

Kevin's guide to the Internet:

A bit like a bucket of manure thrown at you, full of intricate lumps but at the end of the day you'll remember being hit by the bucket!

This is an aide memoire and a potted review on 'some' of the more important parts of the thing we call 'The Internet'.

Enjoy

What is the internet

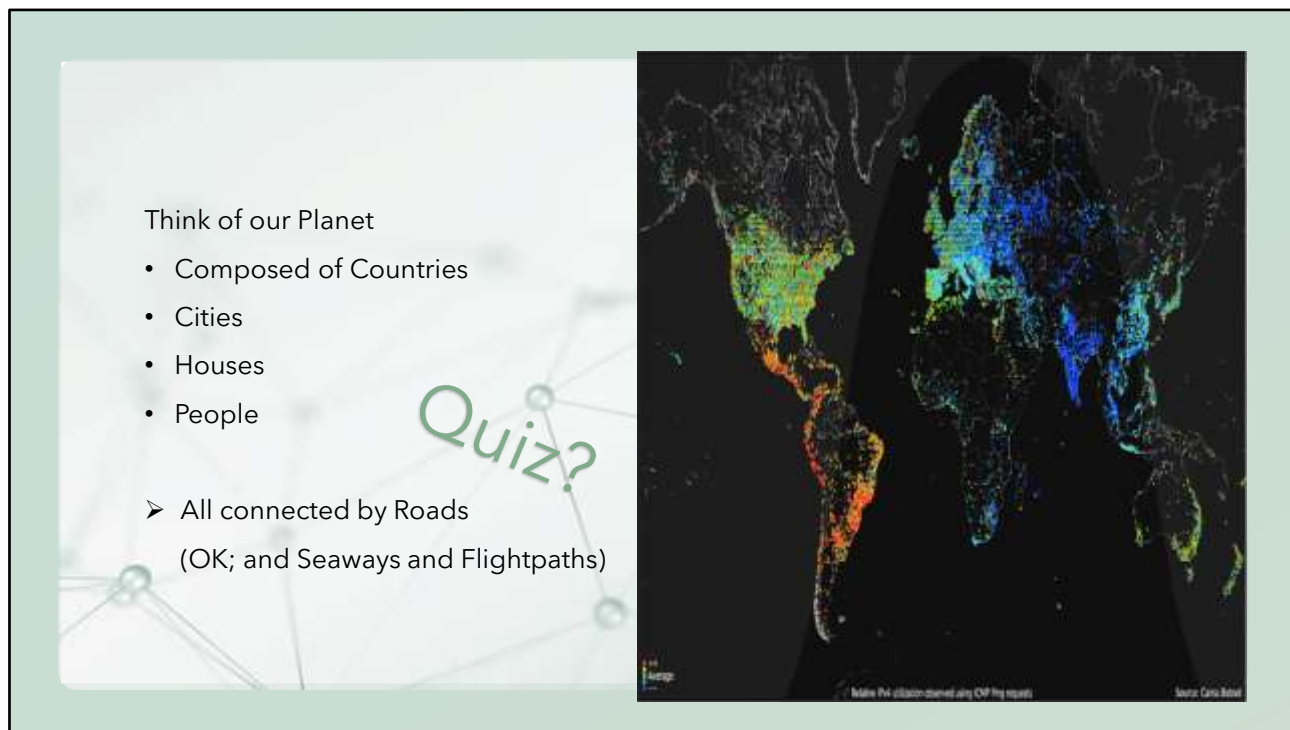


This section is a top level view, broken down into:

- Domains
- Servers
- Sites
- Pages

There is an awful lot more, but these four suffice 😊

What is the internet



Think of our Planet

- Composed of Countries
- Cities
- Houses
- People

➤ All connected by Roads
(OK; and Seaways and Flightpaths)

Quiz?

I like to compare the Internet with the real world, composed into distinct components

Guess how many of each?

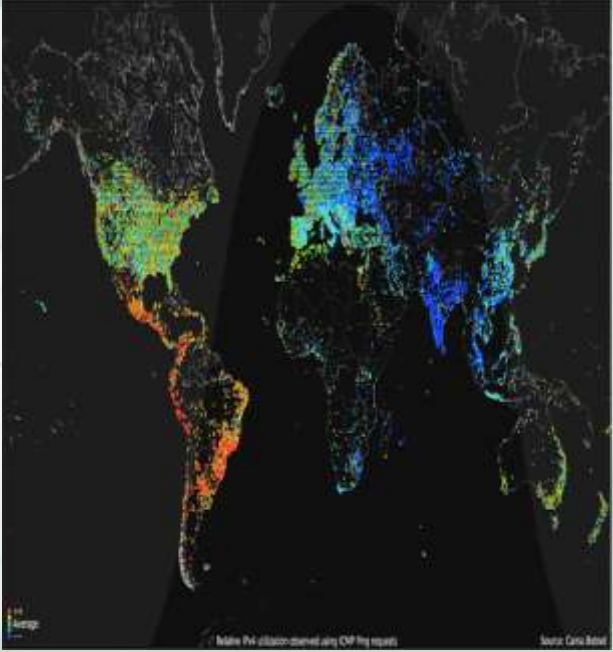
Answers next page.

What is the internet

Think of our Planet

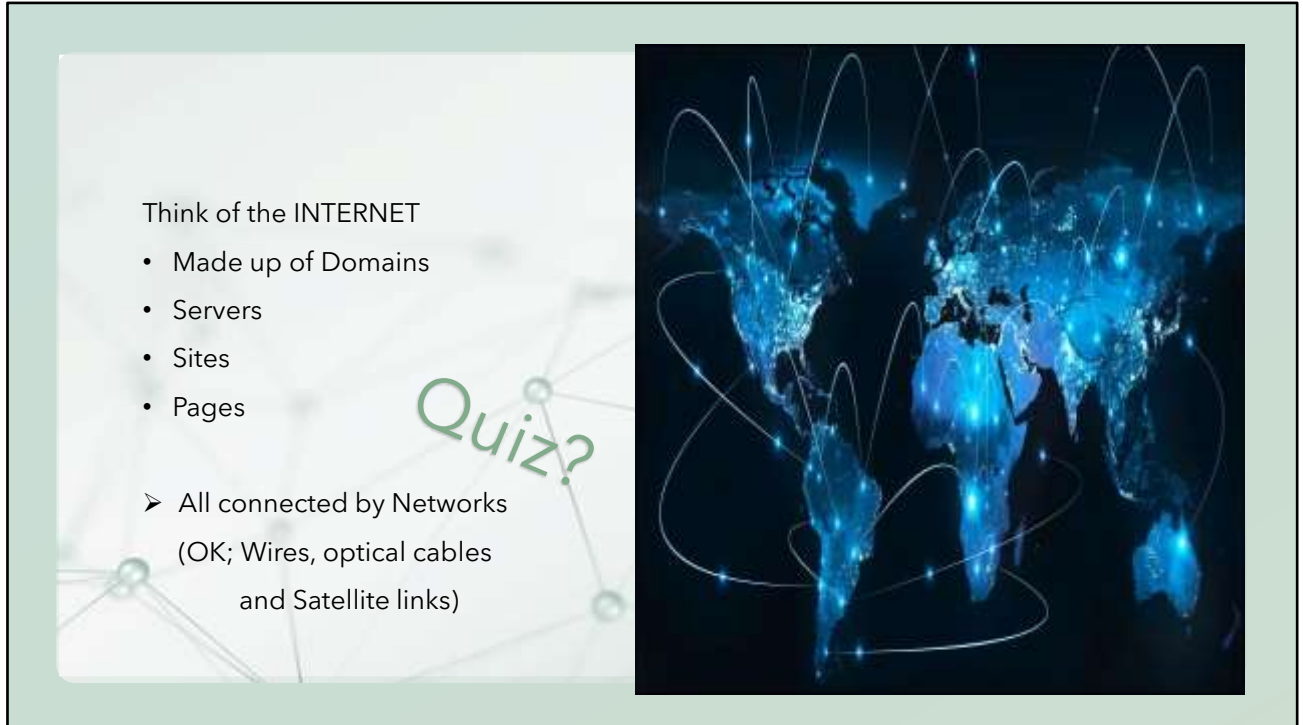
- Composed of Countries **~195**
- Cities **~10,000**
- Houses **~1.2 Billion**
- People **~7.8 Billion**

➤ All connected by Roads
(OK; and Seaways and Flightpaths)



Did you get it right?

What is the internet



Think of the INTERNET

- Made up of Domains
- Servers
- Sites
- Pages

➤ All connected by Networks
(OK; Wires, optical cables
and Satellite links)

Quiz?

Back to the Internet, and (some of) its components.


Again can you guess how many of each, answers next.

What is the internet

Think of the INTERNET

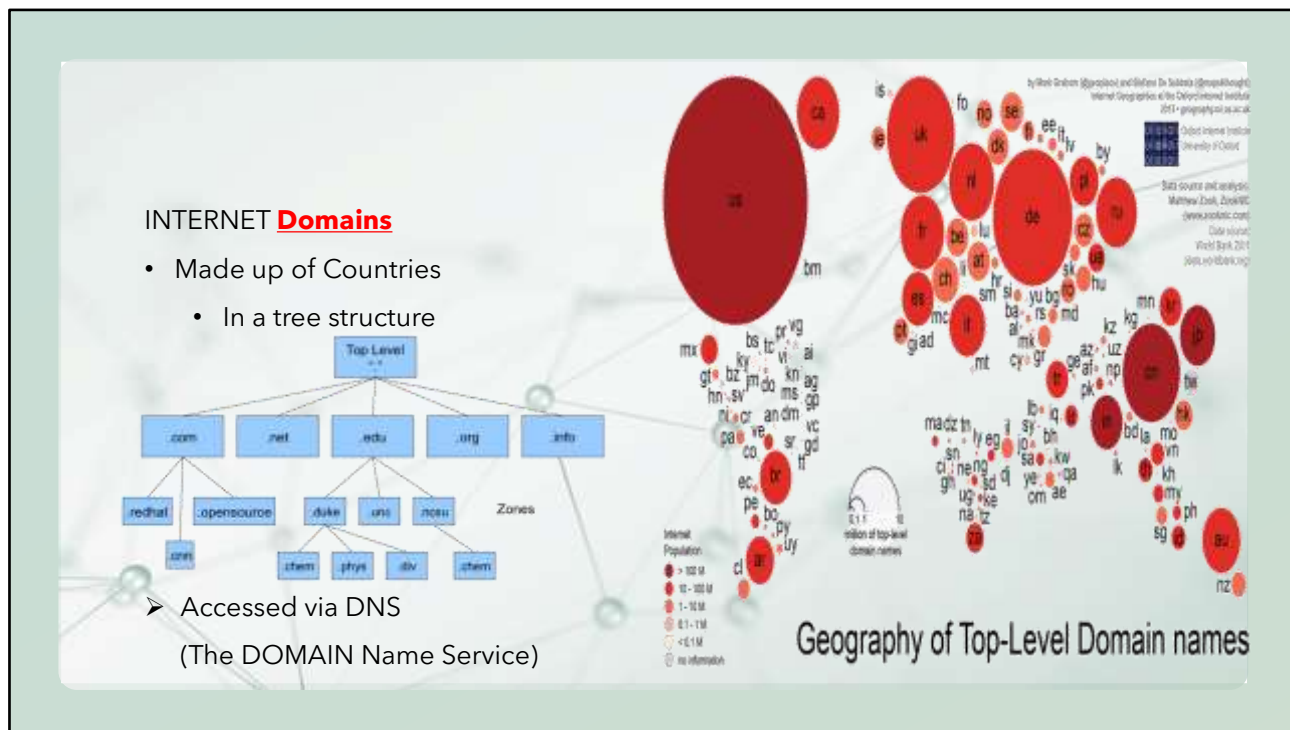
- Made up of Domains **~3.66 Million**
- Servers **~75 Million**
- Sites **~1.72 Billion**
- Pages **~55 Billion ++**

➤ All connected by Networks
(OK; Wires, optical cables
and Satellite links)



Note that by the time you read this these numbers will have grown significantly.

What is the internet

**Domains:**

I have, for the purpose of this explanation, put them first.

Domain Names provide you with the address of the website you want to visit.

Domain names, are read the wrong way around – right to left, tell you a lot about the web sites you visit.

For example:

- www.gov.uk UK's Government
- www.australia.gov.au Australian Government
- www.gouvernement.fr French Government
- www.usa.gov American Government, Note no county id

Other levels include:

.uk , **.au** , **.fr** etc. can give you the country of origin (but not the USA)

.com (or **.co**) = Commercial

.net = Network (provider)

.edu = education

.org = Non profit making organisation

... there are many others (now not only limited to 3 characters)

What is the internet



The **SERVER**

- Just a BIG PC (really)
Often designed to fit into a 19" rack
- They can run a whole range of programs:
 - Web Server
 - Database
 - etc ...

➤ Servers often live in Data Centres

The Server

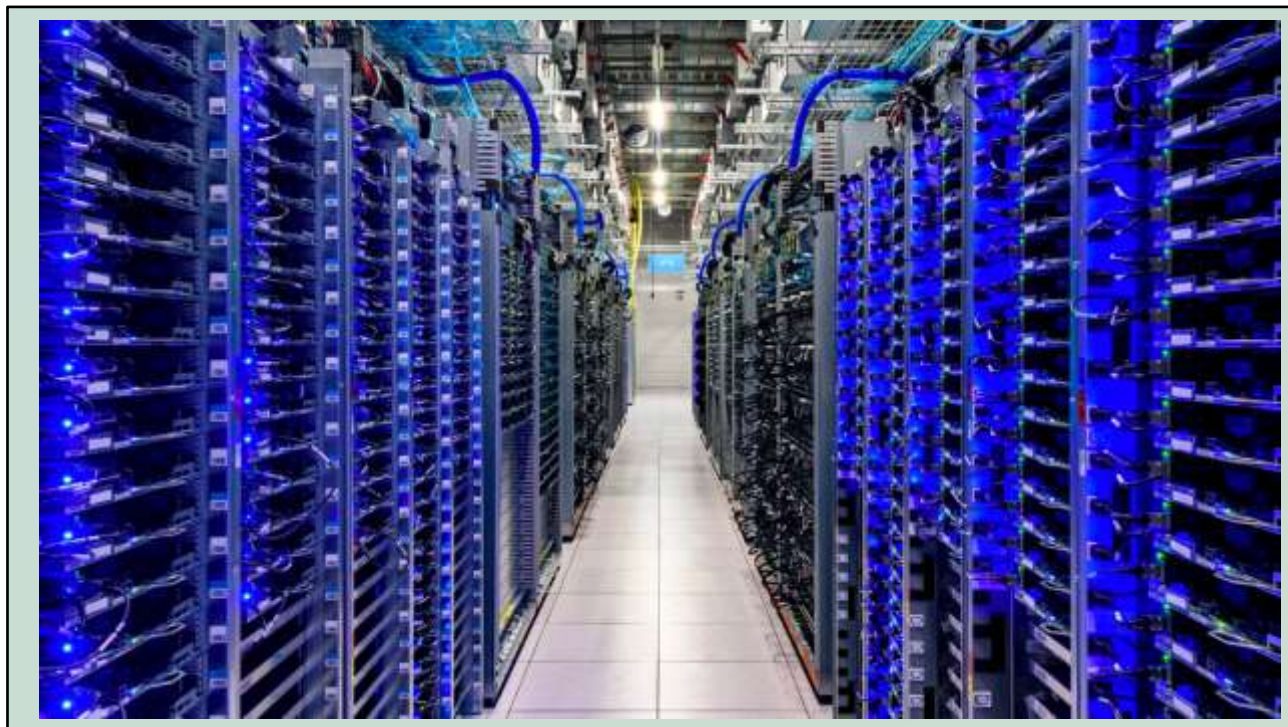
Just a big PC really, some (the Microsoft ones) even run Windows (which may have the same look and feel as your laptop) but underneath supports many simultaneous users/applications.

Other common servers run an operating system called Unix (usually a bespoke hardware manufacturers' variant), which has no (user friendly) graphical command level interface – just strange commands:

```
ps -ef|grep "ora_"|grep $ORACLE_SID|-v grep| \awk '{ print $2 }'| -exec rm ?f {} \;
```

This one kill's all Oracle database processes ☺

What is the internet

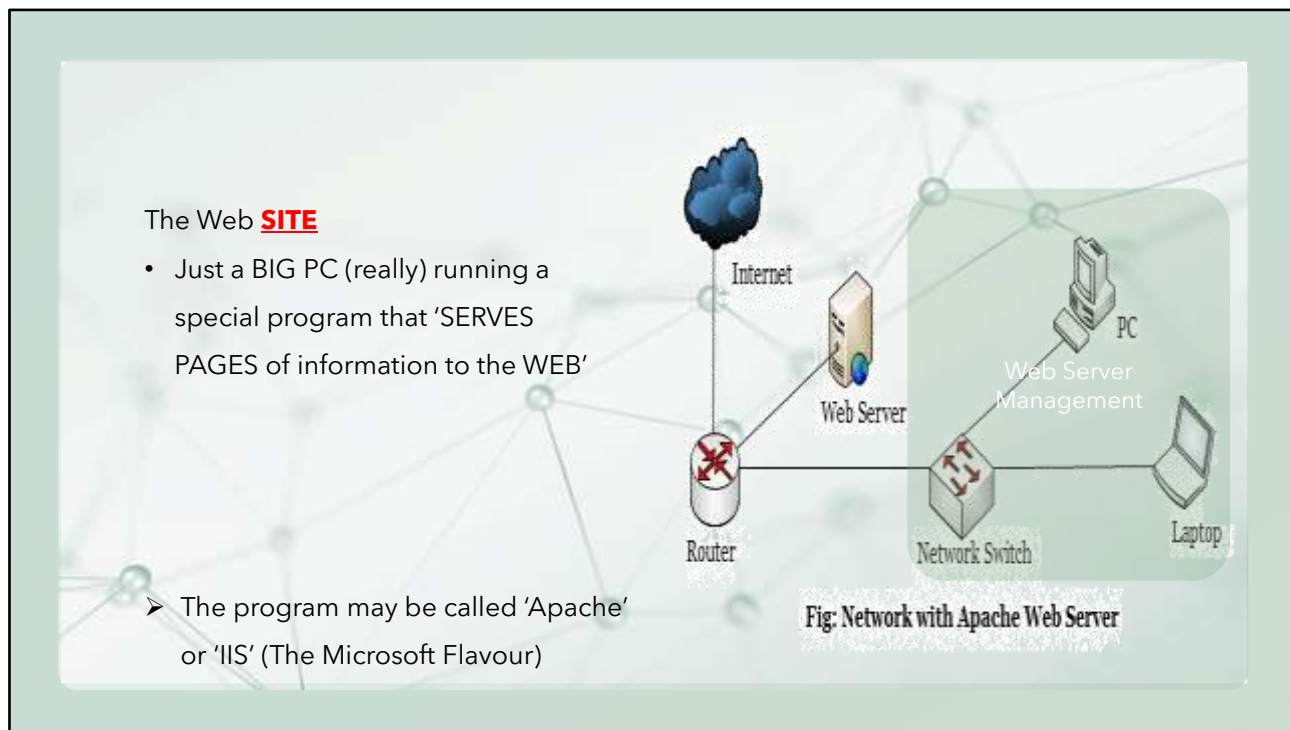


The only commonality for servers are their form factor = They almost always fit into a 19" rack. No keyboard, mouse or monitor, the system manager has one that either connects via the network of a KVM (Keyboard, Video, Mouse) switch. Hundreds and hundreds of servers live in racks in data warehouses, like this.

