Networking fundamentals Ben Foster

Created by Ben Foster as part of a long terms IT certification journey. Please do not redistribute without credit.

Contents

1. Basics of networking

1.1 What is networking?

Networking is everything that makes it possible for computers, servers and other devices to communicate.

This includes the design, construction, use of a network, management, maintenance and operation.

Networks are made of nodes, which are anything that can share, send or store data on a network.

Networks can be wired, or wireless, but both forms generally require physical infrastructure such as routers and switches, as well as software and firmware.

They rely on protocols like HTTP, TCP and IP that manage how it all works.

1.2 Network models

Network models provide a framework for how devices communicate. The OSI model and TCP/IP model are the most widely used, each breaking communication into layers to simplify understanding and troubleshooting. This section covers their structure, purpose and relevance modern settings.

What are network models?

Network models are conceptual frameworks that define how data is transferred between devices across a network. Their purpose is to standardize communication protocols, enabling interoperability between varying devices and technologies.

Why are network models important?

- Standardisation: the models provide a universal set of rules that allow diverse systems to communicate
- Troubleshooting: They provide a structured way to diagnose and resolve network issues
- Modularity: the models divide networking into layers, making it easier to understand and design

TCP/IP vs OSI

introduction

Both the OSI and TCP models provide a framework that help us understand how devices communicate across networks. The OSI model is a conceptual guide, whereas TCP/IP is used in real world to power the internet. Both models break down networking into layers, making it easier to troubleshoot problems and understand the process.

What you need to know for the A plus

- The OSI layers and their roles

- Understand how the layers of the TCP/IP model align with those of the OSI, and their practical application
- Match key protocols to their corresponding layers
- Understand and explain the difference between the models

OSI model

The OSI model is a 7-layer logical framework for standardizing network communication where each layer has specific roles and handles a part of the communication system, from user interaction to data transmission over hardware. Its designed to isolate the different functions that occur in a network to make troubleshooting and development easier. Its widely used as a reference model for understanding networking concepts, but isn't directly implemented.

Layer 7: Application layer

This serves as the interface between the end user and the network. Everything the user interacts with, whether browsing the itnernet or sending an email happens in the application layer.

Functions

- Facilitates communication between user application and the network
- Supports protocls
- Ensures properdata formatting, application access and resource availability

Examples: internet browsing using HTTP to retrieve internet pages, file transfer application using FTP for data exchange

Analogy: a customer interacts with a website and places an order for a pizza.

Layer 6: Presentation layer

This layer ensures that data is presented in a format that both sending and receiving applications can understand. In other words, it acts as the networks translator and editor so that the data is in the correct format for both sender and receiver, while keeping it secure and efficient.

Functions

- Data translation: converts data between application specific formats and a standard format for transmission
- Encryption/decryption: secures data for transmission by encrypting and decrypting on receIPt

- Compression/decrompression: reduces the size of data for faster transmission and restores it to its original form at the destination

Examples: encoding multimedia files such as JPEG for images, encrypting sensitive information with SSL/TLS

Analogy: the pizzaria menu is translated to languages that everyone can read; encryption of the menu would only allow prefered customers to view it

Layer 5: Session layer

The session layer manages and controls the connections, also known as sessions, between devices.

Functions

- Session establishment, maintenance and termination: opens, manages and closes sessions between applications
- Synchronization: adds checkpoints to data streams, allowing communication to resume from the last checkpoint if disrupted
- Session control: manages simultaneous data exchanges to prevent conflicts

Examples: logging into a remote server and keeping the session active, video conference applications maintaining an uninterrupted connection

Analogy: a phone call between the customer and pizzeria, the session being established when you call, maintained for the duration of the call and ending when you hang up

Layer 4: Transport layer

Ensures reliable data delivery between devices, providing error correction, flow control and proper sequenching of packets. TCP and UDP are the relevant protols to the transport layer.

Functions

- Segmentation: splits large data in smaller packets for transmission
- Reassembly: reconstructs packets on arrival in correct order
- Flow control: regulates data flow so the receiver isnt overwhelmed
- Error detection: verifies data integrity using checksums

Examples: TCP would ensure all parts of a file are received and dissassembledcorrectly. UDP would ensure data packets in a ive video stream are delivered quickly to maintain real time playback

Analogy: a delivery service ensures the pizza arrives intact with the slices in the right places. TCP would be a careful courier who esnures the pizza is delivered in tact and

with receIPt. UDP would be a fast carrier who just drops the pizza at the door, and it's a bit of a mess.

Layer 3: Network layer

This layer handles the routing of data packets aross different networks, determining the past path for them to tavel from source to destination. Relevent functions here would be IP, ICMP and ARP. Ensures the reliable communication between applications on devices that may be separatedd by multiple networks.

Functions

- Logical addressing: assigns unique IP address to devices for identification
- Routing: finds the most efficient path for data packets to reach their destination
- Packet forwarding: transfers data packets between networks
  Here routing is the planning of a path, and packet forwarding is the actual movement of the data on the chosen path

Examples: the network layer enures that the data packets thatt mmake up an email travel across multiple networks using the best available path

Analogy: postal services route letters betweeen cities. Each letter has a sender and receiver address (the IP). The postal system (routers) determine the most efficient routes and the vans drive the letters about (packet forwarding)

Layer 2: data link layer

Responsible for the reliable transmission of data across a single physical network link, ensuring error detection and correction, and the control of how devices on the same network segment access the medium. Focuses on the local communication between devices within a local network segment, managing the link between them.

Functions

- Framing: encapsulates the network layer into frames for transmission
- MAC addressing: uses hardware addresses (MAC addresses) to identify devices on the local network
- Error detection: identifies and corrects errors that occur during transmission
- Media access control (MAC): determines how devices share and use the physical medium

Examples: when a PC sends a file to a printer on the same network, the data link ensures that the frames are delivered correctly from your PCs NIC to the printers NIC using their MAC addresses.

Analogy: Once the letter has arrived at the city, it is delivered to the specific address (MAC address)

Layer 1: Physical layer

Responsible for the actual transmission of raw binary bits pver a physical medium. It defines the hardware components, electrical signals and transmission techniques needed to move data between devices.

Functions

- Bit transmission: converts data frames from the data link layer into electrical, optical or radio signals
- Media specifications: defines the physical media used for transmission, such as cables or wireless signals
- Hardware standards: includes specifications for devices like network cables, connectors, hubs and transceivers
- Synchronisation: ensures sender and receiver are aligned on timing to interpret signals correctly

Examples: when you plug an ethernet cable into your PC, the physical layer is responsible for transmitting the electrical signals through the cable to the router or switch
Analogy: the physical layer is the road that the van carrying letters of motorbike carrying the pizzas travels on. Without the road, no movement is possible. Standard ensure the road is safe to travel on (WIFI or hardware standards)

TCP/IP model

The transmission control protocol/internet protocol is a practical, simplified model with only four layers. It is the framework used to power the internet and most modern networks.

Network interface layer

This combines the OSI's physical and data link layers. Handles hardware communications and local network connections, such as Wi-Fi, ethernet and ARP.

Internet layer

This maps to the OSI's network layer and handles routing, IP addressing, and packet delivery through IVP4 and IVP6 and ICMP

Transport layer

Directly matches the OSI's transport layer. Ensures reliable or fast delivery through TCP or UDP.

Application layer

Combines the OSI's application, session and presentation layers. Supports user facing applications and protocols through HTTP, HTTPS, DNS and more.

Key differences between OSI and TCP/IP

Aspect	OSI model	TCP/IP model
Purpose	Theoretical framework for learning	Practical framework for real networks
Number of layers	7	4
Usage	Primarily educational and conceptual	Foundational for internet communication
Layer structure	Detailed separation of functions into 7 layers	Combines similar layers into broader groups
Protocols	Mapped to specific layers	Mapped across four layers

The key difference to remember is that TCP/IP simplifies the OSI by combining real layers to focus on real world use.

Real life examples

Browsing the internet

- The application layer uses HTTP to request a internetpage
- The transport layer ensures data reliability with TCP
- The internet layer uses IP to route the data
- The network interface layer handles the physical transmission data

1.3 Peer to peer and client server networks

Introduction

Networking structures define how devices interact and share resources. Peer to peer (P2P) and client-server models are the two fundamental network architectures, each with its own characteristics, advantages and disadvantages. This section explores the architecture, functionality, and real-world applications of these models.

Peer to peer networks

These are networks where all devices (nodes) are treated as equals, meaning every device can act as both client (requesting resources) and server(providing resources). They're typically used in small networks with minimal infrastructure requirements.

Characteristics

- Decentralized: there's no dedicated server, so resources are shared directly between devices
- Simplicity: easy to set up and configure, with minimal technical knowledge required
- Cost effective: requires no additional hardware or software investments for dedicated servers

Advantages

- Quick and easy to deploy
- Low cost for small networks
- Does not require extensive administrative overhead

Disadvantages

- Limited scalability: performance degrades as the network size increases
- Security concerns: less centralized control over data sharing, increasing vulnerability to authorized access
- Data management: no central point of control, making back-ups and administration more challenging

Use cases

- Home networks for file sharing or printer access
- Small offices with basic resource sharing needs
- File sharing protocols like BitTorrent

Client server networks

In this model, devices (clients) request resources from a central server. The server manages resources, authentication and permissions.

Characteristics

- Centralized: dedicated servers handle tasks like file storage, database management and application hosting
- Structured design: designed to support larger networks with many users

Advantages

- Scalability: can support hundreds or thousands of devices
- Security: centralized control over access permissions, updates and data management
- Reliability: servers are usually robust systems designed for high availability

Disadvantages

- Requires investment in server hardware, software and maintenance
- Requires technical expertise for setup, configuration and troubleshooting
- Dependency: if the server goes down, network resources become inaccessible

Use cases

- Medium to large businesses where centralized management is critical
- Email servers, database servers and internet hosting
- Enterprise environments with shared drives, user authentication and application hosting

2. Networking devices

2.1 Introduction

Broadcast domains and related concepts

A broadcast domain refers to a logical division of a computer network in which broadcast messages (messages sent to all devices on a network segment) are transmitted to all devices within that domain. Devices in the same broadcast domain can receive and process broadcast packets sent to all hosts within the domain, such as ARP requests or DHCP discover packets.

A broadcast is a message that is sent from one device to all other devices within the network segment. For example, when a device is looking for another device on the same network, it might send a broadcast message to find the devices MAC address.

How are broadcast domains identified?

On layer 2 devices, such as switches, a broadcast domain is typically defined by VLANs (virtual area local networks). Each VLAN represents a separate broadcast domain, meaning that devices in different VLANs cannot communicate directly via broadcast. Switches segment broadcast traffic within each VLAN, ensuring that broadcasts are contained to the devices in that VLAN.

When it comes to layer 3 devices, such as routers, they separate broadcast domains. When a broadcast packet reaches a router, the router does not forward it to other networks. Routers act as a boundary between broadcast domains, and the traffic must be routed to reach other networks.

Why do broadcast domains matter?

They are important for managing network efficiency. As more devices are added to a broadcast domain, the traffic load increases because devices in that domain must process the broadcast traffic. This can cause a network bottleneck, especially in larger

networks where broadcasts are frequent. For instance, if many devices are all listening to a broadcast, it can result in a high overhead that slows down the network.

Related concepts

- Collision domains: this is a network segment where devices can collide with each other while trying to send data over the same medium. This is mostly an issue for older hub-based networks or ethernet environments, where devices share the same bandwidth. Modern switches isolate collision domains at the port level, meaning each device connected to the switch has its own collision domain, reducing the chance of collisions.
- VLANs: these are used to logically segment a network into smaller, logical broadcast domains. Devices in different VLANs can communicate with each other through routing (often done by a router or a layer 3 switch), but broadcasts from one VLAN are isolated from others. VLANs improve network efficiency and security by limiting the scope of broadcast traffic and separating different types of network traffic.
- Subnetting: while VLANs segment a network at the data link layer, subnetting segments a network at the IP layer. Each subnet is typically its own broadcast domain, as devices within a subnet can communicate directly using broadcasts. Routers separate subnets and prevent the transmission of broadcasts between them.

Conclusion

Understanding broadcast domains is critical for designing networks that are efficient and secure. By segmenting broadcast traffic using VLANs and routers, network engineers can improve performance by reducing unnecessary broadcast traffic and controlling where it's sent.

A broadcast domain defines networks as we know them; they are bounded by routers and can be divided into smaller segments using VLANs to improve network performance. It is essential to understand broadcast domains in order to understand the roles of the hardware in the following chapters.

2.2 Routers, switches and hubs

Hubs

These are basic networking devices that connect multiple devices such as PCs and printers in a network, and broadcasts all incoming data to all connected ports, regardless of the intended recipient.

Key features

- They operate at OSI layer 1, meaning it does not understand or process data packets
- They operate in half duplex mode, meaning only one device can send data at a time, causing potential collisions
- They do not filter data, leading to inefficient use of bandwidth

There are two types of hub: active hubs, which amplify the signal before broadcasting it, and passive hubs, which just forward the signal without any amplification.

They're rarely used in modern networks due to their inefficiency.

Switches

Networking devices that connect multiple devices in a network and forward data only to the intended recipients port using MAC addresses.

