take the number of the next subsequent subnet address and subtract 2—that will give you the range of assignable host addresses for each subnet. The broadcast address (BA) will be 1 value higher than the last assignable host address (191).

1st subnet: 192.168.2.0

2nd subnet: 192.168.2.64

3rd subnet: 192.168.2.128 (first host: 192.168.2.129, last host: 192.168.2.190, BA: 192.168.2.191)

4th subnet: 192.168.2.192

In this example, only the first 8 high order bits of the host address space are being used to indicate the subnet.

The lowest numbered address in an IP network is the network address, where the host field is all 0s.

Where subnetting is implemented, the lowest numbered address is the address of the subnet.

To send a data packet, the client must first determine the destination network/subnet address.

To accomplish this, the client performs a logical AND operation, multiplying the destination host's IP address by the client’s own subnet mask value.

If the destination network is the same as its own network, the client sends a request for the MAC address of the destination host. This is done using the local subnet that the address belongs to.

If the destination network is not the same as its own network, the client sends the packet to its default gateway.

Imagine a Class B network with the network number 172.16.2.0.

After assessing the needs of the network, you decide to borrow 8 bits in order to create subnets.

When 8 bits are borrowed from a Class B network, the subnet mask is 255.255.255.0 or /24 [slash 24].

Someone outside the network sends data to the IP address 172.16.2.120.

In order to determine where to deliver the data, the router looks in its routing table to find an exit interface for the destination network of the packet.

If there are several routers in the path to the destination, each router makes a similar routing decision, routing the packet hop-by-hop on the basis of the network or subnetwork portion of the destination address.

Only the final router notices that the packet should be delivered to host 120 in that subnet. This router sends an ARP request to determine the host's MAC address for the final forwarding stage.

Now, imagine that you have the same network: 172.16.0.0 and you have the same IP address.

This time, you decide to borrow only 7 bits for the subnet field.

The address along with the binary subnet mask for this would be the following.

Again, someone outside the network sends data to host 172.16.2.120.

In order to determine where to send the data, the router looks in its routing table to find an exit interface for the destination network of the packet.

So, what is different compared to the previous example? Everything may appear to be the same.

The difference is in the number of subnets available and the number of hosts in each subnet. This can be seen by comparing the two different subnet masks.

With only 7 bits in the subnet field, there are now only 128 subnets available. The lowest subnet number in this case is subnet 0. The next lowest subnet number will be 2, then 4, 6, 8, 10,12, and so on. This is because the lowest order bit in that 8-bit byte is part of the host address field.

With 9 bits for host numbers, there can be 510 (512 - 2) hosts in each subnet.

This is twice as many hosts as when 8 bits were used for subnetting.

Although the addresses themselves may look the same, the interpretation of the address is quite different based on the value of the mask.

One of the decisions to be made when creating subnets is to determine the optimal number of subnets and hosts.

The number of subnets required determines the number of hosts available.

Here is an example.

The first and last addresses within each subnet cannot be assigned to hosts—one is the network address of the subnet, and the other is the subnet broadcast address.

When you create subnets, you lose quite a few potential addresses. For this reason, network administrators must pay close attention to the percentage of addresses that they lose by creating subnets.

Take a moment to explore an example.

Enter the subnet mask in dotted decimal notation for each IP address.

Identify the address class for the network ID (either Class A, B, or C).

Identify the subnet ID address for the address and mask combination.

Enter the broadcast address for each subnet.

You've completed the lesson. You can now achieve these objectives.

Routing Basics Lesson

Welcome to the *Routing Basics* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

In networking, there are two addressing schemes. The first uses the MAC address and a data link (Layer 2) address. The second uses an address at the Network layer (Layer 3) of the OSI model. An example of a Layer 3 address is an IP address.

A router is a type of internetworking device which passes data packets between networks, based on Layer 3 addresses. A router makes intelligent decisions regarding the best path for delivery of data on the network.

A router's attachment to a network is called an interface, also referred to as a port.

In IP routing, each interface must have a separate and unique network (or subnet) address.

Routers connect two or more networks, each of which must have a unique network number in order for routing to be successful. The unique network number is part of the IP address assigned to each device attached to the network.

Consider the example on screen.

A network has a unique network number which is A. It has four devices attached to it. The IP addresses of the devices are A2, A3, A4, and A5. Since the interface where the router connects to a network is part of that network, the interface where the router connects to network A has an IP address of A1.

Another network with a unique network number is B. It has four devices attached to it.

This network is also attached to the same router, but on a different interface.

The IP addresses of the devices on this second network are B2, B3, B4, and B5. The IP address of the router's second interface is B1.

You want to send data from one network to another. The source network is A; the destination network is B; and a router is connected to networks A, B, C, and D.

When a data frame, coming from network A, reaches the router, the router performs the following functions.

Pop-up 1 text

The router strips off the data link (OSI Layer 2) header, carried by the frame. The Layer 2 header contains the MAC addresses of the source and destination.

Pop-up 2 text

The router examines the network layer (OSI Layer 3) addresses and compares them both with the subnet mask on the router's inbound interface. If the source address and destination address in the inbound IP packet are in the same subnet, then the router drops the packet. If the two addresses are in the same subnet, the destination station must already have received the packet at Layer 2 based on MAC addresses, and no routing is necessary.

Pop-up 3 text

If the source address and destination address in the inbound IP packet are not in the same subnet, then the router will attempt to route the packet, assuming it finds an entry in the routing table for the destination address.

Pop-up 4 text

The router consults its routing table to determine which of its interfaces leads to the destination network. If there is no entry in the table for the packet’s destination address, the router drops the packet.

In this example, the router determines from its interface with address B1 that it should send the data from network A to network B.

Before actually sending the data out from interface B1, the router encapsulates the data in the data link frame appropriate for the outbound interface.

IP is a Network layer protocol, and because of that, it can be routed over an internetwork, which is a network of networks. Protocols which provide support for the Network layer are called routed or routable protocols.

Protocols such as IPv4, IPv6, and legacy Internetwork Packet Exchange/Sequenced Packet Exchange known as IPX/SPX (Novell) and AppleTalk provide Layer 3 support and are, therefore, routable.

There are legacy protocols which do not support Layer 3; these are classed as non-routable protocols. The most common of these non-routable protocols was NetBIOS Extended User Interface (NetBEUI). NetBEUI was a small, fast, and efficient protocol, limited to running on one segment. It is unlikely that any machine currently runs NetBEUI.

In order for a protocol to be routable, it must provide the ability to assign a network number, as well as a host number to each individual device.

Some protocols, such as IPX, only require that you assign a network number, because these protocols use a host's MAC address for the host number.

Other protocols, such as IP, require a network address which includes a host, plus a subnet mask. The network address is obtained by performing a logical AND operation on the IP address with the subnet mask.

Routing protocols determine the paths routed protocols follow to their destinations.

Here are a few examples of routing protocols.

Routing protocols enable routers to create a map of paths through other routers in the network or on the internet. This allows routing—in other words, selecting the right path—to occur.

Routers are capable of concurrently supporting multiple independent routing protocols and of maintaining routing tables for several routed protocols. This capability allows a router to deliver packets from several routed protocols over the same data links.

