The Internet

The internet and World Wide Web

There is a common misconception around the definition of the internet and WWW. People tend to use the terms interchangeably but there is a distinction.

The **internet** is a global network of interconnected computers. It allows communication between computers and devices using TCP/IP. It has globally unique IP addresses. Individual devices are attached to a local area network, which in turn are connected to wide area networks. The internet is a network of wide area networks and is itself is a wide are network. The internet refers to the physical hardware components such as computers, cables, routers, gateways and so on. The internet has be around since the 1960s.

The **World Wide Web (WWW)** is a service that has web pages and other content that runs on the internet. It is made up of interlinked hypertext documents and uses mainly the HTTP protocol. The WWW was developed in the early 1990s by Tim Berners Lee.

Local Area Network (LAN)

- Individual devices are attached to a local area network.
- A local area network is a network that covers a relatively small geographical area typically extends over the range of a single organisation such as a university campus, school site.
- LANs are typically managed by a single organisation.

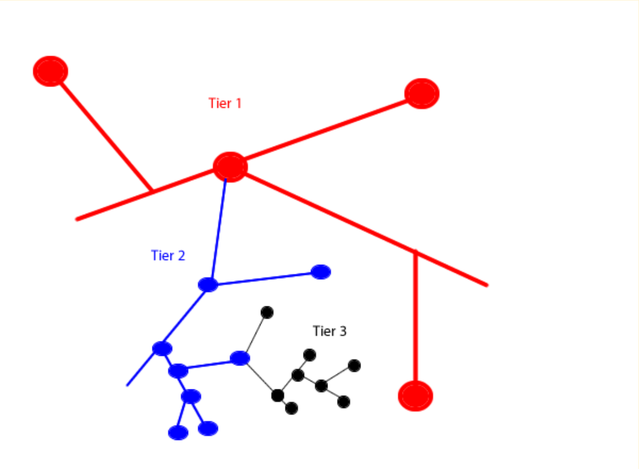
Wide Area Network (WAN)

- A wide area network is made up of many local area networks and covers a much wider geographical area.
- The internet the ultimate wide area of networks. It is a network of networks and allows billions of devices to be interconnected.

Structure of the internet

- **Tier 1** networks are considered the backbone of the internet. These are country wide networks that are connected to other networks by fibre optic cables that link different parts of the world and include cables that cross ocean floors including across the Atlantic ocean.
- **Tier 2** networks are regionally based and allow connectivity between local area networks
- **Tier 3** networks and LANs and enable internet access for homes and small businesses and individual organisations.

Internet tiers



Packet switching

Transfer of data across a network relies on the principal of packet switching. Packet switching is the process of data being broken into packets before being sent over the network and then reassembled at the other end. Packets are forwarded through a series of routers between their source and destination. This allows data to be transfer across a network efficiently. Large files would otherwise clog up the network. Small packets can choose different routes through the network, however packets do get lost during transfer. A typical packet size is 1500 bytes.

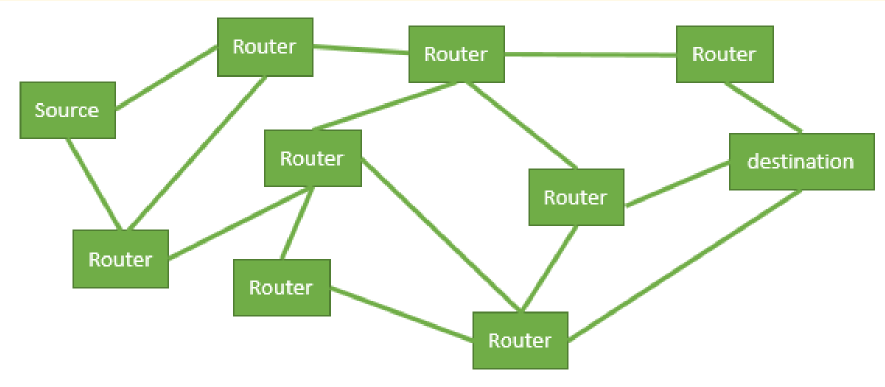
Packet contents

Header	Source IP address	Where the packet has come from
	Destination IP address	Where to send the packet
	Packet identification	Necessary so the computer knows how and in what order to reassemble the packets
	Destination and source MAC Address	This is the address of the network card
	Destination and source port numbers	
	Protocol	Which protocol is being used (eg HTTP, FTP)
Body	Data/Payload	Part of the file that you want to send
Footer	Error control bits	Check for errors in the packets to make sure they have not been corrupted in transport

Gateways and routers

- The purpose of a router is to connect networks together. The router directs internet traffic to destination along the quickest and least congested routes. Normally a packet will be routed through multiple routers before it reaches its destination. Routers have at least two network cards so that traffic can be directed in different directions.
- A gateway, like a router, connects networks together. A gateway provides a single access point to a network. Networks may have different protocols so gateways allow communication to be translated between different protocols.

Possible routes for packets



Uniform Resource Locator

The Uniform resource locator is the address of a resource on the world wide web. The resource can be a web page or other file such as mp3, pdf for instance. It contains both the protocol and domain name and takes the form of:

Protocol://Fully qualified Domain Name/Path

- Protocol: These tell the browser what to do with the web address (eg HTTP, HTTPS)
- FQDN: This is the name of the website
- Path: Points to where the specific page is on the website.

