

# The internet and its uses

## IGCSE Computer Science 0478 · Topic 5 revision notes

Everything from the 7 lessons of this topic on a few pages: the key idea of each lesson, the diagram to remember, the points to learn and the traps examiners set. Read it, cover it, test yourself.

- |                                                        |                                                   |
|--------------------------------------------------------|---------------------------------------------------|
| 1. <u>The internet, the world wide web and URLs</u>    | 4. <u>Session and persistent cookies</u>          |
| 2. <u>HTTP, HTTPS and web browsers</u>                 | 5. <u>Digital currency and blockchain</u>         |
| 3. <u>How a web page is located and retrieved: DNS</u> | 6. <u>Cyber security threats</u>                  |
|                                                        | 7. <u>Keeping data safe from security threats</u> |

# 1 The internet, the world wide web and URLs

## Syllabus 5.1

The **internet** is the **infrastructure**: a worldwide network of interconnected networks, made of computers, routers, cables and wireless links. The **world wide web (WWW)** is the **collection of websites and web pages** that are accessed using the internet. A **uniform resource locator (URL)** is a **text-based address** for a web page; it can contain the **protocol**, the **domain name** and the **web page or file name**.

the parts of a URL

**https://**   **www.example.com**   **/shop/boots.html**

**protocol**

rules used  
to transfer it

**domain name**

the web server,  
found using DNS

**web page / file name**

which page or file  
on that server

a URL is a text-based address that people can remember

Fig. 1: the three parts of a URL

### Worked example: Split a URL

Identify the parts of `https://www.riverside-school.org/sports/fixtures.html`

1. The protocol comes first, before the colon and two slashes: `https`.
2. The domain name runs up to the next single slash: `www.riverside-school.org`.
3. Everything after that is the path and file name: `/sports/fixtures.html`.

**Answer:** Protocol `https`; domain name `www.riverside-school.org`; file name `/sports/fixtures.html`

### Use the key words.

- ✓ **Internet:** the infrastructure, a global network of interconnected networks (hardware).
- ✓ **World wide web:** the collection of websites and web pages accessed using the internet.
- ✓ **URL:** a text-based address for a web page.
- ✓ **URL parts:** protocol, domain name, web page / file name.

### Watch out!

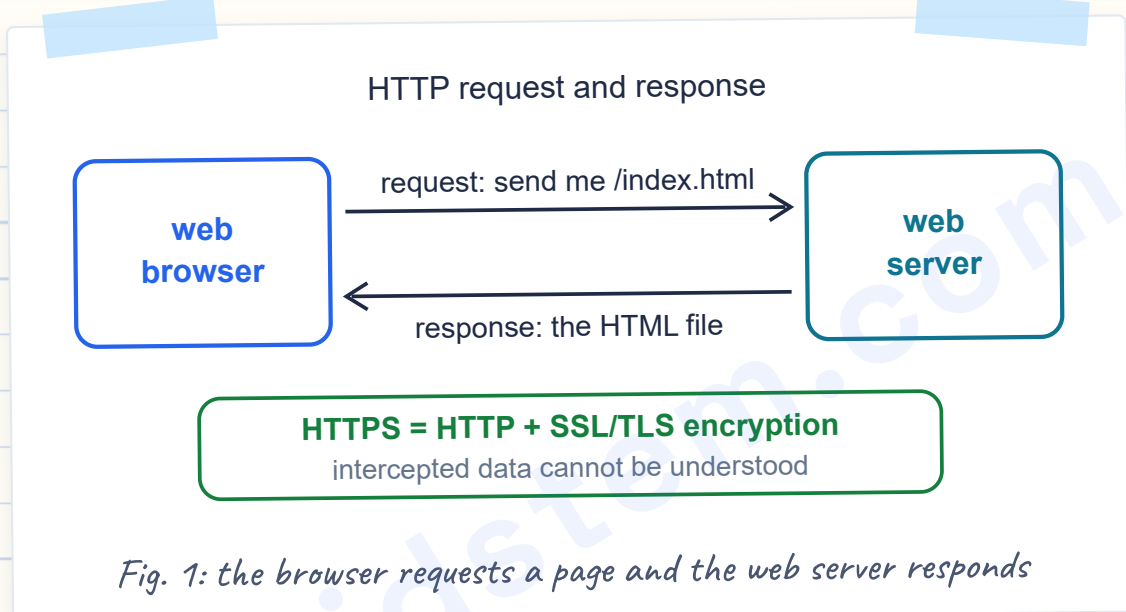
"The web is the internet" or "the internet is a website" earns nothing. Use the key words: the internet is the **infrastructure**; the web is the **collection of web pages** accessed **using** the internet.

