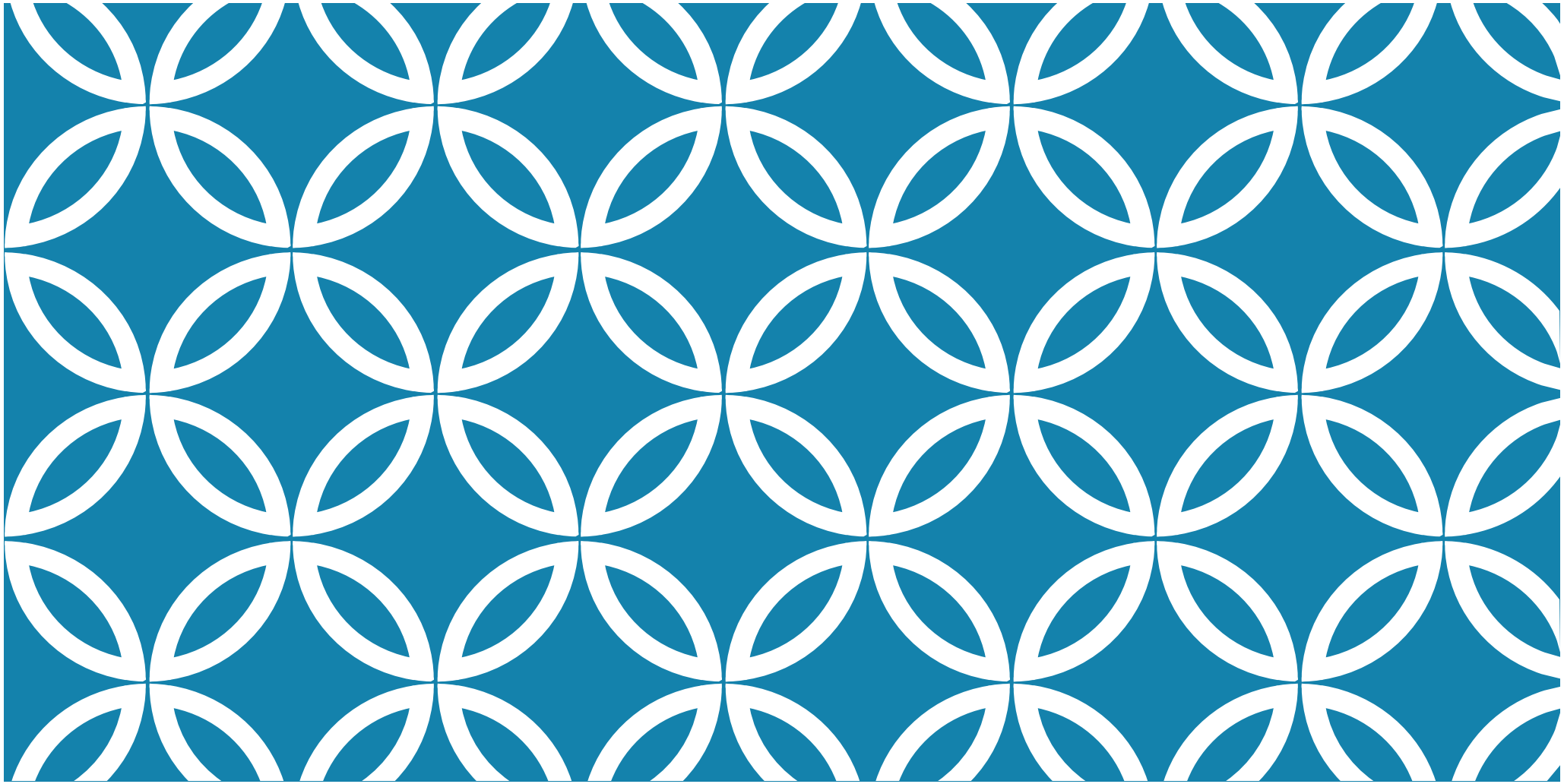


COMPUTER AND NETWORK SECURITY

It needn't be like this

Peter Chapman. October 12, 2021. peter.chapman2@sympatico.ca

OUR COMPUTERS ARE INSECURE

The general purpose computer was conceived by Alan Turing in his seminal paper dated 1936. He envisaged a “universal machine” and knew that it could be programmed to perform any logical, mathematical or algorithmic function.

He conceived it in order to show that some problems were not provable as had been theorised by Kurt Gödel some three years earlier.

Turing did not set out to create a computer, but to prove it could be done and use it to prove the theorem.

Turing’s genius was realized in code breaking machines during WW2.

General purpose computers as we know them today appeared in 1948, in the UK at Manchester and Cambridge Universities and in the U..S. at Harvard.

They were unwieldy using a few hundred vacuum tubes, which are prone to failure.

ARCHITECTURE OF THE 1940S

The idea of storing the program in the computer rather than have it read as needed from some external medium was the first major breakthrough.