Path determination occurs at OSI Layer 3, also called the Network layer. It enables a router to evaluate the available paths to a destination and to establish the preferred handling of a packet. Routing services use network topology information when evaluating network paths. Path determination is the process a router uses to determine the next hop in the path for the packet to travel to its destination. This process is also called routing a packet.

Path determination for a packet can be compared to a person driving a car from one side of a city to the other. The driver has a map which shows the streets between the current location and the destination. The drive from one intersection to another is a hop. Similarly, a router uses a kind of map which shows the available paths to a destination.

Routers can also make their decisions based on the traffic density and the speed of a link (bandwidth), just as a driver may choose a faster path (a highway) or use less crowded back streets.

The basis upon which a router makes its decision how to forward packets is the contents of its routing table. A routing table usually lists the following items.

Click each tab to learn more.

Destination / next hop mappings

These mappings tell a router that a particular destination is either directly connected to this router or can be reached by another router—this is the “next hop” router on the way to the final destination. When a router receives an incoming packet, it checks the destination IP address and attempts to match this address with either a directly connected network or a next hop router. The routing table also indicates which outbound interface on the router itself leads either to the directly connected network or to the next hop on the other end of the cable.

Routing metric

Routing metric is the measure of how far away a destination network is and is used to determine the most favorable route—different routing protocols use different routing metrics. For example, RIP uses hop count as a routing metric. A hop represents an intermediate router a packet must go through before reaching the destination. A route having a lower total hop count is more favorable than another route with a higher total hop count. The lower the hop count, the fewer intermediate routers network traffic has to pass.

Routing updates

Routers communicate with one another and maintain their routing tables through the exchange of routing update messages between one another—these are also known as routing updates. Depending on the routing protocol, routing updates can be sent periodically or only when there is a change in the network topology. Information contained in the routing update includes the destination networks that the router can reach and the routing metric to reach each of these destinations. By analyzing the routing updates from the neighboring routers, a router builds and maintains its own routing table.

If routers can learn routing information automatically, it might seem pointless to manually enter information into a router's routing table. However, such manual entries can be useful whenever a network administrator wants to control which path a router selects.

For example, routing tables based on static information could be used to test a particular link in the network, or to conserve wide area bandwidth, or to avoid unsecured paths.

Static routing is also the preferred method for maintaining routing tables when there is only one path to a destination network—this type of network is referred to as a stub network. There is only one way to get to this network, so it is pointless for routers to exchange update information seeking alternative paths in case the primary path should fail.

Adaptive, or dynamic, routing occurs when routers send periodic routing update messages to each other. Each time a router receives a message containing new information, it calculates a new best route, and sends the new updated information to other routers. By using dynamic routing, routers can adjust to changing network conditions.

Before the advent of dynamic updating of routing tables, most vendors had to maintain routing tables for their clients. This meant that vendors had to manually enter data such as network numbers, associated metrics or distances, and port numbers into the routing tables of all the equipment they sold or leased. As networks grew larger, this became increasingly cumbersome, time-consuming, and expensive.

Dynamic routing eliminates the need for network administrators or vendors to manually enter information into routing tables. It works best when bandwidth and large amounts of network traffic are not issues. Without dynamic routing protocols, the internet would not function at all.

You've completed the lesson. You can now achieve these objectives.

Networking Devices Module

Media Access or Link Layer Lesson

Welcome to the *Media Access or Link Layer* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

All media attenuate the signals they transmit.

As a result, each type of medium has a maximum range in which it can reliably transmit data.

A repeater is a network device that regenerates a signal from one port onto the other ports to which it is connected.

The purpose of a repeater is to extend the maximum range of the network cabling.

A repeater operates at the OSI Physical layer. A repeater does not filter or interpret. It merely regenerates a signal, passing all network traffic in all directions.

A repeater does not require any addressing information from the data frame, because it merely regenerates the physical signal representing binary data.

This means that even if data is corrupt, it will be repeated. Repeaters will even repeat random binary caused by a malfunctioning network adapter.

The advantage of repeaters is that they are simple and inexpensive.

While repeaters cannot connect networks with dissimilar data frames—for example, a DSL network and an Ethernet network—some repeaters can connect segments with similar frame types, but dissimilar cabling.

Some repeaters simply amplify signals. Although this increases the strength of the data signal, it also amplifies any noise on the network. In addition, if the original signal has been distorted in any way, an amplifying repeater cannot clean up the distortion.

Ideally, it would be advantageous if repeaters could be used to extend networks indefinitely, but all network designs limit the size of the network. The most important reason for this limitation is signal propagation. Every network technology specifies a maximum time delay, which limits the period a signal might be in transit.

This is known as propagation delay—the time it takes for a signal to reach the farthest point on the network.

Hubs, also called wiring concentrators, provide a central attachment point for network cabling.

Hubs can be classified into three types: passive, active, and intelligent.

Ethernet specifications allow the use of multiple repeaters or hubs—however, a data path between any two stations may have no more than four repeaters separating them, due to propagation delay limitations.

One of the major disadvantages of using hubs is that they force half duplex mode on all stations. One consequence of half duplex mode is that if two stations transmit simultaneously, then a collision occurs on the line or hub. Because hubs duplicate all signals, they also duplicate collision fragments, thereby flooding the network with rubbish.

Hubs and repeaters have become obsolete in corporate and industrial networks, and have largely been replaced by switches.

One of the significant disadvantages of using repeaters and hubs to extend the network is that it also extends the segment of the network that stations are sharing. Sharing the network in this way increases the likelihood of collisions, as well as the amount of broadcast traffic that the attached stations are required to process.

Collision domain

The term "collision domain" applies to network segments with shared media, typically using hubs or repeaters. All stations connected to one or more hubs or repeaters are said to occupy a collision domain, because they must all share the use of the physical media. Because of the shared media environment, multiple devices may attempt to transmit at the same time and their signals may collide. Hubs and repeaters propagate all signals, including collision fragments. Collision domains are no longer applicable in switched networks running in full duplex mode because no collisions can occur on a non-shared, full-duplex line.

Broadcast domain

Broadcast domain applies to network segments with hubs, repeaters, wireless LAN access points, bridges, or switches, because all of these devices propagate broadcast frames. Broadcast frames are addressed to every station on the local network, requiring the data to be processed by all stations. Broadcast domains are used in many network protocols.

You've completed the lesson. You can now achieve these objectives.

Layer 2 Lesson

Welcome to the *Layer 2* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

Bridges were designed to overcome the limitations of extending networks by means of hubs and repeaters.

Bridges are more intelligent and flexible than hubs and repeaters. Bridges operate at the MAC sublayer of the OSI Data Link layer.

A repeater physically forwards signals out of a port, other than the inbound port where the signals were received; whereas, a bridge uses addressing information and programming logic to forward and filter data to the port where the receiving station is attached.

A bridge determines the location of the port, based on the MAC address of the target device.

The forwarding process works as follows:

Frames on LAN A that are addressed to devices on LAN A, are not forwarded by the bridge. The same is true for frames from LAN B addressed to devices on LAN B. Frames are kept local on the LAN where they originated, unless the frames are addressed to devices on another LAN.

The frames from stations on the local LAN will have reached their destination, without the help of the bridge.

Frames from LAN A addressed to devices on LAN B, are forwarded to LAN B for delivery. Similarly, frames from LAN B are forwarded to LAN A.