A URL is not the same as an IP address. A URL is **text-based**; the IP address is the numeric address of the server, found using the DNS.

## 2 HTTP, HTTPS and web browsers

### Syllabus 5.1

**Hypertext transfer protocol (HTTP)** is the set of rules for transferring web pages between a **web browser** (the client) and a **web server**: the browser sends a **request**, the server sends a **response** containing the files. **HTTPS** is HTTP made secure: the data is **encrypted** using the **SSL/TLS** protocol, so if it is intercepted it cannot be understood. The main purpose of a **web browser** is to **render HTML** and **display web pages**.



### Protocol, then program.

- ✓ **HTTP:** protocol for requesting and transferring web pages between browser and web server.
- ✓ **HTTPS:** HTTP with encryption by SSL/TLS; intercepted data cannot be understood.
- ✓ **Browser purpose:** render HTML and display web pages.
- ✓ **Browser functions:** bookmarks, history, tabs, cookies, navigation tools, address bar.

### Watch out!

HTTPS does not stop data being intercepted. It means that intercepted data is **encrypted**, so it is meaningless without the key. Write "cannot be understood", not "cannot be stolen".

A web browser is not a search engine. The browser displays pages; a search engine is a website that you visit using a browser.

### 3 How a web page is located and retrieved: DNS

#### Syllabus 5.1

Computers on the internet are found by **IP address**, not by name. When a URL is entered, the **browser** sends the **domain name** to a **domain name server (DNS)**, which looks up and returns the matching **IP address**. The browser then sends an **HTTP** request to the **web server** at that IP address. The web server sends back the page as **HTML** files, and the browser **renders the HTML** to display the page.

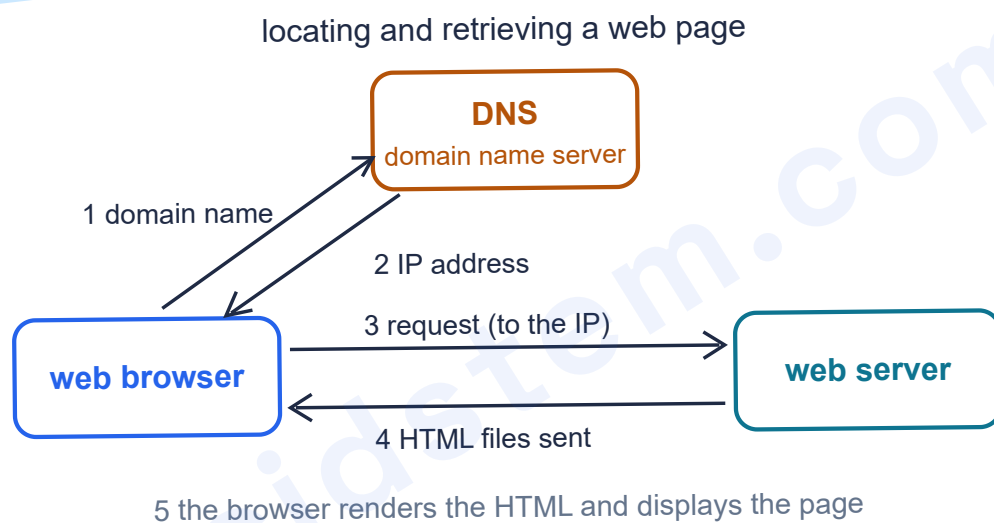


Fig. 1: the DNS turns the domain name into an IP address, then the web server sends the HTML

#### Worked example: Exam-style answer

Describe the role of the DNS when a user enters `www.example.com`.

1. The browser sends the domain name `www.example.com` to the DNS.
2. The DNS looks up the domain name in its database to find the matching IP address; if it does not hold it, it asks another DNS.
3. The DNS returns the IP address to the browser, which can now contact the web server.