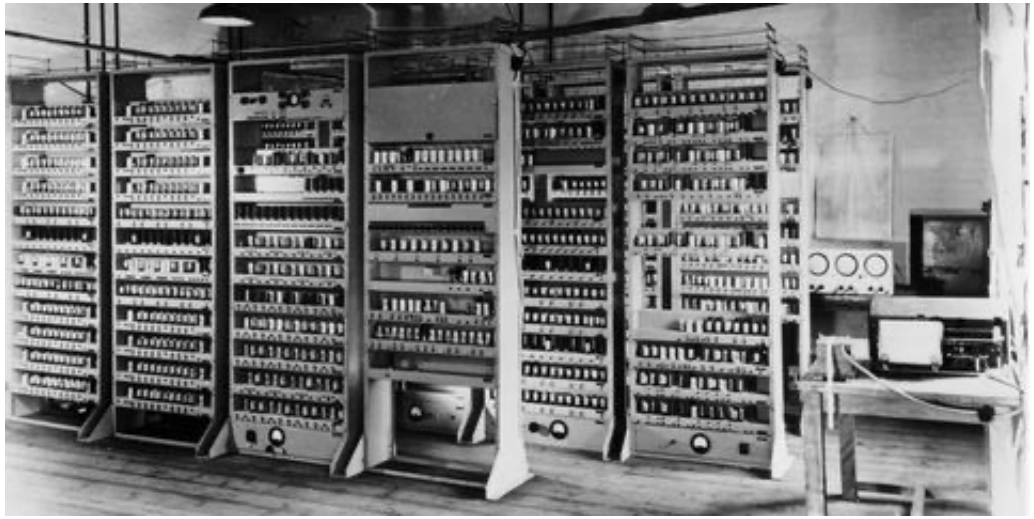
To do this an electronic memory needed to be created.

No-one knew how to do this then

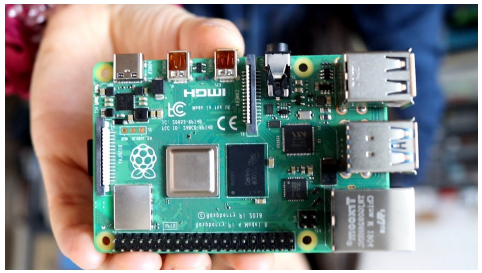
Various techniques were attempted,

- A pair of vacuum tubes for each bit, such that one of the other is conducting and by doing so holding the other off. A short duration electrical signal applied can switch their states over. This known as a bi-stable circuit and is in widespread use today as a component made from gates in an integrated circuit.
- Mercury delay lines were developed which used an acoustic pulse to propagate and was then refreshed or nt depending on the state.
- Accordingly memory was expensive and difficult to create.

COMPUTERS FROM THE 1940S



EDAC at Cambridge University, 1949



Raspberry Pi today.
Complete computer $35 CDN.

SHARED MEMORY
THE GENIUS AND THE ACHILLES HEEL

Because memory is so critical and it was then difficult to produce, means were found to optimise it.

First and foremost was the idea of shared memory.

Memory that is not used for the program is released and made available for data.

Memory is “paged” meaning that it is overwritten by what is currently being used. To optimise this it is separated into “pure” and “impure”. Pure means it has not changed since it was retrieved so It does not have to be saved again. It can simply be overwritten.

Operating systems of the 1970s used the idea of pure and impure code, obviating the need to write pure code back when closing.

TODAY MEMORY IS SO CHEAP

About $5 a billion bytes today.

A processor costs less than $5 today (e.g. ARM core, Broadcom chip)

There is absolutely no reason to share memory, which by doing so makes our computers unsafe.

Why?.

Because we have always done it this way.

Software needs to be reconfigured, though this is not very difficult, because all software is constructed using a compiler and loader. A compiler is software which translates high level language to executable code and organizes the layout of that code. A loader links pre compiled modules together.

HOW SAFE IS TODAY'S SOFTWARE?

Software from a reputable company is safe. It is protected during transit and installation by a signed checksum

- A number is created algorithmically, which will not match if the program has been changed in any way.
- The matching number is checked against a digitally signed number issued in the public domain by the creator of the software.
- The Windows installer asks if you want to install and declares the corticate name as being valid.

So where is the risk?

- Software can be changed, after installation, by alien code.
- The compiler must be totally trustworthy – usually they are from established companies.
  - Unless they deliberately introduce back doors as in cryptographic systems.

WHAT'S ON YOUR HARD DRIVE?

Typically 1 trillion or more bytes

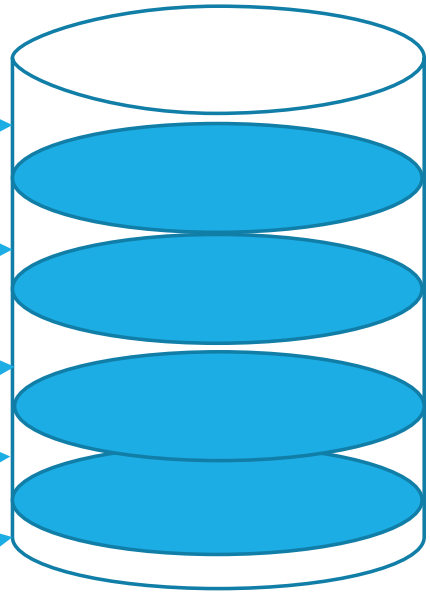
Temporary files,
cookies etc, not readily visible.

