Unit XXVIII Assignment I

By Nathan Windisch

PI: Web Architecture and Components

Internet Service Provider

An Internet Service Provider is a company that provides access to the internet to their customers. This is done by selling equipment to their customers. This equipment can be hardware such as routers and cabling, or software such as localhost webserver access to control the hardware remotely. Internet Service Providers have many different types of customers, including private citizens, small businesses and large corporations. Because all of these different groups have different needs, multiple packages must be issued. For instance, a private individual will not need the same internet speeds and bandwidth as a datacenter, meaning that they will not expect to pay the same price. As a result, the structure of the hardware and software that the Internet Service Provider leased must be flexible to accommodate all these different needs.

Web Hosting Services

A Web Hosting Service is a company that sells access to web servers in order to allow their clients access to publish websites on the World Wide Web. Web Hosting Services traditionally only sell access to small parts of dedicated servers, along with access to a web panel for file access and updating contact details and the like. Most Web Hosts will also give access to FTP, or File Transfer Protocol, as external programs are normally better than in-house, online file transference tools due to their extended capabilities.

Domain Structure

A Domain Structure is mainly constituted of Top-Level Domains, or TLDs, which are used to allow the user access to the website. Common TLDs are `.com`, `.net` and `.org`, but other domain level types exist. An example of this are country-code top level domains, such as `.uk`, `.de`, `.fr` and `.us`. There is also Second-Level domains, which are normally used to designate a region in a country. The main use of this is official state sites in America. The default CCLD in the United States is `.us`, but if the state of Ohio requires official websites then they will use `<website>.oh.us`, to prove it's legitimacy.

Domain Name Registrar

A Domain Name Registrar is a company that sells access to domain names. Domain Name Registrars are allowed to request that new domain names are made and, for the right price, can have exclusive access to them. An example of this is when Google, the search engine and software giant, purchased the rights to the `.app` domain for $25M. Other custom domain names which have been requested and subsequently purchased by companies, rather than ICANN generating them and then giving everyone access, include: `.axa`, `.cisco`, `.htc` and `.windows`. All of these domains can only be used by the companies that own them, or any persons or businesses that are allowed access.

World Wide Web

The World Wide Web is a series of internet servers that allow users to access different documents and files. Users may navigate the World Wide Web via hyperlinks. The World Wide Web should not be confused with the **Internet**. The World Wide Web is the framework or structure of the Internet. The World Wide Web is only one part of the Internet, and other protocols such as `FTP` and `SMTP` use the Internet but are not classified as the World Wide Web.

Internet

The Web, or the Internet, is different from the World Wide Web in that it is a "network of networks". The Internet is formed of many protocols, including `WWW`, `FTP`, `SSH`, `SMTP` and more. The Internet is a global network of servers and systems that allow users to access information from many of the aforementioned protocols.

Email

Email is a main method of communication via electronic means. Messages sent via Email can have more than one recipient. Many different providers allow their users to access their emails in browser, such as Outlook, Gmail and Yahoo. Companies such as Microsoft also provide desktop applications for their users to use, such as Microsoft Outlook. Finally, some people write their own email clients using the specific protocols provided such as POP3, IMAP and SMTP.

Proxy Server

A proxy server is a server that acts as a "middleman" to some networks. Proxy servers are used to relinquish strain on infrastructure via spreading the load if lots of users are using it, but they can also be used to hide the user's location due to privacy concerns, the reason normally being that they either do not want the website to track them or because the website that they are viewing is illegal in their country. Proxy servers can be used by servers, as aforementioned, to prevent the user accessing the site's server location as they do not connect to the server itself, they are routed through a proxy. This means that home webservers are secure from DDoS attacks, along with doxing.

Browser

A browser, or web browser, is a GUI, or graphical user interface, which parses HTML files into their appropriate output, meaning that users can view them when accessing the World Wide Web. Web Browsers can also do other things such as access other protocols like FTP, run code injected in HTML pages such as Javascript and PHP, and can also be used for sending emails via the `webmail` Protocol.