**Answer:** The DNS translates the domain name into the IP address of the web server and sends it back to the browser.

### One job each.

- ✓ **Browser:** sends the domain name to the DNS, requests the page, renders the HTML.
- ✓ **DNS:** looks up the domain name and returns the matching IP address.
- ✓ **Web server:** at that IP address; sends back the HTML files.
- ✓ **Order:** URL, DNS, IP address, request, HTML, render.

### Watch out!

The DNS does not send the web page. It only returns the **IP address**. The **web server** sends the HTML. Mixing these up loses most of the marks.

thelucidstem.com

## 4 Session and persistent cookies

### Syllabus 5.1

A **cookie** is a **small text file** sent by a **web server** and **stored by the browser** on the user's computer. It holds data about the user's visit; the browser sends it back to the site on later requests, so the site can **recognise the user**. A **session cookie** is stored in temporary memory and is **deleted when the browser is closed**. A **persistent cookie** is saved on the hard drive and **remains until its expiry date** or until the user deletes it.

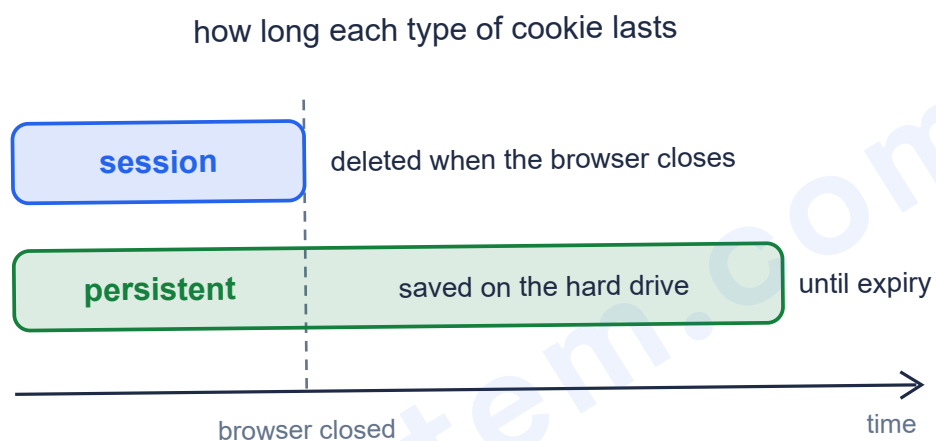


Fig. 1: a session cookie ends with the browser session; a persistent cookie lasts until it expires

### Worked example: Pick the type

A user logs in to a bank website. The site must log the user out as soon as they close the browser. Which type of cookie should store the login?

1. The login must not survive after the browser is closed.
2. A persistent cookie would keep the user logged in until it expired, which is a security risk on a shared computer.
3. A session cookie is deleted when the browser is closed.

**Answer:** A session cookie

## How long does it last?

- ✓ **Cookie:** small text file sent by a website, stored by the browser on the user's device.
- ✓ **Session:** temporary memory; deleted when the browser is closed.
- ✓ **Persistent:** hard drive; kept until it expires or is deleted.
- ✓ **Uses:** personal details, preferences, shopping cart, login details.

## Watch out!

A session cookie is not deleted when the user leaves the page or the website. It is deleted when the **browser is closed**.

thelucidstem.com

## 5 Digital currency and blockchain

### Syllabus 5.2

A **digital currency** is one that **only exists electronically**: there are no physical notes or coins. It is used to pay for goods and services and to transfer money between people over the internet. Many digital currencies (**cryptocurrencies**) track every transaction with a **blockchain**: a **digital ledger**, that is a **time-stamped series of records that cannot be altered**. Each block stores its transactions, a timestamp, its own **hash** and the **hash of the previous block**, and copies of the chain are held by every computer in the network.

each block stores the hash of the block before it



edit Block 1 and its hash changes, so Block 2 no longer matches

Fig. 1: each block is linked to the one before by storing that block's hash

### Worked example: Tracing a change

Block 1 has hash 7a3f and Block 2 stores previous hash 7a3f. A criminal changes the amount in Block 1. Explain what happens.