Data files (created by
programs, or downloaded)

Programs (Excel,
Word, Edge)

OS (windows)

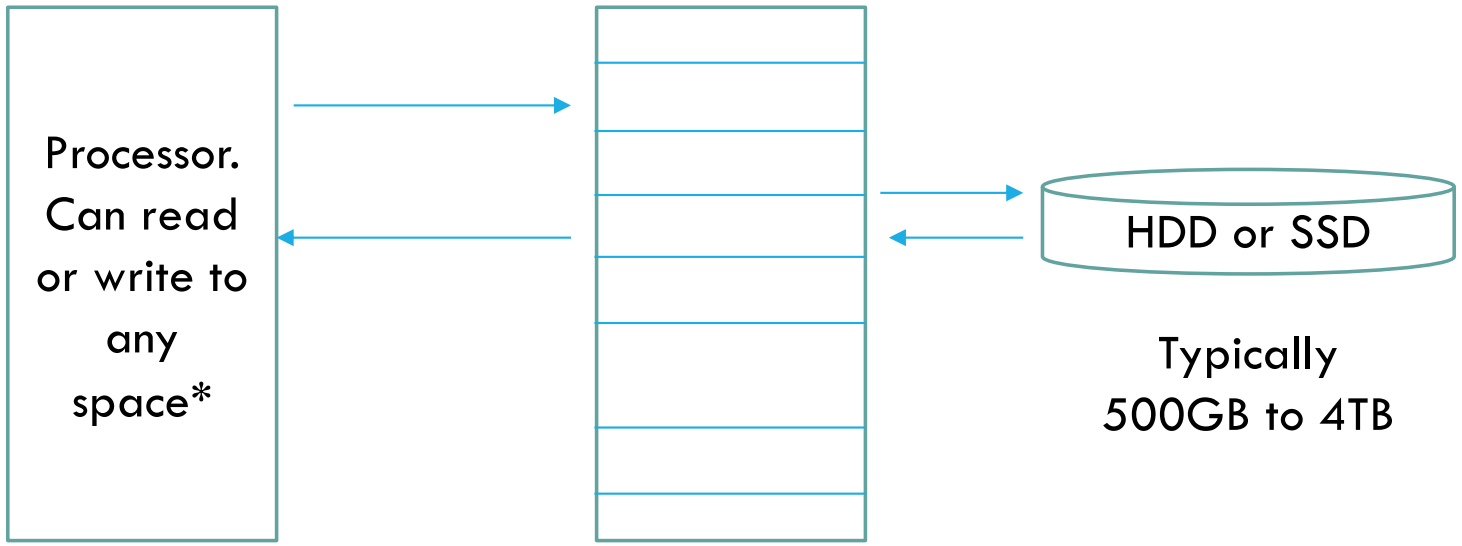
File allocation table



All files structures look the same.

DATA, APPS, OS ALL SHARE THE SAME SPACE

All software shares the same memory space.



Typically
500GB to 4TB

Typically
1GB to 8GB

Semiconductor memory
"DRAM"

*Windows hides the security file (SAM) but the processor can still write to it or read it.

OS,

The OS is the operating system. It contains a “kernel”, software that manages which apps or user programs are running, it manages reads and writes to the HDD, interactions with the Internet, which files are in memory and where they are located et.

But today’s OS do more, much of which we do not desire. They send our data to the “cloud” (i.e to them) they watch what we do, they can see all of our secrets. They offer many services we don’t need:

Games and offers for fees

Quasi secure mode which is really a closed group.

The OS schedules background tasks such as checking for program updates or scheduled events, usually without us being aware.

They are suspected of sending our crypto passwords to government agencies.

APPS, BROWSERS ETC

Many apps are highly desirable and serve us:

- Excel, Word, Photoshop, Zoom etc. (though Zoom is strictly a web app).

A browser is the key to all network activity. It enables web pages to be accessed and viewed and interacts with them.

Most websites require you to enable tracking cookies - small files that identify you to the web page source and its advertisers.

Even benign looking websites will not deliver if you block cookies or advertisements, even if you do not see them.

- It's all about follow the money: there is no free lunch.

SOFTWARE IS INTRODUCED AS A DATA FILE

The challenge for the malicious creator is to get the processor to start executing within their data file.

Often done through a "leaky" stack which can overflow

- The program is fed data which causes it to expand its "stack" (list of temporary registers) beyond the space allocated. One of these items of data is identical to a jump instruction and when the program reaches this instruction, which has overwritten the genuine instruction, it then directs the processor to start executing malicious code.
- It can then read your passwords, corrupt files, delete files or encrypt files as in ransomware.

It is up to the programmer to manage the expansion of the stack through other checks. - Often not done.

THIS IS WHAT FILES LOOK LIKE…

A DATA FILE IS INDISTINGUISHABLE FROM PROGRAM CODE

Which is code and which is data?

01110001	01100001
11001001	00011001
11001001	01110001
01001001	00101001
11001110	11111110
10010101	11010101
01100101	01011101
01001011	01101011
00010101	11010101
10101010	10110001
00001101	10010110
00100101	10101010

This is the key to all malicious software. The computer does not know. The two are inextricably combined in common memory.

Malicious code is introduced into memory as data and then the processor is directed to execute it.

