AQA Computer Science A-Level

4.9.4 The Transmission Control Protocol / Internet Protocol (TCP/IP)

Concise Notes

Specification:

4.9.4.1 TCP/IP:

Describe the role of the four layers of the TCP/IP stack (application, transport, network, link)

Describe the role of sockets in the TCP/IP stack

Be familiar with the role of MAC (Media Access Control) addresses

Explain what the well-known ports and client ports are used for and the differences between them.

4.9.4.2 Standard application layer protocols:

Be familiar with the following protocols:

- FTP (File Transfer Protocol)
- HTTP (Hypertext Transfer Protocol)
- HTTPS (Hypertext Transfer Protocol Secure)
- POP3 (Post Office Protocol (v3))
- SMTP (Simple Mail Transfer Protocol)
- SSH (Secure Shell)

Be familiar with FTP client software and an FTP server, with regard to transferring files using anonymous and non-anonymous access

Be familiar with how SSH is used for remote management

Know how an SSH client is used to make a TCP connection to a remote port for the purpose of sending commands to this port using application level protocols such as GET for HTTP, SMTP commands for sending email and POP3 for retrieving email

Be familiar with using SSH to log in securely to a remote computer and execute commands

Explain the role of an email server in retrieving and sending email

Explain the role of a web server in serving up web pages in text form

Understand the role of a web browser in retrieving web pages and web page resources and rendering these accordingly

4.9.4.3 IP address structure:

Know that an IP address is split into a network identifier part and a host identifier part

4.9.4.4 Subnet masking:

Know that networks can be divided into subnets and know how a subnet mask is used to identify the network identifier part of the IP address

4.9.4.5 IP standards:

Know that there are currently two standards of IP address, v4 and v6

Know why v6 was introduced

4.9.4.6 Public and private IP addresses:

Distinguish between routable and non-routable IP addresses

4.9.4.7 Dynamic Host Configuration Protocol (DHCP):

Understand the purpose and function of the DHCP system

4.9.4.8 Network Address Translation (NAT):

Explain the basic concept of NAT and why it is used

4.9.4.9 Port forwarding:

Explain the basic concept of port forwarding and why it is used

4.9.4.10 Client server model:

Be familiar with the client server model

Be familiar with the Websocket protocol and know why it is used and where it is used

Be familiar with the principles of Web CRUD Applications and REST:

- CRUD is an acronym for create, retrieve, update, delete
- REST enables CRUD to be mapped to database functions (SQL)

Compare JSON (Javascript object notation) with XML

4.9.4.11 Thin- versus thick-client computing

Compare and contrast thin-client computing with thick-client computing

TCP / IP

- Stands for transmission control protocol / internet protocol
- Used in all parts of the Internet to enable different devices to communicate
- Formed from four distinct layers that form the TCP / IP stack
    - Application
    - Transport
    - Network
    - Link
- Each layer is responsible for a seperate part of communication over the Internet

Layer	Role
Application	• Selects and uses the correct protocol to transmit data • Interacts with the user with application software like a web browser
Transport	• Establishes a virtual path between the sender and the receiver • Splits the transmission into packets • Each packet has a sequence number which identifies a packet's position in relation to other packets that form part of the same message • Also contained in each packet is the port number to be used which identifies the protocol in use
Network	• Provides the correct IP addresses for each packet's source and destination • Routers work within the network layer, using the IP addresses on a packet to send it to its destination.
Link	• Controls physical connections between pieces of hardware in a network • Adds MAC addresses to packets which it receives from the network layer • The MAC addresses identify the hardware to which a packet should be sent • These MAC addresses change with every hop through a network

At the receiving end

- Once a packet has been received by its intended recipient, it is stripped of its extra information by reversing the TCP / IP stack
    - The link layer removes MAC addresses from the packet
    - The network layer removes IP addresses
    - The transport layer uses the packet's port number to determine the correct application to send the packet to
    - The transport layer also uses the packet's sequence number to ensure that it is in the correct position relative to other packets in the same transmission
    - The application layer receives the packets and displays the information to the user accordingly