The URL for http://www.bbc.co.uk/news/world-middle-east-23691571

Protocol:	http
FQDN:	www.bbc.co.uk
Path:	news/world-middle-east-23691571

IP Address

Every device on the internet needs to have a unique IP (internet protocol) address. Packets contain the sender's and receiver's IP address so that routers know where to direct the packets. Just like every house in the country has a unique postal address.

Domain name

Each web server has an IP address, and the IP address can be used to request a web page. However, IP addresses are hard to remember so domain names are used to identify IP addresses and are much easier to remember. Multiple IP addresses can be associated with a single domain name.

Example

Domain name	www.google.co.uk
IP address	216.239.238.120

The Domain name identifies the location of the resource on the internet. It is structured hierarchically with domains and sub-domains.

Generic top level domains	.com	.org	.net	.gov		
Country top level domains	.fr	.nz	.au	.uk	.tv	de
Second level domains for UK	.co.uk	.org.uk	.sch.uk	.ac.uk	.gov.uk	.nhs.uk

Fully qualified domain name (FQDN)

The fully qualified domain name contains the complete domain name and hostname of the web server.

Examples of FQDN

access a webserver	www.bbc.co.uk
access a webmail server	mail.google.com

The domain name hierarchy for www.bbc.co.uk.

root	.
Top level domain	uk
Second level domain	Co
Local domain name	bbc
Hostname of server	www

Domain Name server (DNS)

When a domain name is requested the domain name server searches through huge databases to find the corresponding IP address. If the server cannot find the corresponding IP address it then requests the IP address from the DNS system (other DNS servers).

Internet registries

Internet registries allocate domain name to one or more IP addresses. This is a complex task that is overseen by an organisation called ICANN (Internet Corporation for Assigned Names and Numbers). Regional internet registries are responsible (e.g. RIPE NCC in Europe) for allocating a set of IP addresses to domain names. This ensures that domain names and IP addresses are globally unique.

Internet Security

Why do we need network security?

- To prevent unauthorised access to our electronic devices
- To protect our data eg to prevent sensitive data being stolen, to prevent personal data from being stolen.

Firewall

A server is dedicated to acting as the firewall. It has two network cards one for the LAN and one for the internet. A firewall prevents unauthorised access to a network and represent the first line of defence for a network. Networked computers have lots of incoming and outgoing data packets. Whilst most data packets are harmless, some data packets may be harmful and contain malware. It is the role of the firewall software to identify and prevent these packets getting on to the LAN/computer in the first place.

Static packet filtering Incoming packets via the internet are monitored and inspected. The information in the packet headers including the source and destination IP addresses, ports and protocols are checked. Any packets that do not meet the filtering criteria (eg only accept packets with specified source IP addresses) are blocked, otherwise packets will be passed onto the LAN.

Stateful inspection – dynamic packet filtering

- Stateful inspection is form of packet filtering that monitors both outgoing and incoming packets.
- In static packet filtering only information in the header is examined, but in dynamic packet filtering the contents of the packet are also examined.
- Stateful inspection monitors the state of the connection for a particular communication. Static packet filtering is based in a set of predefined rules, whereas dynamic packet filtering considers the context of the connection based on previous packets. For instance if a request to a web server is made, then response packets from that server will be expected and allowed to pass to the LAN.
- Stateful inspection offers a better level of protection than static packet filtering.

Proxy server

When a proxy server is used there is no direct connection between the LAN and internet. Traffic is routed through the proxy server. The proxy server will have a different IP address to the devices on the LAN allowing the IP address to be hidden outside the network. The proxy server then acts as the firewall as required.

Encryption

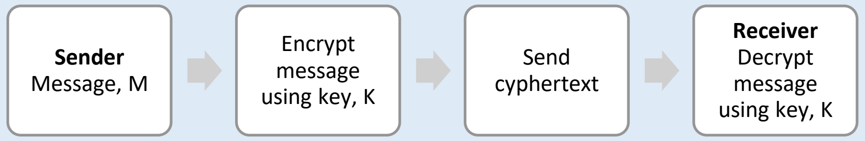
When passing sensitive data over the internet such as credit card numbers you need to ensure that the data are encrypted. This means that the message is garbled up so if the message gets intercepted as it is being transmitted to its destination it will be almost impossible for anyone without the key to read the original message. Ensure that you are using the secure hypertext transfer protocol (https) and also on most web browsers a little green padlock should appear on the URL bar.