I spent many 'happy' hours at <math><18^{\circ}\text{C}</math> in this noisy environment, sitting on the floor with my faithful laptop.

BTW: The walls of many of these data centres have the letters A,B,C ... along one side and 1,2,3 ... on the other. So users like me had to search for server 957 K19 (Rows K and 19).

What is the internet



A Web Site

A computer can run any set of programs, but typically a web site is run on a dedicated server running a webserver program, like Microsoft IIS or the Unix preferred Apache. Other servers in the data centre assist the webserver by running databases etc.

Even our (your) laptops can do this, Laptop 6 has been setup as a WAMP server – running **Windows**, **Apache**, **MySQL** a database and **PHP** (or Python or Perl ...) a program language(s) that are communally used on web pages. All of these programs are free 😊.

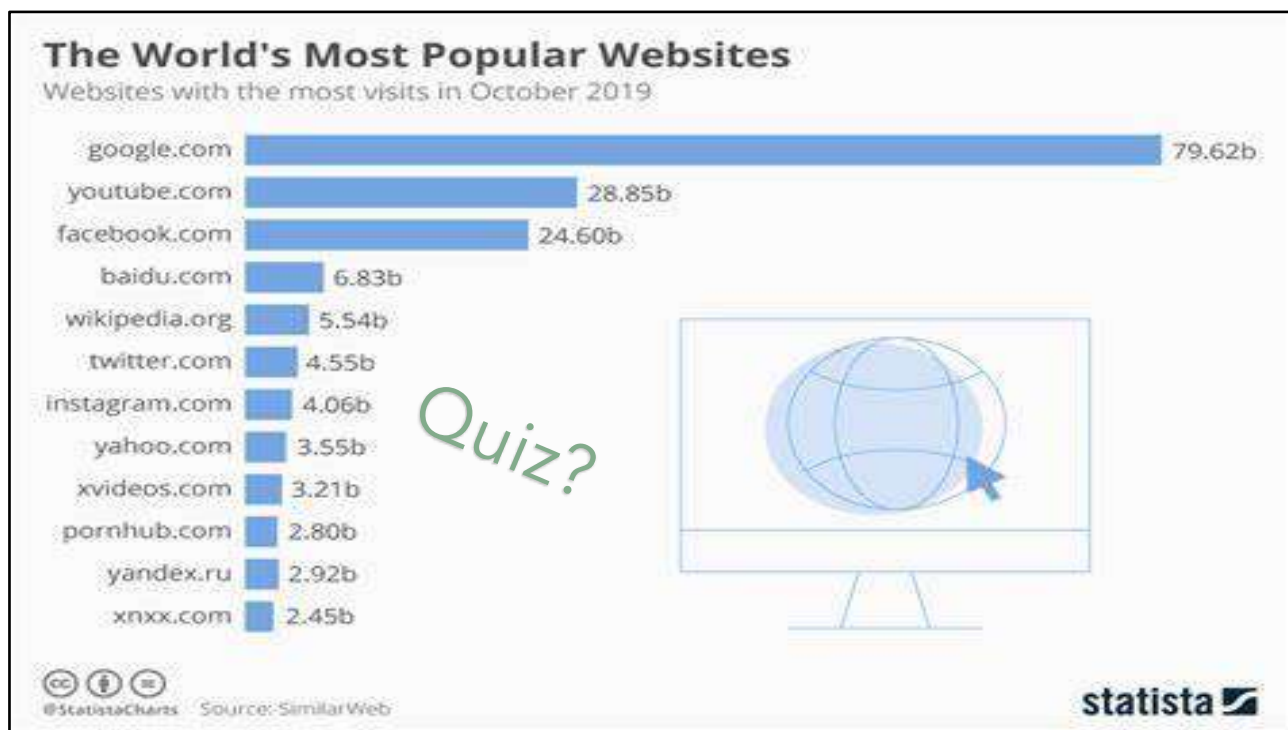
If you want to play let us know.

What is the internet



A list of today's top 50 websites worldwide ...

What is the internet



Can you guess who they all are... I was a bit shocked when I looked up the ones I didn't recognise.

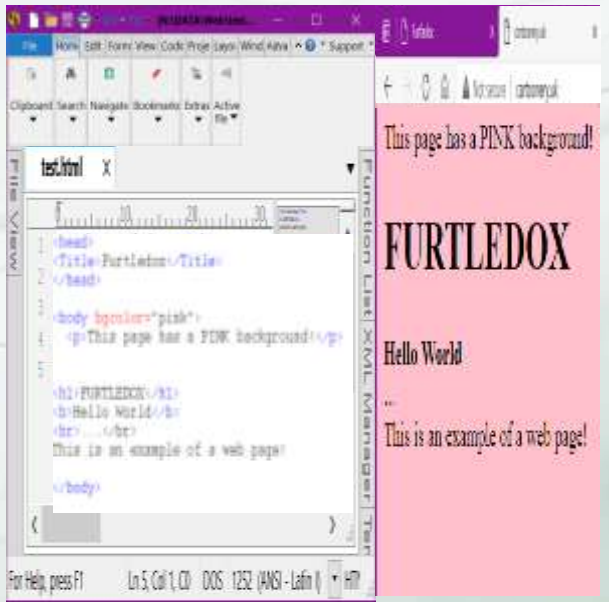
- Google, Youtube, Facebook, Wikipedia, Twitter, Yahoo...
- Baidu = Chinese equivalent of Google (search engine)
- Yandex = Russian search engine
- ...?
- The others are all PORN sites 😞

What is the internet

The Web **PAGE**

- Written in a MARK UP language
- HTML
- A Web Browser 'understands' this language and displays it in a human readable form

➤ The later versions of HTML tries hard to separate FORM from CONTENT.



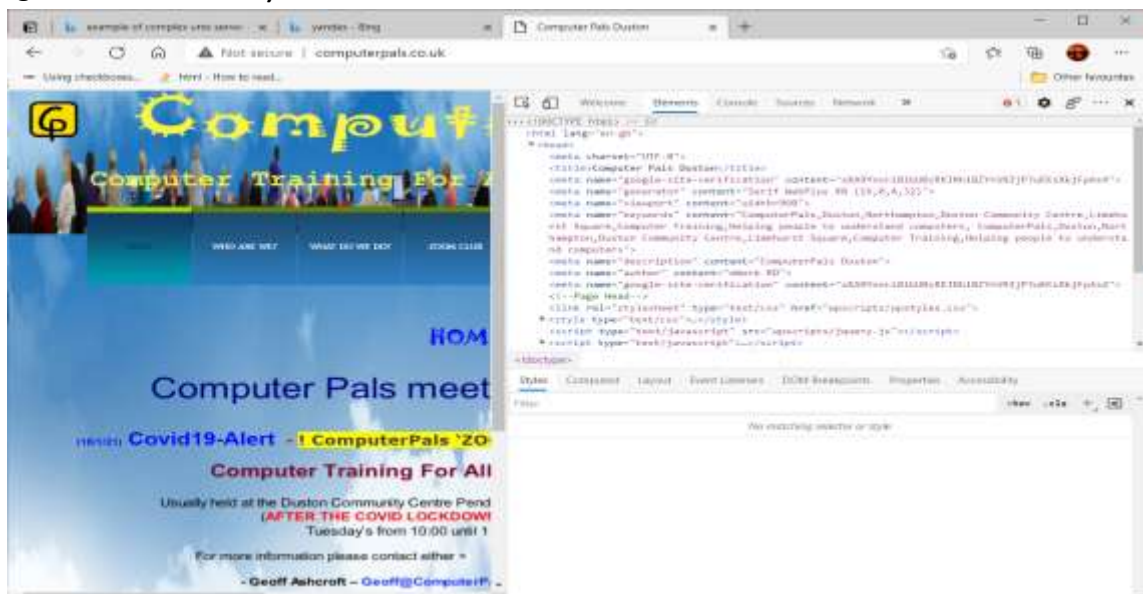
```

<head>
<title>Furledorx</title>
</head>
<body bgcolor="pink">
<p>This page has a PINK background!</p>
<h1>FURLEDORX</h1>
<p>Hello World</p>
<br />
<p>This is an example of a web page!</p>
</body>

```

Web Pages

Written in HTML (High Level Markup Language). Just press [F12] on your keyboard when looking at a web page to see... (and [F12] again to switch it off)




They often also include scripted program commands in 'Javascript, PHP, Perl, ...' See later for more details/ examples.

What is the internet

How do you find a Web **PAGE**

- Simple ... 😊
 - Providing you already know what it is!
- Just enter it's URL into the box on the top of your browser (see next slide)
- This is it's URL
 - Universal Resource Locator


Universal Resource locator: URL





Finding a single page, from the 55 billion out there ... ?


If you know its **URL** (Unique Resource Locator) its easy. For example this document can be found on ...


<http://computerpals.co.uk/furtledox.pdf>



 Hyper Text Transport Protocol


 ComputerPals


 Commercial


 UK


 Document Name


 Portable Doc Format

What is the internet

Universal Resource locator: URL

The URL - a way to find a web page

- It is made up from
 - http:// - to tell your browser how to interpret it
 - The Domain Name - to let your browser find it
 - The name of the page to display
- Your Browser is a program on your device: Chrome, Edge, Firefox, Safari ...

**URL**

NB: Some extra 'potentially useless' information:-

The first part, before the `://`: (known as the scheme) does not have to be `http`

- `http` – Hyper text Transport Protocol
- `https` – Secure Http
- `file` - <file:///host/path> A network file
- `ftp` – File transfer protocol (see later)
- `spotify` - `spotify:<artist|album|track>:<id>`
- `about://` - tells you about your browser

Many, many more even:

- `gopher` – Another pre web mechanism of trolling the web
 - `fish` – Access a file using ssh (heavy Unix stuff)
- `fish://[<user>[:<password>]@]<hostname>[:<port>]`

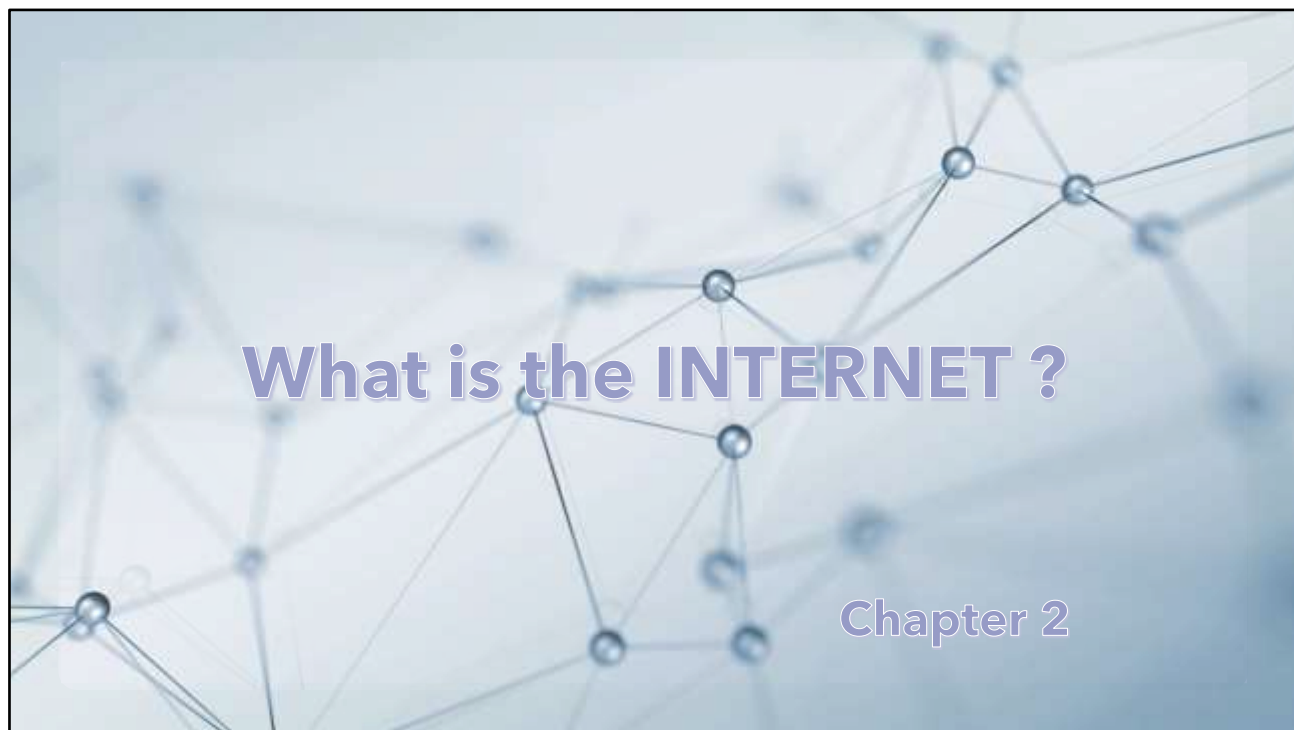
What is the internet

NEXT TIME:

- How do you create a web Page
- How does the web page get to the server
- What is the 'connection' between the Domain Name and the Web Server
- How does a Search Engine find a page
- How does your device connect to the Internet

➤ Only if YOU want, that is?...





What is the internet

This Time

- How do you create a Web Page?
- How do you buy a Domain?
- How does the web page get to the server
- What is the 'connection' between the Domain Name and the Web Server



What is the internet



How to create a Web Page

- It is in plain text, designed for the days of a green screen
- Written in a 'MARK-UP' language called HTML

➤ HTML - Hyper Text Mark-up language

Or even ...

HTML: The key to the Internet

The Hypertext Markup Language, the first version of HTML was written by Tim Berners-Lee in 1993, created on a Green Screen Terminal (ie: no graphics, as illustrated).

Since then, there have been several different iterations of HTML, we are now on HTML version 5.

The main benefit of HTML over all of the other 'Internet' communication mechanisms at the time was the "ANCHOR" (<a>) construct which allowed links written in HTML to connect to another page, potentially located anywhere on the Internet – creating a **web** of interconnected pages.

ie:

```
<a href="http://www.ComputerPals.co.uk">Visit  
Duston's Computer Pals Club</a>
```

What is the internet

How to create a Web Page

- It is in plain text, designed for the days of a green screen
- Written in a 'MARK-UP' language called HTML

➤ HTML - Hyper Text Mark-up language

You can make it **bold**.

`bold`

You can underline.

`<u>underline</u>`

You can *italicise* it.

`<i>italicise</i>`

You can subscript.

`_{subscript}`

You can superscript.

`^{superscript}`

You can strike-through.

`<strike>strike</strike>`

You can mono space.

`<tt>mono</tt>`

You can change the size, colour and style of your font. HTML5 can't read the old `` tag, now you need to use a little CSS:

Red Text

`Red Text`

Big Text

`Big Text`
Tip: You can swap out 'x-large' for pixel size like, say, '12px'.

Font Family

`Font Family`

Line breaks are good. Type `
` to start a new line, or use `<p>content</p>` around your paragraphs.

Example of (the original) HTML Mark Up Language

The MARK-UP part of the language allowed non-graphical terminals to create 'effects' like **bold**, underline or *italic* text or even coloured backgrounds, etc.

Today's version of HTML (v5) now separates "content" and "form", ie: The words used in the web page, and the way they are displayed. Version 5 uses style sheets to describe the way the content is displayed. Just to add more complexity these style sheets can themselves contain style sheets; hence they are named css – Cascading Style Sheets!

```

<style>
body {background-color: powderblue;}
h1  {color: blue;}
p   {color: red;}
</style>

```

Heading 1

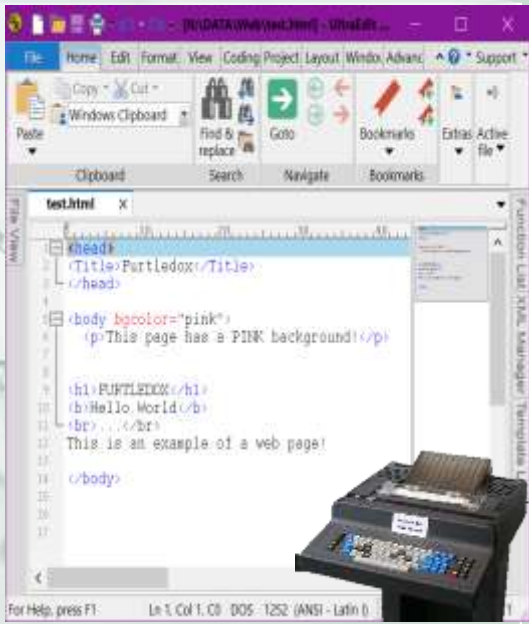
Paragraph

What is the internet

How to create a Web Page

- Originally it was typed on a 'Green Screen terminal, or teletype
- Hence the need for a 'Mark-up' language

➤ 1st Example: Plain Text editor



I typed this 'Web Page' using a TEXT Editor, like Microsoft's Notepad. NB: An improved version, free and safe to download is called Notepad++. There is a copy of this on Laptop6. Get it free from:

<https://notepad-plus-plus.org/downloads>

The downside of 'hand-crafting' all of your webpages using a simple text editor is that there is no prompting for the extensive HTML/Styles code and only one page at a time can be written. That means that creating a template to create a series of linked pages is difficult, time consuming but not impossible.