TCP/IP

TCP/IP is a group of rules that are the basis for communication protocols on the Internet. TCP/IP stands for Transmission Control Protocol/Internet Protocol. TCP/IP can also be considered for use in a communication protocol in a private network, such as intranet or extranet. TCP/IP is useful as it enables two different hosts to exchange data streams via a connection which has been established. TCP/IP also gives users the guarantee of data delivery in the same order that it was sent. TCP/IP is better than IP as it allows for data streams that are not based around packets, whereas IP is based solely around the transference of packets.

Application Layer

The Application Layer is the seventh layer in the Open Systems Interconnection (OSI) System. It contains process-to-process protocols and well built interfaces for communication across an IP network. It can also be used for end-user services. It is named as such because it is the final layer of the OSI system and is the layer that actually interfaces with the user, ergo it is called the Application Layer.

Flow Diagram: Definition

A flow diagram, or flow chart, is a method of displaying the sequence of actions of items within a system. This can include the movement or communication between people and also the dynamic between smaller machines. It is called a Flow Diagram because it shows the transition between steps, similar to how a waterfall transfers water from one place to another.

Flow Diagram: Example

The following is a flow diagram of how the the data flow within networks works. It is also a good example of a flow diagram.

PII: Client & Server Side Features

Client Side Factors

Browser

A browser, or web browser, is a GUI, or graphical user interface, which parses HTML files into their appropriate output, meaning that users can view them when accessing the World Wide Web. Web Browsers can also do other things such as access other protocols like FTP, run code injected in HTML pages such as Javascript and PHP, and can also be used for sending emails via the `webmail` Protocol. There are many different browsers, but the most widely used browsers are Google Chrome, Firefox, Microsoft Edge, Opera and Internet Explorer. Many web browsers, such as Google Chrome, Firefox and Opera, are open-source whereas Microsoft Edge and Internet Explorer are close-source. Chrome is so open-source, Opera is actually based off it.

Cache

Cache, also known as Memory Cache; Cache Store or RAM Cache, is a method of storing data which will be used over and over again. Cache is stored in High-Speed Static Random Access Memory, or SRAM, as it is is much faster but more expensive. This means that Cache is relatively small and is used for programs which will use the same data again and again.

Memory

Memory is a large definition which covers many different components of a computer. One type of Memory is Random Access Memory, or RAM, which is volatile memory which is wiped when power is disconnected. Another type of Memory is Read-Only Memory, or ROM, which is data which is written onto the chip and cannot be changed. Unlike RAM, ROM is non-volatile and cannot be changed.

Processor Speed

Processor Speed is an important component in Client Side Systems as processors with higher speeds can run more programs. Now a days many programs have a minimum requirements in order to actually run, and this number is measured in megahertz and gigahertz, or MHz and GHz.

Server Side Factors

Web Server Capability

Web Server Capability is the amount of hardware that the physical server has. A Web Server is a computer system that is connected to the internet and contains software that allows it to deliver web pages. A few examples of the software that are used are Apache and IIS, and they allow access to the web pages via the HTTP protocol.

Available Bandwidth

Bandwidth is the amount of data that can be sent over a specific static amount of time. Bandwidth is normally expressed via bits/second or bps, but it can also sometimes be expressed in cycles/second, or Hertz if the device is analog. If the Bandwidth on a server is higher, then the web page will be able to be delivered to more people at a faster rate, meaning that more customers will be serviced. Ergo, higher bandwidth may cost more but will mean for that more "hits" are generated.

Number Of Hits

A common misconception of Web Hits is that each hit is a visit, but this is not the case. Web Hits are the amount of files downloaded on your website. This includes images, graphics, videos and audio. If a website has many buttons and each one is an image, then they will get more hits. On average, a good website will get about 15 hits per view.