Symmetric encryption

In symmetric encryption both the sender and receiver use the same key to encrypt the decrypt the data.

$$C = f[M,\ K]$$
$$M = f^{-1}[C, K]$$

where C is the cyphertext, M is the message, K is the key and f can be any symmetric encryption algorithm



Sender Message, M

Encrypt message using key, K

Send cyphertext

Receiver Decrypt message using key, K

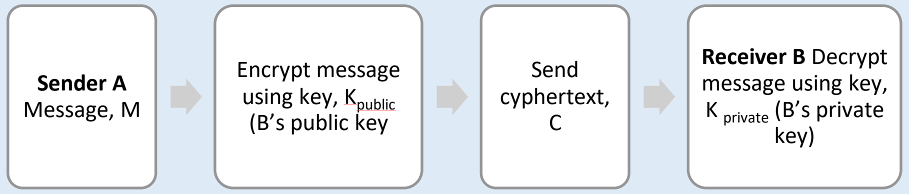
Asymmetric encryption

In asymmetric encryption a different key is used to encrypt and decrypt the data. This is a one way function, you cannot use the same key to encrypt and decrypt the data.

$$C = f[M, K_{public}]$$
$$M = f^{-1}[C, K_{private}]$$

where C is the cyphertext, M is the message, K_{public} is the public key, $K_{private}$ is the private key and f can be any symmetric encryption algorithm.

The sender A encrypts the message using the receiver B's public key. The receiver then decrypts the message using the private key that is not shared with anyone.



Sender A Message, M

Encrypt message using key, K_{public} (B's public key

Send cyphertext, C

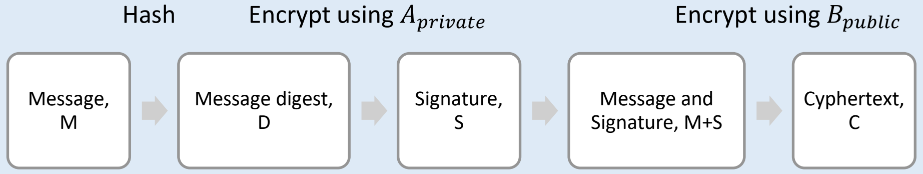
Receiver B Decrypt message using key, K $_{private}$ (B's private key)

Supposing Alice wants to send a message to Bob. Alice uses Bob's public key that is made available to all in order to encrypt the message. She then send the message and Bob decrypts the message using his own private key that is known only to him. To reply Bob uses Alice's public key to encrypt his message. Bob sends the message and Alice decrypts the message using Alice's private key.

Digital signature- Sending

A digital signature is a way of verifying that a message has been sent from the correct source. To send a message from Alice to Bob:

1) Apply a hash (one way encryption), H to the message, M to produce the message digest, D such that:
   $D\ =\ H[M]$
2) Encrypt the message digest D, using the Alice's (the sender's) private key ($A_{private}$) and an asymmetric encryption algorithm E to produce the digital signature, S.
   S $= E(H[M], A_{private})$
3) Append the signature to the original message and encrypt using Bob's (the receiver's) public key (B_{public}) to produce the cyphertext C to be sent.
   C $= E(M + S, B_{public})$

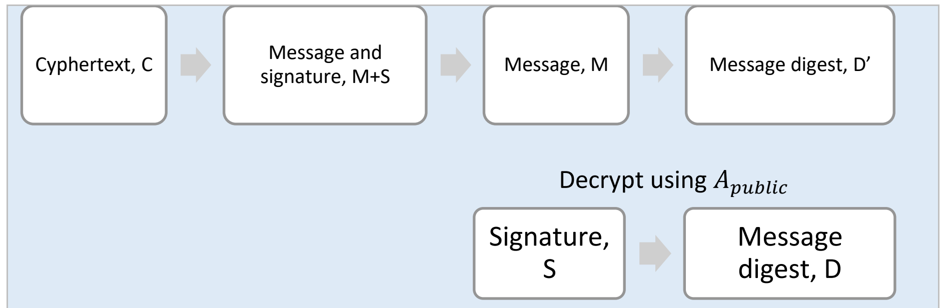


Hash
Encrypt using $A_{private}$
Encrypt using B_{public}
Message, M
Message digest, D
Signature, S
Message and Signature, M+S
Cyphertext, C

Digital signature - Receiving

To receive a digital signature sent from Alice to Bob:

1) Decrypt the cyphertext using Bob's private key to extract the message and signature:
   $M + S\ = E\ [C, B_{private}]$
2) Apply the hash, H to the message, M to produce the new message digest, D' such that:
   $D'\ =\ H[M]$
3) Decrypt the signature to get the message digest D using Alice's public key, A_{public}
   D $= H[M] =\ (S, A_{public})$
4) Evaluate D versus D'. If they are the same then the message is authentic and has not been tampered with



Decrypt using $B_{private}$
Hash
Cyphertext, C
Message and signature, M+S
Message, M
Message digest, D'
Decrypt using A_{public}
Signature, S
Message digest, D

Digital certificate

Digital certificates are another way of verifying internet communications. Digital certificates are issued by certification authorities. The certificates contain information about the owner the public key and also the digital signature of the issuing body. Digital certificates are typically used by banks and e-commerce websites.

Malware

Malware is short for malicious software. Malware is software that has been purposely developed to damage, disrupt or take control of computer systems.

Types of malware

Trojan software is malware that gains access to a computer by pretending to be legitimate software. The Trojan allows hackers unauthorised remote access to your computer without the user being aware. From there the hacker can control your computer and use the machine for nefarious purposes such as installing key loggers to record passwords and pin numbers or launch attacks on other computers thereby obfuscating the original source of the attack.