Key features

- Operate at OSI layer 2
- Support full duplex communication, allowing devices to send and receive data simultaneously
- Uses a MAC address table to learn which devices are connected to which ports, improving network efficiency and reducing collision
- Modern switches often have features like VLANs and QoS (quality of service) configurations

There are two types of switches: unmanaged, which are 'plug and play' with no configurations needed. These are for simple networks; and managed switches which allow administrators to configure features like VLANs, port mirroring and security settings.

Advantages over hubs

- Intelligent data forwarding reduces unnecessary traffic
- Can support larger networks and higher bandwidths
- Data is only sent to the intended device, improving security

They're found in almost every modern network, from small office set ups to large enterprise businesses.

Routers

These are networking devices that connect multiple networks and route data between them. They are the traffic controllers of the internet.

Key features

- Operate at OSI level 3
- Use IP addresses to determine the best path for data to travel across networks
- Can perform network address translation (NAT) allowing multiple devices on a private network to shar a single public IP address
- Many modern ones include built in features like DHCP, firewalls and wireless access points

Advantages

- Enable communication between different networks
- Support security through firewalls and traffic filtering
- Allows for dynamic IP configuration through DHCP

Used in homes to connect LANs to the internet, as in businesses.

2.3 Access points, gateways and bridges

Introduction

Access points (APs) and gateways are critical components for facilitating connections across the internet and various networked resources. Access points extend network coverage, allowing wireless devices to connect to a local network. They are essential for creating Wi-Fi networks in homes, offices and public spaces. Gateways are any devices that connect two different networks and facilitate communication between them. They can be in the form of routers, proxy servers, firewalls, cloud-based security gateways or VPN gateways. They act as intermediaries between different networks, translating data formats, enforcing security policies and managing traffic between local and external networks.

An understanding of these devices is crucial for designing and maintaining scalable, secure and efficient networks. This chapter will cover the role of access points, types of access points, gateways and their functions, access point and gateway configurations.

Access points

Types of access points

APs come in various types, with each serving a different network environment and requirements. While they all provide wireless connection, their deployment, management and capabilities vary based on factors such as network size, scalability, security needs and administrative control. The two main categories are:

- Standalone access points: independent devices that function on their own, typically in small networks

- Controller based access points: managed centrally by a wireless LAN controller (WLC) used in enterprise environments for scalability and advanced management

Standalone access points

These are independent networking devices that provide wireless access to clients without requiring a central management system. They're commonly found in homes, small offices, and retail stores.

Key features

- Configured individually via a internet browser of a command line interface (CLI)
- Operate independently, without requiring a central controller
- Best suited for small scale networks with few access points
- May include basic security features like WPA/2/WPA3 encryption and MAC address filtering
- Limited scalability as each AP must be configured and managed separately

Expanding on practical understanding

These basically function like basic network switches, except they operate wirelessly. In small networks, the router often acts as the AP, but some environments have 'dead' zones that require extended coverage so extra APs can be added to expand the network. Unlike switches, which forward traffic between wired devices, standalone APs allow wireless devices to join a wired LAN. This understanding may be helpful when troubleshooting; if a wired device works but a wireless one doesn't, the issue may be isolated to the access port or its links.

Access points and the OSI/TCP models

An AP operates mainly at levels 1 and 2 (physical and data link) of the OSI.

- Layer 1: the AP receives and converts wireless signals using radio waves, serving as the bridge between wireless devices and the wired architecture of a network. They also convert ethernet signals into Wi-Fi
- Layer 2: the AP handles MAC (media access control) addresses, ensuring data is delivered to the correct wireless devices

In the TCP model, APs function at the network interface layer, handling device communication at the local network level without direct involvement in higher level routing decisions

Gateways

Role of a gateway in a network

A gateway acts as a crucial network device that serves as the point of entry and exit for data traveling between different networks. It functions as a translator between networks that use different protocols or have different communicate standards, enabling them to communicate effectively.

In simple terms, a gateway connects two distinct networks, such as those defined by broadcast domains. It often operates at different layers of the OSI model (usually 3 or 4). It enables devices one network to communicate with devices with another, such as a local network communicating with the internet or between different subnerts.

Functions of a gateway

- Routing traffic: a gateway routes data between networks. It directs data packets between the local network and an external network
- Protocol translation: converts traffic between different network protocols, such as IPV4 to IPV6, or HTTP to HTTPS
- Security and filtering: implements firewall rules, packet inspection and access control
- Network address translation (NAT): allows multiple devices to share a single public IP address
- VPN termination: acts as an endpoint for secure remote access

In home networks, routers typically serve as the default gateway and a firewall. In larger networks, dedicated security gateways may handle traffic filtering, VPN connections and advanced routing. Switches keep traffic local, the gateway determines what traffic can leave the network and how it gets there.

Gateways and the OSI

Gateways are more complex devices that typically function at multiple levels.

- Layer 3/network layer: They often act as routers, determining the best path for traffic between networks, they use IP address to forward packets between local networks and the internet
- Layer 4/transport layer: some gateways, like firewalls and VPN concentrators, inspect and manage transport layer traffic (TCP/UDP) for security and optimization
- Layer 7/application layer: advanced gateways such as proxy servers and cloud based security appliances operate at the application layer by filtering content, translating protocols and optimizing internet traffic.

AP and gateway configurations

Configuring AP and gateways is crucial for network stability, security and performance. Both devices must be properly set up to ensure that users can connect seamlessly, communicate across networks and access external resources without issues.

Configuring an AP

To be deployed effectively, the following configurations are usually required

SSID and security setup

- Define the SSID (service set identifier) which is the name of the wireless network
- Apply WPA2/3 security to prevent unauthorised access
- Configure MAC filtering for additional security (optional)

VLAN and network integration

- Assign VLANs for different wireless networks
- Configure trunk ports on the switch if multiple VLANs are used

IP addressing

- Choose between static IP and DHCP
- Ensure the AP has an IP in the correct subnet to communicate with the network

Roaming and load balancing

- Optimize channel selection to reduce interference
- Enable seamless roaming for users moving between APs

Configuring a gateway

Gateways must be configured to correctly route traffic between internal and external networks, ensuring security and efficiency

Default gateway setup

- Assign the correct IP address to the gateway interface
- Ensure client devices use this IP as their default route

Network address translation (NAT)

- Configure NAT to allow multiple devices to share a single public IP
- Enable port forwarding if internal services need external access

Firewall and security policies

- Implement access control lists (ACLs) to block unauthorized traffic
- Set up VPNs for secure remote access

Routing configuration

- Define static routes or enable dynamic routing protocols
- Ensure correct routes between VLANs if inter-VLANcommunication is needed

Troubleshooting APs and gateways

Even with proper configuration, APs and gateways can experience connectivity issues that affect network performance. Common problems include weak signals, incorrect IP configurations and misconfigured VLANs or security settings blocking traffic.

Common access point issues

Issue	Possible Cause	solutions
Wi-Fi clients can't connect	Incorrect SSID/password, MAC filtering enabled	Verify SSID, check authentication settings
Slow or unstable Wi-Fi	Signal interface, too many clients	Change Wi-Fi channel, reduce congestion
AP not reachable	AP has no IP or is in wrong subnet	Assign a correct static IP or check DHCP
VLAN not working	AP is on an access port instead of a trunk	Change switch port mode to trunk

Common gateway issues

Issue	Possible cause	Solution
Devices can't access the internet	Default gateway is missing or incorrect	Check IP settings, ensure correct gateway
Some websites don't load	DNS settings incorrect	Verify DNS server settings
Intermittent connection loss	Overloaded router, firewall blocking traffic	Reduce load, check firewall rules
One network can't reach another	Missing static route or misconfigured routing protocol	Add a correct route (IP route command)

Bridges

A bridge is a network device used to divide a network into segments while keeping traffic local to each segment. It works at layer 2 of the OSI model (data link), using MAC addresses to make forwarding decisions. Unlike a switch, which has multiple ports, a bridge typically connects just two segments.

Bridges help reduce collisions by filtering traffic and forwarding only what's necessary, which improves overall performance on busy networks. They're less common today due to switches taking over their role with greater efficiency and flexibility, but understanding bridges helps build a solid foundation of how networks evolved.

Legacy networks where you want to connect two different LAN segments without creating a broadcast storm, would be where you'd find a bridge.

2.4 Modems and NICs

The role of modems and NICs in networking

Modems and NICs are essential components that enable devices to connect to networks and communicate over different types of infrastructure. Modems translate digital signals from PCs into a format suitable for transmission over telephone, cable or fibre lines. NICs provide the physical and logical interface between a device and a network, handling both wired and wireless communication (Wi-Fi and ethernet).

NICs operate at layer 1 and 2, handling MAC addressing and network access.

Modems operate mainly at layer 1, converting signals for transmission.

In TCP/IP, both operate in the network layer

Practical understanding

Very device, wired or wireless, needs an NIC to connect to the network. Modems are needed when a network connects to an ISP. No modem, no internet. NIC configs affect VLANS, DHCP and MAC filtering, so they're critical for networking performance and security. They link everything we've covered so far.

Types of modems and their functions

Modems (modulator-demodulators) are the interface between a LAN and an ISP, converting digital data into a form suitable for transmission over network infrastructures.

Dial up modems (legacy)

- Use public switched telephone network (PSTN)
- Maximum speed of 56kbps (slow)
- Rarely used today

DSL modems (digital subscriber line)

- Uses telephone lines, but allows internet and voice calls simultaneously
- Speed range from 1mbps to over 100mbps
- Asymmetric DSL (ADSL) faster downloads, slower uploads
- Symmetric DSL (SDSL) -equal download, upload speed, used in business settings

Cable modems

- Use coaxial cable networks, same as cable tv

- Faster than DSL, typically 100mbps to 1gbps
- Requires DOCSIS (data over cable service interface specification) compliance for ISP compatibility

Fiber modems

- Uses fibre optic cables for extremely high speeds (up to 10gbps)
- Converts light signals into digital data
- Requires specialized ISPs and infrastructure

Fiber optic broadband networks use an optical network terminal (ONT) device to convert optical signals into electrical signals that can be used in end user devices. The ONT is typically installed at the customer's premises and connects the fibre optic line from the service provider to the customers internal network.

FTTP (fibre to the premises) or FTTH (fibre to the home) refer to high speed fibre optic internet connections that deliver internet services directly to the end user' location. In both FTTP and FTTH, the fibre optic cable runs directly from the ISPs central office to the customers location.

Cellular modems 4g/5g

- Uses mobile networks (LTE, 5G) instead of wired connections
- Used in mobile hotspots, rural areas, or as backup internet connections
- Requires a sim card and a mobile data plan

Satellite modems

- Uses satellite signals for internet in remote locations
- High latency (due to long distance signal travel)
- Common in rural areas, maritime applications, military networks

How do they fit into the network?

- Standalone modems connect directly to the router
- Modem/router combos function as both a router and a modem
- Fibre and cable modems often require an ISP provided modem with specific compatibility

Practical applications

Every network connected to an ISP requires a modem, even if its just integrated into the router

Different modem types require different cables; DSL uses telephone lines, cable modems use coaxial cables, fibre uses fibre optic cables

ISPs control modem compatibility; you cant just plug in any modem and expect it to work

Some business use multiple modems for redundancy; a fibre connection and a 4G backup

Network interface cards and their types

An NIC is a hardware component that enables a device to connect to a network, whether through wired ethernet or wireless Wi-Fi. Every networked device requires a NIC to communicate across a network.

Types of NICs

Wired NICs

- Use RJ45 ethernet ports to connect to wired networks
- Operate at speeds of 10mbps, 100mbpsm 1gbps, or 10gbps
- Found in desktops, servers and some laptops

Variation of ethernet NICs

- Integrated NICs – built into the motherboard, most common type
- PCIe NICs- add on cards for high performance networking
- UDB to ethernet adapters – used for devices without built in ethernet ports

Wireless NICs

- Enables devices to connect to Wi-Fi networks
- Uses radio waves instead of physical cables
- Found in laptops, smartphones, tablets and IoT devices
  Variations of wireless NICs
- Integrated Wi-Fi NICs – built into the motherboard (common in laptops)
- PCIe wireless NICs – for desktops that need Wi-Fi capabilities
- USB wireless NICs - plug and play options for Wi-Fi access
- M.2 wireless cards – internal cards used in ultrabooks and mini PCs

Virtual NICs (software based)

- Used in virtual machines and cloud environments
- Allows VMs to connect to physical networks as if they were separate devices
- Found in hyper v, VMware and cloud computing platforms

NIC functions and performance factors

- Speed: 10/100mbps, 1gbps, 10gbps, or even 40+gbps in enterprise setups

- Full duplex vs half duplex:
    - Full duplex: simultaneous sending/receiving (modern NICs)
    - Half duplex (one direction at a time (older networks)
- MAC addressing: Each NIC has a unique MAC address used in layer 2 communication

Practical understanding

Every device needs an NIC – even wireless devices use interface Wi-Fi NIC

Network performance depends on NIC capabilities -a 1gbps network wont be fast if a device has a 100mbps NIC

Some enterprise systems use multiple NICs for redundancy, such as servers with NICs

NICs impact VLANs, QoS, and security – network admins often configure MAC filtering and LVAN tagging at the NIC level

NIC configurations

NICs must be properly configured to ensure optimal network performance, security and compatibility. Configuration settings can be adjusted via the OS, BIOS or network management tools.