On very old bridges, the network administrator had to manually configure the address tables. This proved completely impractical for large deployments.

Newer bridges are called "learning bridges". Learning bridges automatically update their address tables as devices are added or removed from the network. Learning bridges "learn" by memorizing the source addresses of frames that they forward.

A switch is essentially a bridge with multiple ports and more functions and features.

Like a bridge, a switch learns the MAC addresses of end stations, and makes forwarding decisions based on this information.

Today, hubs, repeaters, and bridges have mostly been replaced by switches, in corporate and industrial networks.

Bridged or switched networks have the following characteristics:

You've completed the lesson. You can now achieve these objectives.

Layer 3 Lesson

Welcome to the *Layer 3* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

Routers function at the Network layer of the OSI model. In addition to forwarding and path determination, routers perform additional functions. Here are a few examples:

Unlike bridges and switches, routers do not forward broadcasts. Each port on the router terminates a broadcast domain. The segments attached to router ports are referred to as networks.

Not only is each network in its own broadcast domain, but each network has its own logical network address.

In addition, each station, on each network, has a logical address. The logical address indicates which network the end station is on, and provides the end station with an individual designation on this network.

Routers make forwarding decisions based on network addresses, rather than MAC addresses.

Routers also have other functionalities.

In theory, each port on a Layer 2 or Layer 3 networking device terminates a physical LAN segment. Hubs, being Layer 1 devices, are considered part of the physical LAN segment. Hubs always have a single collision domain and a single broadcast domain, since they represent a single physical LAN. Layer 2 switches always have one collision domain per port and one broadcast domain for the switch. The same is true for bridges. Routers have one collision and one broadcast domain per port. This router has four interconnecting ports.

In practice, collision domains have predominantly been eliminated from corporate and industrial networks, and have been replaced by switches operating in full-duplex mode. A switch port in full-duplex mode terminates a non-collision domain.

You've completed the lesson. You can now achieve these objectives.

Multi-Layer Switches Lesson

Welcome to the *Multi-Layer Switches* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

As the network grows, Layer 2 switches are not enough.

Layer 2 LAN switches are fast, but they create large, flat networks. There is no filtering by subnetwork, and they offer very primitive security mechanisms.

A flow is a specific conversation, consisting of many packets, between a source on one network and a destination on another network.

In general, each packet of a traditional flow must be processed by the router. A router routes all packets belonging to the same flow by comparing the destination address in the IP header with its routing table.

Hybrid devices combine some of the features of both switches and routers. These are called Layer 3 switches.

A Layer 3 switch can forward packets from one VLAN or subnet, to another VLAN or subnet. A Layer 3 switch incorporates the functions of Layer 2 switching, as well as routing.

A traditional Layer 3 router can forward packets from an Ethernet port to a non-Ethernet port, as well as being able to route IP and non-IP traffic. Traditional routers perform their functions in software; whereas, Layer 3 switches perform their functions in firmware, have only Ethernet ports, and support only IP forwarding. This makes them faster than traditional routers.

Some Layer 3 switches improve their performance by routing only the first packet of a flow, and then switching subsequent packets in the same flow. This is done by associating the next hop in the path with the forwarding port on the switch, assuming a switched path exists.

Routers transport packets through a complex network.

Traditional routers make forwarding decisions by examining a packet’s IP header and using the address information found in the Layer 3 portion of the header.

Traditional routers ignore higher layer headers in the packet, because this information is not relevant to the basic function of the packet.

Ignoring this higher-level information makes sense if the router is only attempting to find a transportation route for the packet through the network. However, the router treats all packets and all packet flows equally.

This is not always desirable though, because some packets or packet flows are more important than others, and should be prioritized.

Determining which packets or packet flows should be given priority requires examining higher layer headers, usually those of the transport layer protocol’s transmission control protocol (TCP) or user datagram protocol (UDP).

While routers can be configured to perform this activity, it is not particularly efficient for routers to do this, because router functions are processed in software.

Layer 4 switches solve this problem.

The objective of Layer 4 switching is to analyze the information available in Layer 4 headers about the packet payload, and then make prioritization decisions based on that information.

Switches perform Layer 4 prioritization decision functions in hardware, rather than in software, like routers do.

Because hardware processing is faster than software processing, Layer 4 switches perform this function more efficiently than routers.

Layer 4 switches prioritize different kinds of traffic and forward packets or flows based on priorities, rather than first in first out. This prioritization is known as quality of service.

You've completed the lesson. You can now achieve these objectives.

TCP/IP Suite Module

TCP/IP Overview Lesson

Welcome to the *TCP/IP Overview* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

A spoken language is governed by rules of structure and syntax. These rules exist to facilitate communication between speakers.

A communication protocol does the same in networking. Protocols create a set of rules that enable different systems to understand and communicate with each other.

Transmission Control Protocol/Internet Protocol (TCP/IP) are standards that allow systems all over the world to communicate, no matter who manufactures them.

The major difference between TCP/IP and spoken languages is that TCP/IP is a single, common protocol used worldwide. For example, a PC in North America can communicate with a server in Asia, using TCP/IP.

TCP/IP is the most common networking standard in the world. The TCP/IP protocol suite gets its name from TCP at the OSI Transport layer and IP at the OSI Network layer.

TCP is the protocol that supports most common applications, such as web browsing, email, file transfer, and more. IP is the protocol that allows traffic to be routed anywhere in the world.

There are a number of key differences between the Open Systems Interconnection (OSI) reference model and the TCP/IP stack.

Take a moment to explore the functions of each layer within the TCP/IP protocol stack.

The Internet Protocol (IP) is the world’s most commonly used protocol, running on virtually every consumer device on the earth.

IP is a communication protocol that can run virtually, over any Network Access or Link Layer protocol.

Protocols used at the Network Access layer include Ethernet LAN, point to point, or PPP WAN protocol, as well as other older protocols, such as High-level Data Link Control (HDLC) and Asynchronous Transfer Mode (ATM) protocols.

IP is the most popular communication protocol and hierarchical addressing scheme used today. When referring to IP, the commonly used term is IPv4, which stands for IP, version 4. IPv4 will eventually be replaced by IPv6—IP, version 6.

IPv4 was originally designed as a hierarchical addressing scheme that allowed for the assignment of over 3.5 billion addresses. IPv6 has a much larger address space than IPv4, and addressing has been expanded to meet the needs of the Internet of Things, also known as IoT.

As Smart networked devices become increasingly commonplace, there is now a need for trillions of new IP addresses. IPv6 will allow for 340 trillion, trillion, trillion IP addresses.

The Transport layer provides the end-to-end transport services that support Application layer services.

There are two transport protocols associated with the TCP/IP protocol suite that are the focused on in this course. These are TCP and UDP.

TCP provides acknowledged end-to-end data delivery control using sequence numbers and acknowledgments. UDP does not provide acknowledged end-to-end data delivery control.

The Application layer provides the network services that support the many user applications available to the world population.

For examples Simple Mail Transport Protocol (SMTP) is the standard that allows Google Mail to send mail to a Microsoft Exchange server.

Secure Shell (SSH) is the protocol that supports a PuTTY client connection to a Fortinet firewall.

You've completed the lesson. You can now achieve these objectives.