Computer viruses are software that replicates themselves and can transfer from one computer to another. They are activated by a user often as email attachments and attachment to other files and programs. Once a virus is on a computer system it can make undesirable and unauthorised changes to a computer system. Viruses require human agency to be activated, eg opening an email attachment.

Worms spread like viruses but do not require human intervention, and attached themselves to network tools to spread themselves. In that sense they are more destructive than computer viruses because they can spread automatically from one computer to another very quickly and can affect whole networks.

Code quality, monitoring and protection

- Common areas of exploitation are buffer overflow and SQL injections.
- SQL injections are a way of hacking into databases. The hacker enters input that is able to access data in the database that they should not have access too (eg passwords of other people) or modify the database.
- Ensuring that code is built with a high degree of protection in the first place will avoid many of the potential ways of exploitation.
- With buffer overflow the hacker deliberately overwrites memory locations using existing code in RAM. In those locations the hacker can place malware.
- It is important that any security patches that come out for software that these are installed as soon as possible.

Antivirus software

- Antivirus software is software that scans the computer and is able to detect known malware against a database and quarantine and remove them from the system.
- Quarantine means that affected files are isolated and are unable to infect a computer system.
- Antivirus software needs to be regularly updated to keep up with new viruses that are continuously being developed.

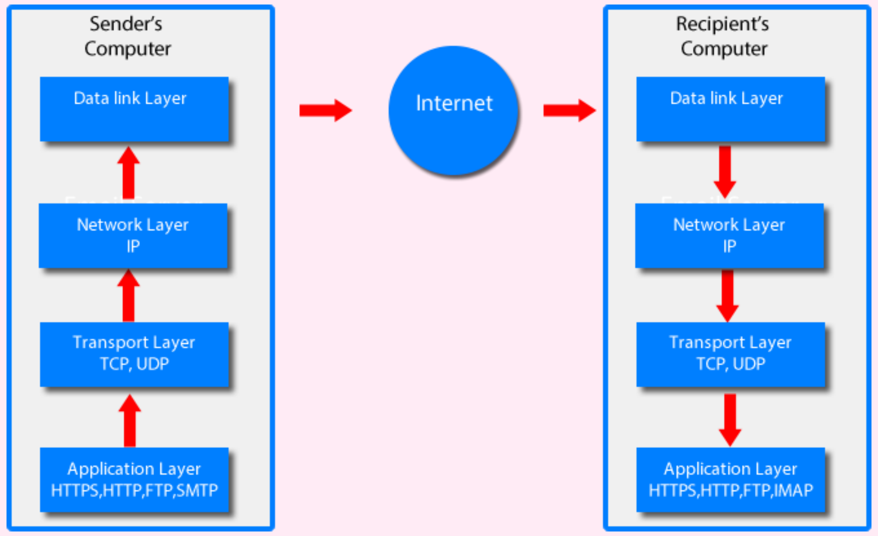
TCP/IP

A **network protocol** is a set of rules that allow computers to communicate and exchange information over a network.

TCP (Transport Control Protocol) When files are sent over the internet they are broken up into small chunks called **packets**. A typical packet size is 1500 bytes. When they arrive at the destination computer they are reassembled back into the original format. TCP handles and controls all this. TCP waits for acknowledgements to verify whether the packets have reached their destination. TCP will also retransmit packets of they have not arrived at the destination or become corrupted.

The **internet protocol** is a set of rules that govern the transmission of data across the internet. The TCP and IP work closely together and are referred to as TCP/IP.

The **TCP/IP stack** is made up of four layers that pass data between each layer.



The **application layer** contains protocols associated to the particular application such as HTTP, HTTPS for web browsers, FTP for file transfer or SMTP and POP3 for email for instance. The application layer interacts with the user via appropriate application software (eg web browser / ftp client / email client).

The **transport layer** establishes the end to end connection. When files are sent over the internet they are broken up into small chunks called **packets**. When they arrive at the destination computer they are reassembled back into the original format. It is the role of the transport layer to split the data into packets and pass the data onto the **network** layer. On the recipient's computer the transport layer reassembles the packets into the original form. TCP and UDP are the main protocols used in this layer. The packets are given an ID by this layer to allow them to be reassembled. The transport layer chooses the port number for sender and receiver.

Consider the following message: "Friends, Romans, Countrymen lend me you ears. I have come to bury Caesar not to praise him". The transport layer will break it up into packets:

Packet 1/4	**Packet 2/4**	**Packet 3/4**	**Packet 4/4**
Source Port: 25 Destination Port: 110	Source Port: 25 Destination Port: 110	Source Port: 25 Destination Port: 110	Source Port: 25 Destination Port: 110
Friends, Romans, Countrymen	lend me you ears.	I have come to bury Caesar	not to praise him

The **network layer** adds the source and destination IP address and route the packets over the network. At the destination the network layer strips out the IP addresses.