NICs can be configured to use either dynamic IP or static IP. Static IPs are needed for devices such as printers, servers, and remote access configurations.

Speed and duplex settings

Auto negotiation: NIC automatically selects the best speed and duplex setting

Manual configuration: admins can manually set: speed to 10mbps to 1gbps and up, or duplex mode to full or half (in full data is sent and received simultaneously)

MAC address filtering and spoofing

MAC filtering is used in network security to allow or block devices based on their MAC address

Spoofing is when NICs allow the MAC address to be changed for privacy or troubleshooting purpose

Wake on LAN (WoL)

Allows a computer to be powered on remotely by sending a special packet

VLAN tagging

NICs can be configured to support VLANs, assigning a device to a specific VLAN for network segmentation. This is common in enterprise environments for isolating departments and improving security

How does it fit into a network?

- Misconfigured NICs can cause speed limitations, connection issues and security risks
- Speed/duplex mismatches can slow performance and drop connections
- Incorrect VLAN settings can isolate devices from accessing network resources
- Static IP misconfiguration can lead to IP conflicts and network instability

2.5 Other network hardware

Injectors, PoE, patch panels, etc

3. Internet connection and network types

3.1 Internet connection types

Network connections are critical for linking devices to the internet and other systems so they can communicate with each other.

Satellite

A satellite connection provides internet access via satellites orbiting the earth. The signal is transmitted from a device to the satellite, which relays it to a ground station connected to the internet backbone. They're ideal for rural locations with limited access where terrestrial services like fibre or cable aren't available, but due to the long distances the data must travel they can experience high latency. They're also weather dependent; rain, snow and heavy cloud cover can affect the signals. Also have limited bandwidth and high costs than other types of connections.

Fibre

Fibre optic internet uses light signals transmitted through glass or plastic fibres to deliver very fast and reliable connections. The speeds are very high, and can be up to 1gbps, along with low latency, so they're ideal for viewing and gaming. Fibre optic connections also provide symmetric speeds, where both download and upload speeds are the same. The downsides are that it is expensive, and requires specialised infrastructure not available everywhere.

Cable

These use coaxial cables to provide internet access, and also carry cable TV signals as well. Popular in residential areas, generally accessible in suburban areas as customers can share infrastructure and it offers decent download speeds of between 100 and 1000 mbps. Has latency in between that of fibre and satellite, and speed can slow down during peak hours as the hardware is shared by many users.

DSL (digital subscriber line)

DSL uses existing phone lines to deliver broadband internet. There are two types :ADSL (asymmetric DSL) and SDSL (symmetric DSL). ADSL has faster download speeds than upload speeds, whereas SDSL has both the same. DSL is good because it is widely available due to existing phone lines, so can be used where fibre and cable aren't available, and its also low in cost. The downside is that its slow, with speeds of 1-25mbps, and its distance sensitive, meaning as you get further from the provider hardware it slows.

Cellular (mobile broadband)

This uses mobile phone networks (3g etc) to provide internet access. It can be delivered through USB modems, mobile hotspots, or built in cellular modules in devices. It has wide coverage, with networks being available in urban, suburban and rural areas. Its also very portable due to its nature of being in built to phones and other mobile devices. Speeds however can vary greatly depending on coverage, and users have to deal with data caps as they are sold as part of sim/phone plans. Signal strength also varies greatly based on location.

Wireless internet service provider (WISP)

In remote or underserved areas where traditional broadband options like fiber, DSL or cable aren't available, businesses may rely or wireless internet service providers (WISPs). These providers deliver internet access using microwave or radio frequency transmissions between fixed antennas.

WISPs typically use line of sight connections between a base station (often mounted on a tower) and a receiver at the customers location. High end WISP services can offer speeds up to 1GBPS with low latency making them suitable for business class use – even VoIP and video conferencing – when configured correctly.

Unlike DSL, which is limited by phone line distance and speed, or cable, which shares bandwidth with nearby users, WISPs offer dedicated or semi-dedicated wireless links. They are especially valuable in rural areas or new construction where digging for fiber or coax isn't practical.

Key identifiers of WISP services:

- Uses microwave or radio transmission (often mentioned in A+ questions)

- Suited for areas outside traditional broadband service zones
- May require line of sight between towers and receivers
- Can deliver low latency and high bandwidth, depending on service tier

3.2 Network types

Understanding network types is crucial for selecting the right type of connection and troubleshooting network issues. Each type has its own scope, advantages and applications.

Local area network (LAN)

This is a network that covers a small geographic area, such as a home or office, connecting user devices like laptops, phones, printers and servers. They're often private, high speed networks using Wi-Fi or ethernet.

Wide area networks (WAN)

This spans a large geographic area, such as a city, country or even continent. It connects multiple LANs and enables comms over very long distances. The internet itself is a WAN, and the largest example of one.

Personal area network (PAN)

This is a small network used for connecting devices in close proximity to each other, usually about 10 meters. Its designed for personal devices such as laptops, smartphones and wearables.

Metropolitan area network (MAN)

This covers a larger area than a LAN but smaller than a WAN, such as a city or large campus. It is designed to connect multiple LANs, providing high speed internet over a city wide area. Would be in use in the case of city wide ISPs, corporate networks and universities.

Storage area networks (SAN)

This is a dedicated high speed network that provides block level data storage access. It connects storage devices like disk arrays and tape libraries to servers, enabling high speed data transfer for storage applications.

Wireless local area network (WLAN)

This is like a LAN but wireless. Provides flexible and convenient network access for mobile devices and is used in homes, businesses and offices.

4. Network cables

4. Internet protocols

4.1 Introduction

IP is the protocol by which internet addresses are assigned to devices. It is the foundation of modern networking, allowing devices to identify and communicate with each other across local and global networks. The two primary versions in use today are IPv4 and IPv6. Servers, PCs, mobile phones all have IP addresses which are used as destinations to tell your own gateway device where to send data in order to reach it's target.

4.2 Public vs private IPs

Introduction

Not all IP addresses are created equal. Some are used to identify devices within a private network, while some are used to identify devices across the broader internet. Understanding the difference between public and private IP addresses is essential for network configuration, security, and internet communication.

Public IP addresses

Definition

A public IP address is a unique identifier assigned to a device on the internet. These addresses are globally routable and allow devices to communicate across different networks worldwide. Public Ips are assigned by internet service providers (ISPs) and must be unique across the entire internet.

Features of public IPs

- **Globally unique**: no two devices on the internet can have the same public IP
- **Assigned by ISPs:** internet service providers distribute public IPs to customers
- **Routable on the internet**: can be accessed from anywhere in the world
- **Used by**: internet servers, email servers, VPNs, and home routers with direct internet access

Types of public IPs

- Static public IP: remains constant and does not change, such as for internet servers or remote access
- Dynamic public IP: changes periodically and is assigned via dynamic host configuration protocol (DHCP)

Examples of public IPs

Public IP ranges are assigned by IANA and include

- 8.8.8.8 (googles public DNS server)
- 142.250.190.78 (Google's internet site IP)

Private IP addresses

Definition

A private IP address is used within a private network, such as at home. These addresses are not routable on the internet and are designed to be used behind a router or gateway that translates private addresses to a public one when needed. Private IP addresses are designed by certain ranges and are often reused across different private networks.

Features of private IPs

- Not globally unique: can be used by multiple organizations in separate networks
- Not assigned by ISPs: managed internally within private networks
- Requires network address translation (NAT): to communicate with the internet, private IPs must be translated into a public IP
- Used by: home networks, corporate LANs and office networks+

Private IP addresses ranges

IANA has reserved specific IP ranges for private use

Address range	CIDR notation	Common usage
10.0.0.0 – 10.255.255.255	10.0.0.0/8	Large corporate networks
172.16.0.0 – 172.31.255.255	172.16.0/12	Mid-size company networks
192.168.0.0 – 192.168.255.255	192.168.0.0/16	Home and small business networks

Example of private IPs

- 192.168.1.1 (common router IP)
- 10.0.0.1 (enterprise network gateway)
- 172.16.5.50 (private network device)

Why do we need private IPs?

- Conservation of IPv4 addresses: IPv4 has a limited umber of addresses (roughly 4.3 billion) and private IPs help conserve them by allowing multiple devices to share a single public IP through NAT
- Improved security: private IPs cannot be accessed directly from the internet, reducing exposure to cyber threats
- Local network communication: devices in an office or home can communicate with each other using private IPs without needing a public address

4.2 IPv4 and subnetting

IPv4 and binary

IPv4, or internet protocol version 4, is the foundation of addressing in most modern computer networks. It provides each device on a network with a unique address, allowing communication across local and wide area networks. Every time you browse a website, stream a video, or send an email, IPv4 addresses are directing that data where it needs to go.

An IPv4 address is a 32-bit binary number, but it's written in a more human friendly format called dotted decimal notation, such as:

192.168.1.1

This format breaks the 32 bits into four 8-bit blocks, called octets, separated by full stops.

Each octet can range from 0 to 255, because 8 bits can represent 8 values.

Subnet masks and CIDR

Every ipv4 address is divided into two parts:

1. Network portion – identifies the specific network, or broadcast domain
2. Host portion – identifies the device (host) on that network

How many bits belong to the network and how many to the host is defined by the subnet mask. This is done by the placement of 1s and 0s in the binary conversion of the subnet mask. The 1s represent the network portion, 0s represent the host portion. This placement then applies to the IP address itself.

Example:

- IP address: 192.168.1.10
  Subnet mask: 255.255.255.0

Because the binary conversion of this subnet mask is 11111111.11111111.11111111.0000000, this means the first 24 bits (3 octets) of the subnet mask represent the network, and the last 8 bits (last 1 octet) the host. This goes for the IP address also. This is essential for setting up network devices, because when there are mismatches between IPs and subnet mask communication doesn't work.

Due to the first 24 bits of the IP being for the network, we know that if the first three octets of a device's IP differ from 192.168.1, it is in a different subnet to 192.168.1.10, and so will need a device such as a router to communicate. A device with these portions of its IP address the same can communicate without a router (given hardware set up,

connection medium etc). So 192.168.1.20, 192.168.1.30 and 192.168.1.40, would all be in the same broadcast domain as the device with the address in the example, because the first three octets (192.168.1) are the same, and we know it is these first three octets that must be the same due to the binary conversion of the subnet mask.

In other words, the subnet mask determines the size of the network. A subnet mask is traditionally written like:

- 255.0.0.0
- 255.255.0.0
- 255.255.255.0

But we also use CIDR (classes inter-domain routing) to express the same thing in short form. One example would look like this:

- 192.168.1.10/24

The /24 means that 24 bits (the first three octets) represent the network portion of the address (such as in the previous example), and the remaining 8 bits are for the hosts.

Block sizes

Lets look at how many usable host addresses are available for different subnet masks:

CIDR	Subnet mask	Hosts per subnet
/8	255.0.0.0	16,777,214
/16	255.255.0.0	65,534
/24	255.255.255.0	254
/25	255.255.255.128	126
/26	255.255.255.192	62
/27	255.255.255.224	30
/28	255.255.255.240	14
/29	255.255.255.248	6
/30	255.255.255.252	2

The total number of hosts is always 2 less than the CIDR notation would indicate, because an address is reserved for the network ID and for the broadcast address.

4.3 IPv6

IPv6 (internet protocol version 6) is the successor to IPv4, developed to solve the problem of IPv4 address exhaustion. While IPv4 uses 32 bit addresses (like 192.168.1.1), IPv6 uses 128 bit addresses, allowing for trillions of unique Ips – more than enough for every device on earth.

The format looks like:

2001:0db8:0000:0042:0000:8a2e:0370:7334

It consists of eight groups of four hexadecimal digits, separated by colons.

IPv6 addresses are abbreviated by removing leading 0s from each group. For example, 0db8 becomes db8. A set of consecutive 0s can be replaced by a double colon. So the above example becomes:

2001:db8:0:42:0:8a2e:370:7334

Where there are consecutive groups of all 0s, such as 0000:0000, this can be compressed to ::. But only one of these abbreviations is allowed per address.

The loopback address for IPv6 addresses is ::1.

5. Ports and protocols

Ports and protocols are another fundamental aspect of networking that is absolutely crucial to the communication process. They are essential in how data is transmitted and received between devices. Protocols define the rules for communication, ensuring that data is sent and received correctly over a network. Each protocol typically uses a specific port number to differentiate between different types of traffic on the same network, enabling multiple services to run simultaneously.

5.1 TCP vs UDP

PCs and network devices communicate over the internet and local networks using protocols. These protocols operate on different ports, which act as assigned communication channels for specific services. In networking, two main transport protocols handle data transmission. These are TCP (reliable, connection oriented) and UDP (faster, but connectionless and unreliable.

Each protocol uses port numbers to determine how network traffic is directed. Some ports are reserved for specific services, called well-known ports. This chapter covers the most important TCP and UDP ports, their functions, and key differences between connection oriented and connectionless comms.

TCP and UDP

TCP and UDP are transport layer protocols. Ports are endpoints used to identify specific apps or services. Some use TCP, some use UDP, and some can use both. They are like delivery methods, with ports the mailboxes assigned to different apps.