Transport Layer Lesson

Welcome to the *Transport Layer* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

The TCP/IP protocol stack has two very different protocols that are used at the Transport layer. Though very different in their operation, both Transmission Control Protocol (TCP) and User Datagram Protocol (UDP) are designed to fulfill the data delivery needs of the applications they support.

The main functions of the Transport layer is to provide host-to-host delivery of data. The Transport layer also prepares the application data for transport inside of an IP packet.

For example, an application may generate massive amounts of data. The Transport layer breaks up that large amount of data into smaller, more manageable segments, and hands those segments down to the IP protocol to encapsulate into packets.

The Transport layer also manages the separation of all the data coming from the many applications running on the same system.

While providing for host-to-host delivery of data across the network, the Transport layer must handle data from many different application processes on the same host system.

This handling, also known as multiplexing, of the application data is managed by the generation and tracking of port or application numbers. This multiplexing capability allows the Transport layer to provide support for many different applications at the same time.

Common internet applications use well-known or standard destination port numbers.

The sending host system generates a random source port number to represent the client side of the application data exchange. The server side of the exchange uses standard port numbers.

It is the combination of the sending and receiving host port numbers that allows the end systems to handle many sessions of the same application on the same system.

To help clarify this concept, consider a web browser with multiple tabs open at the same time. Each of the tabs represents an exchange of data between the local browser and the distant end system with which it is communicating.

When a connection is being established, the following occurs:

The sender sends a synchronization request to a well-known application port number. The request contains a starting sequence number which is generated locally.

The receiver replies with an acknowledgment to the sender’s synchronization request and a synchronization request that contains the receiver’s own, locally generated, starting sequence number.

The sender then acknowledges the receiver’s synchronization request. This back and forth exchange of synchronization requests and starting sequence numbers is referred to as a three-way handshake.

At this stage, the two endpoints have an established connection. The endpoints each have their own sequence numbers for transmission and acknowledgments for reliability.

Data can now be transferred bidirectionally with reliability, and be retransmitted, if required.

As a connection-oriented Transport layer service, TCP ensures end-to-end data reliability and first establishes a connection with the destination host, before application data is sent.

The TCP header is prepended to the segment of application data by the Transport layer.

In the case of UDP, data is sent without prior verification of whether the receiver is present or can accept the data.

No acknowledgment or retransmissions are included in the exchange.

UDP is an example of a connectionless communication protocol. This is because UDP doesn't establish logical connectivity before transmitting the data.

There are many applications which run over UDP and do not require reliable delivery of their data. Each of these applications has corresponding UDP port numbers for the source of the data and the destination.

These port numbers are carried in the UDP header prepended to the application data handed down for transport over the network. The only other information in the UDP header is an indication of the data segment length, and a checksum to verify if the header information is correct. Notice, there is no information in the UDP header for ensuring reliable delivery of the data.

Take a moment to explore an example.

Protocols that make use of UDP for delivery include the following: TFTP, SNMP, network file system (NFS), as well as client requests and server replies for DNS.

Take a moment to explore an example.

The destination port number is defined by the application that sent data down to the Transport layer.

In this example, the application was Telnet. The Transport layer sets the port number to 23, indicating it is sending Telnet data to the destination.

The source port number is a random number ranging from 1024 to 65535. This number is used to uniquely identify a particular session. The destination will respond to this session by reversing the destination and source port numbers.

You've completed the lesson. You can now achieve these objectives.

Establishing a TCP Connection and Sending Traffic Lesson

Welcome to the *Establishing a TCP Connection and Sending Traffic* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

As part of the Transport layer of the TCP/IP stack, TCP is a Transport Layer protocol as mapped to the OSI model. Therefore, TCP is responsible for accepting data from upper-layer applications and preparing that data to be sent across the network to a distant host.

As a connection-oriented protocol, TCP is most commonly used to support applications that transmit large amounts of data and require reliable delivery and error recovery for that traffic.

To accomplish this task, TCP establishes a logical connection to the destination host before beginning to send the actual data to the network for forwarding. In this way, TCP can ensure that both hosts involved are connected, available, and ready to begin the exchange.

As part of its functionality as a connection-oriented protocol, also referred to as a service, TCP ensures that the transmitting host knows that the data has been successfully received, and that any missing or errored data has been recovered and put into proper order for processing by the application.

TCP also provides mechanisms to avoid congestion in the network and to prevent the receiving host from being overrun by data.

As a connection-oriented protocol, TCP is required to establish contact with the remote host and verify both its connectivity as well as ability to exchange traffic before any actual segments of data can be sent.

To establish this exchange with the remote host, TCP uses a three-way handshake approach to exchange connection establishment and data transmission variables.

In this example, an application on Host A needs to send data to Host B using a TCP connection. The data is passed down to the TCP protocol, where it is divided into small segments which are sized to fit into a packet.

Then, to start the data exchange with the destination host, TCP begins the handshake process.

In Step 1, TCP sends a synchronization request message, or SYN request message, with the control field SYN bit set to 1, and a starting sequence number in the TCP header. In this example, the starting sequence number that is generated is 10.

Host B receives the SYN request, and starts a session socket to the application or protocol indicated by the TCP destination port number.

At the receiving end of the exchange, the destination host, Host B, must check if the application or protocol process can be started or is already in place. Host B also must make sure that it is able to accept an incoming connection. If the remote host is ready and able to accept the connection, the three-way handshake continues with the next step.

In Step 2, Host B sends a SYN and an acknowledgment (ACK) message back to Host A.

In the message from Host B, both the control SYN and ACK bits are set to 1, indicating that Host B is accepting the incoming connection SYN request from Host A, as well as sending its own connection SYN request to Host A.

Host B generates and sends its starting sequence number, which in this case is sequence number 100. And acknowledges the receipt of data from Host A.

Host B acknowledges that it has received the starting sequence number generated by Host A, sequence number 10. Host B does this by sending an acknowledgment and incremented value for the sequence number submitted by Host A, in this case a sequence number value of 11.

This incremented value expresses the next sequence number that Host B is expecting to see from Host A. This process is called expectational acknowledgment.

At this point in the exchange, both hosts are actively communicating and have exchanged their starting sequence numbers. So, the final exchange of the three-way handshake is for Host A to acknowledge the receipt of the connection request and the starting sequence number, sent from Host B.

In Step 3, Host A acknowledges receipt of the SYN ACK from Host B by sending a TCP header with the control field ACK bit set to 1.

Host A also sends a sequence number of 11 and an expectational acknowledgment number of 101. This represents an incremented value from the starting sequence number originated by Host B.

At this stage, both Host A and Host B have established an end-to-end TCP session, and are ready to transmit data in both directions.

During the exchange, both hosts will indicate the successful transmission and receipt of data, with the exchange of ever increasing sequence numbers and acknowledgment numbers.

TCP protocol is a primary example of the connection establishment and data exchange behaviors of a connection-oriented protocol.

When you look closely at the exchange behavior between connection partners as they transmit data, you can see that each host is aware of the sequence number in the segment that it is sending.

Each host will look for an acknowledgment from the other host, to indicate the successful receipt and processing of that segment.

As each host sends data to the other host, it embeds the acknowledgment of all successfully received and sequentially processed segments into the TCP header of the segment that it is currently sending.

The host does this by sending an acknowledgment sequence number that indicates the number of the next sequential segment in the exchange that is expected from the other host.