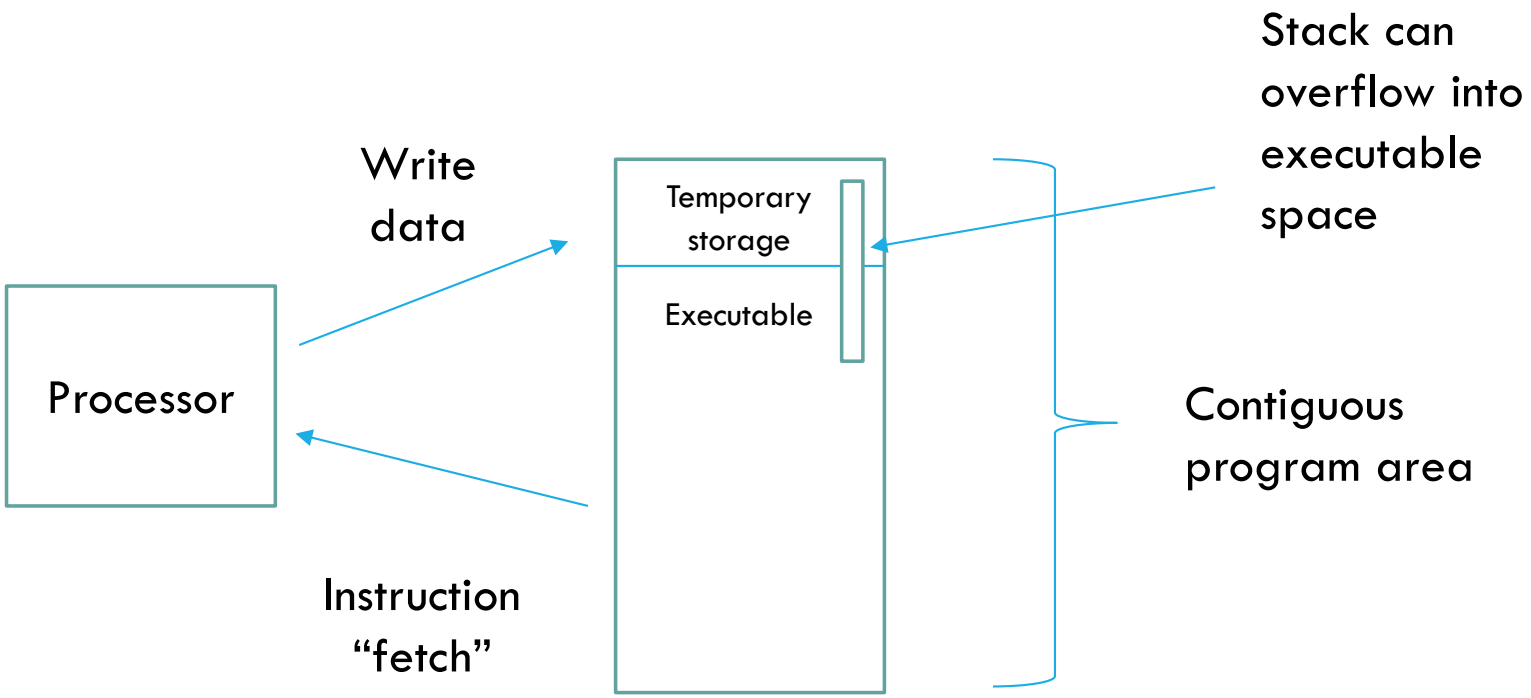
ASSEMBLER CODE – AS USUALLY WRITTEN

Raw binary code		address	binary code	instruction	comment
		0205		check_keyboard:	; name of function
10110100	00000001	0205	B4 01	mov ah,status	; check buffer status
11001101	00010110	0207	CD 16	int keybd	
01110100	00100111	0209	74 27	jz check_win	; skip checks if buffer empty
		020B	B4 00	mov ah,keyread	; otherwise get character
		020D	CD 16	int keybd	
		020F	80 FC 01	cmp ah,1	; check for escape scan code
		0212	75 08	jne sp_chk	; no, skip over escape handler
		0214	C6 06 000B R 00	mov hit_key_flag,0	; restart with "hit Key" message
		0219	E9 000C R	jmp initialize	; restart
		021C	3C 20	sp_chk: cmp al,' '	; test for ASCII space
		021E	75 07	jne brk_ck	; not space keep checking
		0220	B4 00	mov ah,keyread	; get any key
		0222	CD 16	int keybd	
		0224	EB 0C 90	jmp check_win	; and continue on
		0227	0A C4	brk_ck: or al,ah	; test for AL=0,AH=0, Ctrl-Break
		0229	75 07	jne check_win	; no, skip keyboard checking
		022B	B0 03	abort: mov al,3	; yes: set 80x25 b/w
		022D	B4 00	mov ah,setmode	
		022F	CD 10	int video	
		0231	CB	ret	; and return to OS