What is the internet

How to create a Web Page

- Although today it often is created by a program that allows you to type as you would in WORD and it generates HTML
- Often showing you the HTML results below

➤ 1st Example: Plain Text editor

➤ **2nd Example: Microsoft Expression Web**

The image shows a screenshot of the Microsoft Expression Web 4 software interface. The interface is split into several panes. The top pane shows a visual design of a web page with a purple header and a white body. Below this, there are panes for 'Carbons' and 'Addition-Remove'. The right side of the interface has a 'Tools' pane with various icons. The bottom pane shows the HTML code for the page, with tags like <h1> and <h2> visible. The overall background of the slide is a light green with a faint network diagram.

A number other 'free' HTML editors are available online, like Microsoft's Expression Web. Expression Web is a full-featured professional tool for designing, developing, and publishing compelling, feature-rich websites that conform to web standards.

Get it here:

[Microsoft Expression Web 4 Download | TechSpot](#)

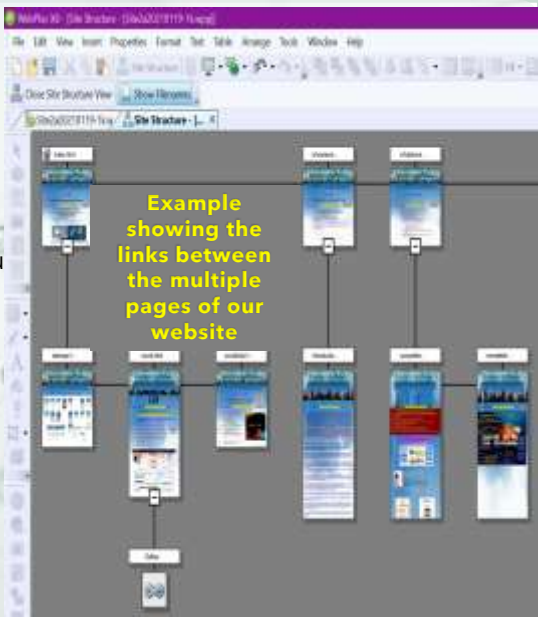
This second generation HTML editor has the advantage of being able to 'automatically' create several interlinked webpages, if not a complete Website

What is the internet

How to create a Web Page

- Although today it often is created by a program that allows you to type as you would in WORD and it generates HTML
- Or build with a full graphical interface

- 1st Example: Plain Text editor
- 2nd Example: Microsoft Expression Web
- 3rd Example: **Serif Web X8**



Example showing the links between the multiple pages of our website

Most 'modern' Web site editors, either provide a total graphical interface, in which the author designs the web site without ever having to use the underlying HTML code, or a series of pre-written templates. The program creates all of the HTML for you automatically.

Note that in addition to pure HTML/CSS Style code most sites also use a scripted program language to make the page more 'dynamic'. These languages include Java Script, Perl, Python and PHP. This simple example prints 'Hello World' using PHP. They however need to be hand crafted, ie:

```

1  <!DOCTYPE html>
2  <html>
3  <head>
4  <title>Hello, World! Page</title>
5  </head>
6  <body>
7  <?php
8  echo 'Hello, World!';
9  ?>
10 <h2>test</h2>
11 </body>
12 </html>

```

What is the internet

How to create a Web Page

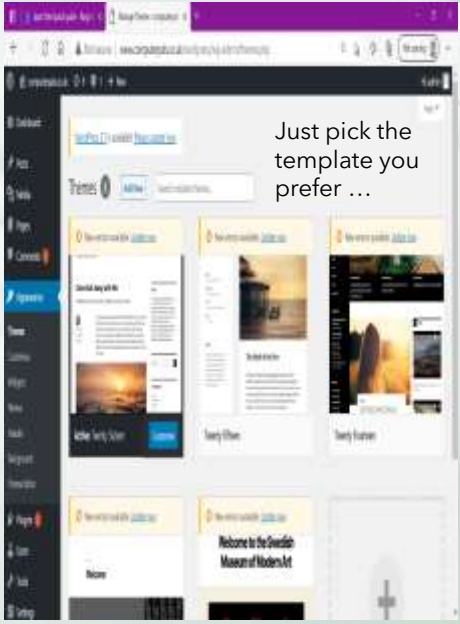
- Although today it often is created by a program that allows you to type as you would in WORD and it generates HTML
- Or build with a full graphical interface
- And create using pre-written templates

➤ 1st Example: Plain Text editor

➤ 2nd Example: Microsoft Expression Web

➤ 3rd Example: Serif Web X8

➤ **4th Example: WordPress**



Many Web Hosting companies sell a complete package:-

- The Domain Name
- The Web Server
- A web site authoring tool

The most common authoring tools use templates, that only require you to fill in your details, colour scheme etc.

Wordpress is an example of this genre.

What is the internet

How does the web page get to the server

- Remember this slide...
- The 'author' of the web page creates it on their laptop (using one of those programs)
- When finished they simply copy it to the web server using **FTP**
- Just another program

➤ FTP - File Transfer Program

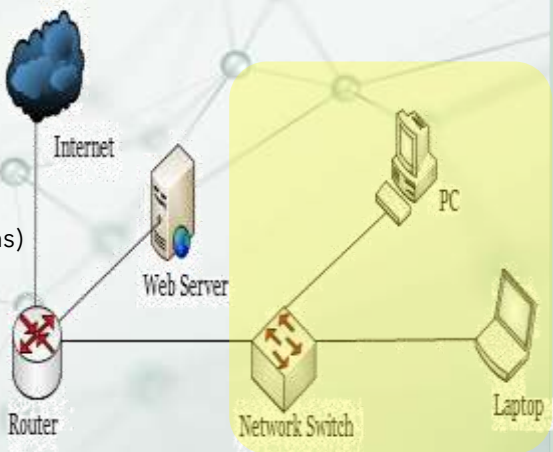


Fig: Network with Apache Web Server

Assuming that you do not work for the company that hosts your web site, and even if you do. It is unlikely that you create your web pages on the same machine that serves them to the web. Thus you need a mechanism to copy the webpages to the server, the most common of these is called FTP. File Transfer Protocol. Again, originally based on command line technology:

```
C:\>ftp ftp.mywebsite.net
Connected to ftp.mywebsite.com.
220 Mywebsite FTP Service
User (ftp.mywebsite.com:(none)): anonymous
331 Anonymous access allowed,
send identity (e-mail name) as password.
Password:
230-Welcome to ftp.mywebsite.com.
230 User logged in.
ftp> put Index.html
200 PORT command successful.
125 Data connection already open;
Transfer starting.
226 Transfer complete.
ftp: 55056 bytes received in 2.55Seconds
21.57Kbytes/sec.
```

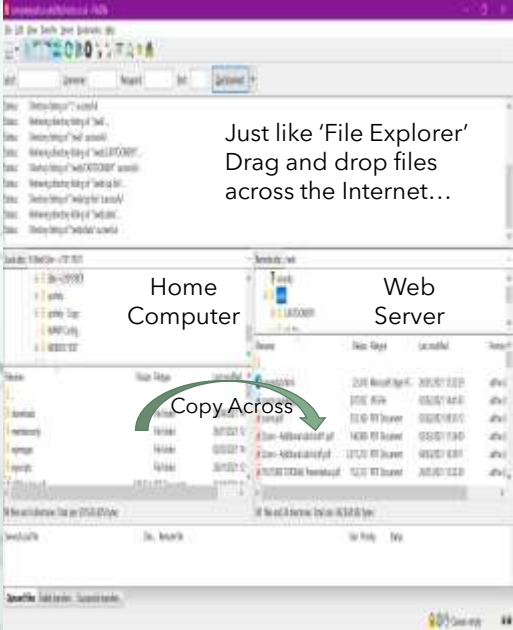
What is the internet

How does the web page get to the server

- Remember this slide...
- The 'author' of the web page creates it on their laptop (using one of those programs)
- When finished they simply copy it to the web server using **FTP**
- Just another program

➤ FTP - File Transfer Program

Just like 'File Explorer'
Drag and drop files
across the Internet...



The screenshot shows the FileZilla Client interface. On the left, the 'Home Computer' local drive is visible. On the right, a remote 'Web Server' is connected. A green arrow labeled 'Copy Across' points from a file in the local drive to the remote server. The interface includes a top toolbar, a file list, and a detailed view of the selected file.

In the 'real (Windows) world' graphical versions of FTP are used.

Filezilla, illustrated above, is a free program that I use to copy the web pages I create on my laptop to the Webserver owned by the hosting company.

Obtainable from here:


[Download FileZilla Client for Windows \(64bit x86\) \(filezilla-project.org\)](http://filezilla-project.org)

What is the internet

What is the 'connection' between the Domain Name and the Web Server

- Step 1: Buy your DOMAIN NAME

➤ ie: LCN.COM



So ... How to create your own website:

1. Buy your Domain Name

Many companies will sell you one of these. See above example.

2. Create your web site, as previously discussed

3. Find someone to 'host' your website.... ie: provide the web server.

You 'could' actually do this at home yourself, but with a couple of proviso's:

- You need a Webserver program, like Apache
- Your (laptop) server needs to be on 24*7*365
- You need to (pay for) a fixed IP address from your ISP

4. Copy your created web page to the web server using FTP

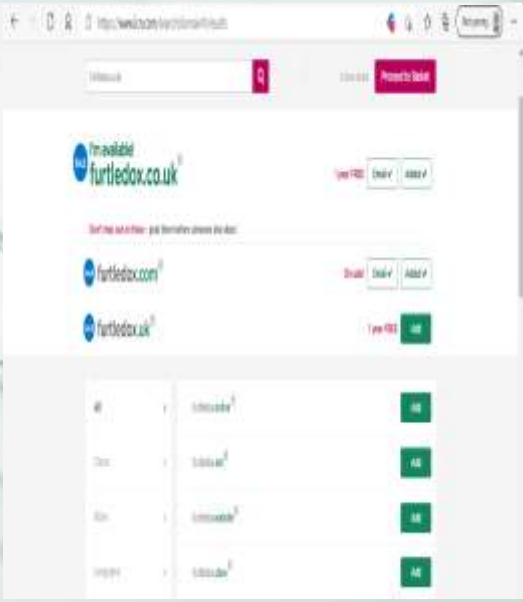
5. Connect the purchased Domain name to your web page

What is the internet

What is the 'connection' between the Domain Name and the Web Server

- Step 1: Buy your DOMAIN NAME (sometimes they are even FREE 😊)

ie: LCN.COM



Domain names can be quite cheap if not a .com

They also provide you the ability to receive emails against the domain you have purchased:

- Tesco@Furtlebox.com
- Bananas@Furtlebox.com
- Any word you want@your purchased domain.xxx

Now that you know how easy, and cheap is it to buy a Domain name, you can see how 'simple' it is for a hacker to mimic your bank etc. ie:

- MoneyBack@Barclays.Fraud-refunds-department.com

 fraud-refunds-department.com

On sale! £35.95
1 year for £1.95

Email ✓

Added ✓

- Now create and host your 'Bank' website ...
- Leave the country and avoid the police

What is the internet

What is the 'connection' between the Domain Name and the Web Server

- Step 1: Buy your DOMAIN NAME
(sometimes they are even FREE 😊)
- Purchase HOSTING
(Much more expensive)

➤ ie: From say £1.00 per month

The screenshot shows a comparison of three web hosting providers:

Rank	Provider	Rating	Price
1	Bluehost	Outstanding 9.5	From £11/mo
2	HostPapa	Excellent 9.1	From £7.0/mo
3	GoDaddy	Good 8.7	From £10/mo

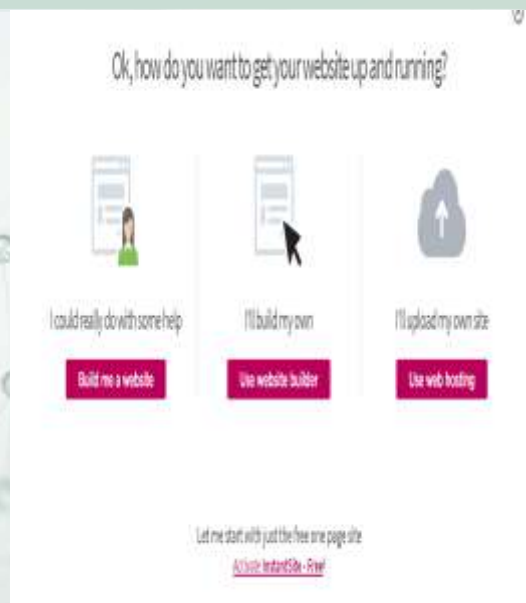
The CONNECTION part is to link the web pages you have created to the web server hosting them.

Most seller's of Domains provide this service for you.

What is the internet

What is the 'connection' between the Domain Name and the Web Server

- Step 1: Buy your DOMAIN NAME (sometimes they are even FREE 😊)
- Purchase HOSTING
 - They may even supply the tools
- ie: They do it, or use their tools, or DIY




As previously discussed, the Hosting / Domain name company often provides the tools for you to create your webpages.

What is the internet

What is the 'connection' between the Domain Name and the Web Server

- Step 1: Buy your DOMAIN NAME (sometimes they are even FREE 😊)
- Purchase HOSTING
 - They may even supply the tools
- Then connect the two -

DOMAIN to SERVER SERVICE Web page



Select the type of forwarding you'd like to change

URL to forward to
<http://www.FURLEDOX.com/index.html>

Forwarding method

Keep domain name in address bar

"Refresh" to forwarding URL

Display banner

Yes No

Web Forwarding Email Forwarding

[Edit](#) [Edit](#)

Now make the connection...

Visit your Domain Name site and connect your purchased to the Webserver's URL. Remember those? – They identify the HTML page on the webserver that should be 'identified' when the client enters that on their browser.

The lookup of the URL to your webserver is performed using a technology called **DNS**

DNS: Domain Name Service. This adds yet another layer of complexity to our story. Ignore this now, I will cover it later.

What is the internet

USEFUL TIP

Domain Name LOOKUP !

- If you receive some iffy looking email/txt copy the domain name, the bit after @
- Visit www.WHOIS.COM
- Enter the domain name and see who it is registered to...
 - If very recent then ... ☹
 - Or in an odd country ... ☹
 - Etc ...

To FIND OUT WHO OWNS a Domain, use a LOOKUP Service

To see who owns a Domain, a lookup can be performed using sites like:

- www.whois.com

Remember our 'fake Bank site' ...

We looked up one that was used in a scam and found two useful facts

1. The date it was registered was less than 24 hours prior to the scam being sent
2. The 'owners' address (in this case a Bank) was very odd... see next page

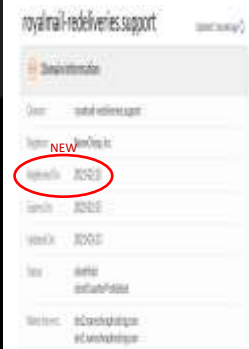
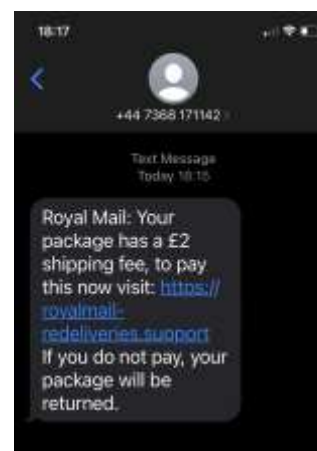
What is the internet



Two 'example' actual messages received with a 'spooF' bank domains: admin-verify-secure.com & online-remove-newpayee.com
Both when looked up, have the same registered address:-
The Icelandic Phallogological Museum in Reykjavik.



Text Message Today 17:42
NATIONWIDE ALERT: Request for NEW payee B FARROW has been made on your account. If this was NOT done by you, visit: nationwide.online-remove-newpayee.com



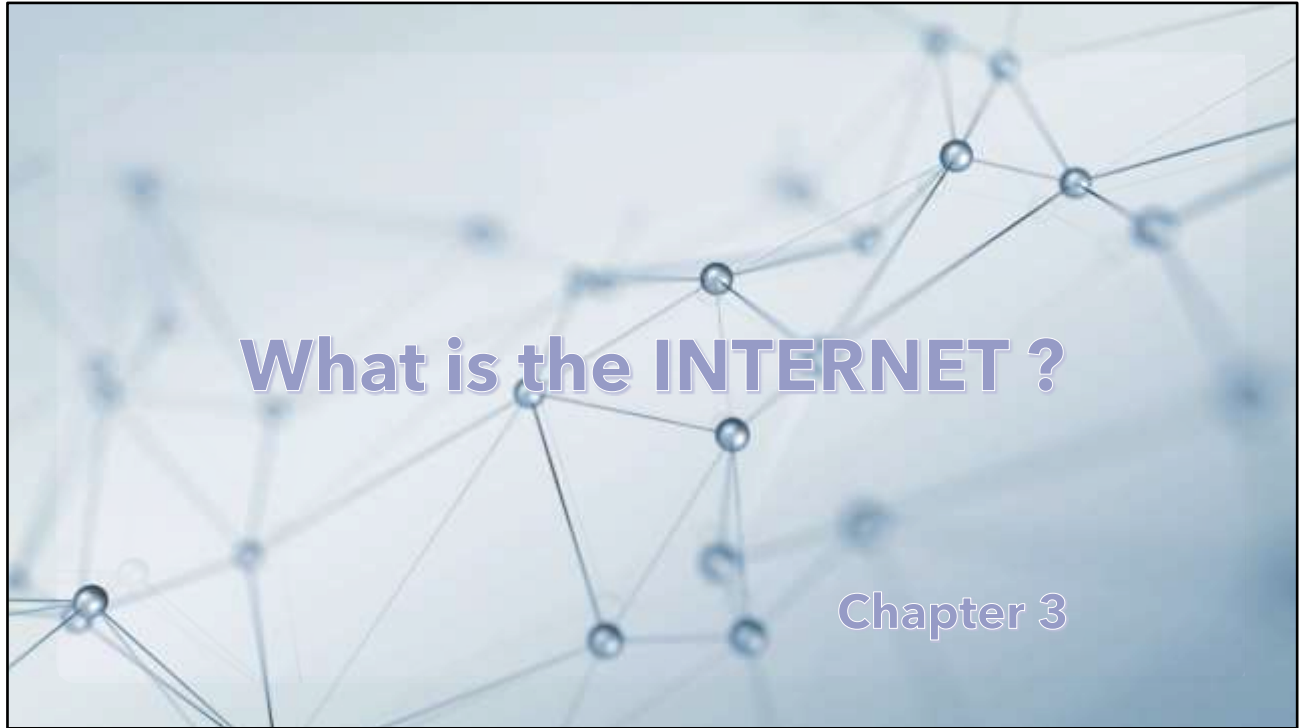
MORAL: If you get an odd message purporting to be from a Bank etc. use this technique to see if they are genuine.