TCP (transmission control protocol)

- Reliable – ensures all packets arrive in order
- Connection oriented – establishes a session before data transfer

- Slower but secure – used when accuracy is more important than speed
- Examples – internet browsing (HTTP), email (SMTP, IMAP, POP3), file transfers (FTP)

UDP (user datagram protocol)

- Unreliable but fast – no error correction, some packets may get lost
- Connectionless – no session setup before sending data
- Used for real time apps – when speed matters more than accuracy
- Examples – video streaming, VoIP calls, online gaming

Well known TCP and UDP ports

Protocol	Port number	Transport type	Purpose
FTP (file transfer protocol)	20/21	TCP	Transfers files between pcs. 21 handles control commands, while port 20 is used for data transfer
SSH (secure shell)	22	TCP	Encrypted remote login for secure admin of network devices
telnet	23	TCP	Remote command line login (insecure, usually replaced by SSH)
SMTP (simple mail transfer protocol)	25	TCP	Sends mail between mail servers
DNS (domain name system)	53	Usually UDP, sometimes TCP	Rsolves domain names into Ips, using UDP for fast lookups and TCP for large responses
DHCP	67/68	UDP	Assigns IP addresses to devices. 67 is used for the server, 68 for the host
HTTP (hypertext transfer protocol)	80	UDP	Transfers internet pages over the internet (insecure)
POP3 (post office protocol v3)	110	TCP	Retrieves emails and deletes the copies on the

			server (used by modern email clients)
NetBIOS/netBT	137-139	UDP/TCP	Legacy protocol for local network file and printer sharing
IMAP (internet message access protocol)	143	TCP	Retrieves email while keeping a copy on the server (used by modern email clients)
SNMP (simple network management protocol)	161/162	UDP	Used for network monitoring and managing devices like routers and switches
LDAP (lightweight directory access protocol)	389	TCP/UDP	Used for directory services, such as active directory authentication
HTTPS (hypertext transfer protocol secure)	443	TCP	Secure internet browsing using SSL/TLS encryption
SMB (server message block)	445	TCP	Windows file and printer sharing. Used in network drives
RDP (remote desktop protocol)	3389	TCP/UDP	Allows remote control of windows desktops and servers

How to identify open ports on a network

Network admins often need to check which ports are open or closed for security and troubleshooting.

On PowerShell this can be done with netstat -an. This shows active TCP/UDP connections and listening ports.

On Linux/MacOS you can enter sudo netstat -tulnp. This lists all open ports and services using them.

6. Wireless networking

6.1 introduction

Wireless networking is essential for modern communication, allowing devices to connect without physical cables. These chapters introduce the basics of wireless networking and their role in enabling seamless comms. Understanding wireless networking is crucial for effective setup, management and troubleshooting of modern networks.

6.2 Wireless networking concepts

Wireless networking allows devices to communicate over the air, eliminating the need for physical cabling. Its what powers Wi-Fi in homes, businesses, cafes and airports. But despite its convenience, wireless networking has a unique set of considerations: limited frequencies, interference from other signals, and the need for carefully chosen channels and configurations to ensure stable connections.

Wireless frequencies and bands

Wireless communication uses radio frequency (RF) signals to transmit data. These are categorized in bands (using the unit GHz). Each frequency band is broken into channels, which are smaller frequency slices used by access points to communicate with devices

The main frequency bands for Wi-Fi are:

2.4 GHz band

- This offers greater range and better wall penetration
- Has fewer available channels and is mor prone to interference from other household devices like microwaves, cordless phones and Bluetooth
- Contains 11 channels (in North America), but only 3 non-overlapping: 1, 6 and 11
-
- Common in older routers and useful for wide coverage in small spaces

5 GHz band

- This band offers higher speeds and less interference
- Has 23+ non overlapping channels, so there is more room to avoid interference, making it ideal for high density areas like apartments or offices
- Shorter range and doesn't penetrate walls as well
- Ideal for high performance applications like gaming and video streaming

6 GHz band

- Introduced with Wi-Fi 6E
- Offers cleaner spectrum, less interference, and support for high speed devices

- Still relatively new but growing in enterprise and high end home networks

Channel widths

- Standard channel widths are 20, 40, 80 and 160 MHz
- Wider channels offer faster speeds, but are mor elikely to overlap with others, leading to potential interference, so 20 MHz is more stable in crowded areas whereas 80 or 160 MHz is used for performance in clean, open spaces.

Interference

Wireless networks are subject to a range of interference sources

- Co-channel interference: two or more routers using the same channel. Devices must wait their turn, causing delays
- Adjacent channel interference: routers on overlapping channels. Causes noise and poor performance
- Non-Wi-Fi interference: other devices (like microwaves or baby monitors) emit signals in the 2.4 GHz range and disrupt Wi-Fi
- Physical interference: walls, metal surfaces, and even water can weaken signals

SSIDs and broadcasts

Every wireless network is identified by its SSID (service set identifier) – the name you see when connecting to Wi-Fi.

By default, routers broadcast their SSID, making the network visible. This can be disabled to make the network hidden. If you can't find a network you're trying to connect to, it might be the case that its hidden.

Signal strength and placement

Signal strength affects connection stability, speed and range.

Key factors:

- Distance from the access point: signal weakens the further you are
- Obstacles: walls, furniture and floors can reduce signal power
- Access point placement: central, elevated positions are ideal. Avoid placing routers near thick walls or metal objects

Signal strength is measured in dbm (decibels milliwatts)

- -30 dbn = excellent signal
- -50 to – 60 dbn = good signal
- -70 dbn or lower = weak or unreliable signal

Roaming and coverage extension

In larger areas, a single access point may not provide enough coverage. This can be remedied with wireless repeaters/extenders, mesh Wi-Fi systems, where multiple access points work together, sharing the same SSID for seamless roaming, and wired backhaul, which is mesh nodes connected via ethernet.

5 GHz band

6.2 Wi-Fi and long ranged fixed wireless

Wi-Fi standards are defined by the IEEE 802.11 family. Over time, each new version of 802.11 has improved on speed, frequency support, range and technologies. Most Wi-Fi standards operate on either the 2.4 ghz band, the 5 ghz band, or both. Newer standards also support 6 ghz.

Wi-Fi

IEEE Standard	Year of Release	New Name	Data Rate	Frequency Band	Range (Indoors; Outdoors)
802.11a	1999	Wi-Fi 1	54 Mbps	5 GHz	35m; 120m
802.11b	1999	Wi-Fi 2	11 Mbps	2.4 GHz	35m; 120m
802.11g	2003	Wi-Fi 3	54 Mbps	2.4 GHz	38m; 140m
802.11n	2009	Wi-Fi 4	600 Mbps	2.4 GHz and 5 GHz	70m; 250m
802.11ac	2013	Wi-Fi 5	1.3 Gbps	5 GHz	46m; 92m
802.11ax	2019	Wi-Fi 6	Up to 10 Gbps	2.4 GHz and 5 GHz	9.1m

802.11b

- This was the first widely adopted standard
- Slow but had good range
- The 2.4 ghz band means this standard had to deal with interference from devices such as microwaves and Bluetooth networks

802.11a

- This improved on 802.11b by using the 5 ghz band, meaning it was faster and had less interference, but at the cost of range

802.11g

- This combined the speed of a with the frequency of b (2.4 ghz)
- Backwards compatible with b due to using the same frequency

802.11n

- The first standard to use both 2.4 and 5 ghz bands
- Introduced MIMO technology (multiple input multiple output), boosting the speed
- Used up to 4 antennas

802.11ac

- 5 GHz only
- Introduced MU-MIMO (multi user MIMO) and beamforming
- Much faster due to MU-MIMO

802.11ax

- Works on 2.4, 5 and 6 ghz
- Introduced OFDMA (orthogonal frequency division multiple access)
- Better performance with many devices at once due to OFDMA
- Backwards compatible with all previous standards

Long range fixed wireless

LRFW is a separate category of wireless comms. Its designed for long distances rather than local networking. While it shares some similarities with Wi-Fi, it operates on different standards, regulations and use cases.

LRFW is a high speed wireless connection over long distances without needing physical cables. Its used in rural areas where broadband is needed but fibre isn't available, point to point links between buildings, cell towers or business locations, and back up networks for when fiber or wired connections fail. Its designed for stability over very long distances, so power and frequency are more tightly regulated.

Licensed vs unlicensed

Unlicensed wireless frequencies can be deployed without government approval, but they come with trade offs. They're easier to deploy and cheaper but have more interference (because multiple users share the same frequencies) and have limited power output due to government restrictions. Unlicensed use 2.4 and 5GHz like Wi-Fi, and in the case of IoT and some rural broadband 900MHz. ultra fast ones in point to point links use 60Ghz.

Licensed wireless frequencies require govt approval. They have no interference from other networks and high power limits allowing for stable long distance links. Its used in cell towers, emergency services and critical infrastructure. Its expensive because it requires licensing fees and approval and slower to deploy because of the licensing process.

Licensing bands include the 3.5GHz (for private LTE and 5g), 28 and 39GHz (for high speed wireless backhaul), 6,11 and 18GHz for telecom backhaul and enterprise fixed wireless.

A city would use a secure wireless network for emergency responders to use as there will be no interference.

Wireless power limits and regulatory requirements

Maximum power output of wireless systems is strictly regulated to prevent interference. These are usually set using EIRP (effective isotropic radiated power) which factors in both transmitter power and antenna gain. Licensed bands generally allow higher power. The limits exist to prevent interference, ensure safety (too much power can cause frequency exposure risks), and to protect licensed spectrum users from unlicensed interference.

Bodies that regulate wireless power and spectrum:

USA – FCC

Uk-Ofcom

France – ANFR

Australia – ACMA

6.2 Wireless security

Wireless networks introduce security risks that wired networks would not. Because Wi-Fi signals extend beyond physical boundaries, attackers can intercept, manipulate or disrupt comms. Understanding wireless security protocols and best practices is essential for securing networks.

Definitions

- SSID (service set identifier): a Wi-Fi network name that devices use to ID and connect to a wireless network. Every wireless network has an SSID, which is broadcasted by the router or access point
- Encryption: the process of scrambling data so that only authorized devices can read it

- Authentication: the process of verifying that a user or device is allowed to connect to a Wi-Fi network. Methods of authentication include PSK, RADIUS, SAE
- Open network: a Wi-Fi network that does not require a password to join. They don't involve encryption, so they're insecure and easy to intercept
- Captive portal: a internet page that appears when you try to connect to certain public Wi-Fi networks where you must enter a password, accept terms, or log in before accessing the internet
- MAC filtering: a security feature that allows or blocks devices from connecting to a Wi-Fi network based on their MAC address. Not foolproof, because MAC addresses can be spoofed

Wireless security protocols

These are security protocols designed to encrypt and protect Wi-Fi comms. Over time, newer and stronger encryption methods have replaced outdated ones due to vulnerabilities.

WEP (wired equivalent privacy) (weakest)

This was the first Wi-Fi security protocol, introduced in 1992. It used RC4 encryption with a static key, meaning it was the same key for all devices, it was vulnerable to IV (initialisation vector) attacks, allowing multiple attackers to crack the key in minutes. Its not secure and shouldn't be used.

- Oldest and weakest
- Uses a static encryption key, meaning all devices on the network share the same key
- Easily cracked with modern tools so is highly insecure
- No longer recommended and almost all device have dropped support for it

WPA (Wi-Fi protected access) (better)

This was introduced as a temporary fix for WEP, and uses TKIP (temporary key integrity protocol) for better encryption. Its stronger than WEP but still vulnerable to attacks.

- Introduced as a quick fix for WEP's flaws
- Uses TKIP, which dynamically changes encryption keys to improve security
- Still weak by todays standards

WPA 2 (Wi-Fi protected access 2) (strong)

This is the most commonly used security standard today, and uses AES (advanced encryption standard) for strong encryption. It can operate in two modes:

- WPA2 personal (PSK), which uses a shared password
- WPA2 enterprise (802.1x), which uses a radius server for authentication

Need to know

- Most commonly used standard today, until WPA3 takes over
- Uses AES, which is stronger than TKIP
- Used in most modern routers

WPA3 (Wi-Fi protected access 3) (strongest)

This is the newest security protocol and uses SAE (simultaneous authorization of equals) instead of PSK, making it resistant to brute force attacks. It protects unencrypted public Wi-Fi by encrypting each users traffic, and provides stronger defence against deauthentication attacks.

Need to know

- Latest standard, replacing WPA2
- Uses SAE instead of PSK
- Prevents offline password cracking
- Offers better security on public networks
- Slowly becoming common in modern routers

Authentication methods

Wi-Fi networks require authentication to prevent unauthorized access. There are two main methods.

PSK (pre shared key) authentication

This Is used in homes and small offices, where a single password is shared by all users. Its weakness is that if the password is leaked, anyone can connect.

Enterprise authentication (802.1X + RADIUS)

This is used in businesses and organization where instead oif a shared password, users authenticate with individual credentials, it uses RADIUS server to verify users and is more secure, as passwords can be managed per user.

Common wireless security threats

Evil twin attack

This is where an attacker creates a fake Wi-Fi network with the same SSID as a legit one. Users unknowingly connect to the fake network, allowing the attacker to steal login credentials and intercept data.

Deauthentication attacks

The attacker sends fake deauthentication frames to disconnect a device from Wi-Fi. Often used in password cracking attacks. WPA3 protects against this.

War driving

This is here hackers drive around scanning for open or weakly secured Wi-Fi networks using tools like kismet to map vulnerable networks.