EXAMPLE PROGRAM CODE AS WRITTEN

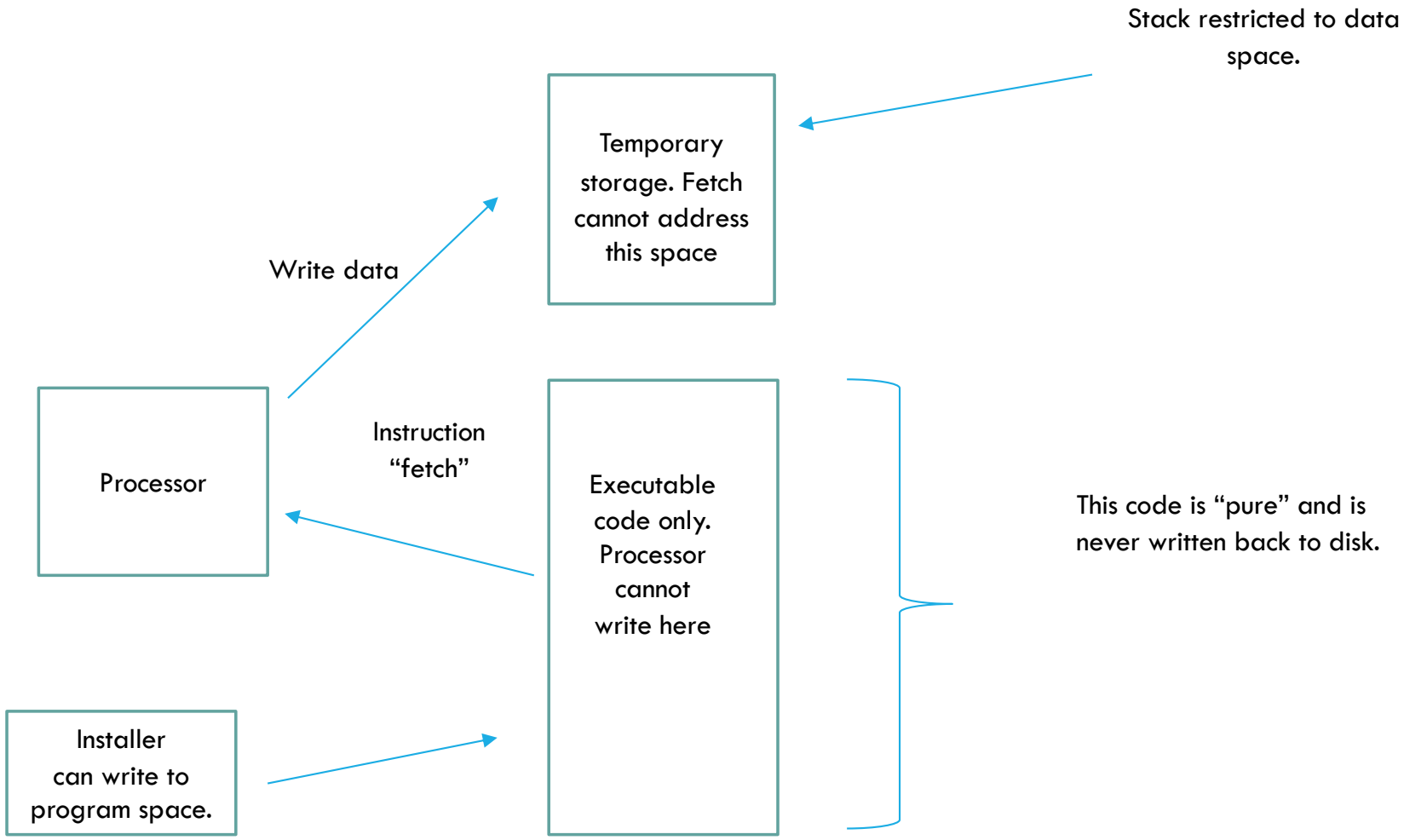
```
include <iostream>
using namespace std;
int main()
{
    int firstNumber, secondNumber, sumOfTwoNumbers;
    cout << "Enter two integers: ";
    cin >> firstNumber >> secondNumber;
    // sum of two numbers in stored in variable sumOfTwoNumbers
    sumOfTwoNumbers = firstNumber + secondNumber;
    // Prints sum
    cout << firstNumber << " + " <<  secondNumber << " = " <<
sumOfTwoNumbers;
    return 0;
}
```

PROGRAMS MIX EXECUTABLE AND DATA



There is no good reason for this architecture.
It is the biggest source of vulnerability.

NEW DESIGN OF SAFE COMPUTER



THE PROCESSOR WILL START TO EXECUTE FROM WHEREVER IT IS POINTED

Generally Windows will recognize any file with the extension ".exe" as executable and will start to execute.

Any file can be renamed to have the exe extension

Programs access other files, typically ".dll" ("dynamic link library") which comprise executable code

Execute means it will interpret the first byte as an instruction to do something.

- That can mean read from any other file, corrupt data structures or any other nefarious action.

HOW DO BAD GUYS INTRODUCE INSECURE CODE.

They get you or your browser to download it.

All browsers can introduce executable code into the computer.

- This needs to stop.

Windows requires permission to install code but there is no check on the safety of that code. It is up to the creator to make it safe (and it often isn't).

HTTPS and the "lock" symbol do not guarantee that the web site is trusted.

- It only means that the connection to it is secure.

WHAT IS THE INTERNET?

The Internet is not the same as the World Wide Web, though the terms are used somewhat too interchangeably.

The Internet is the packet switched network that routes all telephone and data traffic throughout the world.

- All voice calls today use Internet Protocol once beyond the first switch.

That is all it does. It takes packet (explained in a moment) data from one node to another.

From a security perspective it is very insecure.

- Any node can submit a packet and it will arrive at its destination, assuming the address is valid.
- It does not have to originate at an authenticated or even identified origin. It is like the mail service, anyone can post to anyone with no repercussions.

THE INTERNET USES "PACKET" SWITCHING

Blocks of data add the source and destination address to the block and then submit it to the network.

A hierarchical routing architecture routes the packet through a matrix of routers to arrive at the destination.

Each packet is independent though they are often identified as being part of a stream. They sometimes get lost due to failure to deliver within their allocated "time to live" parameter, or even "random discard" due to overload at a node.

Internet provides no guarantee of delivery. TCP which is an Internet overlay protocol provides this by "acknowledge and resend".

SOME INTERNET MYTHS…

The internet is not broken but it is being used outside of its original aims.

It was designed as a file transfer network between knowledgeable users.

It does not support real time services, authenticity or any network layer security.

Resilience is poorer than the telephone network it replaced.