What is the internet

NEXT TIME:

- How does a Search Engine find a page
 - How does your device connect to the Internet
- Only if YOU want, that is?...





What is the internet

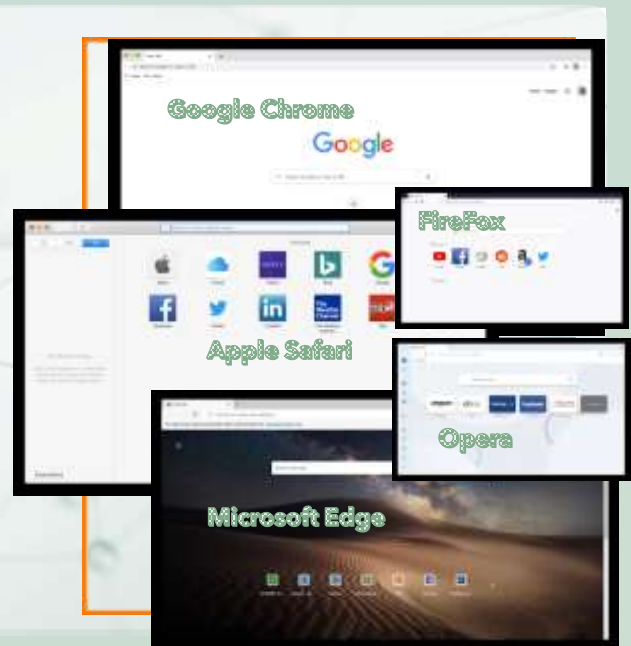
This Time

- What is a Browser?



Your BROWSER

- In order for you to 'see' the internet you have to use an Internet Browser - just another program.
- This 'program' understands HTML and is able to display the full glory of the marked up text
- It also understand how to view a huge range of other types of data



How Internet Browsers Work:

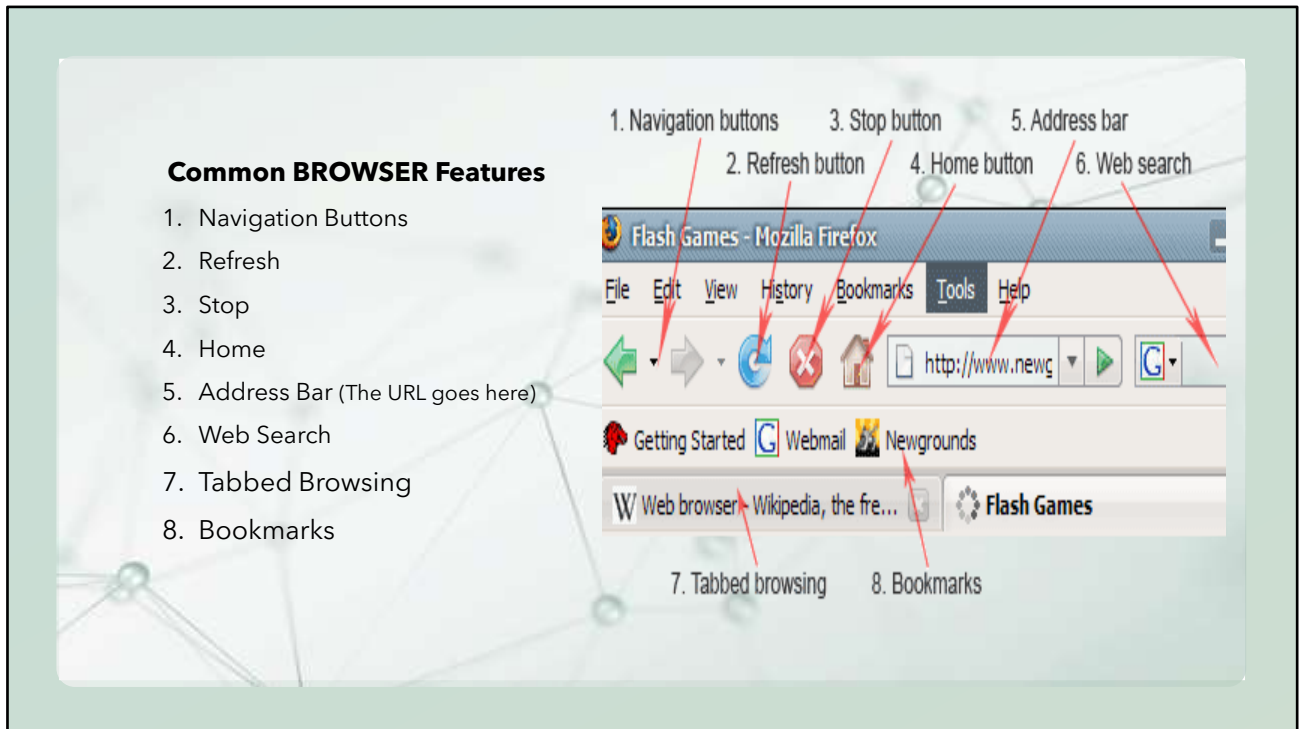
Here's a very quick overview of how browsers work:

1. You type a website's URL into your browser's address bar; "http://www.computerpals.co.uk" is an example of a URL.
2. The browser locates and requests that page's information from a web server.
3. The browser receives a file in a computer code like HTML or Javascript, which includes instructions about how to display the information on that page.
4. The browser interprets that file and displays the page for you to read and interact with. And it does all of this in just a few seconds, usually.

If you want a more detailed technical breakdown of how browsers work, check out [Behind the Scenes of Modern Web](#)

[Browsers](#) by Tali Garsiel and Paul Irish on HTML5Rocks.com.

The important thing to know is that different browsers have slightly different ways of working so that if you are having trouble getting pages to display or sweepstakes to load on one browser, it makes sense to try another. Install two or three of your favourite browsers and switch between them if something doesn't work.



Browser user interfaces have a lot in common with each other.

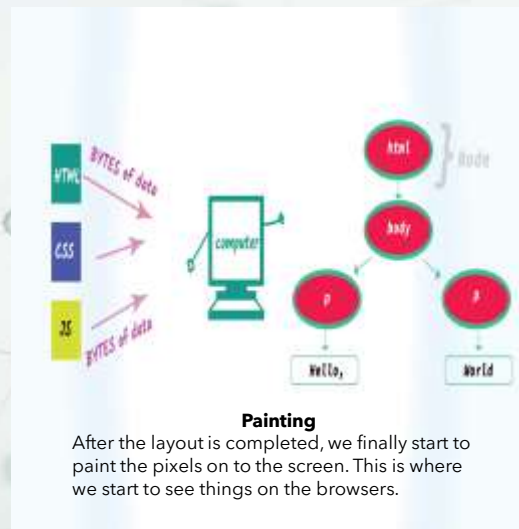
The user interface:

- Address bar for inserting a URI
- Back and forward buttons
- Bookmarking options
- Refresh and stop buttons for refreshing or stopping the loading of current documents
- Home button that takes you to your home page

Every part of the browser display except the window where you see the requested page.

Your BROWSER

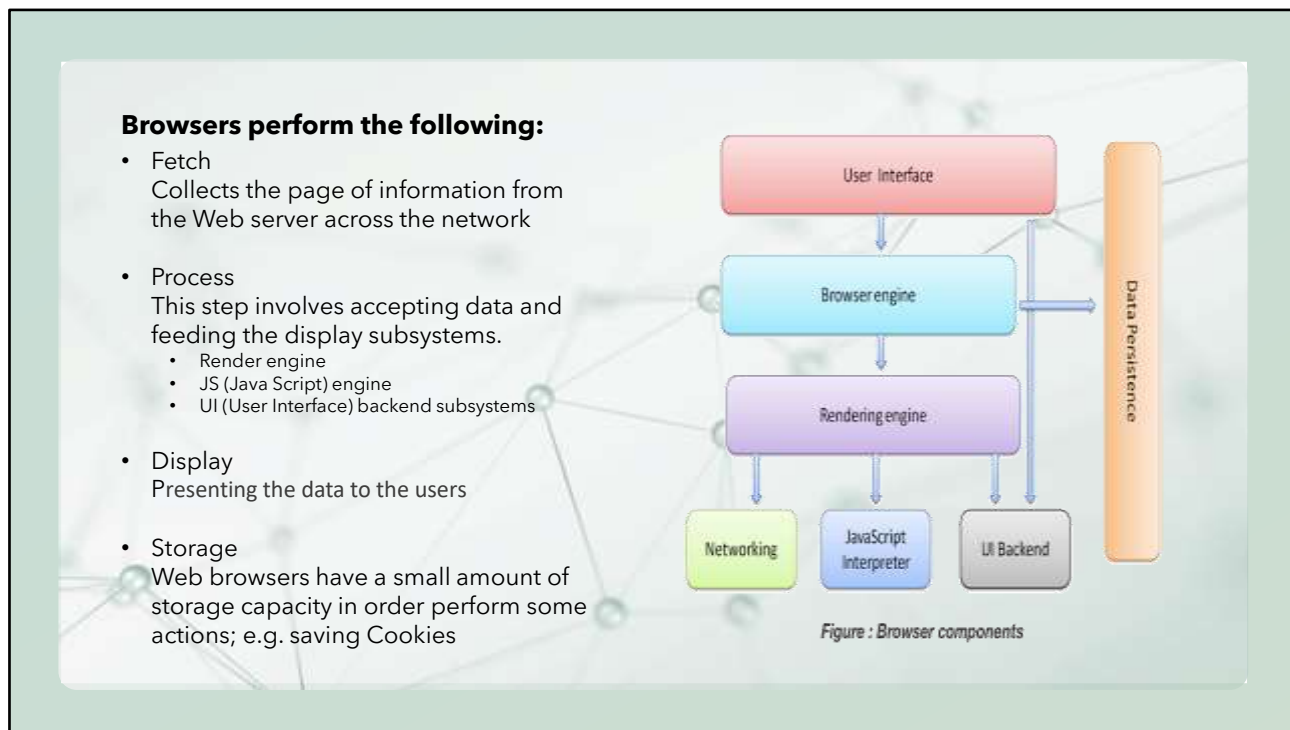
- It understands HTML
- It is able to display pictures:
 - jpg, jpeg, png, gif, tif, svg, bmp, psd, tiff, ai, lsm
- Video:
 - 3gp, avi, flv, gvi, m2v, m4v, mkv, mov, mp4, mpg, ogv, webm, wmv, mpeg, liveslide.zip
- Audio:
 - mp3, wav, aif, aiff, wma
- And much much more...
 - This is the latest battleground where the technology companies are fighting for your attention.



This is a list of ‘some’ of the file types interpreted by a typical browser (note I have vastly truncated the Text file type).

viewer type	thumbnail type	controls	extensions
archive	svg glyph		tgz, zip, rar, iso, tar, bz2, gz, xz, bzip, 7z
audio	svg glyph		mp3, wav, aif, aiff, wma
dataset	svg glyph		csv, sav, fsv, xls, xlsx, xlsx
document	generated	zoom, pagination	doc, docx, dxf, odp, ods, odt, pages, pdf, rtf, ttf, xps, epub
image	generated	zoom	jpg, jpeg, png, gif, ttf, svg, bmp, psd, tiff, ai, lsm
jupyter	svg glyph		ipynb
kml	svg glyph		kml
molecule	svg glyph		cif, pdb, xyz
presentation	generated	pagination	ppt, pptx, pptm
video	generated		3gp, avi, flv, gvi, m2v, m4v, mkv, mov, mp4, mpg, ogv, webm, wmv, mpeg, liveslide.zip
viewer3d	svg glyph		obj, stl, ply, u3d
txt	svg glyph	font zoom	1c, accesslog, armail, arm, avram, actionscript, as, apache, apacheconf, applescript, asascript, ascript, adoc, aspectj, autohotkey, autor, awapta, bash, sh, zsh, basic, brainfuck, bf, c#, csharp, cpp, c, oc, h, c++, h++, hpp,
graph	svg glyph		gephi, gexf
fits	svg glyph	zoom, pagination, histogram	fits
dicom	svg glyph	zoom, pagination, histogram	dicom, dcm

What is the internet



The browser engine: marshals actions between the UI and the rendering engine.

The rendering engine : responsible for displaying requested content. For example if the requested content is HTML, the rendering engine parses HTML and CSS, and displays the parsed content on the screen.

Networking: for network calls such as HTTP requests, using different implementations for different platform behind a platform-independent interface.

UI backend: used for drawing basic widgets like combo boxes and windows. This backend exposes a generic interface that is not platform specific. Underneath it uses operating system user interface methods.

JavaScript interpreter. Used to parse and execute JavaScript code.

Data storage. This is a persistence layer. The browser may need to save all sorts of data locally, such as cookies. Browsers also support storage mechanisms such as localStorage, IndexedDB, WebSQL and FileSystem.

What is the internet

This Time


- How does a Search Engine find a page




What is the internet

Search Engines


- Web based servers that assist you in finding the stuff you want on the internet.
- There are loads, but Google is the most popular at ~ 94% of all searches
- Some are different:
 - DuckDuckGo - no adverts and private
 - Dogpile / Metacrawler both use other search engines to do their job, and rank them in order



Click on these (six) icons to visit ☺



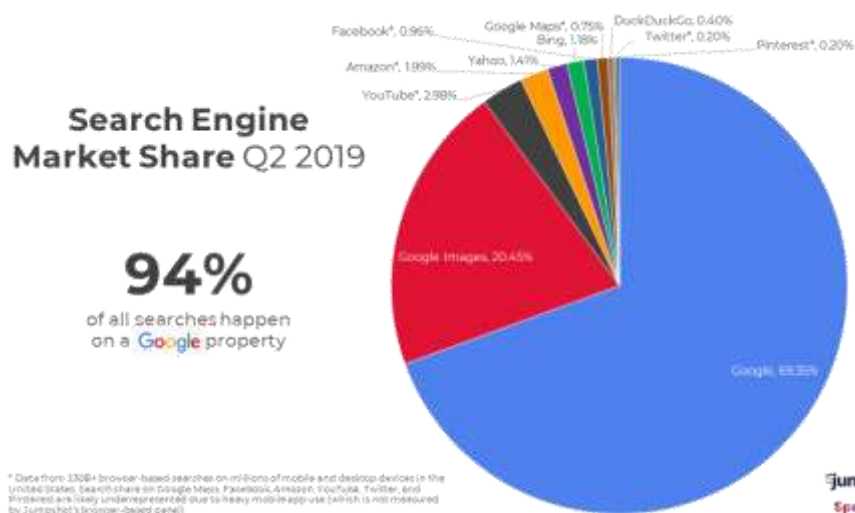
Some of the others ...



There are many search engines out there, each looking at the Internet in a different manner.

As a useful tip, if you don't get the result you want from one search engine, try another. For example 'FURTLEDOX' my created key word took 10 days before Bing recognised it, Google still (at time of writing) give no results.

Using this theory, technical questions about Microsoft products are better researched on Bing (Microsoft), than Google.



What is the internet

Search Engines

- Just another server ... but big!



This illustrates one of the many Google Search Engine server farms. The servers themselves are HUGE, and specially built to run Google's search engine programs.

What is the internet

Search Engines

- Step 1:
 - The Crawler
- These continually trawl the Internet for web sites
- When they find one, they look for a file called 'robots.txt' to see if they are allowed
- Then they look for the web page's 'META TAG' for Key words

Example of Meta description tag code:
`<meta name="description" content="This is a meta description. This text will show up in Google's search engine results page.">`

How does the Search Engine actually find the pages of Internet information we are looking for?

ANSWER: They use CRAWLERS.

A crawler, just another server/program effectively starts with Internet Computer number 1.1.1.1 and then 'crawls' its way through 1.1.1.2 ... 223.255.255.254.

When it 'finds' a computer on the web with a valid address, it looks to see if it's owner wants its contents to be crawled. There is a file on most web servers called 'robots.txt' that either grant or refuse a search engines requests.

Once it finds a willing web server it will then crawl its way through the available web pages, primarily looking for some non-printable HTML code called <meta tags> - These hold "key words" that the pages creator thinks the searching public may use to find their page.

What is the internet

Search Engines

- Step 1:
 - The Crawler
- This is our site:

www.computerpals.co.uk
- Look at the visits from robot crawlers
 - Notice the we can see who visits us!