1. The data in Block 1 has changed, so its hash is recalculated and is no longer 7a3f.
2. Block 2 still stores previous hash 7a3f, which no longer matches Block 1.
3. The link is broken, so every block after Block 1 becomes invalid, and the copy no longer matches the copies held by the other computers.

**Answer:** The change is detected and rejected; the ledger cannot be altered unnoticed.

### *Linked by hashes.*

- ✓ **Digital currency:** only exists electronically.
- ✓ **Blockchain:** a digital ledger, a time-stamped series of records that cannot be altered.
- ✓ **Block:** transactions, timestamp, its own hash, previous block's hash.
- ✓ **Tampering:** hash changes, chain breaks, other copies disagree.

### *Watch out!*

Do not describe a digital currency as "money paid with a bank card". The key point is that the currency **only exists electronically**.

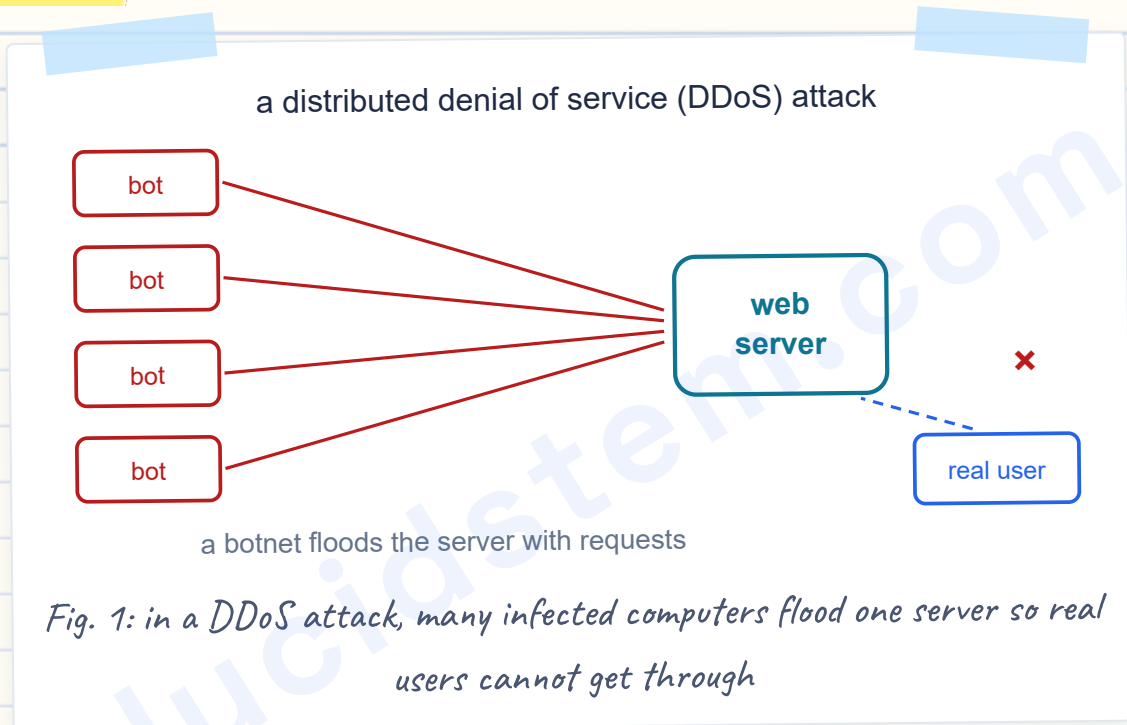
Blockchain does not stop transactions being made, and it is not encryption of the money. It makes the **record** of transactions tamper-evident and permanent.

thelucidstem.com

## 6 Cyber security threats

### Syllabus 5.3

A **cyber security threat** is any attempt to damage, steal or gain unauthorised access to data or computer systems. For every threat, an exam answer needs the **process** (how the attack is carried out) and the **aim** (what the attacker wants). Some threats attack the **technology**: **brute-force attacks**, **data interception**, **DDoS attacks**, **hacking** and **malware**. Others attack **people** or redirect them: **phishing**, **pharming** and **social engineering**.



### Worked example: Brute-force arithmetic

A password is 6 lowercase letters. Software tries 1 million passwords a second. What is the longest a brute-force attack could take?