- The telephone system was so good it could tolerate a fibre cut and users would not notice:
  - Switchover in 200 msec or less.

Internet uses a hierarchical routing system but routing tables have to be created for every node.

DARPA did not create the Internet for resilience.

INTERNET PROTOCOL

In order to be connected to the internet, you need an Internet Protocol address. This isa 32 bit (IPv4) or 128 bit (IPv6) number. In Ipv4 it is written as a sequence of four 8 bit, (0 to 255) numbers. You can display your IP address by typing ipconfig in a command window.

It will almost always be something like: 192.168.1.2.

This in fact your IP address within your home. It is of little interest to the bad guys. They need your external IP address. It is translated by your home modem to the IP address interface you externally to the Internet.

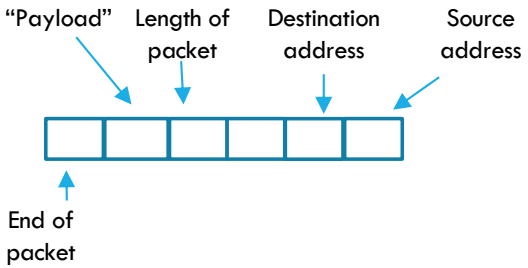
An external IP address is something like: 70.48.47.42

Your Ipv4 address is not permanent, it is renewed from time to time, invisibly to you.

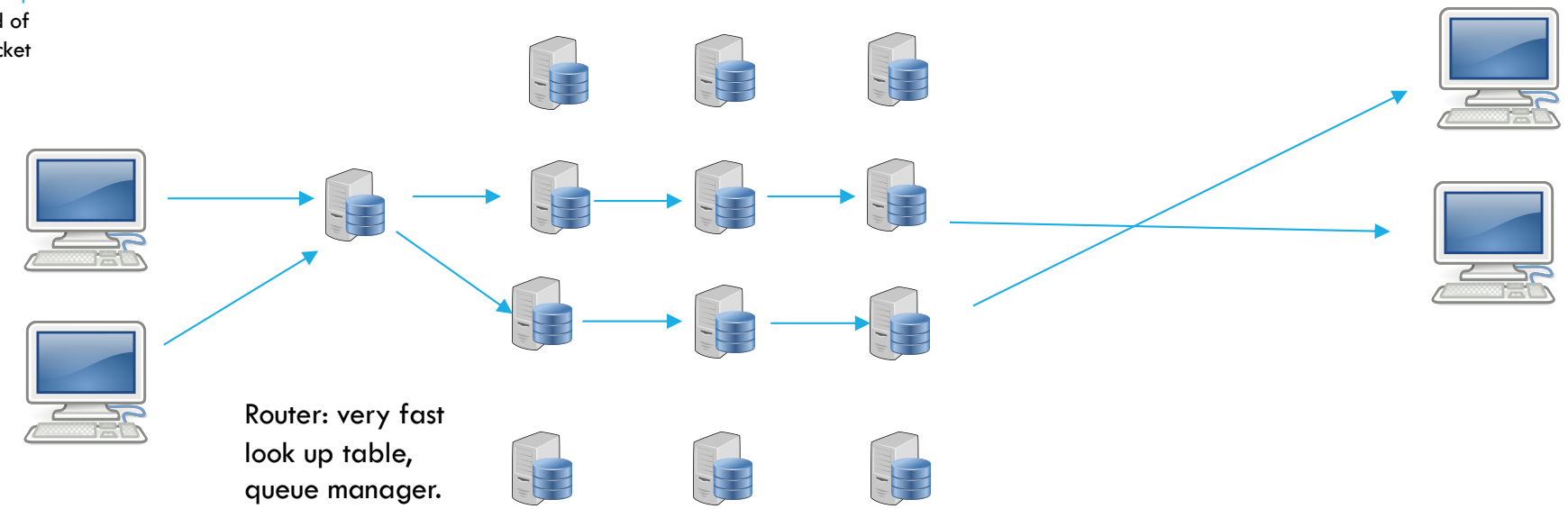
You might have more than one if you have a TV service.

INTERNET OVERVIEW.

VERY MUCH SIMPLIFIED



Packet can be up to 64K bytes, but is usually up to 1492 (8 less than 1500).



Router: very fast look up table, queue manager.

Internet is a "best effort" collision system.

WHAT IS THE WORLD WIDE WEB?

The world wide web made the Internet useable for all.

Collision based packet switching forced it to be a “store and forward system”, which turned out to be the key to useability.

You don’t have to be connected to the other end in real time.

Instead of needing to know the IP address of the recipient, which is a 32 bit number (IPv4) or 128 bit number (IPv6) all you need is the url (universal resource locator)