As each host continues to send traffic during the exchange, each side indicates the number of the current segment it is sending, and continuously acknowledges the receipt of traffic from the other host.

Each end will only acknowledge the segments that are received in proper sequence and can be successfully sent up to the application for processing.

If a segment goes missing or arrives out of sequence, the receiving host will not send back an acknowledgment, until the missing segment arrives and can be put back into proper order for the application to process.

If the transmitting host does not receive a timely acknowledgment for all of the segments it has sent, then it will begin retransmitting from the indicated point of the last successfully received acknowledgment.

If the receiving host successfully receives the missing segment and can process it in proper sequence, the receiving host will send back an acknowledgment to indicate the next segment it is expecting, and the exchange continues from there.

This process of error recovery ensures that the exchange can recover from the possibility of lost segments and that all of the sent traffic is accounted for.

Sometimes, a receiving host will start to run short of resources to process the amount of traffic being sent. This often occurs if the host is a server servicing many clients.

Rather than allowing itself to become overwhelmed, the receiving host may signal back to the transmitting host instructing it to pause before sending additional data segments.

Once processing resources are available again, the receiving host will signal back to the transmitting host, instructing it to resume transmitting.

This practice is called flow control and it is common in connection-oriented exchanges.

Rather than halt the exchange entirely, the receiving host may implement flow control by indicating the amount of traffic it can accept for processing at any given point in the exchange. The transmitting host will use this information to determine how much traffic to send before waiting for the acknowledgment of segments, before it continues.

This practice is called windowing and it is another form of flow control.

Windowing determines how many segments can be in transit across the network at one time, when awaiting acknowledgment.

To end the exchange, one host signals to the other host that it has finished sending all of the traffic that it had to send and is ready to shut down the connection.

The other host may continue sending traffic or it may also be finished sending traffic for the exchange.

In order to shut down the established connection, a closing handshake, that is very similar to the three-way handshake that opened the exchange, takes place.

Host A may start the closing handshake by sending a TCP header to Host B with the control field FIN, or Finished, bit set to the value of 1.

Host A will send its final TCP segment with a sequence number of 14 and an expectational acknowledgment number, which in this case is 300—an incremented value from the sequence number last received from Host B.

At this stage, Host A has indicated that it has no more traffic to send.

Host B will acknowledge this by sending an acknowledgment of the final segment from Host A and an acknowledgment of the control field FIN bit that was sent.

If Host B is also finished with the exchange, it will send a FIN bit in the control field of the same TCP header it is sending to Host A.

To complete the process of shutting down the connection, Host A will send a TCP header with a sequence number incremented to indicate the receipt of the last segment sent from Host B. Host A will also set the control field ACK bit to 1, to indicate receipt of the FIN bit sent from Host B.

At this stage, both sides of the exchange have neatly closed the connection after verifying that all transmitted data has been successfully received by the remote host.

TCP provides a primary example of the connection establishment, data exchange, and connection closing behaviors of a connection-oriented protocol.

You've completed the lesson. You can now achieve these objectives.

Design Issues Module

Extension and Flexibility Lesson

Welcome to the *Extension and Flexibility* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

When working with shared LAN environments, there are some network architecture design issues that you need to consider carefully. The first issue that you need to consider is the nature of the LAN itself.

A large shared LAN has several characteristics that require careful planning and mitigation.

For example, in a shared LAN, all the devices are connected to a single broadcast domain. So, when a workstation sends out a broadcast, such as an ARP request, every device in that broadcast domain is interrupted. If a broadcast storm occurs, the network performance of all the devices degrades.

Furthermore, because all the devices are in a single broadcast domain, all the devices are also in a single IP network. This makes it very difficult for you to apply any security between the different organizations within this network.

As a result, it is very difficult for you to make any changes to the network.

VLANs offer network segmentation, security, and flexibility. They resolve some of the issues associated with very large LANs.

VLANs are referred to as virtual because the relationship between the stations connecting to the network is defined in the configuration of the switch to which the stations are connected.

VLANs are logical definitions of LANs as opposed to physical definitions of LANs.

All stations that are connected to the same VLAN can communicate freely with each other, just as if they were physically connected to the same LAN.

However, in order to communicate with stations that belong to a different VLAN defined on the same switch, stations must communicate through a routed connection.

As a result of this separated routed connection, each VLAN is its own broadcast domain. This helps alleviate the issue of congestion due to broadcasts. Also, because each network is a separate entity, there is enhanced security, and it is easier for you to make changes.

Each port on a Layer 2 switch is called an access port. They are given this name because the attached stations obtain access to the network through these ports.

For scalability and ease of connectivity, it is possible to extend the VLAN configuration between interconnected Layer 2 switches in a network. This is accomplished by defining the interconnecting links between the switches as trunk ports.

Trunk ports carry the traffic for multiple VLANs across the trunk link between the switches that those VLANs define.

Consider the following example. The VLANs numbered VLAN 10, VLAN 20, and VLAN 30 are associated with the configuration of ports on all three switches. This association allows stations connected to any switch to communicate freely across their respective VLANS, regardless of the switch they are physically attached to. This configuration is the difference between a logical network association and a physical network association.

VLAN data is sent from one switch to another. To differentiate the data from each VLAN, the switches insert a tag into each frame before the frame is forwarded across the interconnecting link. Both switches tag the data on either side, independently. Each switch has to reference the tags in the incoming data to distribute the data to the correct VLAN associated ports on that switch.

This tag includes the unique VLAN number of the source VLAN.

Here is a closer look at an Ethernet frame that has been tagged with the unique VLAN ID of the VLAN where the data originated from.

The VLAN tag is inserted between the Ethernet header and the header that is on the data inside the Ethernet frame. The header on the data is usually an IP header, since most data today is carried in IP packets.

The format of this VLAN tag field includes some bits that indicate the unique VLAN ID, as well as bits indicating the priority handling of the frame across the interconnected Layer 2 network.

This tag also includes a bit that indicates if the tag is inside of an Ethernet frame or a token ring type frame. Token ring is rarely used today, but was quite common when the VLAN tag protocol was first developed.

The tag itself is also referred to as a "dot 1 Q", tag in accordance with the Institute of Electrical and Electronics Engineers (IEEE) protocol standard that defines the format of the tag and how it is implemented.

Given that a large amount of traffic may cross the interconnecting links between switches, it is necessary to ensure that the links have adequate bandwidth, as well as redundancy.

Redundancy ensures that if a link between the switches fails, there is a backup link that can be used instead to forward traffic.

While it may seem straightforward to simply add additional physical links between the switches, doing this can create difficulties because of the way Layer 2 switches work.

Layer 2 switches listen, learn, flood, and forward traffic. Adding another physical connection can cause the traffic to get caught in a forwarding loop across the redundant links.

To prevent this forwarding loop from happening, the IEEE invented a protocol called Spanning Tree.

The Spanning Tree protocol is implemented on the interconnected switches, and selects one of the links to use as the active forwarding link. The other link then becomes a backup link that can be used if the active link fails.

While this improves redundancy, it doesn’t improve bandwidth, because only one link is used at a time.

To improve the bandwidth available between interconnected switches, another method was created to allow the use of multiple physical connections at the same time.

This method is referred to as link aggregation.