1. Each position has 26 possibilities, so there are  $26^6 = 308\,915\,776$  combinations.
2. At 1 000 000 guesses a second:  $308\,915\,776 \div 1\,000\,000 \approx 309$  seconds.
3. 309 seconds is about 5 minutes.

**Answer:** About 5 minutes. Adding the digits 0 to 9 gives  $36^6 = 2\,176\,782\,336$  combinations, about 36 minutes; longer passwords grow much faster.

## How, and why.

- ✓ **Process:** describe how the attack is carried out, step by step.
- ✓ **Aim:** say what the attacker wants: data, access, money or disruption.
- ✓ **Malware:** virus, worm, Trojan horse, spyware, adware, ransomware: know the difference.
- ✓ **Phishing vs pharming:** a link in a message, versus redirection by malicious code.

## Watch out!

"Hacking is when someone gets your data" is too vague. Say **unauthorised access** to a computer system, and give an aim such as stealing or deleting data.

A virus needs a host file and is spread when that file is run; a **worm** is standalone and spreads itself across networks. Do not use "virus" as the word for all malware.

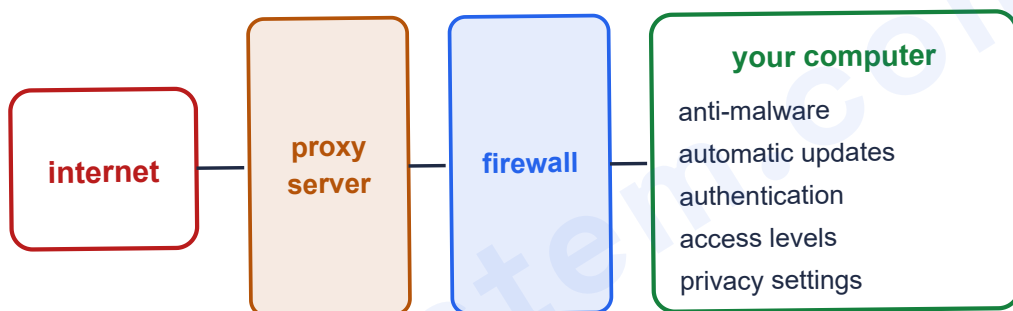
Phishing needs the user to **click a link** in a message. Pharming needs **no click**: the user types the right address and is still redirected by malicious code.

## 7 Keeping data safe from security threats

### Syllabus 5.3

Data is kept safe by **several layers** of protection. **Firewalls** and **proxy servers** control traffic between a network and the internet. **Anti-malware** (including anti-virus and anti-spyware) and **automatic software updates** protect each computer. **Authentication** (username and password, biometrics, two-step verification) and **access levels** control who can see what. **SSL** encrypts data in transit. **Privacy settings**, and users **checking the spelling and tone** of messages and the **URL attached to a link**, protect against people-based attacks. Exam answers must explain **how** a solution keeps data safe.

layers of protection between the internet and your data



plus the user: check URLs, spelling and tone before clicking

Fig. 1: security is built in layers, from the network edge to the computer and its user

### Worked example: Match the solution

An email says: "Dear custmer, your acount is locked. Click here within 24 hours." The email claims to be from Brightbank, but the link shows [www.brlghtbank-login.com](http://www.brlghtbank-login.com). Which checks reveal that it is phishing?

1. Spelling: "custmer" and "acount" are misspelt.
2. Tone: a generic greeting and an urgent deadline put the user under pressure.
3. URL: the domain uses the digit l in place of the letter i, so it is not the genuine Brightbank site.

**Answer:** Checking spelling and tone, and checking the URL attached to the link.

## Name it, then explain it.

- ✓ **Network:** firewall filters traffic; proxy server hides IP addresses, filters and absorbs DDoS.
- ✓ **Computer:** anti-malware scans and quarantines; automatic updates patch weaknesses.
- ✓ **Account:** passwords, biometrics, two-step verification; access levels limit what each user sees.
- ✓ **User and data in transit:** check spelling, tone and URLs; privacy settings; SSL encrypts.

## Watch out!

"Use a password" is not enough for marks. Explain **how** it helps: e.g. a long, complex password has so many combinations that a brute-force attack would take too long to find it.

A firewall does not remove malware that is already on the computer; that is the job of **anti-malware**. A firewall **filters traffic**.