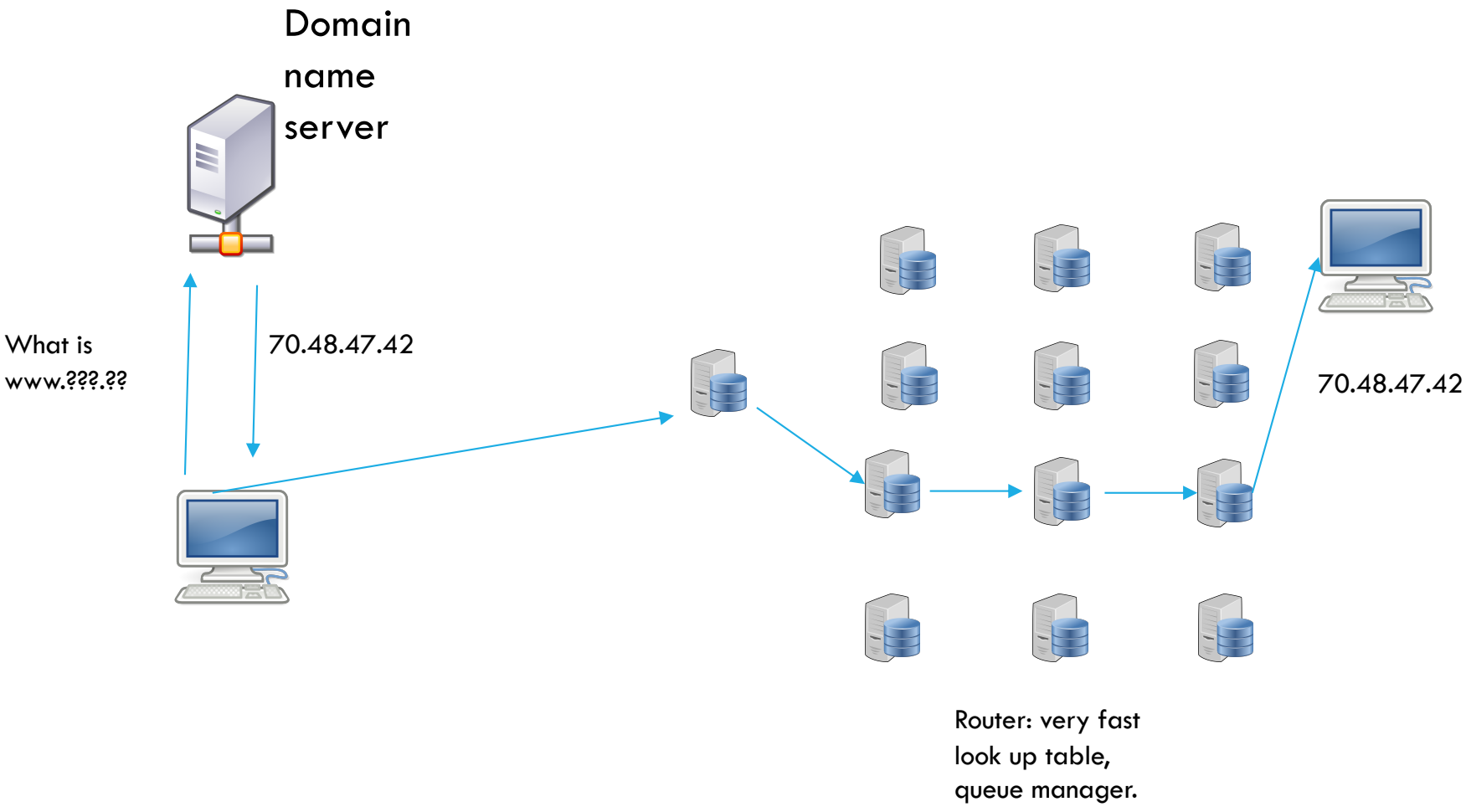
- IPv4 address looks like (as usually written) 192.168.1.1 or 23.41.209.179:
- url looks like www.anyname.ca

The world wide web is a system of Domain name Servers (DNS), operated by the internet Service provider hierarchy. Any url typed into the computer or, clicked on in a browser, sends a request to the DNS server to look up the IP address., which the DNS server returns.

You can get the IP address for any url by opening a command window and typing ping www.anyname.ca (this url is real btw)