To use link aggregation, you must configure the redundant physical connections between the switches on either side as belonging to a bundle known as a link aggregation group (LAG).

Once bundled together, the port connections are treated as if they are a single cable interconnecting the switches on both sides. Then, protocols like Spanning Tree treat the entire LAG as a single active forwarding link.

Link aggregation uses an algorithm to select the appropriate physical link and distribute the traffic across each of the links as evenly as possible.

You've completed the lesson. You can now achieve these objectives.

Internetworking Fundamentals Module

WAN Service Types Lesson

Welcome to the *WAN Service Types* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

The wide area network (WAN) plays a key role in the interconnection of enterprise networks.

WANs provide connectivity around the globe. This could be connectivity for private enterprises with global presence. This could also be connectivity for the global network that we use on a daily basis, for example for web browsing and shopping.

WANs can be privately owned or publicly offered as a service by a service provider.

The key difference between public and private WANs is who owns the equipment and infrastructure, and who is responsible for running and maintaining the network.

Private WANs are owned, managed, and maintained by a single organization.

As you might imagine, purchasing and provisioning the infrastructure connectivity and devices is very costly.

However, once installed, these networks are relatively inexpensive to run and manage.

By owning the entirety of the infrastructure, the private corporation can guarantee a high level of service and security because the network is not integrated with any public network.

Public data networks are owned by service providers, also referred to as common carriers.

Common carriers have been providing public data network services for decades.

Services include high-speed and low-speed data interfaces.

These services are typically offered along with service level agreements (SLAs) either for a flat fee, a contract price, or as a subscription service that you pay for as you use.

Common carriers offer many different types of services, including broadband connections, packet switching, broadband virtual private networks, dark fiber, wavelength-division multiplexing, LTE/5G, multiprotocol label switching, metroethernet, xDSL, and broadband cable.

Take a moment to explore a small subset of the services offered by public data network service providers.

There are many different WAN services available to provide global connectivity for large enterprise networks and even for public networks.

Some WANs are wholly owned by a single entity, while others rely on the connectivity services provided by a third party known as a service provider or common carrier.

You've completed the lesson. You can now achieve these objectives.

WAN Service Technologies Lesson

Welcome to the *WAN Service Technologies* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

To select the WAN service most appropriate for a customer's business requirements and implementation, it is important to have a strong grasp of the characteristics of common WAN services, including the benefits of each.

In this lesson, you will examine the following service offerings: broadband connections, packet switching, internet connectivity, VPN services of different types, and cellular service offerings of different generations.

There are many different types of services offered by public data network service providers. Now, you will explore these different service types in more detail, beginning with broadband connections.

All broadband connections provide dedicated communication links. These dedicated links provide various capacities that are limited only by the underlying physical facilities of an enterprise, and their willingness to pay for these dedicated lines.

Broadband connections incorporate permanent connections using point-to-point links.

A high-bandwidth, point-to-point, connection provides a pre-established WAN communication path from the customer premises, through the provider network, to a remote destination.

The links operate at OSI Layer 1 and Layer 2, providing physical cabling connectivity, as well as protocols for data transport across the link.

Point-to-point connections are simple to implement and provide high-quality and reliable capacity.

Connectivity services are offered by service providers, which include a fixed amount of bandwidth at an agreed-upon flat-rate price and service quality. Therefore, point-to-point connections are generally costly and have fixed capacity, which makes them inflexible.

However, they are also very reliable and very secure, which is why point-to-point connections are a good choice for implementations that require a high level of reliability and security.

Broadband provides high speed network connectivity leveraging multiple types of technologies, such as fiber optics, wireless, cable, and satellite. Broadband offers connectivity services that implement data rates of at least 25 Mbps download and 3 Mbps upload speeds, as defined by the Federal Communications Commission (FCC).

Packet switching is another type of service offered by public data network service providers.

Packet switching networks operate at OSI Layer 1 and Layer 2.

In packet switching networks, messages are divided into smaller blocks for efficiency during transport. Blocks are then encapsulated into packets.

Each packet includes source and destination address information, as well as packet handling instructions. Each individual packet is routed through the network independently.

Packet switching networks do not require a dedicated circuit to be established before the transmission of data.

Packet switching networks also allow many pairs of nodes to communicate over the same network pathways.

The cost associated with the service provided by a packet switched network is based upon using a fixed amount of bandwidth within the network.

Packet switching is a good connectivity option for networks that have several locations that need to be interconnected. Because of the fixed bandwidth pricing, packet switching networks are also a cost-effective way of interconnecting multiple sites and local networks.

Take a moment to explore an example.

In packet switching, the packets that make up a message can take different routes through the internetwork.

As a result, a constant bit rate is not guaranteed. A constant amount of delay, or latency, is expected.

Step 1

Packet services, also known as datagram services, treat each packet as an independent message.

Step 2

Each packet is routed through the internetwork independently, and each node along the forwarding path determines the network segment that should be used for the next step in the packet's route.

Step 3

This independent routing and individual handling of packets enables switches and routers to bypass busy segments, and take alternate paths to forward packets through the internetwork.

Step 4

Network-layer protocols are responsible for delivering the packet to the appropriate network for delivery to the endpoint attached to that destination network.

Step 5

Packet switching meets the need to transmit large messages in a fairly small packet size, so the packet can be accommodated by the physical network.

Step 6

The Network layer is responsible for fragmenting message segments from upper protocol layers into smaller packets, if necessary. These messages are resized for optimal transmission across the Physical layer.

Step 7

The Network layer is also is responsible for reconstructing message segments from the packets, as they are received at the endpoint.

The global internet is a network comprised of many interconnected networks. Similar to packet switched networks, the data is forwarded across the network in blocks, referred to as packets. However, the internet is not a WAN service offered by a single service provider; rather, it is a myriad of networks that are interconnected.

Instead of using a separate WAN infrastructure, enterprises today commonly take advantage of the global internet infrastructure for WAN connectivity. Internet service providers offer "last mile" connectivity to the internet—in other words, connectivity between a local network and the rest of the internet.

In the past, due to the lack of adequate performance guarantees, the internet was not a viable option for a WAN connection, resulting in many security risks and lack of service level agreements (SLAs).

Currently, with the development of cheap and secure virtual private network (VPN) technologies, the internet has become one of the most common connection types.

Internet WAN connection links include various broadband access technologies, such as fiber, digital subscriber line (DSL), cable, and broadband wireless.

These technologies are usually combined with a VPN, to provide security. Other access options are cellular networks and satellite systems.

Invoicing is typically based on a fixed bandwidth flat rate.

Each WAN technology provides advantages and disadvantages for the customer.

When choosing an appropriate WAN connection, it is important to consider whether to use public-based internet connections, or connections implemented within non-public service provider networks.

Internet-based connections are readily available, flexible, and a cheaper option that can be made secure using technologies, such as VPNs. VPNs create private connections through the service provider network.

WAN service connections within a service provider’s network guarantee security and performance through an otherwise public network.

A Layer 3 multiprotocol label switched VPN, or MPLS VPN, provides a Layer 3 service across the backbone. A separate IP subnet is used on each customer site.

When you deploy a routing protocol over this VPN, the service provider needs to participate in the exchange of routes. Neighbor adjacency is established between your customer edge, or CE, router and the provider edge, or PE, router owned by the service provider.

