Networks

definitions are highlighted!

Internet

- A global network of networks that enables computers worldwide to interconnect and communicate
    - Computer network — consists of 2 or more computing devices connected via a transmission medium to enable communication, information exchange and resource sharing
- Benefits:
    - Allows efficient collaboration and data sharing without the need for physical travel
- Physical structure
    - Cables are laid along sea beds as trans-continental links, forming a global backbone.
        - They are primarily made of fiber optics, which allow data to travel long distances at high speeds with very little loss of signal quality.
    - Internet Service Providers (ISPs) connect directly to this backbone, and then distribute the connection to smaller regional providers, who ultimately deliver internet access to homes and businesses.
    - In some areas, satellite communication is used.
        - **Reason:** Physical cables are difficult and/or expensive to install.
        - Data signals are relayed via orbiting satellites so that there is internet access even in remote/rural locations where terrestrial infrastructure is not available.

Internet registries and registrars

- *Registries — organisations that manage critical internet resources to ensure uniqueness and orderly operation*
    - Own the central database and set the rules
    - IP address registries
        - Internet Assigned Numbers Authority (IANA) manages global allocation of IP addresses.
        - Then, Regional Internet Registries (RIRs) distribute IP blocks to ISPs within specific-geographical areas.
        - Lastly, ISPs assign specific IP addresses to end-users.
    - Domain name registries manage the master database for a specific Top-Level Domain (TLD), such as ".com" and ".sg", ensuring that all domain names within that TLD are unique.
- *Registrars — companies accredited to sell domain names directly the public (retail agent)*
    - Provide domain name registration services to individuals and businesses
    - Act as the middleman between customers and domain name registries (which own the domain names)
    - Facilitate the process of acquiring, managing and renewing domain names
    - Maintain records of ownership and hosting details

Internet Protocol (IP) address

- *A unique identifier assigned to a device on a network*
- Specifies the source and destination location for data packets
- IPv6 was designed to solve IPv4 address exhaustion (stemming from the massive growth of internet-enabled devices), as well as improve network efficiency and security support.

- IPv4 vs IPv6

	IPv4	IPv6
Size	32 bits	128 bits
Format	4 decimal numbers ranging from 0 to 255, separated by dots. Each number corresponds to 1 byte. e.g. 192.168.1.1	8 groups of 4 hexadecimal digits separated by colons. Leading '0's and a maximum of 1 consecutive sequence of 0-only groups can be omitted. e.g. 2001:db8::20:17:bad:c0de
Total addresses	Almost 4.3 billion (2^{32})	Effectively unlimited (2^{128})

- Special IPv4 addresses:
    - 127.0.0.1 (i.e. localhost) — refers to your own computer
        - Used for testing
    - 0.0.0.0 — represents an unspecified/unknown address
- Structure of IPv4 addresses:
    - Classless Inter-Domain Routing (CIDR)
        - Allows the division between NetID and HostID to be flexible
            - NetID identifies the network.
                - Used to route data between different networks
            - HostID identifies the specific device within that network.
                - Only checked once the data arrives at the destination network

        - Allows ISPs to allocate exactly the number of addresses an organisation actually needs, preventing waste (as opposed to the classful system)
    - Subnetting
        - The process of dividing a single network into smaller logical networks by borrowing bits from the host portion of the IP address and adding them to the network portion.
        - Divides a large, rigid block of IP addresses into smaller, customised segments that match the actual size of each sub-network
        - The network address identifies the subnet itself, and has all the host bits as 0.
        - The broadcast address is used to send data to all devices in the subnet, and has all the host bits as 1.
- Static and dynamic
    - Static IP addresses are used by devices that require a consistent and fixed network identity.
        - e.g. Server, router
    - Dynamic IP addresses are used by most end-user devices connecting through as ISP.
        - When a device connects to a network, it is assigned an available IP address from a pool, which will be different every time it connects.

IP type	Advantages	Disadvantages
Static	• Reliable — simplifies remote access, hosting and configuration • Secure — easier to implement strict access control	• More expensive (ISPs usually charge more) • Manual configuration is

		time-consuming and harder to maintain.
Dynamic	• Cost-effective — since not all devices are online at the same time, many of them can be served with a limited set of addresses through this efficient IP address usage, reducing costs for both ISPs and users • Easy and scalable — new users can connect to the network without the need for manual configuration due to automatic setup	• Inconsistent IPs can disrupt services that require a fixed address • Harder to diagnose network issues when the address keeps changing

Media Access Control (MAC) addresses

- Used to identify devices and transfer data within a LAN
- 12 hexadecimal digits
    - 48 bits long
- Identifies the specific hardware, while an IP address identifies the logical location on the network

Domain name system (DNS)