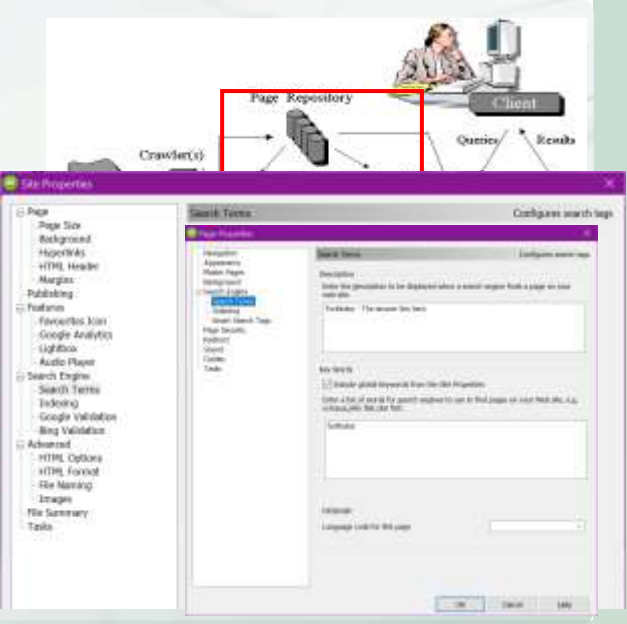
Just to prove these 'Crawlers' actually exist, this log shows the visits made to the www.ComputerPals.co.uk website.

Along with you people, you had to visit our site to get this page to be able to read it – so we know that you were here, by your IP Address (more about that later). We can see a number of Crawlers, most notably the Chinese versions - The PETALBOT ☹️ ?

What is the internet

Search Engines

- Step 2:
 - Copy the page
 - Index the 'Meta Tags'
- Save the information - for quick access
- They then rank the entries
 - Did you tell Google about it?
 - Did the owner pay Google?
 - How many other sites reference it?
 - Are your 'tags' informative?



The diagram shows a 'Crawler(s)' sending data to a 'Page Repository', which is then accessed by a 'Client' to produce 'Queries' and 'Results'. Below this, a screenshot of the 'Site Properties' window in a web browser is shown, with the 'Search Terms' tab selected. The 'Page Properties' section is expanded, showing various meta tags and search engine options.

This shows where I put the <meta tags> for ComputerPals, and this shows the 'hidden' tags on the top of our web page, using [F12] on my browser. See earlier. These are stored and indexed by the search engine to rank their appearance. The more money you pay = the higher ranking you get.

```

<!DOCTYPE html>
<html lang="en-gb">
  <head>
    <meta charset="UTF-8" == $0
    <title>Computer Pals</title>
    <meta name="google-site-verification" content="xXA9Ynnri8lkU0cRKJNh18ZYrU9IjPJuEKsXkjFp454">
    <meta name="generator" content="Serif WebPlus X8 (16,0,4,32)">
    <meta name="viewport" content="width=960">
    <meta name="keywords" content="ComputerPals,Duston,Northampton,Duston Community Centre,Limehurst Square,Computer Training,Helping people to understand computers, ComputerPals,Duston,Northampton,Duston Community Centre,Limehurst Square,Computer Training,Helping people to understand computers">
    <meta name="description" content="ComputerPals Duston">
    <meta name="author" content="eWork RD">
    <meta name="google-site-verification" content="xXA9Ynnri8lkU0cRKJNh18ZYrU9IjPJuEKsXkjFp454">
    <!--Page Head-->
    <link rel="stylesheet" type="text/css" href="wpscripts/wpstyles.css">
    <style type="text/css">...</style>
    <script type="text/javascript" src="wpscripts/jquery.js"></script>
    <script type="text/javascript">...</script>
  </head>
  <body style="height:1000px;background:url('wpimages/wp0a9e5cc5_06.jpg') repeat scroll center top / 100% 100% #ffffff;">
    <div id="divMain" style="background:transparent;margin-left:auto;margin-right:auto;position:relative;width:960px;height:1000px;">...</div>
    <script type="text/javascript" src="wpscripts/jsMenu.js"></script>
    <script type="text/javascript">...</script>
  </body>

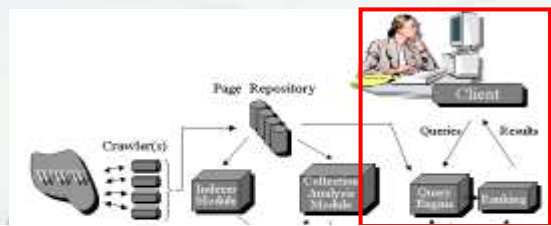
```

What is the internet

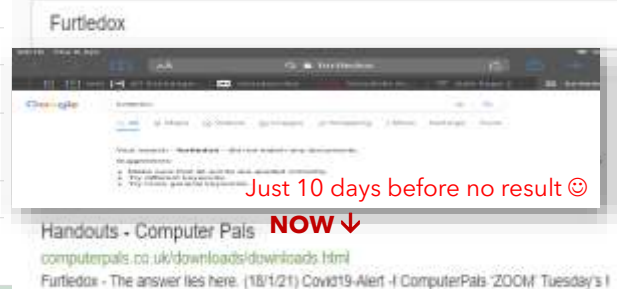
Search Engines

- Step 3:
- The Client Search
 - Did you know you can use 'Special' search terms
- Use a 'META Crawler'

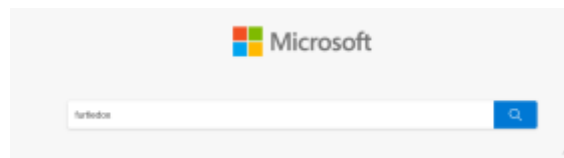
Unlike search engines, metacrawlers don't crawl the web themselves to build listings. Instead, they allow searches to be sent to several search engines all at once. Unlike search engines, metacrawlers don't crawl the web themselves to build listings.
- Try 'METACRAWLER.COM'



metacrawler



The user then issues their search phrase from their browser's search engine: ie BING



Note that search engines 'understand' human phrases like:

- "tell me how to"
- "how do I"

But also can be made to understand other special terms:

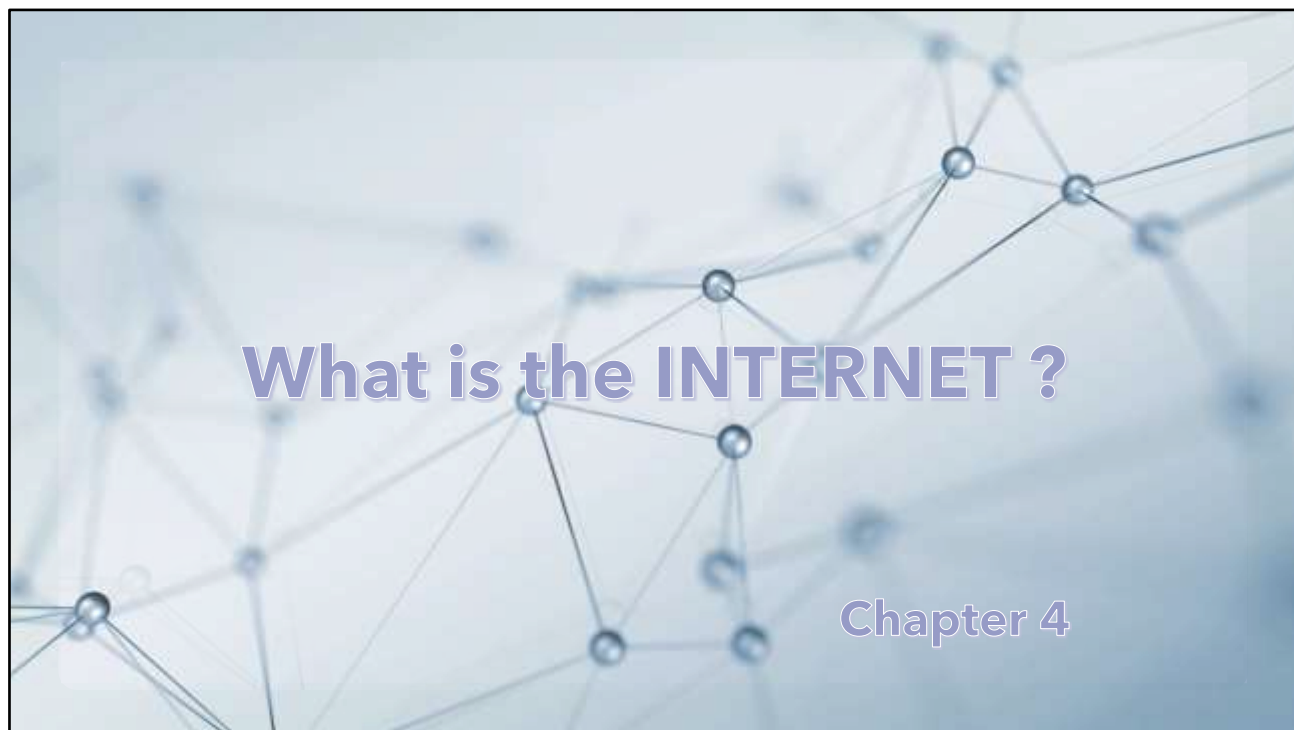
Command	Example	Explanation
This and That	new iPhone deals	Search all words: <i>new</i> , <i>iPhone</i> , and <i>deals</i> ; similar to using AND between the terms
This or That	sailing OR boating	Search <i>sailing</i> or <i>boating</i>
<u>Exact Match</u>	"love me tender"	Search this phrase as a whole
Exclude Words	printer -cartridge	Search <i>printer</i> but hide any results that include <i>cartridge</i>
Definitions	define:serendipity	Definitions for <i>serendipity</i>
<u>Partial Search</u>	san * california	Search all words but allow for another word between them

What is the internet

NEXT TIME:

- How does your device connect to the Internet
- Only if YOU want, that is?...





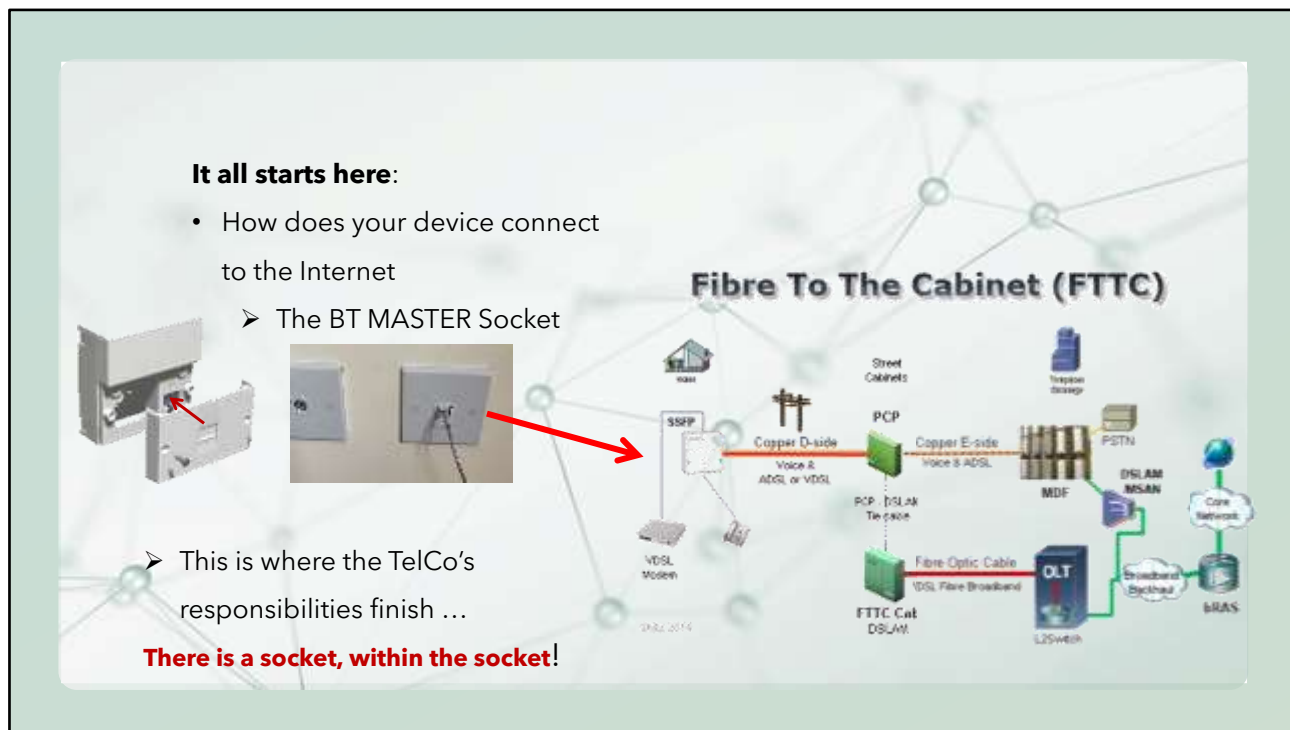
What is the internet

This Time

- How does your device connect to the Internet



What is the internet



The route to the Internet, from your home:

1. The BT (or other companies) Master Socket

Note: It is from that point the SPEED you purchase for the Internet is measured. This is to disregard other devices / WiFi, extensions etc the user may have in their house.

To check a BT line for errors, plug a wired phone into the socket within the socket (unscrew the two screws on the front) and dial 17070. Pick Option two, from the recorded message and listen – This is a QUIET LINE TEST, so if you hear frying eggs, *THERE IS NOISE ON THE LINE (a serious fault, that will effect your internet use)* and if you have used this internal socket the telephone call out will not charge you, as this is their responsibility (if you test on the outside socket, it *could* be your internal wiring!).

2. The wire (or fibre, see next slide) from your house then goes to the green box in the street (or Grey if it is Vodaphone etc)

What is the internet

A quick aside The Fibre Network:

- Under our pavements / roads
 - Loads of ... TUBES
 - Full of wires/optical fibres
- Tubes/wires are Pulled
- Optical Fibres are blown
- To the Green/Grey Cabinet



The collage contains several images: a worker in orange overalls pulling a cable on a street; a close-up of a fibre optic cable being pulled through a tube; a BT cabinet with a 'Cable pulling!' sign; and a diagram of 'Mid-Point Cable Blowing with Floater'.

FIBRE, rather than copper to the house, known as FTTP (Fibre to the Premises) travels through plastic tubes to a patch panel and a concentrator.

If you look closely you can see the light blue rope in the pipes that is used to pull new cable/fibre bundles from manhole to manhole to cabinet.



What is the internet

Back to how your house connects:

- How does your device connect to the Internet
 - The BT MASTER Socket

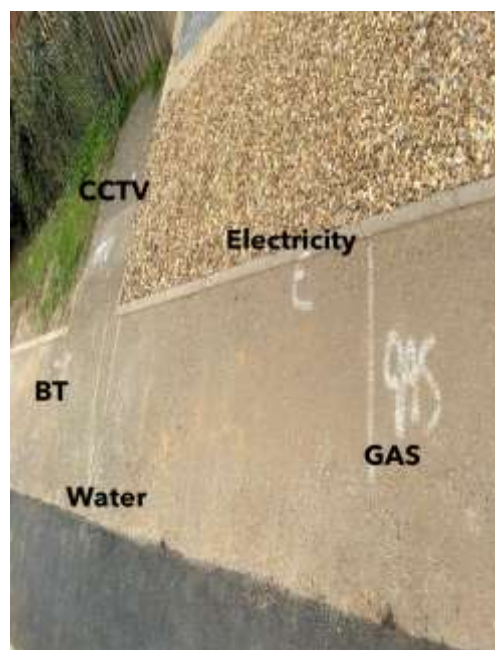
➢ This is where the TelCo's responsibilities finish ...

AND YOUR Network Speed is measured from here!

The diagram illustrates the connection path from a house to the internet. It shows a house connected to a SSFP (Service Selection Point), which connects to a Copper D-side (Voice & ADSL or VDSL) and then to a PCP (Primary Connection Point). The PCP connects to a DSLAM (Digital Subscriber Line Access Module) and then to an FTTC Cabinet. Photos show a worker at a PCP and a worker at a DSLAM.

The wire (fibre) from your house then goes to a rats nest of connections in the PCP (Primary Connection Point), this then connects to a second green box (that usually hums – electric cooling fans) that concentrate all of the individual houses phone/internet lines into a single set of lines that go to the telephone exchange.

These lines are now almost always fibre.



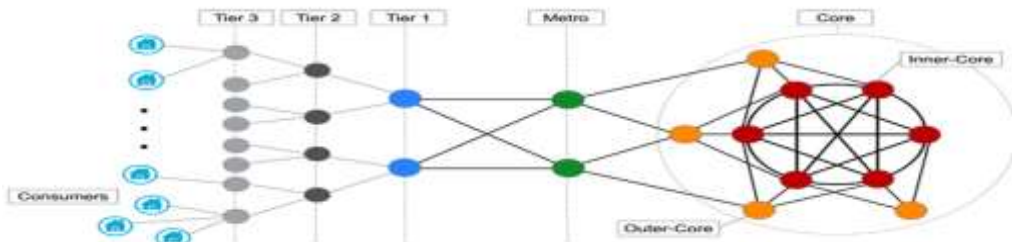
It all starts here:

- How does your device connect to the Internet
 - The BT MASTER Socket
- Then off into a world of wires ...

[TO SEE MORE: https://telephonesuk.org.uk/exchanges/exchange-inside/](https://telephonesuk.org.uk/exchanges/exchange-inside/)

The exchange has a similar set of rats nests full of cables:

- More connection points
- More concentrators
- This technology converts the 'simple' home to exchange connection protocols to a series of high speed MPLS networks. Northampton = Metro, Milton Keynes = Core



- Your data is routed to an inner core exchange, such as LINX which houses network switches from each of the ISP's (BT, Sky, Vodaphone, Plusnet, Talktalk etc.) and connections to similar exchanges around the world.
- Your packets of information are routed to from your house to the destination though many more rats nests of cables ... A world of cables / satellite links

What is the internet

Lets start at home:

- The key is your Hub, which plugs into:
 - The BT MASTER Socket
 - An then connects all of your other devices
 - Usually via WiFi
- WiFi uses CSMA/CA !

Back to your home.

The core component here is your network hub, this connects every thing you have (apart from your mobile phone's telephone calls that use the mobile 4/5G network) to the Internet. Even your phones are usually set up to use your local hub for the internet.