Sender IP address: 211.17.23.3 Receiver IP address: 143.27.98.21	Sender IP address: 211.17.23.3 Receiver IP address: 143.27.98.21	Sender IP address: 211.17.23.3 Receiver IP address: 143.27.98.21	Sender IP address: 211.17.23.3 Receiver IP address: 143.27.98.21
Packet 1/4	Packet 2/4	Packet 3/4	Packet 4/4
Source Port: 25 Destination Port: 110	Source Port: 25 Destination Port: 110	Source Port: 25 Destination Port: 110	Source Port: 25 Destination Port: 110
Friends, Romans, Countrymen	lend me you ears.	I have come to bury Caesar	not to praise him

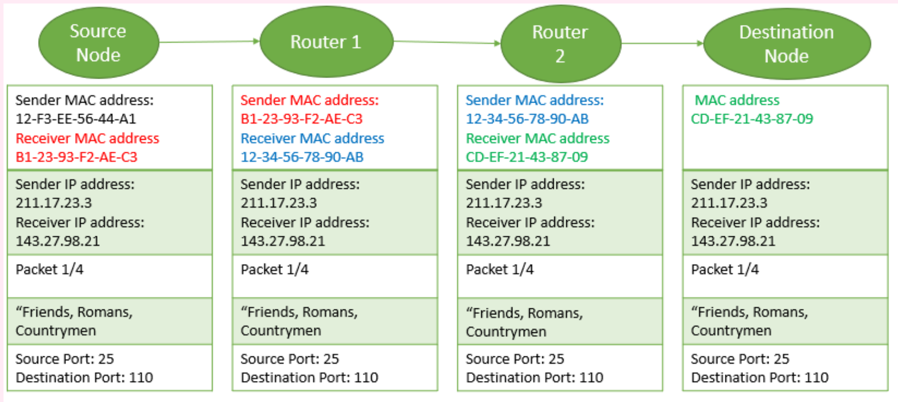
The **data link layer** has a network card and deals with the physical connection and adds the physical addresses (MAC address) of the hardware to the packets that it receives from the network layer. Each network interface card has a unique MAC address that is a 12 digit hexadecimal code (e.g. 12-F3-EE-56-44-A1). For each step the sender and receiver MAC address is removed then a new sender and receiver MAC address is added. The receiver MAC address becomes the sender MAC address.

The purpose of a **MAC address** is to provide a unique *hardware address* every *node* on a network. A node is a point at which a device (e.g., a computer, printer or router) is connected to the network.

MAC address and hopping

When data are sent over the internet they pass through a number of routers. The packets contain the destination IP address but the packets need also to include information on the route it takes across the internet. Packets need to know which routers to hop between. So packets need to know the MAC address of the next router. At each step the MAC address is stripped out, and the source MAC address is replaced by the destination address of the previous hop and the destination MAC address is replaced by the MAC address of the next router.

At each step the source MAC address is replaced by the destination address of the current node and the destination MAC address is replaced by the MAC address of the next node



Source Node	Router 1	Router 2	Destination Node
Sender MAC address: 12-F3-EE-56-44-A1 Receiver MAC address B1-23-93-F2-AE-C3	Sender MAC address: B1-23-93-F2-AE-C3 Receiver MAC address 12-34-56-78-90-AB	Sender MAC address: 12-34-56-78-90-AB Receiver MAC address CD-EF-21-43-87-09	MAC address CD-EF-21-43-87-09
Sender IP address: 211.17.23.3 Receiver IP address: 143.27.98.21	Sender IP address: 211.17.23.3 Receiver IP address: 143.27.98.21	Sender IP address: 211.17.23.3 Receiver IP address: 143.27.98.21	Sender IP address: 211.17.23.3 Receiver IP address: 143.27.98.21
Packet 1/4	Packet 1/4	Packet 1/4	Packet 1/4
"Friends, Romans, Countrymen	"Friends, Romans, Countrymen	"Friends, Romans, Countrymen	"Friends, Romans, Countrymen
Source Port: 25 Destination Port: 110	Source Port: 25 Destination Port: 110	Source Port: 25 Destination Port: 110	Source Port: 25 Destination Port: 110

A port is the end point between a networked device and other networked devices through which packets enter and leave. By convention depending on the protocol a specific port number will be used. A client port number is temporarily assigned for the duration of a connection. At the server end the service continually listens waiting for instructions from clients. The port numbers are added to the packet in the transport layer and determine which application layer protocol to use

Port number	Protocol
20	FTP
22	SSH
25	SMTP
80	HTTP
110	POP3
443	HTTPS

A **socket** is a connection between two applications over a network that allows data to be exchanged. A socket includes the IP address and the port number. The port number and IP address together are called the socket. e.g. 127.0.0.1: 5000 where 127.0.0.1 is the IP address and 5000 is the port number.

Application Layer Protocols

SSH allows remote login of a machine. It allows a secure connection where data being sent over the network are encrypted. A series of commands can be used to perform tasks on the remote machine. Allows remote login access by a network administrator to servers that could be at another location. Servers may not have their own terminal (no keyboard, mouse, or monitor) so the only way to access the servers is via SSH. SSH allows tunneling through which other applications eg SMTP can operate more secure

Some common shell commands

Command	Example	Explanation
pwd	pwd	States current directory (folder)
ls	ls	List contents of a folder
touch	touch filename.txt	Create file
mkdir	mkdir dirname	Make a folder
cd	cd dirname	Change directory (folder)
rm	rm filename.txt	Delete file
mv	mv file1.txt file2.txt	Rename file
cp	cp file1.txt file2.txt	Copy file
head	head file1.txt	Show the contents of the top of the file
echo	echo “Hello World” > file.txt	Print message
cat	cat file1.txt file2.txt	Concatenate (append) files.