- A domain name identifies the location or site of a resource on the internet.
- Domain names are structured in a hierarchy:
    - TLDs — generic (e.g. “com”) and country (e.g. “sg”)

    - Second Level Domains (2LDs)
        - e.g. “google”
    - Third Level Domains (3LDs)
        - e.g. “www”
- A complete address specifying the exact location uniquely identifies resources and allows DNS to resolve them to IP addresses.
    - Multiple IP addresses can be associated with a single domain name, for load balancing and redundancy.
- Domain names are stored in publicly accessible databases upon registration.
- DNS
    - Allows the usage of domain names instead of IP addresses to make websites easier to access
    - Operates using a distributed hierarchical structure
    - Each domain is associated with authoritative DNS servers that store the domain name to IP address mappings.
    - Process:
        - The browser sends a request to the local DNS server, which checks its cache to see if it has the IP address.
        - If it does not, it recursively queries higher level servers — TLD and the authoritative DNS server for the domain.
        - The IP address is returned to the user when found so that their device can visit the requested domain.
        - The local DNS catches the result and returns it to the browser.
        - The browser uses the IP address to connect to the web server.

LANs and WANs

- Local Area Networks (LANs)
    - Computer systems are situated relatively close to each other (e.g. same cluster of buildings).
    - Features:
        - Uses wire cables, wireless links or fibre optic cables
        - Allows the sharing of data files and hardware resources e.g. printers, scanners
            - Data transfer is fast.
        - Fast and highly reliable data transmission
        - Low setup cost
        - Easier maintenance and security
    - Usually follows a client-server model — user computers are the clients, servers provide central storage and services
    - Usually owned and operated by a single organisation
    - **Issue:** Difficult to scale for very large networks
- Wide Area Networks (WANs)
    - Used to link multiple separate LANs together (e.g. multinational organisation), as it connects computers in geographically dispersed locations
        - Bridges the long distances between countries
        - Enables communication across large areas like cities, countries or even continents
    - Uses public infrastructure (e.g. telephone links, undersea cables, satellites)
    - Issues:
        - Higher costs
        - Slower and less reliable data transmission
        - More complex security management (due to public connections)

Intranet

- *A private network that operates exclusively within an organisation*
    - Uses the same technology as the Internet but is restricted to authorised users only
        - Provides the same services e.g. internal web pages, email, file sharing and collaborative tools
- Features:
    - Controlled, secure and private access
        - Access is limited to employees/members, with protections like user accounts and passwords being used to ensure this.
        - Different levels of access rights/passwords can be used to ensure that only specific people can access specific facilities or data.
    - Better content quality — given that there is a smaller volume of content compared to the Internet, it is more likely to be correct, relevant to the organisation and up-to-date.
    - More secure
        - Isolation from the public Internet reduces the risk of external attacks.
        - Often protected by firewalls
    - Better performance — faster and more reliable due to higher bandwidth and less traffic congestion
- Virtual Private Network (VPN)
    - A technology that creates a secure, encrypted “tunnel” between a user’s device and the organisation’s private network, over a public network (e.g. Internet)
    - Allows users to access the intranet safely remotely (e.g. traveling, working from home)
    - Encrypts data during transmission to prevent it from being read or intercepted by unauthorised people

- Overall connection structure
    - Built using the organisation's internal LANs and WANs
    - Remote (off-site) access is securely enabled via a VPN
    - When parts of the Intranet are made available to selected outside users (e.g. suppliers, partners), it is called an Extranet.

Data flow modes

- Simplex — data flows in 1 direction only
    - e.g. Keyboard to computer, TV broadcasting
- Half duplex — data flows in both directions, but not at the same time; devices take turns sending and receiving
    - e.g. Walkie-talkie
- Full duplex — data flows in both directions simultaneously
    - Requires separate channels for sending and receiving
    - e.g. Mobile phone calls, Modern Ethernet

Message types

- Broadcast — one-to-all communication
    - The message is sent to every device on the network.
    - Simple, but inefficient as it wastes bandwidth
- Multicast — one-to-many communication
    - The message is sent only to devices that have joined a specific group.
    - Efficient way to send data to selected recipients
- Unicast — one-to-one communication
    - The message is sent directly to a specific device.
    - Used for private and direct communication

Data packets

- Data to be transmitted across a network is broken down into manageable chunks called packets.
- Packet size can be fixed or variable.
- Each packet contains:
    - Header — control information
        - IP addresses of the sender and recipient
        - Protocol used
            - Protocols ensure the proper transfer of data between devices by specifying the data format and signals required to start, control and end the transfer.
        - Packet number in sequence (only in TCP) — to reassemble the packets in order at the destination
        - Time to live (TTL) — limits the lifespan of a packet
            - Each router decreases this value by 1.
            - The packet is discarded TTL reaches 0 to prevent infinite loops.
        - Hop limit
    - Payload — the actual data being sent
    - Trailer — error-checking information
- Error detection (only in TCP)
    - The trailer contains a checksum or Cyclical Redundancy Check (CRC) to detect transmission errors and ensure data integrity.
    - Process:
        - A value is calculated based on the data at the sender's end and added to the trailer.
        - The same calculation is performed on the receiving end.
        - If the values match, the data is likely correct. If not, the data is corrupted, so the packet is requested again.