They communicate with WiFi using a technology called CS MA / CA

- MA = Multiple Access, all can join in
- CS = Carrier Sense, but they wait until they don't hear another device chatting
- CA = Collision Avoidance. As they can never know if another device will start at the same time as them they all wait for the Hub to acknowledge their call

These Wifi connections use multiple radio frequencies on both 2.4GHz and 5GHz networks to avoid 'contention'. ie: Too many devices using the same frequency at the same time.

What is the internet



The diagram illustrates the components of an ADSL network. It shows a computer connected to an ADSL modem, which is connected to a splitter. The splitter is connected to a telephone socket (phone line) and a telephone. A power socket is also shown. A hub is connected to the telephone socket and a power socket. A router is connected to the hub and a power socket. A laptop is connected to the router. A network switch is connected to the router and a power socket. A server is connected to the network switch. A network switch is connected to the router and a power socket. A server is connected to the network switch. A network switch is connected to the router and a power socket. A server is connected to the network switch.

Your Equipment

- The Master Socket
- The ADSL Filter
- Cat 5 cable & RJ45 Plug
- Your Hub
 - A Modem
 - A radio transmitter/receiver
 - An Ethernet Switch/Router
 - A DHCP server
 - A Web Server ...

You Hub is a clever bit of kit with many separate functional components ...

Some Definitions

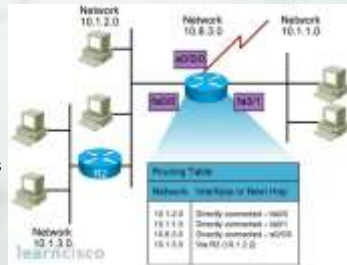
Your Equipment

- Your Hub
 - A Modem
 - A radio transmitter/receiver
 - An Ethernet Switch/Router
 - A DHCP server
 - A Web Server ...



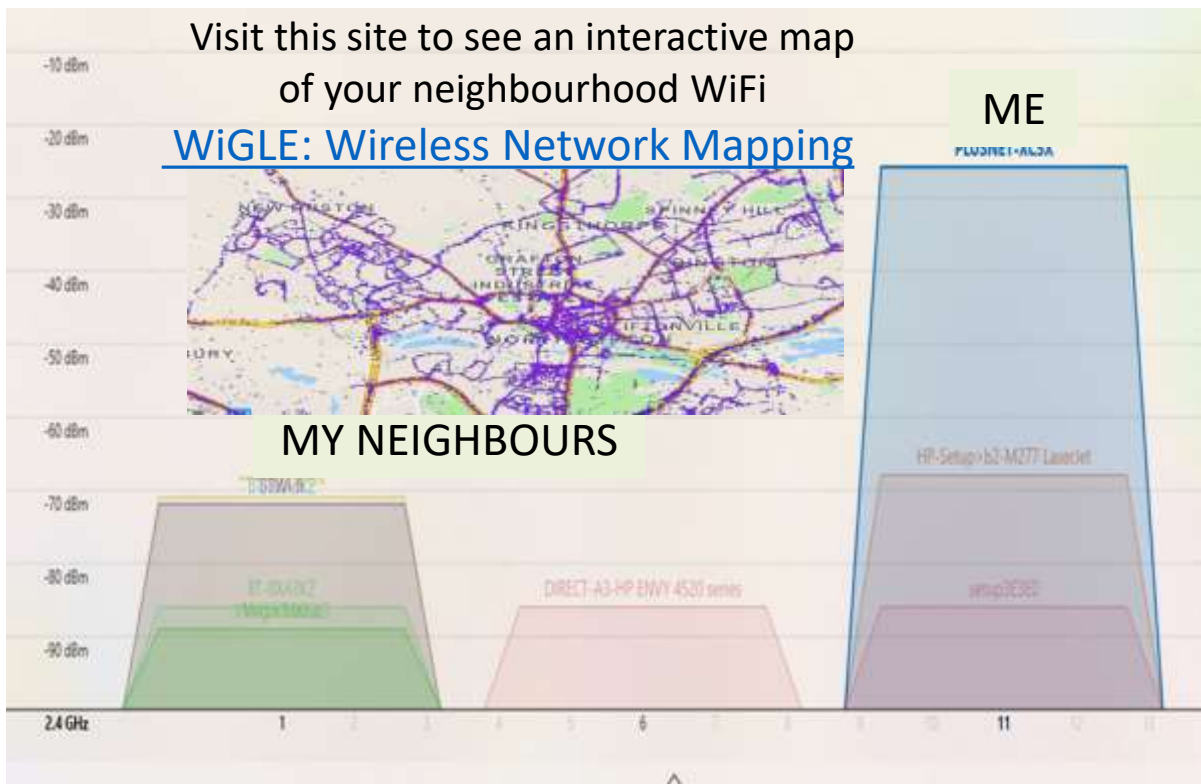
ROUTER: A network device that knows the routes to network destinations, such as within your property and outside (the Internet)

MODEM: A modulator-demodulator, a hardware device that converts data from a digital format, to another. Such as your network to the BT line.



A **DHCP Server** automatically assigns IP addresses to client computers.

A better example of multiple radio frequencies being used for WiFi communication, note you can see the neighbours too ...(see bar 1)



What is the internet

Your HUB

- It is ...
 1. A Cable Modem
 2. A radio Transmitter/Receiver
 3. An Ethernet Switch/Router
 4. Ethernet Controller
 5. A Computer
 - A Web Server
 - A DHCP server
 - Firewall
 - System Logger

The functional components are implemented in hardware as distinct computer chips, as shown above. The metal 'loop' around (2) is the WiFi aerial.

What is the internet

All my 27 Devices!

Your Network:

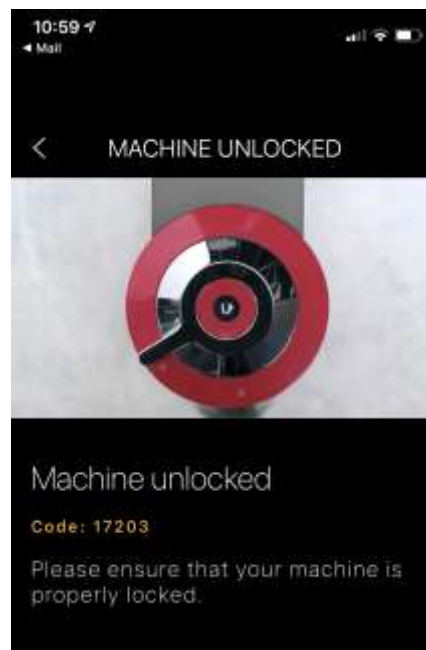
- The Hub controls your network
An then connects all your devices
- Laptops
- Tablets
- Phones
- Printer
- Disk Drive / Play Station, X Box
- TV's & DVD
- Sky /Virgin Box
- Radio / Network Streamer
- Alexa / Smart Speakers
- Fan, Coffee Machine, Lamp ...?



My network has 27 devices, including a FAN, Standard Lamp and a Coffee Machine. This is not a problem (apart from possible contention) until I change my hub / ISP* when I need to re-configure each device with the new network name and password – OUCH 😞

*ISP – Internet Service Provider

My Coffee machine
Talks to me over
the Internet →

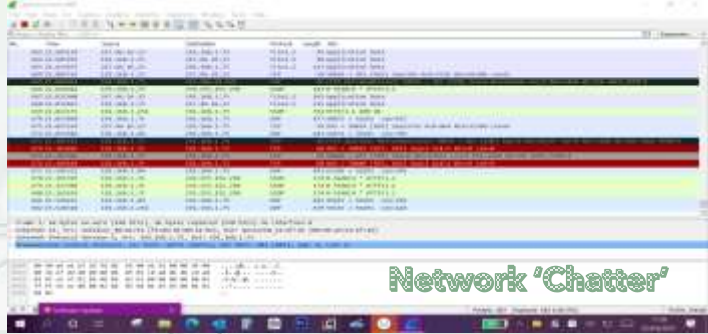



What is the internet

NEXT TIME:

- How does your home talk to the Internet

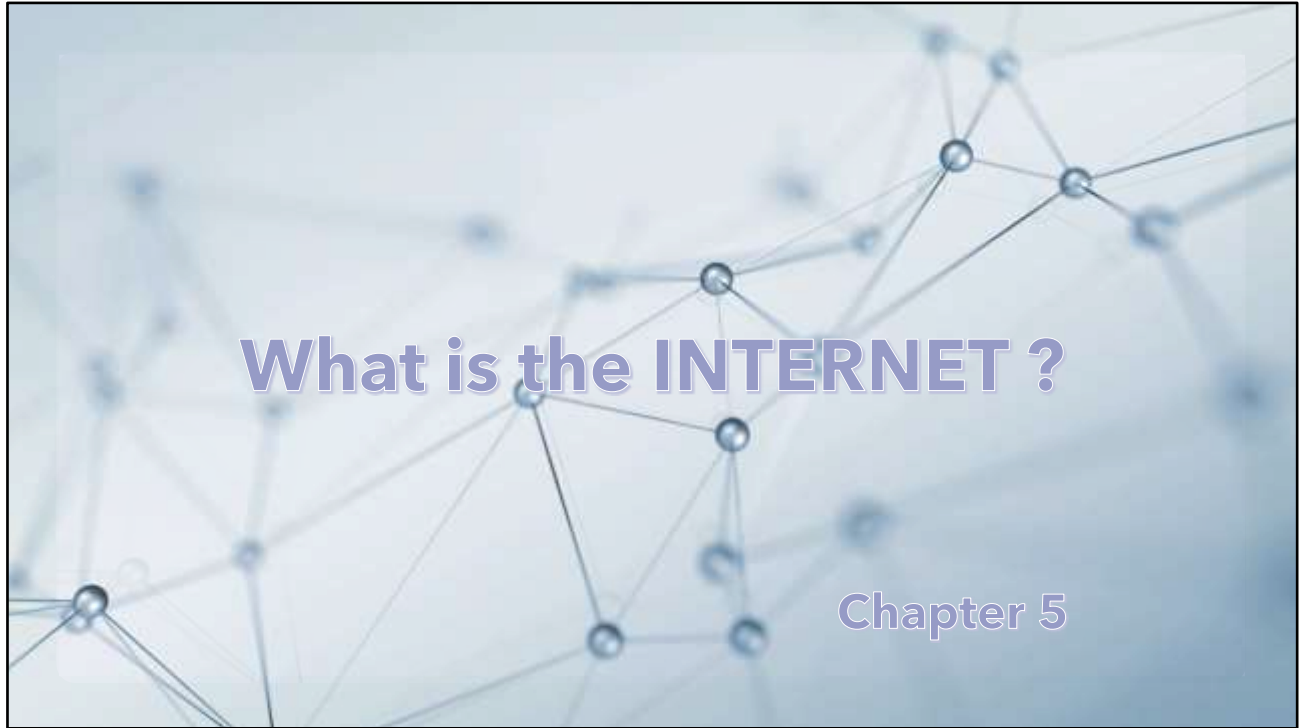
➤ Only if YOU want, that is?...

This 'teaser' is for next time shows the network chatter.

This was generated by a free tool called 'Wire Shark', who for the initiated can "listen" and capture all network traffic ... Mmm!

This is yet another reason why you should not use public WiFi (local coffee shop etc) to do your Banking, as the guy sitting opposite 'could' be capturing your data!



What is the internet

This Time

- IP Addresses !

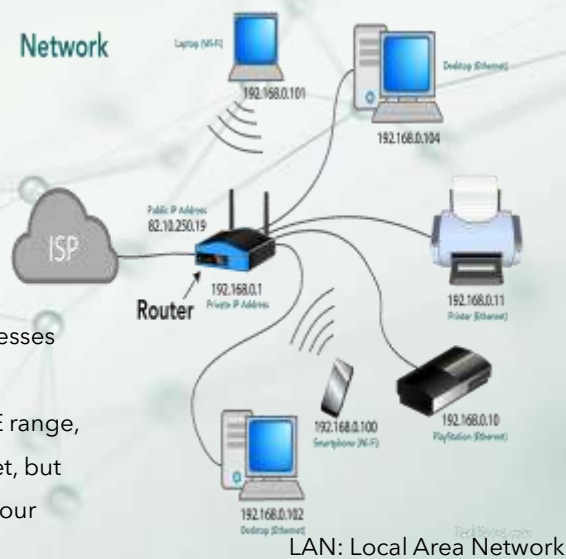


What is the internet

This is where it gets (slightly) ComplicatedThe IP Address

- Every device on a network, has an address
- It is made up from four sets of numbers
 - abc.def.ghi.kjm
 - Each set can range from 0 - 255
 - They are set up in classes, known as A,B,C,D
- This describes the original IP version 4 scheme, as you may guess we have run out of unique v4 addresses and a new IP v6 scheme is being introduced
 - The 192.168.x.y range is set up as a PRIVATE range, one that cannot be routed across the internet, but remains local - Your Hub uses this range in your home

Network



NB: When data leaves your 'local' network your 'local 192.168' address is replaced by your public address (supplied by your ISP) ie: 82.10.250.19

They are set up in classes A, B, C, D which classify NETWORKS (ie: collections of systems) and Hosts (the computers etc. themselves).

Originally designed to be used by Companies / Corporations.

For Example, the following Class A Network IP addresses belong to:-

12.0.0.0/8	AT&T Services
17.0.0.0/8	Apple Inc.
19.0.0.0/8	Ford Motor Company
38.0.0.0/8	PSINet, Inc.
48.0.0.0/8	Prudential Securities Inc.
56.0.0.0/8	US Postal Service
73.0.0.0/8	Comcast Corporation

- A: Values 1-126 in the first set – ie: 126 A networks, 16,777,214 hosts
- B: Values 128.0.0.1 to 191.255.0.0; Some 16,382 B networks, 65,534 hosts
- C: Value 192.0.0.0 to 223.255.255.0 and 2,097,150 C Networks, 254 hosts
- Where: 192.168.0.0 to 192.168.255.255 Class C Private Range

As you may guess, we are running out of IP Addresses

If you are in to Networking, you may recognise the PRIVATE range 192.168.x.y - This is what YOUR hub has allocated to your home network

What is the internet

This is where it gets (slightly) ComplicatedLocal IP Address talk directly

Lots of network traffic occurs where each host (computer) works out its own list of locally reachable systems, building its own table of Physical MAC addresses against IP addresses

- Once this has been created it knows who it shares the local network with
- They can then talk directly
- Without asking the router (your hub) to forward the messages

\$> arp -a The Windows command to show this

```
Interface: 192.168.1.75 --- 0xa
Internet Address      Physical Address      Type
192.168.1.65          90-48-9a-75-6e-26    dynamic
192.168.1.69          90-cd-b6-44-da-b2    dynamic
192.168.1.71          00-90-a9-e1-6f-26    dynamic
192.168.1.72          c8-ff-77-d6-21-4b    dynamic
192.168.1.76          68-57-2d-4a-38-6e    dynamic
192.168.1.84          60-02-b4-55-cd-93    dynamic
192.168.1.97          14-0a-c5-35-27-cf    dynamic
192.168.1.99          f4-03-2a-66-11-d2
192.168.1.100         f4-03-2a-16-cd-ee
192.168.1.104         1a-0b-bc-b5-e6-74
192.168.1.111         5c-93-a2-78-9d-3b
192.168.1.118         fe-27-ea-c9-2d-f2
192.168.1.119         c0-4a-00-ab-fe-f2
192.168.1.254         40-c7-29-0d-80-03
224.0.0.2             01-00-5e-00-00-02
224.0.0.22            01-00-5e-00-00-16
224.0.0.251           01-00-5e-00-00-fb
224.0.0.252           01-00-5e-00-00-fc
239.255.255.250      01-00-5e-7f-ff-fa
```

This is an example of your computers 'Address Resolution Protocol (ARP) table, that lists all local adjacent devices, so that they can talk directly without the intervention of the router (hub)

This is an example of a HUB's DHCP settings.

- It provides the range of addresses it is permitted to lease out
- And the lease period
- Here are some of the leased addresses, and their associated MAC's

The screenshot shows the BT Smart Hub Manager interface. The top navigation bar includes 'Home', 'Advanced settings', and 'My network'. Under 'My network', there are tabs for 'My devices', 'IPv4 configuration', and 'Address table'. The 'IPv4 configuration' tab is active, showing DHCP settings. The 'Hub gateway IP address' section has 'IP address' set to 192.168.1.254 and 'Subnet mask' set to 255.255.255.0. The 'DHCP server' section has 'Enabled' set to YES and 'Authoritative DHCP' set to ON. The 'Server address range' is 192.168.1.64 - 192.168.1.253, and the 'Lease time' is 1 Day(s) and 0 Hour(s). Below this, the 'IP4 devices' section lists several devices with their IP addresses and MAC addresses.