WPS (Wi-Fi protected setup) attacks

WPS is a feature that allows easy devices connection via a PIN or button press. Hackers can brute force the PIN, and gain network access. Best practice is to disable WPS

Rogue access points

A rogue AP is an unauthorized wireless AP on a network. Employees may set up their own Wi-Fi without security, creating an entry point for attackers.

Packet sniffing

Attackers use tools like Wireshark to capture unencrypted wireless packets. Encrypting traffic with WPA2 or WPA3 prevents this.

Best security practices for wireless security

- USE WPA2 or WPA3 encryption
- Change default SSIDs and passwords as attackers target default settings
- Disable WPS, preventing brute force PIN attacks
- Enable MAC address filtering, restricting Wi-Fi access to known device MAC filtering
- Hide SSID broadcasting as this prevents people from seeing the Wi-Fi network
- Implement firewalls and intrusion detection systems; these prevent unauthorized traffic from entering the network

6.3 Bluetooth

While Wi-Fi is designed for high-speed, long-range networking, Bluetooth is designed for short range, low power comms between devices. It operates using radio waves in the 2.4GHz band just like Wi-Fi, but instead of sending data over a large area, Bluetooth is optimized for device-to-device connections within a short range (usually 10 meters or less). Because they both use the 2.4 band, Wi-Fi and Bluetooth can interfere with each other. The up-to-date version is Bluetooth 5.3 (2022). It reduces latency in gaming

headsets, improves audio quality and makes connections more stable despite interference. To solve this (if router allows for it), you can change the router to the 5 GHz band so you don't get interference.

When to use Wi-Fi vs Bluetooth?

If you need high speed internet, Wi-Fi is the better choice. But if you just want two devices to connect without the internet, Bluetooth is the one. Bluetooth is perfect for low power, stable audio connections. Things like smart watches that need constant connection without draining the battery too.

6.4 GPS, RFID and NFC

GPS (global positioning system)

GPS is a satellite based navigation system that provides location and time info in all weather conditions, anywhere on earth, 24 hours a day. It is widely used in wireless networking and mobile devices to track locations, facilitate navigation and provide location based services.

It relies on a constellation of 24 to 32 satellites, which use signals to track the earths position and the exact time each signal was sent. Earth stations use the time relay to track the distance to each satellite. By doing this from multiple satellites the GPS receiver can calculate its precise position.

RFID and NFC

RFID

RFID (radio frequency identification) is about wireless tagging and tracking. Its used in things like inventory tracking, toll passes and pet tagging.

How does it work?

RFID tags contain a small chip that stores data and RFID readers scan radio waves from a distance. The tag doesn't need power, it gets activated by the readers signal (passive RFID)

NFC (near field communication)

NFC is a type of RFID that only works at very short distances, like 4cm or less. Its designed for secure and quick interactions like touch payments.

How does it work?

NFC chips store small amounts of data. When two NFC enabled devices come close, they exchange data. Like with RFID, the NFC tags don't need power as they are powered by each other.

Bluetooth, RFID and NFC summary

Bluetooth is for wireless gadget connections. RFID is for long range, automatic scanning like store security tags and microchips. NFC is for short range secure transactions.

6.5 Wi-Fi antenna types

Wi-Fi antennas come in different shapes and designs, each suited for specific network setups. The key differences are how the signal is distributed and the range they cover.

Omni directional antennas

These have 360 degree signal coverage around the antenna. They're ideal for home networks or small offices where you need a broad, even coverage area. They're easy to set up and provide coverage in all directions, but have a shorter range compared to other types.

Directional antennas

These focus signal in a specific direction, providing a narrow beam. They're used in long distance connections, such as between buildings. They're ideal for point to point connections but require precise alignment.

MIMO antennas

These use multiple antennas to send/receive data simultaneously. They're used by Wi-Fi 5 and 6 (ac and ax) for improved performance as they have high data throughput and better network efficiency in busy networks. The only downside is that they require compatible devices and have a higher cost.

Parabolic antennas

Provide an extremely narrow, focused beam for very long distance connections

Sector antennas

Focus signal in a specific sector such as from cell towers.

7. Network services

Servers play an essential part in networking. They provide services to devices and the network such as the provision of IP addresses, a logbook with translations of IP addresses to website names, security services, centralised stores of mail and files, and more.

7.1 DHCP

DHCP overview

DHCP automates the process of assigning IPs to a network. When a device is set to dynamic and not static IP settings, the DHCP server hands the device an IP, subnet mask, default gateway and DNS server info. This saves time and avoids IP conflicts.

- Port 67 is used for the server receiving DHCP requests
- Port 68 is uses by the client for receiving DHCP responses
- UDP is typically used for this as it doesn't require a connection

The device's address is on lease; it is a temporary assignment of an IP to a device (client) on a network. The lease is granted by the server to the client for a specific period of time. When a client connects to the network (if DHCP enabled) the server allocates an IP from its scope for a designated time, after which the lease must be renewed if the device is still active on the network.

Lease process:

- Discovery; device sends a broadcast message to discover an available DHCP server
- Offer; DHCP server responds with an offer, including an available IP and other configuration details
- Request; the device then requests the offered address
- Acknowledgement; the server sends an acknowledgement confirming assignment of the IP

This process is also known as DORA.

Reservations

A DHCP reservation is a mechanism by which a device on the network always receives the same address. It's useful for important devices such as routers, printers or server that require a constant IP for reliable comms. Its set up on the DHCP server and the reservation is associated with a MAC address.

Scope

This is the range of Ips a server can allocate. This also includes the subnet mask, default gateway and DNS servers. The scope includes lease time, IP addresses excluded from the scope, and reservations.

Troubleshooting

Sometimes a device might have an IP address starting in 169. This means it attempted to reach a DHCP server, but couldn't receive an address from it, so gave itself an APIPA

address. These addresses can only be used to communicate with devices on its local network. Common solutions may be:

- **Disabled DHCP server**: check on the server if DHCP is enabled, and on the pc. You can also check a PC's IP settings via ipconfig and checking if DHCP is enabled on the appropriate adapters.
- **DHCP scope**: this defines the range of Ip addresses available for assignment. This can be checked on the DHCP server console. Make sure the start and end IP ranges are right and that there are enough available IPs for the networked devices. If the range is exhausted and there are no available IPs, consider expanding the scope by increasing the IP range.
- **Address conflicts**: multiple devices on the network may have the same IP. This can be checked by pinging an address you suspect has a conflict. If two devices respond then there is a conflict. You can also check with ARP -a for IP conflicts as the mac resolution table aligns physical and internet addresses. Once you've found which devices have conflicts, they can be given new addresses using ipconfig/release and then ipconfig /renew.

7.2 DNS

DNS is the phonebook of the internet. It translates human readable names into IP addresses so pcs can understand it. It uses port 53 UDP for speed but can use 53 TCP if the request is too large.

If an internet site isn't loading it may be a DNS issue. You can check this by running nslookup followed by the internet site name. if you receive an IP for the site, it's not a DNS issue. If you don't, then the issue may be DNS. I could be any other number of connectivity issues, so a ping is a valid tool here too.

DNS records

A and AAAA

DNS uses records to map domain names to IPs. One such is an A record (address record). It links a human readable name to an ipv4 address. It's the most common form of DNS and used for most websites and internet services. AAAA is what's used for IPv6 addresses. They work the same, but one uses A records and one uses AAAA records.

MX

Mail exchanger (MX) is another type of DNS record, used to route emails to the correct server for a domain. It's a key component in the email delivery process. An MX record specifies the mail servers responsible for receiving email messages on behalf of a domain. When you send an email to someone, your email server uses the recipient's

domain to query DNS for the corresponding MX record. This record tells the senders mail server where to deliver the message.

TXT

A TXT record is another type of DNS record. This allows domain owners to store text-based info within the DNS database. These records are used for verification, security and configuration purposes. They provide a way to associate text with a domain name, allowing external services to check for specific info about that domain.

Summary:

- A record – maps a domain to an IPv4 address
- AAAA – maps a domain to an IPv6 address
- MX record – specifies the mail servers responsible for receiving emails on the domain
- TXT record – holds plain text data; commonly used for SPF, DKIM, and DMARC

Spam management

SPF, DKIM and DMARC are DNS based mechanisms that help ensure the authenticity and integrity of email messages. They verify if the email is really from the sender it claims to be from, preventing malicious activities like fraudulent emails and impersonation. These records are stored in your domains DNS settings, meaning that DNS servers are responsible for checking these records when an email is received.

SPF (sender policy framework)

This is a DNS record that helps to verify the senders email server. It allows the domain owner to define which mail servers are authorized to send emails on behalf of their domain, to help prevent email spoofing, a common tactic in phishing attacks, where an attacker sends an email that appears to come from a legitimate domain.

DKIM (DomainKeys identified mail)

This DNS based security mechanism focuses on ensuring the integrity of an emails content. It uses cryptographic signatures to ensure that the content of the email has not been altered during transit. When an email is sent, the senders mail server adds a DKIM signature to the email headers, which is verified by the recipients server using a public key stored in the DKIM record in DNS.

DMARC (domain-based message authentication, reporting and conformance)

DMARC is a policy framework built on top of SPF and DKIM. It allows domain owners to specify how emails that fail the SPF and DKIM checks should be handled. In addition to providing a policy, DMARC also allows the domain owner to receive reports about emails that fail authentication checks.

Summary

- SPF checks if the senders IP is allowed to send mail for the domain
- DKIM adds a digital signature to verify the email hasn't been tampered with
- DMARC uses SPPF and DKIM results to decidewhat to do with suspicious mail (such as reject or quarantine) and sends reports to the domain owner

7.3 Other server roles

File share servers

File share services allow devices on a network to remotely access files. They use protocols like SMB (server message block) and NFS(network file system) depending on the OS.

SMB uses port 445 and works well for file sharing in windows-based networks. NFS uses port 2049 for file sharing in Linux/Unix environments.

Troubleshooting

If users can't access shared files:

- Ensure the file sharing server is running
- Check the firewalls isn't blocking the necessary ports
- Ensure correct permissions are set on shared folders for relevant users

Print servers

These manage printing tasks over a network, acting as a middleman so devices don't have to directly connect to the printer. This makes it easier to share a printer across a group of pcs. This role is often built into the printer.

Ports 515 and 613 are commonly used.

Troubleshooting

- Check if the printer service is running
- Ensure the printer is properly connected to the printer
- Check for IP conflicts
- Try printer specific problems and solutions

Mail servers

Mail servers manage the sending, receiving and storing of emails. They use protocols like SMTP (25) for sending, pop3 (110) or IMAP (143) for receiving.

Troubleshooting

If emails aren't sending or receiving:

- Check if the mail server is running
- If outgoing mail isn't working, there may be an issue with port 25, or an SMTP configuration issue
- If emails aren't being received, there may be an issue with 110 or 143
- Server Ips also need to be checked for issues

Syslog servers

These servers collect and store log messages from various devices on the network such as switches, routers and firewalls. The logs are essential for troubleshooting and security auditing. Syslog comms generally use 514 UDP, although TCP is an option for more reliable log transmission.

Troubleshooting

Sometimes logs wont appear in the syslog server. In this event:

- Make sure the syslog service is configured to receive logs from network devices
- Make sure devices are configured to send logs to the correct syslog server IP
- Check that the firewalls allow traffic through port 514

Internet servers

An internet server is a software and hardware system that serves internet content like HTML files, images and scripts to users over the internet using HTTP and HTTPS. When a user types in a URL to their internet browser, a request is sent to the server to retrieve the internet page. The server processes the request and returns the requested content to the browser.

They server content like pictures, text files, videos (static content), scripts (dynamic content), server-side requests such as getting data from databases.

Popular internet servers include Apache HTTP, NGINX.

Common ports are 80 (unencrypted) and 443 (encrypted).

Troubleshooting

If a website isn't loading (and it's not a DNS issue)

- ensure the internet server software is up and running
- check on the firewall settings that ports 80 and 443 are open and allowing inbound traffic
- use the ping and curl commands to check connectivity with the server

if it's an HTTPS site, then ensure the SSL/TLS certificate is valid and not expired using online tools. Also make sure the SSL config is set up properly.

You might also get error messages

- 403 forbidden error: check internet server has proper read permissions for the requested files
- 404 not found error: ensure that the requested file exists at the expected location in the servers document root. Also check the server is configured correctly to root URLs to the appropriate files
- 500 internal server error: review internet servers error logs for specific details on what went wrong

File share servers

These allow multiple users to access and share files over a network. These are essential in environments where users need to collaborate and share resources, such as in an office or small business setting. The server manages files, permissions, and user access.

Protocols involved in file sharing

SMB (server message block)

This uses port 445, allowing for file sharing and printer sharing among other things, in windows environments. It can also be used for remote file access and network browsing. SMB enables devise to share files on the network, and it supports advanced features like file locking, caching and security mechanisms for file access.

NFS (network file system)

This uses port 2049. It's like SMB but for Unix and Linux environments. It allows a system to access shared files on a network just like local storage.