- Parity check
    - An extra bit is sent with each byte.
    - Only detects errors, but cannot fix them
    - Types:
        - Even parity — ensures that the total number of 1-bits is even
        - Odd parity — ensures that the total number of 1-bits is odd

Transmission methods

- Circuit switching
    - A dedicated communication path is set up between 2 devices before data transfer begins.
    - Bandwidth is reserved exclusively throughout the connection, even if no data is being sent.
        - Leads to inefficient resource usage
    - Data is transmitted in order and follows the same path throughout the connection
    - e.g. Traditional telephone systems
- Packet switching
    - *Transmission of data in small packets, where each packet takes the best available route at a given time, and where packets may not arrive in sequence a t the destination, so they must be reassembled there*
    - Each packet travels independently and can take different routes (depending on traffic).
    - Connectionless service
        - **Protocol:** UDP
        - Packets are sent without establishing a connection.
            - No setup is required.
        - **Benefit:** Low overhead, which increases speed

        - Problems:
            - Packets can arrive out of order.
            - Packets may not be delivered.
        - e.g. Live-streaming, online gaming
    - Connection-oriented service
        - **Protocol:** TCP
        - A virtual circuit is established before transmission.
        - Benefits:
            - Packets are reassembled into the correct order at the receiver even if they arrive out of order.
            - Packets are delivered reliably.
        - Problem: Slower (due to error checking and connection setup)
        - e.g. Web browsing, file transfer, email

Network devices

- Routers
    - Route data packets to their correct destination across 2 or more networks that use the same protocol (e.g. TCP/IP)
    - Serve as intermediate devices that facilitate the transfer of data packets between networks
        - Routers read the recipient's IP address in each packet and forward it via the fastest, least congested route in the global network to the next router.
            - A routing algorithm calculates the optimum path.
            - Each movement of the packet from a router to another is called a "hop".
    - Use routing tables to store device locations and efficient paths (updated dynamically)

    - Have a globally unique and registered external IP address upon being connected to the Internet
- Gateways
    - Route data packets to their correct destination across 2 or more networks that use different protocols and/or data formats
    - Strip away the original header and trailer, translate the data into a format understood by the destination network, then repackage the raw data with new headers before sending it onward
- Modem — converts between digital and analogue signals for internet transmission
- Network Interface Card (NIC) — allows a device to connect to a network
    - Contains the MAC address
        - MAC addresses are local, with access limited to individual networks.
        - In comparison, IP addresses are global, so they can be accessed across different networks.
- Switch — forwards data to the correct device within a LAN using MAC addresses
    - Operates on the data link layer
    - Connects multiple devices within a single network, enabling communication among them

TCP/IP model

- Enables communication between devices on different networks regardless of their locations
- Encapsulation — when, during communication, data is passed down the stack and each layer adds its own header (and sometimes trailer)
- Decapsulation — each layer removes its own header
    - The reverse process of encapsulation

- Has 4 layers
    - Application
        - Provides network services to user applications
        - Selects the appropriate protocol for data transmission and formats the data for communication between the application and network, ensuring compatibility and readability
        - e.g. HTTP/HTTPS for web browsing, FTP for file transfer
    - Transport
        - Establishes end-to-end communication
        - TCP (reliable)
            - Splits data into packets/segment and adds packet/sequence numbers for correct ordering
            - Adds port numbers to ensure that the data reaches the correct application on the receiving device
            - Ensures reliability — requests retransmission if packets are lost (detected through acknowledgements that verify successful reception)
        - UDP
            - Unreliable as there is no acknowledgement of receipt and no retransmission
    - Network (aka IP layer)
        - Determines the path packets take between networks
        - Adds the IP addresses of the source and destination
        - Routers utilise these addresses to forward packets to their destinations.
        - The IP address and port number form a socket that specifies the specific application/process running on a particular device.
    - Data link
        - Enables hardware identification for packet delivery

        - Handles physical connections between network nodes
        - Also includes the physical medium (e.g. cables, signals)
        - Adds the MAC addresses of the source and destination
            - The destination MAC address identifies the next hop (i.e. physical device receiving the packet next), not the final destination device.
            - The MAC addresses are stripped and re-added after every hop (whereas the source and destination IP addresses remain constant throughout).
- On the receiving end:
    - As the packet travels up the stack, each layer removes its own header and trailer.
        - First, the Data Link layer removes the MAC addresses.
        - Then, the Network Layer removes the IP addresses.
        - After that, the Transport Layer removes the port numbers and reassembles the data in order.
        - Finally, the Application Layer presents the final message to the user using the correct protocol.