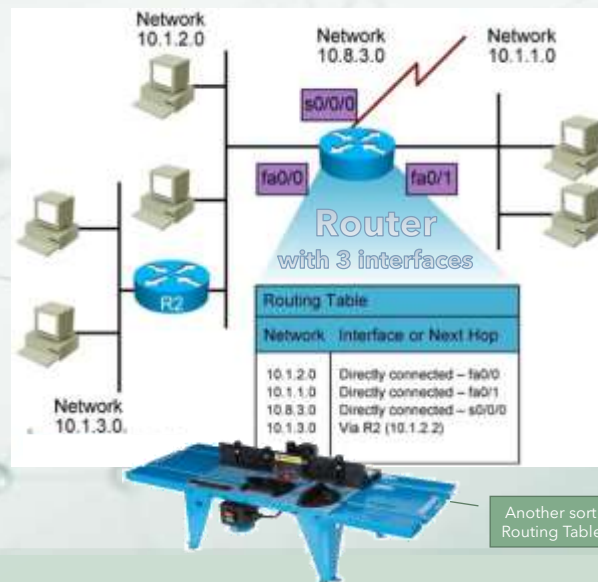
Device Name	IP Address	MAC Address	Lease Time Left
Iceberg	192.168.1.75	MAC0090A9E16F20	0:15:55:37
DESKTOP-FCS3TDD	192.168.1.100	MAC000000020053	Expired
192.168.1.99	192.168.1.99	MACF4032A6611D2	0:22:23:41
NM6-UK-WHAD415A	192.168.1.72	MAC88FF77062148	0:16:52:04

What is the internet

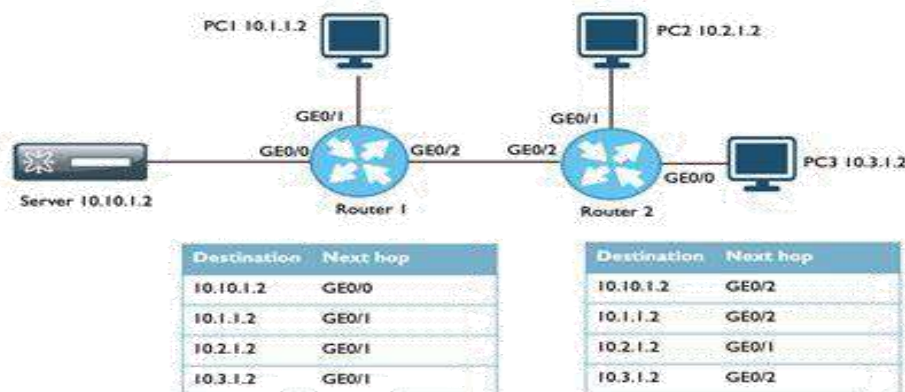
This is where it gets (slightly) ComplicatedIP Address that are ROUTED

If the IP address is not on the same local network, it needs to be **routed** to be able to reach it

- Your computer has a **Default Gateway** configured in its network settings
- The default gateway is usually your hub
- Any non-local address request is then sent to the default gateway
- This either knows where to forward it, or sends it to its own default gateway

**Routing Tables:**

So how does your data traverse the Internet? The answer lies in routing. Dedicated devices (called 'Routers', what else) talk to



each other to create their own local 'routing table' which tells them which physical interface to use to get to the requested destination. The only component missing from this diagram is the default gateway where it sends all the data it doesn't have a destination listed. "Is this the way to Australia?" – "I don't know, just carry on down the road" sort of technology.

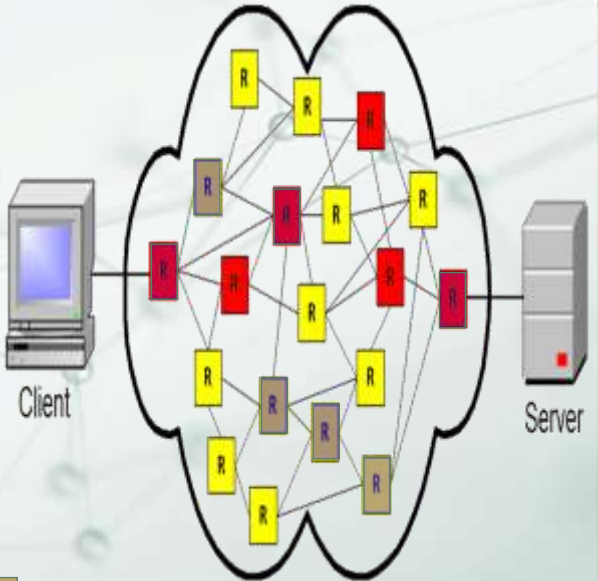
What is the internet

This is where it gets (slightly) Complicated

IP Address that are ROUTED

Exactly the same mechanism occurs across the internet, which is as we have seen made up of many devices and connections

- As your message arrives at each internet router the same question is asked:
 - Do you know how to connect to ...
 - If the answer is on, the message is just passed to that routers **Default Gateway**
 - Until the destination is found
- It is a bit more complex than that, as routes have associated costs, and the cheapest path usually wins. *So the return path may be different*



There and back may not be the same

This diagram shows the path from client to server. However the return path may not be the same. Routers are clever devices, as they often assign a cost on each interface. The route selected uses the least cost path. So for the return journey the path selected may be different.

Add another level of complexity. Different network traffic (like bicycles, cars and lorries – video, web, ftp etc) can be assigned different routes, or even be delayed to allow more time critical traffic to pass swiftly (like voice and video).

What is the internet

This is where it gets (slightly) Complicated

PROOVE IT!

Step 1: <https://whatismyipaddress.com>

This website will tell you what your external (public) IP address is. The one provided by your ISP.

NB: They may change it from time to time


- Mine is currently : **91.125.1.22**

NOTE: This web site ALSO tells you where the address is (probably) located. See later.

<https://lookup.icann.org/>

IP Details For: 91.125.1.22

Decimal:	1534919132
Hostname:	220.1.125.91.dyn.plus.net
ASN:	6871
ISP:	Plusnet
Organization:	Plusnet
Services:	None detected
Type:	Unrouted
Assignment:	Likely Dynamic IP
Continent:	Europe
Country:	United Kingdom
State/Region:	Northamptonshire
City:	Northampton



Latitude: 52.2318 (52° 13' 54.48" N)
Longitude: -0.9171 (0° 55' 1.56" W)
Postal Code: NN5

[CLICK TO CHECK BLACKLIST STATUS](#)

How do I know that this is true?

I can check my EXTERNAL IP address by looking it up, as demonstrated.

This lookup shows us another interesting piece of information – the location of an IP address. Companies / Countries use this to ‘GEO-Fence’ Internet activity. If you have ever tried to get the BBC iPlayer when you were abroad you would be told that it is forbidden. They looked up your IP address and saw that you were not in the UK.

There is a way around this, as always. Use a VPN.

A VPN (Virtual Private Network) is a service that you purchase that grants you access to a computer located somewhere else, it creates a ‘pipe’ from your machine to that remote system. I use one in France that pops out on a Server in Rugby, so the iPlayer believes I’m local and I can watch the BBC. NB: If they get wise to this they could blacklist my rugby server, so don’t tell anyone 😊

What is the internet

This is where it gets (slightly) Complicated


PROOVE IT!


Step 1: Find your external address:
(Mine is currently : 91.125.1.22)

Step 2: Select a destination:
I picked the Western Australian water Corporation, and googled it.

Step 3: collect it's URL (remember that?)
watercorporation.com.au

Step 4: use the built in Windows tool
\$> tracert (Trace route) followed by the URL
NB: iNetTools free for iPads does this too.



There are a number of built in network tools on a Windows 10 computer, available from the command prompt. (Type 'cmd' in the Search  box). Or from iNetTools ([Inet Tools Overview](#)) for tablets.

Three of these include:

PING: to measure connectivity

```
C:\>ping 8.8.8.8
Pinging 8.8.8.8 with 32 bytes of data:
Reply from 8.8.8.8: bytes=32 time=22ms TTL=117
Ping statistics for 8.8.8.8:
    Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
    Approximate round trip times in milli-seconds:
        Minimum = 9ms, Maximum = 91ms, Average = 32ms
```

TRACERT: to trace the route

```
C:\>tracert 8.8.8.8
Tracing route to dns.google [8.8.8.8]
over a maximum of 30 hops:
  0  1 ms  2 ms  1 ms  BTHUB [192.168.1.254]
```

NSLOOKUP: DNS Name service lookup, convert name to IP Address

```
C:\>nslookup watercorporation.com.au
Server: dns.google
Address: 8.8.8.8
Non-authoritative answer:
Name: watercorporation.com.au
Address: 20.188.217.99
```

What is the internet

This is where it gets (slightly) ComplicatedTrace Route

This is a standard network tool for tracing the route taken by (your) data packets and measuring the time taken for each hop.

- It gropes the Internet (uses another tool called PING {Packet Internet Groper})
- It sends 3 pings 1 hop and sees who is there
- It measures the time taken (I edited the output)
- It repeats for 2 hops, etc. up to a max of 30
- The output is shown on the right, annotated using whatismyipaddress.com to add missing locations.

NB: Some Internet equipment has the ability to respond to ping switched off, hence "Request timed out", due to a 'ping 'o death hack ☹️

Tracing route to wawa.com.au [110.142.127.20] over a maximum of 30 hops:

```

1 766 ms      BTHUB [192.168.1.254] Duston - my house
2 *          91.125.1.22 External Address - link to ISP - 220.1.125.91.dyn.plus.net Northampton
3 *          Request timed out.
4 389 ms      136.hiper04.sheff.dial.plus.net.uk [195.166.143.140] Sheffield (Plusnet)
5 10 ms      peer7-et-3-1-2.telehouse.ukcore.bt.net [194.72.16.102] London [the LINX]
6 9 ms       166-49-214-194.gia.bt.net [166.49.214.194] London (BT)
7 9 ms       134.159.95.25 i-0-4-0-7-peer.ulco-core02.pr.telstraglobal.net London (Telstra Global)
8 13 ms      i-1001.ulhc-core02.telstraglobal.net [202.84.148.102] Hong Kong (Telstra Global)
9 104 ms     i-1022.ny8a-core01.telstraglobal.net [202.84.141.229] Hong Kong (Telstra Global)
10 137 ms    i-10748.paix-core02.telstraglobal.net [202.40.148.102] Hong Kong (Telstra Global)
11 289 ms    i-10748.paix-core02.telstraglobal.net [202.40.148.102] Hong Kong (Telstra Global)
12 289 ms    i-37.sydco-core03.telstraglobal.net [202.84.247.46] Sydney (Telstra Global)
13 282 ms    bundle-ether3.oxf-gw11.sydney.telstra.net [203.50.13.97] Sydney (Telstra)
14 289 ms    bundle-ether1.chw-core10.sydney.telstra.net [203.50.6.92] Sydney (Telstra)
15 296 ms    bundle-ether8.exi-core10.melbourne.telstra.net [203.50.11.125] Melbourne (Telstra)
16 307 ms    bundle-ether2.way-core10.adelaide.telstra.net [203.50.6.230] Adelaide (Telstra)
17 335 ms    bundle-ether3.pie-core10.perth.telstra.net [203.50.6.235] Perth (Telstra)
18 335 ms    ae10.pie-ice301.perth.telstra.net [203.50.61.225] Perth (Telstra)
19 *        Request timed out.
20 352 ms    110.142.127.20 Water Corporation Perth

```

The more astute of you may have noticed I cheated here, changing the facts to make the story more interesting. When I worked for the West Australian water Authority they were known as WAWA. The traceroute to this address was much more enlightening than the one to watercorporation.com.au, sorry.

Tracing route to watercorporation.com.au [20.188.217.99] over a maximum of 30 hops:

```

1      13 ms      17 ms      3 ms      BTHUB [192.168.1.254]
2      *          *          *          Request timed out.
3      *          *          *          Request timed out.
4      15 ms      14 ms      23 ms     140.hiper04.sheff.dial.plus.net.uk [195.166.143.140]
5      24 ms      18 ms      20 ms     peer3-et0-0-5.redbus.ukcore.bt.net [194.72.16.102]
6      11 ms      13 ms      9 ms      ae64-0.lts-96cbe-1a.ntwk.msn.net [104.44.15.134]
7      9 ms       9 ms       10 ms     ae25-0.icr01.lon22.ntwk.msn.net [104.44.239.101]
8      251 ms     250 ms     250 ms     be-100-0.ibr01.lon22.ntwk.msn.net [104.44.21.87]
9      250 ms     250 ms     250 ms     be-5-0.ibr01.par30.ntwk.msn.net [104.44.17.76]
10     249 ms     356 ms     248 ms     be-8-0.ibr01.mrs21.ntwk.msn.net [104.44.29.167]
11     342 ms     251 ms     250 ms     be-21-0.ibr01.mrs20.ntwk.msn.net [104.44.28.163]
12     249 ms     249 ms     294 ms     be-14-0.ibr01.sg2.ntwk.msn.net [104.44.17.66]
13     247 ms     247 ms     246 ms     be-7-0.ibr01.per30.ntwk.msn.net [104.44.19.163]
14     249 ms     249 ms     250 ms     be-8-0.ibr01.syd23.ntwk.msn.net [104.44.7.123]
15     247 ms     270 ms     250 ms     ae100-0.icr01.syd23.ntwk.msn.net [104.44.11.102]
16     *          *          *          Request timed out.
17     *          *          *          Request timed out.
18     *          *          *          Request timed out.
...
28     *          *          *          Request timed out.
29     *          *          *          Request timed out.
30     *          *          *          Request timed out.

```

Trace complete.


These 104.44. networks all belong to Microsoft Azure server farm in Kansas???

What is the internet

USEFUL TIP

Domain Name to IP ADDRESS LOOKUP

- You can look deeper into the source of an iffy email/txt
- On Widows (or via an app on your tablet) Get a command prompt CMD and type
 - **'nslookup'** to get the **Domain name to IP address**
- Then use **'www.whatismyipaddress.com'** to find where the IP comes from ...
 - In an odd country ... ☹
 - Etc ...




You have to be quick to spot the gangsters. After doing the nslookup on the spoof admin-verify-secure.com I just got a warning about a scam site provided by the domain registration company. When I originally did it, but forgot the screenshot the server was in Venezuela.

Just to prove it works 😊 ... www.bancodebogota.com (the Bank of Bogota Columbia) gives:

```
C:\>nslookup www.bancodebogota.com
Server: dns.google
Address: 8.8.8.8
Non-authoritative answer:
Name:   webav.bancodebogota.com
Address: 200.14.232.18
Aliases: www.bancodebogota.com
```

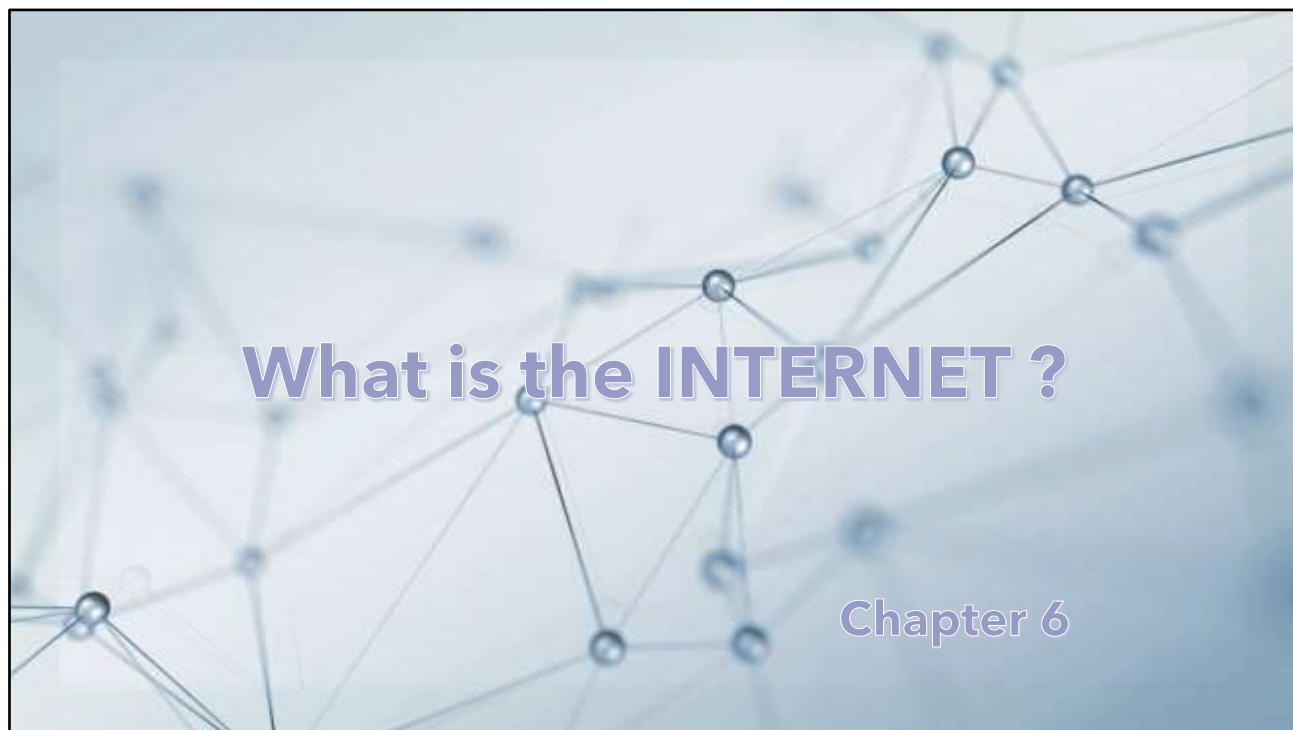
IP Details For: 200.14.232.18

Decimal:	3356420114
Hostname:	200.14.232.18
ASN:	264714
ISP:	Claro Chile
Organization:	A Toda Hora S.A
Services:	None detected
Type:	Corporate
Assignment:	Static IP
Continent:	South America
Country:	Colombia
State/Region:	Bogota D.C.
City:	Bogotá



Latitude: 4.6351 (4° 38' 6.36" N)
 Longitude: -74.0703 (74° 4' 13.08" W)
 Postal Code: 111311

[CLICK TO CHECK BLACKLIST STATUS](#)



This Time

- Is Big Brother actually watching you?



What is the internet

Where can we be Spied upon?

1. Your laptop
 - a) The operating system
 - b) Opening hidden pixels
2. The Web Browser
 - a) The browser itself
 - b) cookies
3. Your Hub
4. The exchange equipment
5. The search engine
6. The remote servers
7. Email servers / Twitter / Facebook / Instagram ...etc.

There is nowhere to hide (on the Internet)...

You have almost certainly given permission on your own computer/tablet/phone for 'diagnostic' data to be collected by the operating systems manufacturer (Microsoft, Apple, Google Android). This will provide them with details of the applications you use, your location, your equipment and a lot more.

They probably have created a unique identification number for you, which although will not name you as such will provide them (and the companies they sell it to 'useful' marketing data).

Once you start using a program, such as a Web Browser or an application like Facebook, Twitter, Instagram even more data will be collected from you.

This is even before you actually go anywhere on the internet.

What is the internet

The Operating System

With Windows 10 Microsoft has introduced a policy where the collection of basic data from machines running the operating system, data that it calls telemetry, is mandatory. Part of the deal of being allowed to run Windows 10 is handing this data over to Microsoft.

• For legal reasons

We will share personal information with companies, organizations or individuals outside of Google if we have a good-faith belief that access, use, preservation or disclosure of the information is reasonably necessary to:

- meet any applicable law, regulation, legal process or enforceable governmental request.
- enforce applicable Terms of Service, including investigation of potential violations.
- detect, prevent, or otherwise address fraud, security or technical issues.
- protect against harm to the rights, property or safety of Google, our users or the public as required or permitted by law.