Problems and solutions

Common issues

- **File access denied**: usually due to incorrect permissions
- **Slow file transfer**: normally caused by network congestion or hardware limitations
- **Network visibility issues**: sometimes clients can't see network files due to firewall settings, IP issues, subnet mismatches

Troubleshooting

- Make sure the server is running. This can be done in server manager
- Check permission for users

- Verify port accessibility; ensure the appropriate ports aren't blocked; 2049 in Linux/Unix systems, 445 in windows
- Check network configurations to make sure device and server can communicate
- Check for file locks
- Check for DNS issues

7.4 AAA

AAA stands for authorization, authentication and accounting. It's a conceptual framework that helps to manage and secure access to network resources. AAA utilizes various protocols and practices to ensure that only authorized users gain access to network resources, their actions are properly monitored, and usage is accurately tracked.

Authentication

This verifies the identity of users or devices that attempt toa access the network. The goal is to ensure that they are they are who they say they are. Methods for authentication include username and password, biometrics, 2 factor authentication and certificates.

Authorization

This involved permissions and privileges, based on role within the organisation. Role based access lists (RBALs) assign access based on user's role, such as admin or guest. Access control lists (ACLs) define what each user can access, whether its files, apps, or network devices.

Accounting

This involved tracking the actions of authenticated users to ensure compliance, and for the sake of troubleshooting. It includes a log of what users accessed, what actions they took, how long they were on the network. This is essential for security audits, billing and troubleshooting.

Common AAA protocols

RADIUS (remote authentication dial in user service); this is typically used in authentication such as for Wi-Fi access or VPNs, as well as remote network access. It sends authentication requests where user credentials can be verified.

TACACS+ ; terminal access controller access control system plus); this is typically used for management of network devices. Its more secure than radius because it encrypts comms, and separates the AAA process into their individual aspects, giving more control over each.

LDAP (lightweight directory access protocol); this is used mainly for accessing and managing directory services, like AD. It helps with centralizing user and resource management so its easier to implement AAA functions. It helps in the authentication process as requests are sent to its server and it checks them against its saved data. It doesn't manage authorization, but it does store user info such as permissions or groups a user belongs to.

Summary

Together, these elements ensure that network security is tightly controlled and that network administrators can track usage, enforce policies, and respond to unauthorized access or suspicious activity.

7.5 network performance and optimization concepts

A well-functioning network is more than just connected – it needs to be responsive, consistent and efficient. In this chapter, well look at behind the scenes technical factors that influence network performance, such as packet handling, traffic prioritization and device negotiation. These concepts are essential for both identifying performance issues and optimizing networks for demanding workloads like gaming, VoIP, or video streaming.

Latency (delay)

This is the time it takes for a packet to travel from source to destination.

- Measured in milliseconds
- High latency results in laggy connections in VoIP, gaming and streaming
- Affected by distance, congestion, routing, and processing delays

How to reduce latency

- Use faster network routes and optimized routing protocols
- Upgrade to fiber optic cables (these have lower latency than copper)
- Reduce network congestion by balancing traffic

Bandwidth

This is the maximum capacity of a network link to transfer data

- Measured in bps (bits per second), Mbps, or Gbps
- Doesn't measure the actual speed – only the potential data flow
- Affected by network congestion, interference, and hardware limitations

Optimizing bandwidth

- QoS (see page 55) – prioritize critical traffic like VoIP and video conferencing

- Traffic shaping and throttling – limit bandwidth usage for non-critical applications
- Upgrade network infrastructure – use higher speed connections where possible

Throughput vs bandwidth

Bandwidth is the theoretical maximum rate of data transfer. Throughput is the actual amount of data successfully transferred over time. Throughput is affected by:

- Protocol overhead (TCP headers, encryption)
- Interference from collisions
- Packet loss and retransmissions
- Hardware bottlenecks

You can have a 100 mbps connection and still experience poor throughput due to these factors.

Jitter

This is the variation in latency between packets and is critical in real time applications such as VoIP and video calls. It causes choppy audio, buffering and lag.

Reducing jitter

- Implement QoS to prioritise VoIP and streaming traffic
- Use a dedicated internet connection for real time apps
- Upgrade to better routers that handle packet buffering efficiently

MTU (maximum transmission unit)

The mtu defines the largest size (in bytes) of a single packet that can be transmitted over a network segment. If a packet exceeds the mtu, it may be fragmented (split into smaller parts) or dropped entirely if fragmentation is not allowed.

- Typical mtu size for ethernet is 1500 bytes
- Vpns and tunnelling protocols add overhead, which may require reducing mtu to prevent dropped connections
- Incorrect mtu settings can result in slow performance, especially over vpns or when connecting to cloud services

Troubleshooting mtu settings often involves ping tests with specific sizes and the Do Not Fragment flag to find the largest allowable packet size.

QoS (quality of service)

QoS is a way to prioritise certain types of traffic over others to ensure that time sensitive data – like VoIP, video conferencing, or gaming – gets through with minimal delay.

- Routers and switches can identify traffic based on type (voice, video, bulk file transfer)
- QoS rules can be applied to give higher priority to voice packets, for example, and to lower torrent downloads
- On consumer networks, QoS is often configured via router settings, but business networks may apply it via managed switches or firewalls.

Limitation: QoS becomes less effective when traffic is encrypted (as in vpns), since the router cant see what type if data is being sent.

Duplex settings

Duplexing controls how data flows between two network devices:

- Half duplex: data can only flow in one direction at a time (like in a walkie-talkie)
- Full duplex: data can flow both ways simultaneously (like in a phone call)

Most modern devices use full duplex, but mismatches (one side set to full, the other set to half) can cause:

- Collisions
- Retransmissions
- Slow connections

These issues often appear in improperly configured switches, or when older hardware is mixed with newer gear.

Link speed negotiation

Network devices auto-negotiate the maximum supported speed (1000mbps, 1gbps etc) and duplex mode when a connection is established. However, this negotiation can sometimes fail or be misconfigured manually, causing:

- Link speeds to fall back to lower settings
- Duplex mismatches
- Connectivity that appears stable but performs poorly

Its important to verify both ends of the connection (e.g., NIC and switch port) are set to either auto-negotiate or matching manual speeds.

Packet loss

Packet loss occurs when one or more packets of data fail to reach their destination. Causes include:

- Network congestion
- Faulty cables or hardware
- Wireless interference

- MTU related issues

Symptoms include:

- Choppy audio/video
- Sluggish performance
- Timeouts or failed connections

Tools like ping, tracert, or packet capture utilities can help identify where the loss is occurring.

ICMP blocking and rate limiting

Tools like ping and tracert rely on ICMP (internet control message protocol) to test connectivity and trace routes. However, some firewalls or routers block ICMP or rate limit responses to avoid abuse.

This can cause confusion when:

- A device is online and functioning, but doesn't respond to ping
- Traceroute appears to fail mid path but the final destination is reachable.

Be aware that ICMP blocking may be the cause of devices being unresponsive.

NAT and double NAT

Network address translation (NAT) maps private Ip addresses to a public ip. In some cases, multiple layers of nat (such as one router behind another) create a double nat situation.

Issues with double nat:

- Difficulty hosting game servers or accessing external services
- Vpn conflicts
- Port forwarding complications

Solutions include bridge mode, DMZ configuration, or using UPnP (universal plug and play) to allow dynamic port assignments.

DNS resolution speed

While often overlooked, slow or unreliable DNS servers can lead to:

- Delayed website loads
- 'page not found' errors
- Perceived slow internet

DNS benchmarking tools can help identify which dns provider offers the fastest response times in your region. Switching to provides like google or Cloudflare can

improve perceived performance, especially if your ISP's DNS servers are overloaded or slow.

Network monitoring tools

To track network performance, you can use monitoring tools such as

- Ping – this measures latency and checks connectivity
- Tracert (windows) / traceroute (Linux) – tracks the path of packets through the network
- Netstat – displays active network connections
- Wireshark – analyses packet traffic in detail
- SNMP (simple network management protocols) – monitors devices

Troubleshooting network performance issues

7 identify the issue using tools like ping and tracert to check connectivity
8 check network traffic by looking for bandwidth heavy applications
9 analyse latency and jitter by monitoring VoIP or streaming performance
10 optimize routing and QoS, prioritizing critical traffic

upgrade network infrastructure by replacing outdated cables, switches or routers

7.5 Internet appliances

Spam gateways

This is a security appliance that filters and blocks spam emails before they reach the users inbox. It checks keyword, suspicious attachments, blacklisted senders and blocks appropriately. They use spam filters for this, and the process is integrated with email security processes like anti-virus and anti-malware.

Unified threat management

This is the multi-layer security of a network. It includes firewalls, intrusion detection, anti virus, spam filtering and content filtering. This is a single appliance where all of these functions are done. It sits between the internal network and the internet, so it's a single point of entry for all security services in the network, and all internet traffic has to pass through it to get into the network.

Load balancers

This distributes network or app traffic across multiple servers to ensure no single server is overwhelmed, optimizing resource usage, improving response times, and increasing overall availability. They work by monitoring server health and traffic loads, and directing traffic to the least congested or healthiest server. There are three types: round robin, least connections and IP hash.

Proxy servers

This acts as an intermediary between a user's device and the internet. It requests content on behalf of the user, filters requests, caches data and hides the users IP to enhance security, privacy and performance. You have forward and reverse proxies, with forward being the ones that send the users requests to websites, and reverse receiving client requests on the internet server side. Proxy servers act as security too by filtering out malicious sites or content, and ensuring all requests are SSL encrypted

7.6 Legacy/embedded systems, SCADA and IoT

Legacy systems refers to older computing systems, apps or hardware still in use that might be outdated or less effective by modern standards. They continue to operate due to their reliability, effectiveness and the cost of upgrading. They are often deeply integrated into a system.

Embedded systems are specified pc systems designed to perform a specific task or function within a larger system. They're typically not programmable and are optimized for their specific function.

Supervisory control and data acquisition (SCADA)

This is a control system architecture that is used to monitor and control industrial processes and infrastructure. Found in power grid, water treatment, oil and gas. They collect real time data from sensors and devices on the field and feed that data to a central supervisory system.

Internet of things (IoT)

This refers to a network of connected devices that communicate with each other and the internet. They are embedded with sensors, software and various techs to collect and exchange data. This might be smart home devices such as thermostats and lighting systems to industrial equiptment like huge sensors and smart meters. They communicate through Wi-Fi, Bluetooth, cell networks, among others. Their data is sent to the cloud, where it can be stored, processed and analysed. This data can then trigger automatic decisions such as adjusting temperature based on weather conditions.

They are often less secure than traditional computing devices, with weak passwords and unencrypted comms, so changing default credentials and encrypting comms is essential.

7.8 VLAN and VPN

VLAN(virtual area network)

This is a logical subdivision of a physical network. It allows you to create separate broadcast domains within the same physical infrastructure. The purpose of VLANs is to group devices based on function, department or project, regardless of their physical location in the network.

Benefits of VLAN:

- Security; by isolating devices you can limit access between different departments or users
- Performance; it reduces broadcast traffic because devices in different VLANs don't receive the broadcast traffic of other VLANs
- Flexibility; devices can belong to the same VLAN regardless of their physical location on the network

VLANs operate much like a typical subnet defined broadcast domain. A gateway has an address which devices in its VLAN use to forward data packets to for other VLANs. The difference is that this gateway is virtual, set up on the routing device (whether that's a layer 3 switch or a router).

VPN

A VPN, or virtual private network, is a secure connection that creates a private tunnel over a public network – most often the internet. It allows remote users, branch offices, or traveling employees to access internal resources as if they were directly connected to the private network. In personal use, its also commonly used to hide browsing activity, mask IP addresses, or bypass geo-restrictions.

How a VPN works

When a VPN is active, your device establishes a secure connection to a VPN server. All traffic from your device is then encrypted and routed through that VPN server before reaching its destination.

From the outside, all traffic appears to be going to the VPN server. Inside the tunnel, your data is routed as usual, but now its protected from snooping or interference.

VPN use cases include:

- Remote work: securely connect to company resources
- Bypassing censorship or geo-blocks: appear to be in another country
- Public Wi-Fi protection: encrypt your traffic on insecure networks
- Privacy: hide your IP and activity from your ISP or trackers

VPN uses various protocols:

- PPTP (older, fast but insecure)
- L2TP/IPsec (better security, more overhead)

- OpenVPN (widely used, flexible)
- IKEv2/IPsec (stable on modern devices)
- Wireguard (modern, simple, fast)

Each uses different encryption methods and handles tunnelling in different ways.

Potential issues when using a vpn

- Quality of service limitations: some routers prioritize a type of traffic over another; because VPN traffic is encrypted, its harder for the QoS protocol to identify and prioritize it, so VPN traffic may not benefit from QoS rules, leading to lag or jitter
- Latency typically increases using a VPN because the VPN server introduces an extra server to the data path

8. Virtualisation

8.1 Virtualisation basics

What is virtualisation?

At its core, virtualisation is the process of creating a virtual version of something – typically a computer system. Instead of running one operating system directly on physical hardware, virtualisation allows us to run multiple virtual machines (VMs) on the same hardware, each with its own OS and applications.

Think of it like this: you've got one powerful computer, but inside it, you've created several 'pretend' computers. Each of these VMs thinks it has its own cpu, memory, and storage, but really they're all sharing the same hardware underneath.

What are the benefits?