Web server

- A web server is a computer on which are stored all the elements of a website including text images and other multimedia content as well as the HTML and CSS files.
- The web server will be able to understand HTTP requests from clients and respond to those requests.
- The web server will continuously be listening out for requests from clients.

HTTP – Hypertext transfer protocol

HTTP is the protocol used for the world wide web. A request for a web page from a server by a client web browser is made to a web server that is hosting the web site. The server then sends the web page to the client along with a status response.

Example request sent by client. GET is used to request a resource from a server by a client.

```
GET /index.html HTTP/1.1
```

POST is used to send data from the client to the server.

Example status response sent by server

```
HTTP/1.1 200 OK
```

Several GET requests may be needed to download different parts of a page e.g. images and other multimedia content.

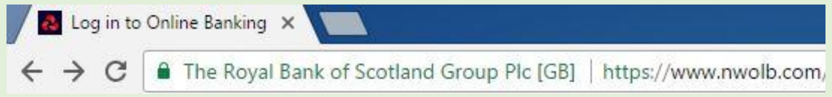
HTTP response status codes

1xx – Information
2xx – Success
3xx – redirection
4xx – Client error
5xx- Server error

HTTPS – Secure Hypertext transfer protocol

HTTPS is a secure way of transferring data between a web browser and a server. During transfer the data are **encrypted.** If the data are intercepted it is very difficult to find out what data are in the messages.

HTTPS is most often used for e-commerce and online banking, where sensitive data such as credit card numbers and passwords are encrypted.



FTP – File Transfer Protocol

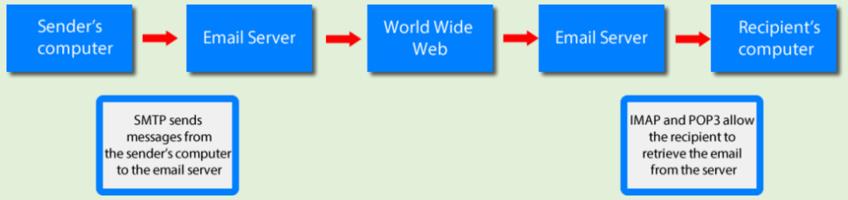
FTP is used to transfer files between two computers. It is usually used to download or upload large files between a client and server.

Common commands

open – start a session
pwd – present working directory
dir - list contents of directory
get – download a file
mget – download one or more files
mput – upload one or more files
put – upload a file
cd – change directory
close – close a session

Email

The email server distributes incoming email messages to users and sends outgoing messages across the internet. Email is sent by the sender client to the email server. It is stored there until the recipient requests access and then the email is forwarded to the receiver’s computer. The email server stores all mail. Outgoing mail uses SMTP to send from the client to the server. The sender mail server sends the message to the receiver mail server. Incoming mail uses POP3.



SMTP sends messages from the sender’s computer to the email server

IMAP and POP3 allow the recipient to retrieve the email from the server

Email protocols (SMTP, POP3)

- SMTP (simple mail transfer protocol): Sends the mail from the user client onto the mail server.
- POP3 (Post office protocol): Retrieves the mail from the mail server to the client (user) when requested and the email is removed from the server. Only allows retrieval of email onto a single network.

IP Addresses

IP addresses

IP address can be stored as dot decimal notation, and this is the notation that is most familiar to us. There are four numbers stored as a value between 0 and 255 that are separated by a point.

192. 168. 33. 22
11000000 10101000 00100001 00010110

IP address structure

IP addresses are split into a network identifier and a host identifier part. The IP address is made up of the network ID plus Host ID. Which part of the IP address corresponds to the network and host ID depends on the mask.

Subnet masking - Network

Subnet masks allow us to identify the network identifier part of the IP address. This is achieved by applying a bitwise logical AND to the IP address with the subnet mask.

Example

address	11000000 10101000 00100001 00010110	192.168.33.22
mask	11111111 11111111 11100000 00000000	255.255.224.0
network	11000000 10101000 00100000 00000000	192.168.32.0

Subnet masking - Host

Subnet masks allows us to identify the host part of the IP address. This is achieved by applying a logical NAND to the IP address with the subnet mask.

Example

address	11000000 10101000 00100001 00010110	192.168.33.22
mask	11111111 11111111 11100000 00000000	255.255.224.0
NOT mask	00000000 00000000 00011111 11111111	0.0.31.255
Host	00000000 00000000 00000001 00010110	0.0.1.22

Reserved IP addresses

- A host cannot have the following IP addresses if we have 8 bits for the host:
- xxx.xxx.xxx.0 Network identifier
- xxx.xxx.xxx.255 Broadcast across the whole network not to a single machine.
- Therefore there are $2^n - 2$ possible host where n is the number of bits for the host identifier.