Within the service provider network, there are many service provider core routers, known as P routers.

The job of P routers is to provide connectivity between PE routers. What this means is that the service provider becomes the backbone of your customer network.

A Layer 3 MPLS VPN is appropriate for customers who prefer to outsource their routing to a service provider.

The service provider maintains and manages routing for the customer sites.

If you look from the customer perspective, with Layer 3 MPLS VPN, you can imagine the whole service provider network as one big virtual router.

Site-to-site VPNs connect entire networks to each other, such as connecting branch offices, home offices, or business partners networks to the main office network.

Each site has a VPN-capable device, called a VPN gateway.

Routers, firewalls, and other security appliances can act as VPN gateways.

The VPN gateways establish the connection between themselves.

A remote-access VPN connects an individual endpoint, over the internet, to the VPN device at the edge of the remote network.

With the advent of VPNs, enterprises can support remote users by leveraging the internet.

A mobile user simply needs access to the internet to communicate with the main office. For telecommuters, internet connectivity is typically achieved through a broadband, DSL, or cable connection.

Long term evolution (LTE) is an evolution of 3G networks, such as universal mobile telecommunications service (UMTS), and High-Speed Downlink Packet Access (HSPDA).

UMTS is a third-generation broadband, packet-based transmission of text, digitized voice, video, and multimedia at data rates of up to two Mbps.

HSPDA is a packet-based mobile telephony protocol used to increase data capacity and speed up download rates.

LTE development started in 2004, but the first networks were only deployed in 2010. LTE networks were developed to offer improved performance for the newer hand-held devices and cell phones that incorporated data, browsing, and multimedia, as well as voice calling capability.

LTE required a completely new network infrastructure, which needed to operate in coexistence with 3G networks. Because it is optimized for IP-based networks, LTE networks offer much better performance than 3G networks.

An LTE network is sometimes called a 3.9G network.

The fourth generation (4G) of mobile telecommunications technologies is also available.

It is called LTE-advanced and is expected to offer about 300 Mbps.

Like its predecessors, 5G networks are cellular networks in which the service area is divided into small geographical areas called cells.

All 5G wireless devices in a cell are connected to the internet and telephone network by radio waves, through a local antenna in the cell.

The main advantage of the new networks is that they have greater bandwidth and provide higher download speeds. Download speeds are expected to reach up to 20 Gbps, as development continues and additional improvements are added.

Due to increased bandwidth, 5G networks do not exclusively serve mobile devices, as is the case with other existing cellular networks. 4G mobile devices are not able to use the new networks, which require 5G-enabled wireless devices. Therefore, existing 4G networks will remain in place for the foreseeable future to accommodate older devices.

5G is being used as a general internet service provider (ISP) for laptops and desktop computers; offering new possibilities in the internet of things (IoT) and machine-to-machine areas.

Today, 5G is competing with existing ISPs, such as cable internet.

You've completed the lesson. You can now achieve these objectives.

Virtualization and Cloud Module

Virtualization Overview Lesson

Welcome to the *Virtualization Overview* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

Virtualization is the creation of a virtual—rather than an actual—version of something, such as an operating system (OS), a server, a storage device, or network resources.

Virtualization uses software that emulates hardware functionality, to create a virtual system.

Virtualization software partitions hardware resources, such as processing, memory, and networking, onto a host machine. The host machine then allocates these resources so that multiple emulated machines can run on a single hardware platform.

IT organizations can use virtualization to operate multiple operating systems, more than one virtual system, and various applications, all on a single server.

The benefits of virtualization include greater efficiencies and economies of scale.

Virtualization supports the efficient use of computing resources such as CPU, memory, and storage. Virtualized networks have no need for excessive hardware and have reduced energy costs, when compared to non-virtualized environments.

Virtualization can be used in production environments, as well as development environments and test environments—for example, to virtualize certain types of services, such as database services, or servers, such as application servers.

Virtualization results in better disaster recovery mechanisms. It supports the quick and easy management of production environments, as well as the isolation of development and test environments.

A hypervisor, sometimes called a virtual machine monitor (VMM), is a type of computer software that enables the deployment and management of virtual machines.

The hypervisor is responsible for partitioning hardware and system resources from the host system, and for managing the allocation of those resources for use by the virtual machines running on that system.

There are many different kinds of hypervisors, but generally, they can be divided into two main categories: type 1 hypervisors and type 2 hypervisors.

There are many differences in how the two types of hypervisors function, but the key difference is the way that they communicate to the system hardware.

Type 1 hypervisors

Type 1 hypervisors are also known as bare metal hypervisors or native hypervisors.

Type 1 hypervisors run on the computer hardware, interacting directly with its CPU, memory, and physical storage.

Type 1 hypervisors take on the role of the operating system and perform the tasks involved with virtualization, such as resource partitioning and management.

Type 1 hypervisors can partition and manage the resources of the bare metal server; so, they make is possible to implement multiple virtual machines to run on the server hardware.

Type 1 hypervisors are often used in data centers and on high-performance server hardware that is designed to run many virtual machines. Type 1 hypervisors run on dedicated hardware and they require a management console.

Type 1 hypervisor vendors and products include, Oracle VM Server for Sparc, ESXi, Hyper-V, Kernel-based VM (KVM), and others.

Type 2 hypervisors

Type 2, or hosted, hypervisors, run as software supported by the host operating system.

A type 2 hypervisor acts as a virtual machine manager on the host system. It is installed as a software application on the system, and runs over the existing operating system on the host.

Type 2 hypervisors run on conventional operating systems, such as Windows, macOS, or Linux.

Type 2 hypervisors support guest virtual machines by coordinating calls for host resources, such as CPU, memory, disk, network, and other resources, through the physical host's operating system.

Unlike type 1 hypervisors, type 2 hypervisors don’t have to run on dedicated hardware. This makes it easy for an end user to run a virtual machine on a PC by installing the hypervisor software to coordinate the use of resources.

Type 2 hypervisor vendors and products include, VMware Fusion, Oracle Virtual Box, Oracle VM for x86, Solaris Zones, Parallels Desktop, VMware Workstation Pro, and others.

There are many well-known virtualization vendors. Their offerings comprise a variety of commercial products for businesses, as well as free-to-use products for the general public.

Shown here are some common examples. Now, you will explore a high-level overview of each one.

The essential VMware server products for the commercial marketplace and enterprise network environment include, VMware ESXi, VMware vCenter, and VMware vSphere for ESXi environments.

Click each product to learn more.

VMware ESXi

VMware ESXi is an enterprise-class, type-1 hypervisor developed by VMware, for deploying and serving virtual computers. As a type 1 hypervisor, VMware ESXi includes and integrates vital operating system components, such as a kernel, which is the core functionality of an operating system.

VMware ESXi server is essential for businesses that are in need of a virtualized production environment. Like all type 1 hypervisors, ESXi is not a software application, and therefore must be installed at the operating system level.

VMware vCenter

VMware vCenter is a software application that makes the centralized management of multiple ESXi hosts through a single interface possible.

It comes with many important features, such as VMware vSphere Motion, VMware vSphere Distributed Resource Scheduler, VMware vSphere High Availability, VMware vSAN, and more. These features make VMWare vCenter a reliable product for production environments.