Virtualisation is a game changer. Here's why:

- Cost efficiency: fewer physical machines = less hardware to buy and maintain
- Isolation: each VM is separated from the others. If one crashes, the others are unaffected
- Resource optimisation: you can run multiple workloads on one system, maximising cpu and memory use
- Snapshot and recovery: you can take snapshots of VMs to roll back if something goes wrong
- Scalability: easy to clone, deploy, or move VMs when demand changes

Hypervisors: the core of virtualisation

A hypervisor is the software layer that enables virtualisation. It manages the physical hardware and creates virtual environments.

There are two types:

- Type 1 hypervisor (bare metal): runs directly on hardware. No OS in between. Examples would be VMware and Microsoft hyper V. They're fast, efficient, and ideal for data centres and enterprise use
- Type 2 hypervisor (hosted): runs on top of an existing operating system. Examples would include VirtualBox and VMware workstation. They're easier for individuals or small-scale use.

Real life usage

You'll find virtualisation used all over the place in the real world:

- IT departments use it to test software in different environments
- Businesses run multiple servers (mail, web, file storage) on a single machine
- Cloud providers like AWS and Azure run thousands of VMs for their clients
- Developers use it to simulate different operating systems on one machine

Server based vs client-based virtualisation

Theres a key difference in where the virtualisation runs – and it affects performance, security, and use cases.

Server based applications (terminal based)

- The app runs on a central server
- Users access it via a thin client (like remote desktop)
- Example: a hospital with one server running patient record software, accessed by many terminals
- Benefits include centralised updates and backups, better security as data stays on the server, and lower client hardware requirements

Client based applications (local VMs)

- The app runs on the user's device via a type 2 hypervisor
- Each user has a fully virtualised OS locally
- Good for development and offline use
- More control per user

8.2 Containerisation

What is containerisation?

Containerisation is a lightweight way to run software in isolated environments, called containers. Think of a container like a self-contained box that holds everything an application needs to run: code, libraries, and settings – but not an entire operating system.

Where a VM simulates a full OS for every instance, a container shares the host OS kernel but still feels like a separate system to the app inside it.

So, what you get is a fast, lightweight, and portable unit that can run just about anywhere.

How does it work?

Containers rely on features built into modern operating systems – mainly namespaces and control groups (CGroups) – to isolate processes and manage resources. Instead of emulating hardware like virtualisation does, containers run directly on the host OS. They have their own file system, processes, and network interfaces, but they share the same underlying OS kernel.

This makes containers:

- Much lighter than VMs (they start in milliseconds)
- Easier to scale
- Faster to deploy and update

Popular container tools include Docker, Podman and Containerd.

Why do we use it?

We use containerisation because it solves a lot of real-world deployment problems:

- Containers behave the same no matter where you run them, so they'll run the same on any machine
- Quicker development and testing cycles
- Easier automation for CI/CD (continuous integration and delivery)
- Much lower resource use than running full VMs
- Easier to isolate microservices in modern app design

Weak point: the underlying OS

Here's the trade-off: while containers are lightweight, they're also more dependent on the host OS than VMs.

All containers on a system share the same OS kernel. So, if there's a vulnerability in the host OS, all containers would be affected. That's why in high security environments, containers are often combined with other protective tools like sandboxing or even used inside lightweight VMs for extra isolation.

Virtualisation vs containerisation

Feature	Virtualisation	Containerisation
OS kernel	Each VM runs its own OS	Containers share host OS kernel
Resource usage	Heavy – simulates full hardware	Lightweight – no hardware emulation
Startup time	Slow – boots full OS	Fast – starts in milliseconds
Isolation	Strong (full OS per VM)	moderate (isolated processes only0
Use cases	Ideal for different OS types, security	Great for microservices, portability

Think of it like this:

- Virtualisation gives you full apartments, each with their own infrastructure
- Containers give you rooms in the same apartment – separate, but sharing core systems

Summary

Containerisation is about packaging apps and dependencies into small, fast, portable units. It relies on the host OS kernel, making it more efficient but less isolated than VMs. It's perfect for modern cloud native development, especially in environments that value speed, scalability and automation.

8.3 Purposes of virtual machines

Introduction

Virtual machines are a foundational part of modern computing. From cloud services to software testing, they allow multiple isolated systems to run on a single physical machine offering flexibility, efficiency and security.

Hypervisors recap

- Type 1 hypervisors run directly on the hardware (bare metal), without an OS in between. As there's no intermediary OS, there's a smaller attack surface – making bare metal hypervisors more secure and reliable for production environments
- Type 2 hypervisors run on top of a host OS. This means they're easier to set up, but are more vulnerable since they rely on the host OS, which can be a point of attack

Containerisation recap

While not the same as VMs, containerisation is closely related. It uses less resources because it doesn't require a full OS per instance – instead, containers share the host OS kernel. This means they're more lightweight and faster, but offer slightly less isolation than VMs.

Virtual machine usage

Cloud computing

Modern cloud services (like AWS, Azure, GCP) rely heavily on VMs. Each customer or service can run on its own VM, isolated from others, but hosted on shared physical infrastructure.

- Physical space is saved by using fewer actual servers
- Higher availability is achieved through redundancy and failover systems built on virtualised hardware
- Dynamic resourcing allows VMs to scale up or down based on demand

Hyper converged infrastructure

This is a system where computing, storage and networking are all virtualised and managed as a single system – often using VMs at the core. It simplifies infrastructure and boosts flexibility. Might be used in data centres, remote business branches, virtual desktop.

Application virtualisation

This allows individual programs to run independently on the host of the OS.

- Software is encapsulated and can execute without being fully installed on the local system
- This separates the program for the OS it runs on, increasing compatibility and reducing conflicts

VDI (virtual desktop infrastructure)

VDI provides non persistent desktops – meaning the desktop resets to a default state after each session. This improves security by preventing persistent malware or misconfigurations. Useful in public or shared environments like schools, libraries and call centres.

Sandboxes

A sandbox is a secure, isolated environment for testing or running untrusted code. A VM acts like a separate computer – it has its own storage, OS and memory. So if you run malware inside the VM, it cant escape and infect the host system (unless there's serious

hypervisor vulnerability). This is why VMs provide a safe environment for testing untrusted code or software.

Cross platform virtualisation

Running software across different operating systems is tricky – but VMs solve that. With a VM, you can run MacOS, windows, and Linux on the same machine – even simultaneously. This is essential for developers, software testers, and IT pros who need to ensure compatibility of software across platforms. VMs simulate the full OS environment, letting you test the real behaviour of an app, not just guess how itll act on another system. Without VMs, you'd need multiple physical machines, which is more exoensive and space consuming.

Training and lab environments

Training in IT often involves risky changes, breaking things, or testing complex setups. VMs are ideal because:

- They let you build full environments (multiple VMs in a network) without needing real hardware
- You can clone, revert and reset VMs easily, so learners can break and fix things without consequence
- Virtual labs are widely used for certification practice, network simulations, and server setup training

No other method offers this level of flexibility, cost efficiency, and safety.

Emulation

Emulation mimics one hardware system on another – like simulating an old console or mobile chip. Its different from virtualisation because it doesn't rely on shared hardware architecture.

- For example, you can emulate a raspberry pi or ARM-based phone on a regular pc
- This is used when software needs to be teste on a system that isn't natively available.

But emulators are slower and less efficient than VMs – why is why for modern OS or application environments, virtualisation is still preferred.

VMs are for general purpose OS use, compatibility testing, and secure environments, whereas emulators are for hardware specific or legacy systems where virtualisation isn't possible.

8.4 Resource requirements for virtual machines

Running VMs effectively requires careful attention to hardware resources. While software plays a key role, the underlying physical components determines how well VMs perform and how many you can run at once.

CPU

The cpu is a central component when it comes to virtualisation, and several factors affect whether a system is suitable for running VMs. First, different manufacturers build in different technologies specifically for virtualisation – intel processors use a feature called VT-x, while AMD processors use AMD-V. these technologies are designed to allow the cpu to handle virtualization tasks more efficiently by offloading certain functions from the hypervisor.

However, even if a processor supports virtualisation, this feature is often disabled by default in the bios or uefi firmware. To use hypervisors like Hyper-V, VMware Workstation or VirtualBox, virtualisation must first be enabled manually in the systems firmware settings. Without this step, the software wont be able to launch VMs at all.

A key enhancement that greatly improves VM performance is SLAT, or second level address translation. This helps manage virtual memory more efficiently, reducing the overhead on the hypervisor. Intel refers to this as extended page tables (EPT), while AMD calls it rapid virtualisation indexing (RVI). Systems with SLAT support will run VMs much more smoothly, especially under load.

Beyond specific technologies, general cpu characteristics are also important. Multi-core processors are essential for virtualisation because each VM requires processing power. The more cores available, the more VMs you can run simultaneously without bottlenecks. Features like hyper-threading, which allows each core to handle multiple thread, can also enhance performance by increasing the efficiency of resource use.

Another consideration is the type of processor – specifically whether its 32 bit or 64 bit. A 32 bit processor is limited to addressing a maximum of 4 GB of ram and cannot run 64 bit guest operating systems, making it largely obsolete for modern virtualisation needs. A 64 bit processor, on the other hand, is capable of running both 32 bit and 64 bit VMs and can support much more memory, which is essential for most contemporary virtual environments. All of these factors combined determine how capable a cpu is for hosting VMs and whether it can meet the demand of modern virtualisation platforms.

Memory

Memory plays a crucial role in virtualisation, ass every virtual machine needs its own chunk of ram to function. When you create or run a VM, you allocate a portion of the hosts physical memory to it – and that allocation is treated as though its dedicated ram

by the guest operating system. This means the more VMs you want to run at the same time, the more physical memory your system needs to have available. If the host runs low on memory, it can start paging to disk, which drastically slows everything down. Some virtualisation platforms use techniques like memory ballooning or dynamic memory allocation to help balance usage across VMs, but these are only useful when there's enough underlying memory to begin with. Ultimately, ram capacity becomes one of the primary limitations on how many VMs you can run, and how smoothly they'll perform.

Storage

Storage is another key consideration because VMs are essentially stored as large files – including virtual hard drives, configuration files, and snapshots. Every VM you create take up space, and depending on what its used for, that space can grow quickly. A basic VM might only need 10 – 20 gb, while more complex setups or server grade VMs could demand 100 gb or more. You also need to factor in the space required for snapshots and backups, which can duplicate large chunks of data if not manged carefully. Performance matters too; traditional hard drives will work, but ssds offer a significant speed boost when it comes to booting VMs, loading applications, and running updates. In high density environments, storage not only needs to be large enough, but also fast and reliable, especially when multiple VMs are accessing the same disk resources simultaneously.

Networking

VMs don't just need computing and storage – they need connectivity. By default , VMs share the host machines network interface card, meaning all virtual traffic is funnelled through a single physical connection. This can quickly become a bottleneck, especially if multiple VMs are handling heavy data transfers or streaming services. For setups requiring higher throughput or redundancy, you can configure NIC teaming -combining multiple network cards into lone logical interface to improve both speed and fault tolerance. Its also worth noting that VMs can be configured with their own virtual network adapters, connected to virtual switches. These virtual switches allow for advanced network segmentation, internal inly communications between VMs, and more granular control of bandwidth and access. Because VMs often operate just like physical machines on a network, proper configuration is essential to ensure they get the connectivity and isolation they need – without bringing the hosts performance to a crawl.

8.5 Virtualisation security

Virt introduces powerful benefits to IT infrastructure – but it also comes with unique security challenges. While virtual machines are isolated by design, they still rely on

shared physical resources, which opens the door to specific types of attacks and misconfigurations if not managed carefully.

VM escape

One of the most critical threats in virt is VM escape. This is when an attacker manages to break out of the virtual machine and interact with the host system or other VMs. In a properly configured environment, a VM should be fully sandboxed – isolated from the host and other guest machines. However, if there are vulnerabilities in the hypervisor or misconfigurations in the system, an attacker may be able to exploit those gaps to gain access to the underlying hardware.

To succeed, the attacker must interface with the underlying resources – cpu, memory, or I/O – in a way that bypasses the virtual boundary. This is why keeping hypervisors updated and patched is absolutely essential. The risk of VM escape is one of the main reasons why many organisations choose type 1 hypervisor over type 2. Since type 1 runs directly on the hardware, there's no underlying OS to exploit. In contrast, type 2 hypervisors sit on top of an operating system, giving attackers another potential layer to target.

VM hopping

Another related threat is VM hopping. This is when an attacker inside one VM tries to move laterally to another VM on the same host. Unlike VM escape, this doesn't necessarily require access to the host OS – instead, it targets weak segmentation, shared virtual networks, or poorly isolated virtual switches.

Effective countermeasures include strict virtual network segmentation, disabling unnecessary VM to VM communication, and ensuring that each VM is firewalled and monitored independently. Just because two VMs are on the same host doesn't mean they should trust each other.

Sandbox escape

In many environments, VMs are used as sandboxes – isolated test areas to run suspicious or untrusted code. However, if the sandbox isn't properly secured, a threat actor can use the code running in the VM to escape into the broader system. This is especially risky when VMs have shared folders, clipboard access, or direct hardware pass through enabled, as these open paths to the host.

To minimise the risk, it's important to restrict shared resources, disable integration features when not needed, and keep the virtual environment fully isolated from production systems. Sandboxes should never have unnecessary access to the host or network unless explicitly intended and secured.