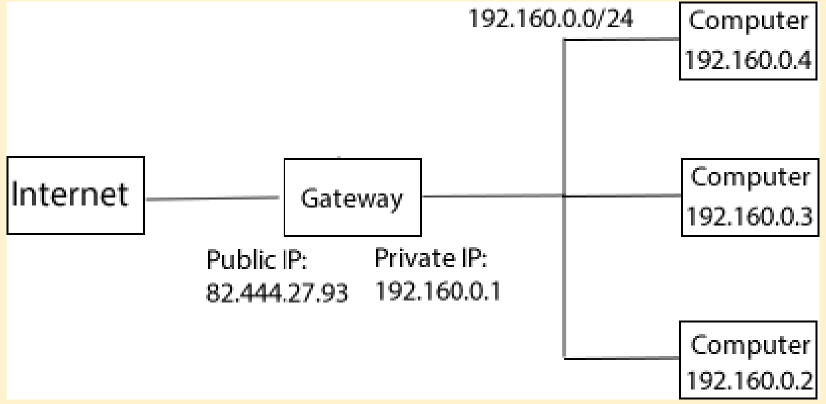
Classless inter-domain routing (CIDR)

Instead of an IP address with a mask you might also see an IP address presented as: 192.168.33.22/19. This means that the first 19 bits represent the network identifier and the remaining 13 bits represent the host part. This is therefore the same as the mask: 11111111 11111111 11100000 00000000 (255.255.224.0)

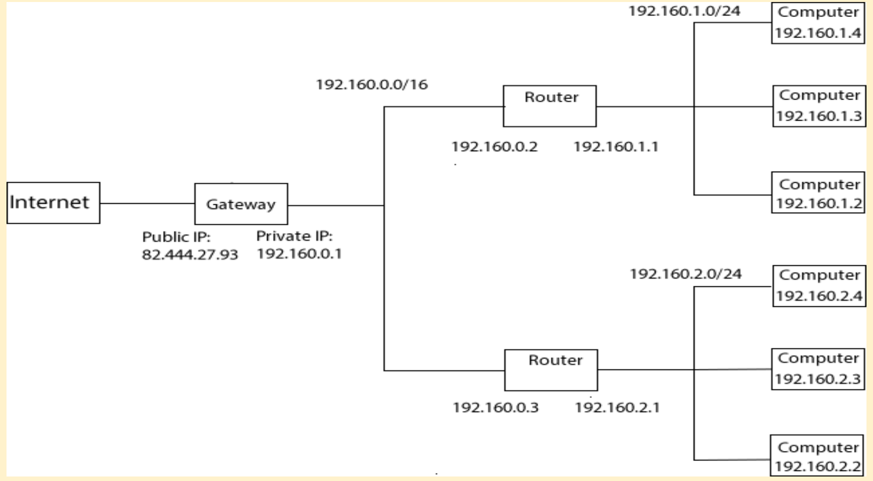
CIDR and subnet

Subnet Mask	Last two octets	CIDR	Number host bits	Number of hosts on each subnetwork
255.255.255.252	11111111.11111100	/30	2	2
255.255.255.248	11111111.11111000	/29	3	6
255.255.255.240	11111111.11110000	/28	4	14
255.255.255.224	11111111.11100000	/27	5	30
255.255.255.192	11111111.11000000	/26	6	62
255.255.255.128	11111111.10000000	/25	7	126
255.255.255.0	11111111.00000000	/24	8	254
255.255.252.0	11111110.00000000	/23	9	510
255.255.128.0	11100000.00000000	/19	13	1022

Schematic of a private network



Schematic of a private network with sub networks



Dynamic Host configuration Protocol (DHCP)

- Static IP addresses never change.
- Dynamic addressed are allocated each time a device connects to a network.
- Each time a host is connected an IP address will be allocated from a list of available addresses and will then be removed from the list.
- When a device is no longer connected the IP address is then added to the list of available addresses ready to be allocated once again.
- A DHCP server is used to perform this task.
- This is a way of preserving IPv4 addresses.

Public and private IP addresses

- Public IP addresses are routable and must be unique. This means that they can be addressed by any other device on the internet.
- Private IP addresses are non-routable and only need to be unique within the local area network in which they reside.
- Private IP addresses have helped conserve IPv4 addresses because they are not unique globally.
- This means that it is not possible to have direct external access private IP addresses.

IP standards

Every device on the internet needs to have a unique IP (internet protocol) address. Packets contain the sender's and receiver's IP address so that routers know where to direct the packets. Just like every house in the country has a unique postal address.

IPv4 has 32 bits and 4.3 billion addresses. All these have been exhausted now but IPv6 was introduced overcome this challenge and designed to run along side IPv4. Each IP address is split into 4 blocks and contains a denary value between 0 and 255. Example of IPv4 address: 172.92.255.01

IPv6 has 128 bits so the number of addresses is 3.4 x10^{38}. This is unlikely to be exhausted anytime soon! IP address is spilt into 8 blocks of up to 4 hexadecimal numbers. Example of IPv6 address: 2001:0db8:0a0b:12f0:f00e:1200:cc00:d001

Network address translation

- Network address translation (NAT) allows external connection to IP addresses within a private LAN
- Translates IP addresses between public and private IP addresses so that private IP addresses can be addressed
- Assigns a public address inside a private network. Limits the number of IP addresses and allows us to preserve IP addresses
- To access the internet outside the private network

Port forwarding

- Computers on a private network cannot be seen by computers outside local network
- Port forwarding is an application of NAT that allows remote computers on the internet to communicate with a specific computer within a private local-area network (LAN).
- A specific port on the gateway/router will be used to forward communications on to a specified computer on the private network
- Eg to access a web server on a private network a port number of 80 is normally used.