Others

It may be necessary – by law, legal process, litigation, and/or requests from public and governmental authorities within or outside your country of residence – for Apple to disclose your personal information. We may also disclose information about you if we determine that for purposes of national security, law enforcement, or other issues of public importance, disclosure is necessary or appropriate.

We may also disclose information about you if we determine that disclosure is reasonably necessary to enforce our terms and conditions or protect our operations or users. Additionally, in the event of a reorganization, merger, or sale we may transfer any and all personal information we collect to the relevant third party.

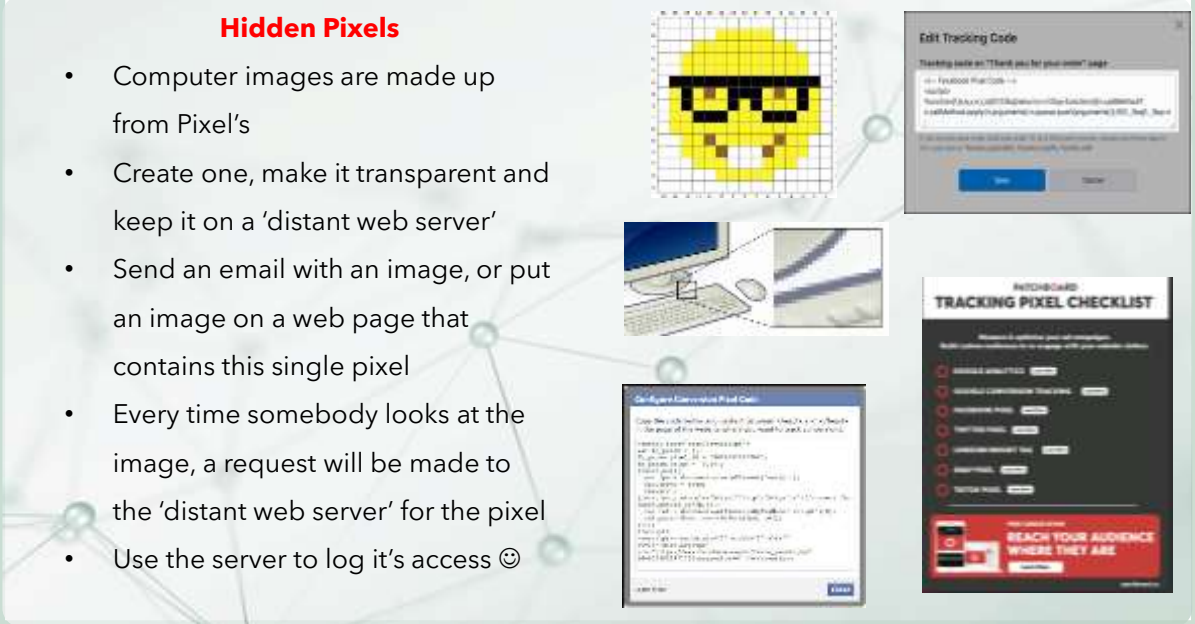
Finally, we will access, disclose and preserve personal data, including your content (such as the content of your emails, other private communications or files in private folders), when we have a good faith belief that doing so is necessary to:

1. comply with applicable law or respond to valid legal process, including from law enforcement or other government agencies;
2. protect our customers, for example to prevent spam or attempts to defraud users of the services, or to help prevent the loss of life or serious injury of anyone;
3. operate and maintain the security of our services, including to prevent or stop an attack on our computer systems or networks; or
4. protect the rights or property of Microsoft, including enforcing the terms governing the use of the services - however, if we receive information indicating that someone is using our services to traffic in stolen intellectual or physical property of Microsoft, we will not inspect a customer's private content ourselves, but we may refer the matter to law enforcement.

What is the internet

Hidden Pixels

- Computer images are made up from Pixel's
- Create one, make it transparent and keep it on a 'distant web server'
- Send an email with an image, or put an image on a web page that contains this single pixel
- Every time somebody looks at the image, a request will be made to the 'distant web server' for the pixel
- Use the server to log it's access 😊



Have you ever been scrolling through Facebook and an ad shows up for something you were *literally* just talking about? Weird, right?

While Facebook isn't actually reading your mind (yet), they *are* using what we call a tracking pixel.

What is a Tracking Pixel?

Marketing pixels, aka tracking pixels, are essentially these tiny snippets of code that allow you to [gather information about visitors](#) on a website—how they browse, what type of ads they click on, etc.

This behavior data helps you, as a marketer, send the user [paid ads](#) that are likely to be most interesting to them. Tracking pixels are also used to measure a marketing campaign's performance, track conversions, and build an audience base.

Now that you have a general overview of what a pixel is, let's talk about the different types of pixels. Don't stress too much though, there are only 2 that you really need to worry about.

Retargeting Pixels

[Retargeting](#) pixels are focused solely on the behavior of your website's visitors.

For example, let's say you're shopping online for new furniture and then you go onto an entirely different website, like Facebook, and notice that all of the pop-up ads are furniture related.

That's how retargeting pixels work. They're basically monitoring your behavior in order to tailor [paid ads](#) they think will catch your attention on other websites.

Conversion Pixels

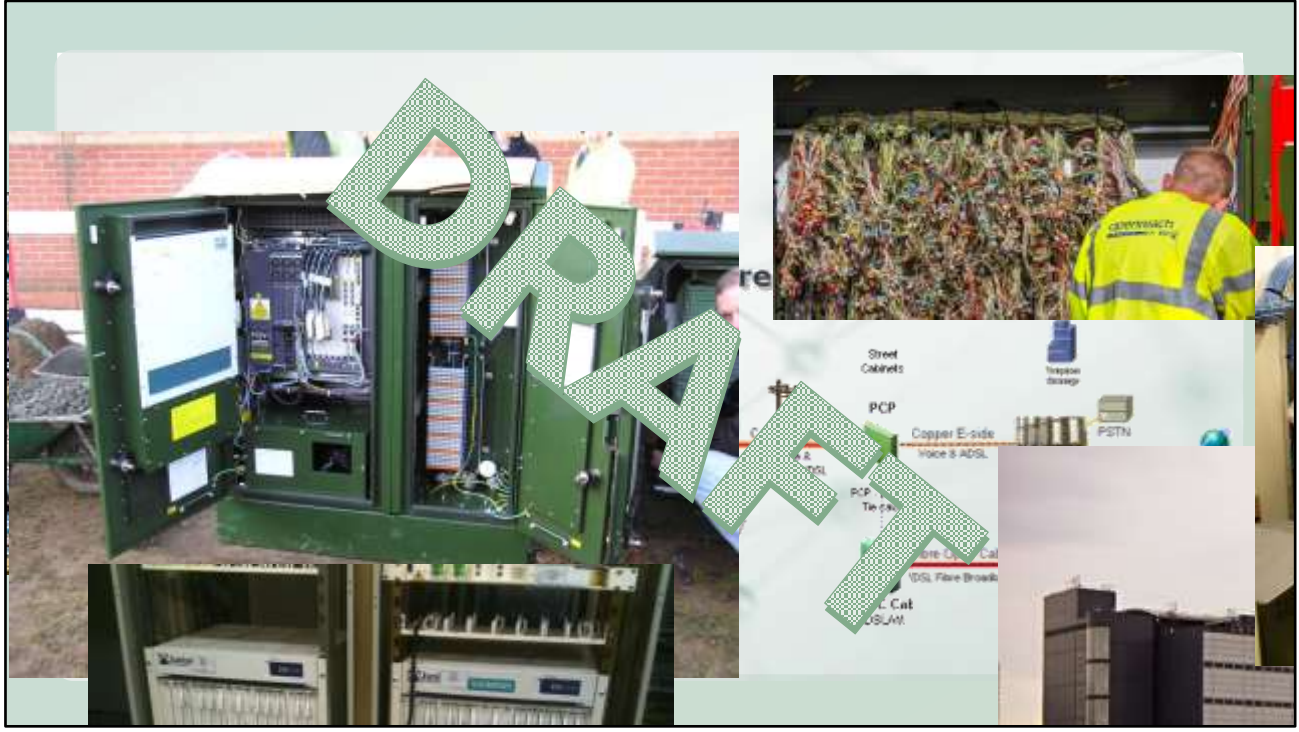
A conversion pixel comes into play once a purchase has actually been made. They are responsible for tracking sales from a specific ad campaign.

In order to gather correct data, conversion pixels need to be placed within the code of an order confirmation page such as an automated "Thank You" you typically find in your inbox after buying something.

Conversion pixels also allow marketers to identify the source of their conversions and measure the success or failure of specific campaigns.

Don't believe me? – [just click here](#) 😊 ...

What is the internet



What is the internet

Your Equipment

- The Master
- The ADSL Filter
- Cat 5 cable & RJ45
- The WiFi Router
- YOUR STUFF !
 - More than 25 in my house inc: Fans, Coffee Machine, TV, Sky box, Phones, iPads, ...
 - You may have a lot of internet Things

Tracing route to wawa.com.au [110.142.127.20] over a maximum of 30 hops:

0		110.142.127.20	Water Corporation Perth
1	1766 ms	BTHUB [192.168.1.254]	Duston - my house
2	*	71.125.1.22	External Address - link to ISP - 220.1.125.91.dyn.plus.net North
3	*		Request timed out.
4	389 ms	136.hiper04.sheff.dial.plus.net.uk [195.166.143.136]	Sheffield (Plusnet)
5	110 ms	peer7-et-3-1-2.telehouse.ukcore.bt.net [109.159.252.230]	London [the LINX]
6	9 ms	166.49.214.194.gia.bt.net [166.49.214.194]	London {BT}
7	9 ms	134.159.95.25	i-0-4-0-7-peer.ulco-core02.pr.telstraglobal.net London {Telstra
8	13 ms	i-1001.ulhc-core02.telstraglobal.net [202.84.178.70]	Hong Kong {Telstra Glob
9	104 ms	i-1022.ny8a-core01.telstraglobal.net [202.84.141.229]	Hong Kong {Telstra Gl
10	137 ms	i-10748.paix-core02.telstraglobal.net [202.40.148.102]	Hong Kong {Telstra G
11	289 ms	i-10748.paix-core02.telstraglobal.net [202.40.148.102]	Hong Kong {Telstra G
12	289 ms	i-37.sydo-core03.telstraglobal.net [202.84.247.46]	Sydney {Telstra Global}
13	282 ms	bundle-ether3.oxf-gw11.sydney.telstra.net [203.50.13.97]	Sydney {Telstra}
14	289 ms	bundle-ether1.chw-core10.sydney.telstra.net [203.50.6.92]	Sydney {Telstra}
15	296 ms	bundle-ether8.exi-core10.melbourne.telstra.net [203.50.11.125]	Melbourne {Te
16	307 ms	bundle-ether2.way-core10.adelaide.telstra.net [203.50.6.230]	Adelaide {Telstra
17	335 ms	bundle-ether3.pie-core10.perth.telstra.net [203.50.6.235]	Perth {Telstra}
18	335 ms	ae10.pie-ice301.perth.telstra.net [203.50.61.225]	Perth {Telstra}
19	*		Request timed out.
20		110.142.127.20	Water Corporation Perth

How We Use Cookies

We use cookies for a variety of reasons detailed below. Unfortunately in most cases there are no industry standard options for disabling cookies without completely disabling the functionality and features they add to this site. It is recommended that you leave on all cookies if you are not sure whether you need them or not in order they are not to provide a service that you use.

Disabling Cookies

You can prevent the setting of cookies by adjusting the settings on your browser (see your browser Help for how to do this). Be aware that disabling cookies will affect the functionality of this and many other websites that you visit. Disabling cookies will usually result in also disabling certain functionality and features of this site. We therefore recommend that you do not disable cookies.

The Cookies We Set

- Site preferences cookies

In order to provide you with a great experience on this site we provide the functionality to set your preferences for how this site runs when you use it. In order to remember your preferences we need to set cookies so that this information can be called whenever you interact with a page is affected by your preferences.

Third Party Cookies

In some special cases we also use cookies provided by trusted third parties. The following section details which third party cookies you might encounter through this site.

- This site uses Google Analytics which is one of the most widespread and trusted analytics solution on the web for helping us to understand how you use the site and ways that we can improve your experience. These cookies only track data such as how long you spend on the site and the pages that you visit so we can continue to produce engaging content. For more information on Google Analytics cookies, see the official Google Analytics page.
- The Google AdSense service we use to serve advertising uses a DoubleClick cookie to serve more relevant ads across the web and limit the number of times that a given ad is shown to you. For more information on Google AdSense see the official Google AdSense privacy FAQ.
- Several partners advertise on our behalf and affiliate tracking cookies simply allow us to see which customers have come to the site through one of our partner sites so that we can credit them appropriately and where applicable allow our affiliate partners to provide any bonus that they may provide you for making a purchase.

Kargo Global Inc.
Cluep LLC
SpearAd GmbH
Playrcart Limited
Pelmorex Corp.
Sparklit Networks Inc
ARKHEUS
Mediascore mbH
WAM Group Ltd, t/a Admix
AREIA (ADLOOP)
Netpoint Media GmbH
Axiom Media Connect
Pulse Innovations Limited
BYTEDANCE PTE. LTD.
Media16 ltd
Habu
SPEAKOL DMCC
QuantumCast digital GmbH
iProspect GmbH
Rapid Performance GmbH
QUARTER MEDIA GmbH
Leadoo Marketing Technologies Ltd
Hashtag Labs Inc.
CMI Marketing, Inc. d/b/a CafeMedia
EXOCLICK, S.L.

Quantly00 GmbH & Co. KG
Hillside (Sports) GP Limited
NumberEight Technologies Ltd
Online Media Solutions LTD (BDA: Brightcom)
Listonic Sp. z o.o.
Dianomi ltd
Remixd Media, Inc.
Prebid.org
GADSME
Admo.tv (Clickon)
Channel Pilot Solutions GmbH
ANINPRO-CREATIVE, S.L.
GeoProve
TIMEONE PERFORMANCE
Threedium
Maiden Marketing Pte Ltd
IPSOS MORI UK LTD
Fido S.r.l
Mobkoi Ltd
digidip Gm

194 Cookies
just from
One Site !

Investigatory powers act ...

As it stands, the main purposes of the law are:

- Security services are legally empowered to bug computers and phones upon approval of a warrant. Companies are legally obliged to assist these operations and bypass encryption where possible (more on this ...)
- Security services can acquire and analyse bulk collections of communications data. For example, this could mean a bulk dataset such as health records. This now can only be undertaken in cases of serious crime, defined as those which could attract a sentence of 12 months minimum.

Unless you have taken steps to prevent it, your ISP is tracking pretty much everything you do online. That means it can see your internet search history, the websites you've visited, what you've downloaded, etc., at any time.

Depending on where you are in the world, your ISP is likely to be required to retain this data on you for a specific period of time. Not only that, but it's highly likely your ISP is profiting off your data by selling it to advertisers, similarly to how [Google](#) operates with its tracking.

<https://www.google.co.uk/search?client=safari&hl=en-gb&sxsrf=ALeKk01Gm-...kbylPQ%3A1620902021342&lei=hQCdYlKvFNma1fAP5aeO6Ag&q=investigatory%20powers%20act%20200&ved=2ahUKEwjyhcJusbwAhVZTRUIHeWTA40QsKwBKAB6BAG7EAE&biw=1103&bih=1415&dpr=2>

Search engine history

<https://www.wordstream.com/blog/2006/06/identifying-search-engines-history>

Identifying search engine users ... the AOL ... mpr ... 2006...

User 491577 searches for "florida cna p... ke... nipa", "em... school training florida", "low calorie meals", "infant seat", and "fisher price roller blades".

Among user 39509's hundreds of searches are: "oklahoma disciplined pastors", "oklahoma disciplined doctors", "home loans", and some other personally identifiable and illegal stuff I'm going to leave out of here.

Among user 545605's searches are "shore hills park mays la... ng ni... nk... m sindoni md", "ceramic ashtrays", "transfer money to china", and "capital gains on sale of house". Considering all of the data, these examples are on the safe side.

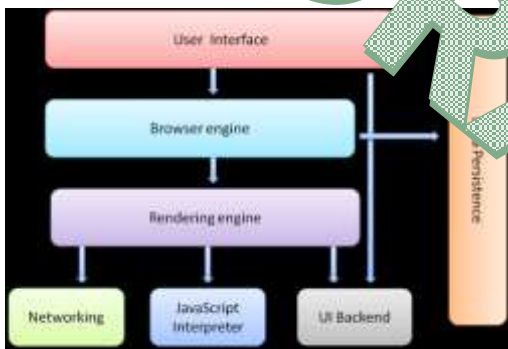
And

Buried in a list of 20 million Web browser cookies created by AOL and recently released on the Internet is user No. 4417749. The number was assigned by the company to protect user's anonymity, but it was not much of a shield. No. 4417749 conducted hundreds of searches over a three-month period on topics ranging from "numb fingers" to "60 single men" to "dog that urinates on everything."

And search by search, click by click, the identity of user No. 4417749 became easier to discern. There are queries for "landscapers in Lilburn, Ga," several people with the last name "Lilburn" and "homes for sale in shadow lake subdivision gwinnett county georgia."

It did not take much investigating to follow that data trail to Thelma, a 67-year-old woman who lives in Lilburn, Ga., frequently researches her friends' medical ailments and loves her three children. "I'm sure my son would be my best friend," she said, after a reporter read part of the list to her.

DRAFT



DOM

The output tree (the "parse tree") is a tree of DOM element and attribute nodes. DOM is short for Document Object Model. It is the object presentation of the HTML document and the interface of HTML elements to the outside world like JavaScript. The root of the tree is the "Document" object.

The DOM has an almost one-to-one relation to the markup. For example:

```
<html>
  <body>
    Hello World
  </body>
  <div> </div>
</html>
```

The above markup could be translated to the following DOM tree:

Figure 1: DOM tree of the above markup