VM sprawl

Not all threats come from malicious actors – some stem from poor management. VM sprawl happens when too many virtual machines are created without proper oversight or tracking. Over time, this leads to unmonitored VMs with outdated software, insecure configurations, or unused data, all of which become security liabilities. To avoid this, organisations need strong VM life cycle management – every VM should have a purpose, an owner, and a clear plan for updates and decommissioning. Automation tools can help enforce policy and limit the spread of forgotten or duplicated VMs.

Summary

Virt offers flexibility and efficiency, but it changes the security landscape. Key threats like VM escape, VM hopping, and sandbox escapes target the shared nature of virtual infrastructure, while challenges like VM sprawl highlight the need for tight operational control. By combining secure hypervisor choices, strong isolation practices and careful VM management, organisations can take full advantage of virtualisation without exposing themselves to unnecessary risk.

9. Cloud computing

9.1 Introduction

In the previous chapter we explored virtualisation in detail – how it creates isolated, software defined environments that mimic physical hardware. Virtualisation forms the foundation of modern IT infrastructure, and it's a crucial prerequisite to understanding cloud computing. Why? Because cloud computing uses virtualised resources to deliver scalable and on demand services over the internet.

Cloud computing is essentially an extension of virtualisation – its how we take virtualised environments and make them available anywhere through a network, typically the internet. Instead of running everything on local hardware, we can access powerful resources like servers, storage, and software from remote data centres. This means were no longer limited by what's physically installed in our own offices or homes.

9.2 characteristics of cloud computing

Cloud computing brings unique characteristics that set it apart from traditional, on premises IT. These characteristics are crucial for understanding how cloud environments operate and how to make the most of them.

Shared vs dedicated resources

One of the core ideas behind cloud computing is that resources can be either shared or dedicated. In shared resource models, multiple customers or tenants use the same hardware and underlying infrastructure – this is common in public cloud (see 9.3) environments, where resource pooling is what enables cost savings and flexibility.

In contrast, dedicated resources are reserved for a single organisation, which might be necessary for compliance or performance needs. Dedicated servers and VMs provide more predictable performance and greater control.

Metered utilisation

A key feature of cloud services is metered utilisation – you only pay for what you use. Usage is tracked across compute (cpu, ram) storage, and network resources.

It's important to understand the cost implications of:

- Ingress: data moving into the cloud (often free or minimal cost)
- Egress: data moving out of the cloud (often more expensive and carefully metered)

Managing these costs means carefully monitoring data flows and optimizing for workloads to reduce unnecessary data movement, particularly for egress.

Elasticity

Cloud computing offers elasticity, meaning resources can scale up or down quickly based on demand. This flexibility allows businesses to handle sudden spikes in traffic or workload without having to invest in permanent hardware. It also prevents over-provisioning, which can waste resources and money.

Availability

Availability refers to how reliable and accessible cloud services are. Most cloud providers offer high availability through redundancy, geographic distribution, and failover options. This ensures services remain online even if hardware fails or data centres experience issues.

File synchronisation

Finally, cloud computing often includes file synchronisation features. This allows data and files to be automatically updated and consistent across multiple devices or users. For example, services like OneDrive or Google Drive enable users to access the same up to data files from any device, ensuring seamless collaboration and work continuity.

9.3 cloud deployment models

When adopting cloud computing, organizations can choose from several deployment models depending on their needs for control, cost and security. Each model has unique benefits and trade-offs.

Public cloud

The public cloud is the most common model. In a public cloud, resources – like servers, storage, and networking – are owned and operated by a third-party provider. Customers

share the same physical infrastructure (multi-tenancy) but access their own isolated environments.

Key providers of public cloud services include:

- Amazon Web Services (AWS)
- Microsoft Azure
- Google Cloud Platform (GCP)

Why would you choose the public cloud?

It provides scalability and flexibility, where resources can be easily scaled up or down to match demand. It has a lower upfront cost, as you avoid buying and maintaining physical hardware. It also has global reach, as services are available in multiple geographic locations for improved performance and redundancy.

Private cloud

A private cloud uses dedicated infrastructure – either hosted in an organization's own data centre or provided by a third party but exclusively for that single organisation. It provides greater control and security because resources are not shared with others.

Why would you choose the private cloud?

It offers enhanced security, as sensitive data and workloads remain within the organisation's control. It offers better compliance, as it's easier to meet strict regulatory requirements. It's very customisable, as hardware and software can be tailored to specific needs.

Hybrid cloud

A hybrid cloud combines public and private clouds to create a flexible, blended environment. Organisations can move workloads between the two as needed.

Why choose the hybrid cloud?

It offers the best of both worlds – you can use the private cloud for sensitive data, and the public cloud for dynamic or less critical workloads. It's also cost efficient, as it allows the offload of some workloads to the public cloud to save on resources. It can also offer improved resilience and flexibility in disaster recovery as not all data is stored locally.

Community cloud

A community cloud is shared environment for multiple organisations that have similar goals or regulatory requirements. For example, several hospitals might share a community cloud for health data.

Why choose a community cloud?

It allows for collaboration between related organisations with similar needs. It also means the costs of cloud services are shared, so it may be more affordable. And due to the shared nature of it, it is built to meet industry specific regulations.

Choosing the right model

When deciding which deployment model to adopt, consider:

- Data sensitivity and compliance requirements – private or community cloud if you need strict control
- Budget and scalability needs – public cloud for cost savings and rapid scaling
- Integration – hybrid cloud for flexibility and workload balancing

9.4 Cloud service models

Once the deployment model has been chosen, the next step is to select the right service model. These models define what is managed by the cloud provider and what remains the responsibility of the customer.

Software as a service (SaaS)

SaaS delivers complete applications over the internet. The cloud provider hosts and maintains everything: the software, the servers, and the data storage. Users simply access the application through a web browser or thin client.

Examples include Gmail, Microsoft 365, Salesforce.

Why use it? It eliminates the need to install or maintain software locally, and ensures you always have the latest updates.

Platform as a service (paas)

PaaS provides a platform for developers to build, test and deploy applications without managing the underlying hardware or os. The provider supplies everything needed for development, like development tools, runtime environments and databases.

Examples would include google app engine, Microsoft azure app service.

Why use it? It speeds up development, since developers can focus on writing code, not managing infrastructure.

Infrastructure as a service (iaas)

IaaS gives customers virtualised hardware resources – like virtual servers, networking, and storage – so they can build their own IT environment from scratch.

Examples would include Amazon EC2, Microsoft Azure virtual machines.

Why use it? It allows for maximum flexibility – you control the os, applications, and network configurations.

8.1 Cloud Networking

Introduction

Imagine you could run an entire business without needing a single physical server, router or data centre on site. Instead, everything – your files, apps, security, the entire network, is hosted in a remote data centre and accessed over the internet. This is cloud networking.

It replaces traditional on premises network infrastructure with internet based, scalable solutions. Instead of managing physical hardware, businesses use cloud providers like AWS, Microsoft azure and google cloud to handle networking, storage and computing needs. This allows companies to expand quickly, reduce costs and improve accessibility. Employees can connect to cloud resources as long as they have an internet connection.

Understanding cloud networking

Cloud networking based on three core principles:

Remote accessibility -users and businesses can access resources from anywhere

Scalability – resources can expand or shrink based on demand

efficiency Cost– no need for on premises networking hardware and maintenance

<u>how cloud networking works</u>

instead of relying on physical network infrastructure, cloud networks use virtualized resources managed by a cloud provider. These include

Cloud based file storage – access files from any device on the internet

Cloud virtual machines (VMs) – run apps without owning any physical servers

Cloud security services – firewalls, encryption, and security tools to protect data

Cloud networking services - virtual routers, load balancers and DNS services

Cloud networking allows businesses to scale instantly, enabling global operations and seamless connectivity across multiple locations

Cloud networking models

Public cloud networking

This is where third party cloud providers are used to host networking services and the services are shared across an organisation. Examples include AWS, google cloud and azure, and use cases would be internet hosting, cloud storage and app development.

Private cloud networking

This is where a cloud infrastructure is dedicated to a single organisation, hosted either on premise or by a cloud provider. It allows for more control over security and compliance. Use cases would be secure banking, healthcare, government networks.

Private cloud models allow for data access during internet outages, as the underlying hardware is on premises, so no internet access is needed for the cloud services.

Hybrid cloud networking

This is a mixture of private and public cloud networking, where some data, such as critical/sensitive data, is stored in the private portion of the cloud infrastructure, while other services run in the public cloud. This would be used by businesses who need both security and scalability.

Multicloud networking

This is where a customer would use multiple cloud providers to prevent being locked into one vendor and the issues that can come with that. It provides higher redundancy and better performance across global locations.

Cloud solutions in practice: how the models are used

Understanding cloud service models isn't just about knowing what the acronyms stand for – its about knowing how companies use them. Each model shifts responsibility for managing hardware, software and development tools in a different way. What matters in questions on this topic may include who needs control, what is being built or used, and how much infrastructure the customer wants to manage.

PaaS (platform as a service) -developer focused testing environments

When the goal is to develop and test applications without managing infrastructure, PaaS is ideal. It provides a full development environment in the cloud, including operating systems, programming frameworks, databases and tools – all pre configured.

- Used by: developers who want to build/test apps quickly
- Why: they don't want to worry about setting up servers or installing dev tools
- Exam clue: creating and testing applications – think paas

IaaS (infrastructure as a service) – Full control over the system

Iaas gives you raw virtual machines and storage in the cloud. You install the OS, middleware and applications. Its like renting a data centre, where you manage almost everything except the physical hardware.

- Used by: sysadmins, IT teams, businesses wanting maximum control
- Why: they need to build custom environments or run existing applications
- Exam clue: replacing on premises servers or needing total control – Think IaaS

Saas (software as service) – ready to use applications

Saas delivers fully functional apps over the internet – email, CRM, office sweets. Users access the service but don't manage anything under the hood

- Used by: end users and businesses
- Why: not setup required, just log in and work
- Exam clue: using a service like google docs, outlook 365 or salesforce – think saas

Bottom line: pick the right tool for the job

Scenario	Best cloud model
Building/testing apps	Paas
Running full servers or VMs	Iaas
Using pre built software tools	Saas

Cloud networking services

Cloud managed networking

Services that are traditionally run by hardware such as routers, firewalls and switches are configured on an online dashboard, so don't require on site management.

Cloud hosted network services

Instead of network functions being on premises, businesses use cloud services

DNS as a service (DNSaaS)- converts domain names to IP addresses without a physical server

Cloud firewalls - protect cloud applications from cyber threats, such as with AWS shield and azure firewall

Software defined WAN (SD-WAN) – optimized cloud traffic, replacing traditional routers

Virtual private cloud (VPC)

A VPC allowed organisations to create isolated private networks within a cloud providers infrastructure. These segment traffic between different cloud-based services and improve security by keeping sensitive data in a private network. Examples would be AWS VPC or azure virtual network

Cloud networking security

Security is critical in the cloud as they operate over the net. Tools are used by providers, but businesses must also take measures to protect their data.

Key cloud security features**:**

Identity and access management (IAM) – controls user permissions for cloud resources

Multi factor authentication (MFA) - adds extra security beyond passwords

Cloud firewalls – filters traffic and prevents cyber attacks

Encryption (TLS/SSL, IPsec VPNs) – secures cloud comms and data

Intrusion detection and prevention (IDPS) – monitors network traffic for suspicious activity

Security challenges in cloud networking

Data breaches - sensitive data stored in the cloud is a target for hackers

Compliance issues - some industries require strict data protection regulations

Shared responsibility - cloud providers secure the infrastructure, but businesses must secure their own data and applications

Best practices for cloud security

- Use strong encryption for all cloud comms and stored data
- Enable role-based access control (RBAC) to limit who can modify cloud resources
- Set up automated backups and disaster recovery solutions

Resource allocation and management

Rapid elasticity

This is the ability of cloud services to automatically scale up or down based on demand. For example, if a website experiences a sudden traffic spike, cloud servers automatically scale up to handle the load and scale down when traffic decreases.

Measured service

This is where cloud providers automatically track and report resource usage (CPU, storage, bandwidth) for billing and optimisation. For example, AWS charges users based on the amount of storage or computing power they use. Measured service tracks the usage of resources.

Metered service

This is where a user only pays for the resources they use, like a utility bill. providers track usage of computing resources and customers are based on an actual consumption number. Metered service uses the data of how much resources were used in order to charge customers accordingly.

On demand self service

Users can provision and manage cloud resources (such as servers or apps) without human intervention form the cloud provider. For example, a company deploy new VMs instantly using a cloud platform without contacting support

Resource pooling

Cloud providers use shared computing resources to serve multiple customers dynamically. For example, a cloud provider allocates and reassigns resources from a pool of data centre servers based on customer demand.

Real world use cases for cloud networking

Cloud networking is widely used across industries for scalability, security and global reach. **Businesses and enterprises** use it for **remote work** infrastructure to allow employee to work from anywhere via secure VPNs and virtual desktops. They also use it for **global internet hosting**, where sites and apps can be hosted across multiple cloud data centres for better speed and uptime. It's also useful in this case for **disaster recovery and backups.** in the case of **content delivery and streaming** its used because having data across multiple data centres **reduces latency**, and is useful for the handling of **large scale data transfers** such as with twitch or Netflix. In the case **of AI and big data processing**, cloud services allow **AI models and data analytics to run** on powerful remote servers instead of local machines. Cloud providers also **offer machine learning platforms to train AI** without massive hardware investments.