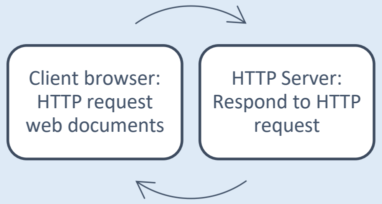
Client server model

Client server model

- **Server:** runs programs to serve applications to other computers
- **Client:** a computer that makes use of a service
- A client will make a request to a server. The server will run processes that are continuously listening for communications on a specific port. A server will serve many clients.
- When the server receives the request it then responds to the request
- Typical servers: file server, email server, FTP server, web server

Web server

- A web server is a computer on which are stored all the elements of a website including text images and other multimedia content as well as the HTML and CSS files.
- The web server will be able to understand HTTP requests from clients and respond to those requests.
- The web server will continuously be listening out for requests from clients.
- A browser requests a file from a web server using HTTP. When the request reaches the web server the file sent back to the browser, via HTTP.


Client browser: HTTP request web documents
HTTP Server: Respond to HTTP request

Websocket

- A websocket is an API that define the protocols between a client web browser and a server.
- A websocket protocol allows a persistent and dedicated full duplex (simultaneous two-way communication) connection between the client web browser and the server.
- Allows continuous transmission of data
- By comparison HTTP is half duplex and has greater overheads so is less efficient.
- Much faster because packets are smaller and contain less information so bandwidth requirement is reduced
- Useful for applications that require continuous real time data transfer ways between the client and server such as online gaming video conferencing, live video streaming anything that requires the constant transfer of data especially both

CRUD

There are four processes needed in a database with full functionality: Create, Retrieve, Update, Delete (CRUD).

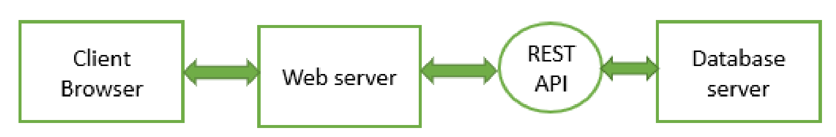
REST (Representational state transfer)

- REST is an API (Application Program Interface (API) that allows programs to work together. The functionality of one program can be accessed from another program
- REST runs on a server and allows clients to communicate with the server
- The database is connected to a client browser using REST API
- The HTTP request methods are mapped to SQL using the REST API following the principles of CRUD
- Javascript which runs on the client can communicate with the server through HTTP and can make calls to the REST API

CRUD, HTTP and SQL mapping

CRUD	HTTP	SQL
Create	POST	INSERT
Retrieve	GET	SELECT
Update	PUT	UPDATE
Delete	DELETE	DELETE

Web database architecture



JSON and XML

JSON (Javascript Object Notation) and XML (Extensible Markup Language) are standard methods of transferring data between a server and a client.

JSON Example

```
{“students":[
    { "firstName":“Thomas", "lastName":“Brown“,
“dateOfBirth":27/3/2001},
    { "firstName":“James", "lastName":“Frank“,
“dateOfBirth":13/4/2002}
   ]}
```

XML Example

```
<students>
 <student>
  <firstName>Thomas</firstName>
  <lastName>Brown</lastName>
  <dateOfBirth> 27/3/2001 </dateOfBirth>
 </ student >
 < student >
  <firstName>James</firstName>
  <lastName>Frank</lastName>
  <dateOfBirth> 13/4/2002 </dateOfBirth>
 </student>
</ students>
```

JSON versus XML

JSON	XML
Very easy to read	Contains tags so is not so easy to read
More compact less code	Lots of tags needed so is less compact
Only set data types can be used	Greater flexibility of data types
Syntax is very simple so easy to create	More complex syntax
Quick to parse	Slow to parse because it contains lots of tags

Thin versus thick client computing

Thin client

- Relies on a server to do much of the processing.
- The server needs to be extremely powerful to in order to be able to process all the requests from all clients on a network.
- A thin client computer can be low specification and does not need much hard disk storage or processing power.
- Much of the application software will be installed on the server.
- Client is essentially a terminal

Thick client

- All the applications are installed on the local machine.
- The processing is performed on the client.
- Very little reliance on the server

	Advantages	Disadvantages
Thin Client	Cheap low spec and old machines can be used for the clients Easy to maintain and manage software updates on a server Data are stored in one central location on the server so can be more secure Make it harder to pirate software	If server goes down the whole network is affected Need expensive very high performing server Need a good quality network to transfer data and requests between clients and server Can be a security risk as data are transferred over the network
Thick client	Do not need such a robust network and there is a lot less network traffic Do not need such high spec servers Different software can be installed on different machines	Need to have high specification clients Software needs to be on all client machines, thereby making maintenance and updating software more difficult