Socket addresses

- Formed of an IP address, followed by a colon, followed by the port number in use
- Identifies which application on the recipient device a packet should be sent to

Well-known ports

Protocol	Port(s)	Purpose
FTP (File transfer protocol)	20 & 21	• Sending files between devices • Clients can access FTP servers anonymously or non-anonymously
SSH (Secure shell)	22	• Remote management of computers • Requires a username / password combination • Encrypts information during transmission
HTTP (Hypertext transfer protocol)	80	• Sending web pages in text form from web servers to clients
HTTPS (Hypertext transfer protocol secure)	443	• Performs the same function as HTTP, but encrypts information during transmission • Keeps information sent by clients secure • Helps to prevent information from being tampered with or modified during transmission
POP3 (Post office protocol 3)	110 & 995	• Retrieving emails from an email server • Checks for and downloads any new messages
SMTP (Simple mail transfer protocol)	25, 587 & 465	• Sending emails • Used between a client and an email server

The structure of IP addresses

- IP addresses are split into two parts:
  - a network identifier
  - a host identifier
- Computers in the same network share the same network identifier
- Networks can be divided into smaller networks, called subnets
- Each subnet has a different network identifier
- The network identifier part of an IP address can be determined with a subnet mask
- To find a device's network identifier:
  - convert its IP address and subnet mask to binary
  - perform a logical AND operation on the two binary numbers
  - convert the result to decimal
- The more bits that are assigned to the network identifier of an IP address, the more different subnets a network can have
- The more bits that are assigned to the host identifier, the more different devices can be connected to each subnet simultaneously

IP address standards

- There are two types of IP address in common use:
  - version four (IPv4)
  - version six (IPv6)

IPv4

- IPv4 addresses are dotted quad numbers
- They consist of four parts that are separated by dots
- Each of the four parts of an IPv4 address is assigned one byte (eight bits) allowing for numbers from 0 to 255 to be represented
- There are slightly over 4 billion (256^4) unique IPv4 addresses
- IPv4 addresses are in short supply
- The number of devices on the Internet that require a routable IP address is increasing so rapidly that a new version of IP address had to be created

IPv6

- IPv6 addresses are formed of eight blocks separated by colons
- Each block contains four hexadecimal characters (a-z and 0-9)
- IPv6 addresses use 128 bits which allows for far more unique permutations than IPv4 addresses

Public and private IP addresses

- IP addresses can be either routable or non-routable (public or private)
- If every device that is connected to a network had its own public IP address, there wouldn't be enough IP addresses to go around
- Each home or business that requires internet access has a small number of public IP addresses
- Most homes have just one public IP address while some businesses will have multiple addresses
- Routable IP addresses are globally unique
- Millions of devices can have the same non-routable IP address
- Global authorities are responsible for assigning routable IP addresses
- This ensures that the same address is never issued twice

Dynamic host configuration protocol (DHCP)

- The number of available private IP addresses within a private network is limited
- DHCP is used to assign IP addresses to devices as they join a network
- DHCP uses a pool of available IP addresses to allocate IP addresses
- IP addresses are allocated for the duration of a device's session
- Once a device leaves the network, the IP address that the device was using is returned to the pool for reallocation

Network address translation (NAT)

- Routers have two network interface cards, one for their private network and one for the Internet
- Therefore, routers have two IP addresses, a private address and a public address
- When a computer on the private network communicates with a device on the Internet, the device couldn't respond to the computer directly because the computer's IP address is not globally unique
- NAT gets around this problem. When a device on the private network needs to communicate with a device on the Internet, it sends packets through the router, which:
    - makes a record of the packet
    - replaces the private IP address of the computer with its own routable IP address
- When a response is received, it is sent to the router's public IP address, which then forwards the response to the correct private IP address by using the record it made when sending the packet