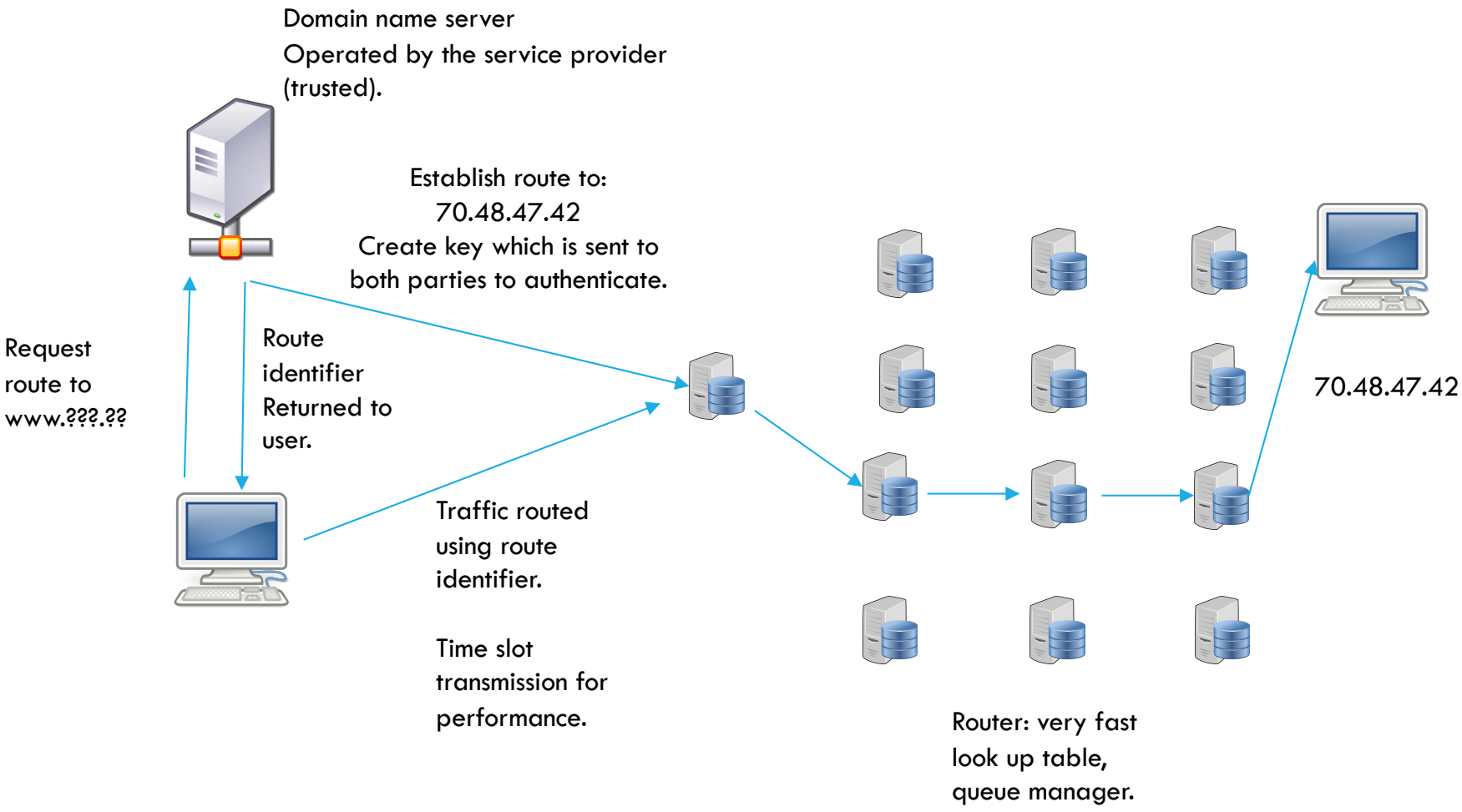
WWW OVERVIEW.

VERY MUCH SIMPLIFIED



NEXT GENERATION INTERNET/WWW OVERVIEW.

VERY MUCH SIMPLIFIED



DNS POISONING

Bad guys try to access the DNS server so that a legitimate look up redirects. They can redirect to any IP address anywhere in the world.

It takes time to trace through the ISP IP address.

RBC was targeted in this way and lost enormous amounts of money. Web site was identical to theirs and provided https response.

- However, it was not RBC. They captured many passwords and emptied accounts.
- There was no way of telling it was not them
- The malicious IP address was ultimately traced – but they had long gone.

DNS servers need to be very well protected!

The trust authority is the Internet Service Provider.

ANONYMOUS OR AUTHENTICATED TRANSFER

Communicating with your bank, you need solid authentication.

Browsing for a product or price comparing, you want anonymity.

Next Generation Internet provices both of these at the network layer.

Anonymity is generally not allowed by web sites as it interferes with the revenue stream.

What are “cookies”?

Small files identified by a browser and related to the accessed web site, that provide information tot eh web site:

- History of browsing
- History of purchases
- Login identity

Acceptance of a cookie policy is required by law in Europe.

THE INTERNET IS JUST A DELIVERY SERVICE WITH NO CONTROLS.

Anyone can submit any IP packet to anyone. Data communicate can use either an acknowledged link called transmission control protocol (TCP) or unacknowledged. Strictly TCP is not part of IP it is an "add on" to confirm error free delivery. Its use is widespread in all data communications.

Unacknowledged is used for applications such as voice telephone, streaming of TV or music etc. where resent packets would arrive too late to be useful. Streaming services use a return path for user control but not for packet stream resilience.

Note though that cellular voice does not use IP over the cellular link phone to base station, nor does the wireline phone to the first "Bell" switch.

WHERE DO YOU GET YOUR IP ADDRESS?

Your Internet Service provider (Bell or Rogers or whoever). It is not permanent and will be renewed from time to time.

- This is invisible to the user.

Initially IP addresses were only provided for the session. Then they went to a 72 hour lease. Today they are essentially permanent (typical duration is several months)

Why does this matter?

- Fishing sites match IP address to your profile (age gender, lifestyle, browsing and purchase history) then sell this to advertisers.
- Your household partner will see targets based on your profile.
  - Embarrassing for surprises etc.

ANONYMOUS HOSTING.

Some nefarious operators have set up secret hosts on services using wifi services:

- A laptop in a parking lot hosting illegal material.
- Public spaces, shopping malls airports etc.
- Wifi can be listened to by anyone, so emails at airports or shopng malls is risky
  - This does not apply to email over the cellular networks.

Wifi operators have been able to control this to some extent.

Every computer network interface adapter has a unique hardware machine identifier known as a MAC address. Wifi can be set to only accept links from known devices identified by this address.

Next Gen version restricts hosting to registered sites.

ISSUES – FOLLOW THE MONEY…

The Internet is now built on “Follow the money’.

Web sites will not allow anonymous browsing because they cannot then sell targeted advertising data.

User will be able to know they are agreeing to targeted advertising based on their shared (with other household members), IP address.

Security as consumers need it conflicts with the money tree.

It is now a business/political issue, nto a technical issue.

DO WE WANT SAFETY OR TOTAL TRANSPARENCY?

The Internet Engineering Task Force, a voluntary industry body, fought vigorously against any network introduced restrictions.

They worried third world governments would use it to censor news.

The challenge is not technical – it is social, political, business, legal.