VMware vSphere (ESXi)

VMware vSphere is a software suite that includes different VMware products, such as VMware ESXi, VMware vCenter, and an HTML5-based vSphere client.

VMware has also developed a line of desktop and end-user computing products, which include VMware Workstation Pro, VMware Fusion, and VMware Player.

Click each product to learn more.

VMware Workstation Pro

VMware Workstation Pro is a type 2 hypervisor that can be used for the professional creation and management of virtual machines on Windows or Linux operating systems. VMware Workstation Pro offers many support features, which include, but are not limited to, the following.

VMWare Fusion

VMware Fusion acts as the macOS counterpart for the Workstation Pro. It supports similar functions to Workstation Pro, such as connecting to vSphere and managing snapshots that allow building and managing of single or multiple virtual machines and networks.

VMware Player

VMware Player builds the free-to-use lightweight counterpart to the Workstation Pro. Because it is free of charge, it supports only very basic management features.

For instance, you cannot connect your virtual machines to VMware vSphere, create snapshots, or customize the virtual networks, which heavily limits the use cases for this application.

Hyper-V is the best-known Microsoft product for the deployment and management of virtual machines. There are three ways to install Hyper-V, each with its own characteristics.

First, there is the Windows Server Hyper-V installation. As the name suggests, this is the installation of Hyper-V as a role on a Windows Server, providing hypervisor service for managing resources on the server. It is the most common way of installing Hyper-V.

Second, there is the Hyper-V Server, which is a standalone version of Hyper-V. Even though this version of hypervisor does not require a license to install, once installed, all virtual machines will need a proper license for their operating system.

Finally, there is Client Hyper-V, which refers to a version of Hyper-V that is included in non-server operating systems. This version of Hyper-V is not meant to be used in production, but rather for development or testing purposes.

As the name suggests, Citrix Hypervisor is a type 1 hypervisor developed by Citrix. This hypervisor supports the products Citrix Virtual Apps and Citrix Virtual Desktops.

Citrix Hypervisor supports the centralized management and deployment of virtual apps and desktop environments. These environments can be customized in a way that lets administrators deploy specific services for a specific group of users.

You've completed the lesson. You can now achieve these objectives.

Cloud Overview Lesson

Welcome to the *Cloud Overview* lesson.

Click **Next** to get started.

After completing this lesson, you will be able to achieve these objectives.

The cloud refers to software and services offered by companies who specialize in making the resources of their data centers available to be utilized and accessed using the internet, rather than installed on your computer or your physical network.

Many end users are in contact with cloud services daily, without even realizing it. One of the most commonly used cloud services is storing data on the internet.

Once uploaded, data that is stored in the cloud can be accessed from anywhere using any device with an internet connection.

Behind the scenes are large, high-capacity data centers. Companies that specialize in the cloud usually offer a wide variety of IT services, including data storage, computing power, whole networks, and much more.

This enables businesses to potentially outsource parts or even all of their IT services.

A significant benefit of cloud services, compared to on-premises solutions, is that the services provided in the cloud are easily accessible by everyone, because a working internet connection is all that’s required to access resources.

Most large cloud providers also offer the option to configure high availability and fault tolerance.

This option makes it possible for services to stay up in the case of a failure, and provide disaster recovery in the case of data loss.

The ability to configure high availability and fault tolerance makes cloud services a reliable and safe solution for processing and storing large amounts of data.

There are three important service categories associated with cloud computing: infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS).

The main difference between these three categories of services, is how much responsibility the customers themselves have for managing the underlying architecture of their products. Compared to an on-premises setup, where a company manages all of their hardware and software by themselves, a cloud solution relieves the customer of a great portion of those responsibilities. These services are accessed through a dashboard or an application programing interface, or API.

The names of the three different models of services suggest the level of influence that the customer will have over their products.

When IaaS is implemented, the customer is supplied with various resources, ranging from basic network and processing power, to storage devices and servers. The resources and networking are fully managed and maintained by the provider. This means that the customer can focus on building the environment using their own virtual machine operating systems and applications.

The advantage of this model, compared to an on-premises solution, is that the customer does not have to worry about the hardware side of things anymore; the hardware is covered by the provider. In addition to this, the customer does not have to spend money on new hardware. Instead, they only pay to use resources offered by the provider.

Examples of IaaS include a Linux virtual machine in AWS or Azure.

In addition to hardware resources, PaaS also puts software, such as the operating system, databases, and application run-time environments, into the hands of the provider.

The addition of this software relieves the customer of even more management responsibility, because the provider is responsible for the management and security of the provided environments.

The advantage of this model is that customers are provided with a reliable environment for testing, development, and hosting purposes—without having to worry about the underlying architecture.

An example of PaaS is hosting a web application or web page.

SaaS is the perfect model for customers who need access to only certain applications, and who do not need a specific architecture or environment.

This service takes almost all of the responsibilities for the underlying architecture, platforms, and environments, away from the customer. It leaves the customer with just the software they want to use.

Applications are usually accessed through a dashboard or API, removing the need to install applications on the currently used device. This model also provides the option to swap devices, without having to worry about migrating applications and data.

An example of SaaS would be access to Dropbox or OneDrive.

Leading cloud providers own multiple data centers across the world. Usually, the provider lets the customer choose the location of their services. This location is a geographical location referred to as a region, for instance, the US East Coast.

A region is a geographical area in which multiple isolated groups of data centers reside. These areas are referred to as availability zones.

An availability zone usually consists of multiple data centers where the provided services are run. Each group of data centers is provisioned with its own, usually redundant, power and networking. Each availability zone also has its own isolated power and network connectivity.

The isolation of groups of data centers into availability zones, provides high availability and failover capability between the groups of data centers.

The redundant design allows for almost 24/7 uptime across the whole year, with minimal downtime.

Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform (GCP) are the most well-known platforms in cloud computing.

Of the three, Amazon was the first to publish its own cloud platform, in 2006. Microsoft and Google followed up in the years 2010 and 2011, respectively.

Together, these three providers make up for more than half of the market share.

Amazon Web Services

Amazon was the first to publish their cloud platform in 2006 and, as a result, gained an immense lead in owning the market share.

Microsoft Azure

Microsoft was second to publish their cloud platform. Microsoft Azure became available to the general public in 2010.

In the years since publishing their platform, the significant annual growth of Azure has allowed them to reach the position of the second largest platform in terms of market share, behind AWS.

Google Cloud

After announcing the Google App Engine in 2008, Google released their cloud platform to the general public in 2011.

Google Cloud Platform is now the third largest global cloud service provider.

Today, AWS, Microsoft Azure, and Google Cloud Platform offer solutions in all kinds of cloud product categories. These categories range from computing power, storage, databases, and networking, to media services, mobile, machine learning, and many more.

Over the years, all providers have grown to be strong and popular platforms, able to provide services for businesses of all sizes. Examples include companies like eBay, PayPal, and Twitter.

Three well-known categories of products and services include compute, storage, and networking.

Compute provides the customer with computing power, such as servers, virtual machines, and containers.

Storage provides access to different kinds of scalable storage systems, ranging from simple raw data object storage, to databases, backups, disaster recovery, and more complex solutions.

The networking category includes a variety of services, such as setting up firewalls, routing, configuring, load balancing, and more.

You've completed the lesson. You can now achieve these objectives.