Port forwarding

- Port forwarding is used when a client needs to communicate with a server that is connected to a private network
- The client sends packets to the public IP address of the router belonging to the server's private network
- The packets sent by the client contain the port number of the application running on the server that the client wishes to access
- The private network's router then forwards the packets to the server using NAT

The client server model

- In a network that uses the client server model, clients send request messages to servers, which reply to the clients with response messages
- These messages may contain
    - requested information
    - a confirmation that a requested action has been completed
    - a message explaining why the requested action hasn't been completed

APIs

- Stands for application programming interface
- A set of protocols relating to how different applications communicate with each other
- Define how interaction between the applications should be carried out
- Allows applications to make use of other applications

The websocket protocol

- An example of an API which operates in the application layer of the TCP / IP stack
- Can be used to provide a constant stream of information between two devices, usually a client's web browser and a server
- The connection created by the websocket protocol is full-duplex, meaning that data can be transmitted in both directions at the same time
- Allows for fast transmission of data by reducing the size of packet headers
- Used in video streaming, online games and instant messaging

Web CRUD Applications and REST

CRUD

- An acronym for **c**reate, **r**etrieve, **u**pdate, **d**elete
- Four commands that can be used to query online databases
- Each of the four CRUD commands has a SQL equivalent

CRUD	SQL
Create	INSERT
Retrieve	SELECT
Update	UPDATE
Delete	DELETE

REST

- Acronym for **re**presentational **s**tate **t**ransfer
- A design methodology for online database applications that are queried with a web browser
- Uses the four HTTP request methods POST, GET, PUT and DELETE to query databases

HTTP	SQL
POST	INSERT
GET	SELECT
PUT	UPDATE
DELETE	DELETE

When a client needs to connect to a database, it follows these steps:

1. Client-server request made by the client to the web browser
2. Web browser responds with the requested web page
3. This text file contains JavaScript which loads an API. The API uses REST to enable the database server to be queried by the client with the use of HTTP request methods.
4. The client sends HTTP requests to the database server.
5. The database server responds to the client's requests using either JSON or XML.
6. The client's browser processes the JSON or XML and displays the response to the user.

XML and JSON

- Database servers deliver responses to queries using either XML or JSON
- XML stands for **ex**tensible **m**arkup **l**anguage
- JSON stands for **J**ava**S**cript **o**bject **n**otation
- JSON is more compact, easier to read, easier to create and faster for computers to process than XML
- XML is sometimes seen to be more flexible than JSON

XML	JSON

<Department>
<Subject>
<Name>Physics</Name>
<Board>Edexcel</Board>
</Subject>
<Subject>
<Name>Maths</Name>
<Board>AQA</Board>
</Subject>
</Department>
	{
	"Department": {
	"Subject": [
	{
	"Name": "Physics",
	"Board": "Edexcel"
	},
	{
	"Name": "Maths",
	"Board": "AQA"
	}
	]
	}
	}

Thin- and thick-client computing

- Networks can be configured as either thin-client or thick-client networks.

Thin-client networks

- The majority of the network's processing power belongs to servers
- Servers provide services and resources including storage and processing
- It's easy to add new clients and the clients themselves are inexpensive machines
- Software updates and security can be managed from the server, allowing for greater centralised control of the network
- Require a powerful server which is expensive and requires expertise to set up and maintain

Thick-client networks

- The clients are powerful enough to provide their own processing power and storage
- This independence eliminates the requirement for a server, although it is possible for thick-client networks can make use of a server
- Require more powerful clients than thin-client networks, making the network expensive to set up
- The cost and expertise required in setting up and maintaining an expensive server is done away with
- Thick-client networks are harder to maintain because there is no facility to issue updates and manage security from a central server
- Compared to thin-client networks which suffer from high volumes of traffic communicating between clients and the server, thick-client networks boast much quieter communication channels which reduces the likelihood of data collisions