Types of networks

- Client-server
    - 1 or more clients connect to and send requests to a central server, which processes said requests and provides resources/data to the clients.
    - Each client holds some of its own files and software, and also accesses shared resources from the server.
    - The server stores, manages and delivers the abovementioned shared resources to the client.
    - In large networks, different servers handle specific tasks.

        - e.g. File server — stores and manages files accessible to clients
        - e.g. Print server — manages print jobs and directs them to printers
        - e.g. Web server — delivers web pages and content to clients
        - e.g. Mail server — manages the sending and receiving of the emails
        - e.g. Database server — stores, organises and secures data
        - e.g. Cloud computing
    - Advantages:
        - Secure — centralised storage of information allows strict control over access rights
        - Easier backup and recovery, since it is done centrally
        - Efficient resource (e.g. hardware, software, data) sharing
        - Easy to manage — clients can be upgraded or replaced easily without affecting the rest of the network
    - Disadvantages
        - Server hardware, software and maintenance are expensive.
        - If there is high traffic that overloads the server, delays occur (poor performance).
        - If the server fails, the entire network becomes unavailable (single point of failure).
        - Requires dedicated resources
- Peer-to-peer
    - All computers have equal status and can act as both clients and servers. Each device shares its own resources directly with others without the need for a central server.
    - e.g. File sharing, content distribution, real-time collaboration

    - Advantage:
        - Simple setup
        - More cost-efficient
        - Distributed resource sharing (no single point of failure)
    - Problems:
        - Less secure
        - Harder to manage in larger networks

Cloud computing

- Delivers computing services (e.g. storage, software, processing power) over the internet
- Follows the client-server model
    - Allows users to access resources on demand, without needing local hardware or infrastructure
- Service models
    - Software as a Service (SaaS)
        - Users access applications through a web browser/interface
        - The provider hosts and manages the software and data.
    - Infrastructure as a Service (IaaS)
        - Provides virtualised computing resources e.g. servers, storage
    - Platform as a Service (PaaS)
        - Offers a platform for developing, testing and managing applications
- Since data is stored on remote servers, it is important to use reputable providers with strong security measures.
- Advantages:
    - Scaleable — resources can be increased or decreased as needed
        - Users only pay for what they use.

    - Accessible — data and applications can be accessed from any device and location
    - Cost-efficient — eliminates the need for expensive hardware + reduces maintenance costs
    - Allows multiple users to work on the same data simultaneously
    - Reliable — provides built-in backup and disaster recovery solutions

Data processing

- A client sends a request to a server, which processes it, generates the required content (e.g. web page) and returns it to the client.
- Client-side processing occurs on the user's device.
    - Handles simple tasks to reduce network traffic and the server's workload
    - Provides faster interactivity and immediate feedback (more responsive)
    - Manipulates the user interface
    - Increases efficiency by reducing the amount of data transferred over the network
    - Performs initial checks to prevent invalid data from being sent
        - Only basic checks that cannot protect against more complex threats like SQL injection
        - The extent of validation may be limited by the client's devices.
- Server-side processing occurs on the central server.
    - *Server-side scripting — when resources/data are generated on the server-side using its resources and processing power (e.g. database interaction, calculations) and then sent over to the client based on its request and needs, such that minimal processing is required on the client-side*

    - Performs further and secure validation to ensure data integrity and security
        - In case the client-side validation is maliciously circumvented
    - The server has the resources and processing power to do things like:
        - Efficiently handle complex calculations and large datasets
        - Directly access, query and update server databases
        - Generate web pages
    - Enforces strict security and protects sensitive information
        - Whereas client-side processing only prevents basic malicious input
    - Returns data to the client based on its requests and needs, such that minimal processing is required on the client's end
- The choice of where to process depends on:
    - Task complexity
    - Device capability
    - Security needs

Thin- vs thick-client computing

- "Thickness" refers to how much processing and storage happens on the client device compared to the server.
- Thin client
    - Minimal hardware; relies mostly, if not completely, on the server for the operating system (OS), applications, data processing and storage
    - Benefits:
        - Easy to update and secure due to centralised management
        - More secure

        - Cheaper and lesser hardware reduces costs
    - **Problem:** Requires a stable connection to work
    - e.g. Remote desktops, terminals in call centres
- Thick client
    - Has its own OS, processing power and storage
    - Benefits:
        - Can run demanding software (high performance)
        - Can function offline, without a network connection
    - **Problem:** More expensive hardware and maintenance
    - e.g. Personal laptops, desktop PCs

@star.notes