File Types

File Types, or File Formats, are a standardized attempt to store information and allow it to be read by specific programs. For example, `.exe` files can only be run on Windows systems, `.app` files can only be run on OSX systems and `.sh` files can only be run on Linux systems. This is because all Operating Systems are different under the hood, and each one performs actions differently. In the example of website development, `.html` files are Hyper Text Markup Language files and are used to display web pages on the internet. HTML files can contain HTML, JavaScript, PHP and others natively, but they can also import code from other files such as `.js` and `.php` files. HTML files can also use Cascading Style Sheets, or CSS, files to stylise the website. Breaking up your website and storing commonly reused code means that it only has to be written once and it can then be imported again and again.

PIII: Security Risks, Protection Mechanisms and Data Protection Act

Security Risks

Hacking

Hacking is when an external force accesses or modifies software, hardware or data, either to achieve their own goals or to disrupt the system. Hacking is considered a crime in most parts of the world, and has punishments similar to breaking & entering, for accessing a system when unauthorized, all the way up to theft, if data is cloned without permission.

Virus

Viruses in computing are malicious pieces of code that infect the system when executed, either by editing files, replicating itself or by opening or closing applications. Sophisticated viruses can not only infect the current operating but can also access the boot sector of the hard drive, effectively making it useless.

Identity Theft

Identity Theft is the acquiring of personal information and using this private information for personal gain. The method that the data is normally used for is for financial gain. This data is now normally gained by the internet, either via keyloggers accessing bank passwords or Trojan viruses installing a backdoor into systems to allow the attackers to gain access to the system.

Protection Mechanisms

Firewall

A firewall is a vital part of all computer systems. It denies access to the computer or network if the data is unauthorized or defined as dangerous. Most firewalls only block incoming information, where as outgoing data is free to travel. Firewalls can do many things, including:

- only allow access from certain IPs and ports, effectively a whitelist,
- denying access based on IP address or port, effectively a blacklist,
- allowing access to all but placing all internal files within a quarantine for later viewing and decision making, effectively a greylist.

Security Sockets Layer (SSL)

Security Sockets Layer, or SSL, is a level of security that is used for generating a link between a client and a server. The link is encrypted and is used typically between a client and either a web server or a mail client, so when a user either accesses a website or their emails.

Strong Passwords

Strong Passwords are much harder to brute force or guess when compared to a weak password. Weak passwords are normally all one case, contain single English words and may contain numbers at the end. A strong password should contain at least 15 characters, a mix of upper and lower case, symbols and numbers and should not be either an English word or in any relation to your personal information. The following is an example of a weak password from a person called Dan:

- `d4n1999`
  This is a bad password as it contains his first name with only one character changed, it is all only one case and it contains his year of birth. The following is a better password:
- `D0P[w,jgRcA|c>z`
  This is a better password because it is a mix of uppercase and lowercase, it contains numbers and symbols, it contains no personal information, it is 15 characters long and it is not a word in any language. The one downside to having secure passwords is that it is very hard to remember all of them, but this can be solved with a secure password manager that has one strong password protecting the rest.

Data Protection Act

The Data Protection Act is a law created by the UK government in 1998 to protect people's information on the internet and offline. The following is a brief summary taken from https://www.gov.uk/data-protection/the-data-protection-act/

> The Data Protection Act controls how your personal information is used by organizations, businesses or the government.
>
> Everyone responsible for using data has to follow strict rules called 'data protection principles'. They must make sure the information is:
>
> - used fairly and lawfully
> - used for limited, specifically stated purposes
> - used in a way that is adequate, relevant and not excessive
> - accurate
> - kept for no longer than is absolutely necessary
> - handled according to people's data protection rights
> - kept safe and secure
> - not transferred outside the European Economic Area without adequate protection
>
> There is stronger legal protection for more sensitive information, such as:
>
> - ethnic background
> - political opinions
> - religious beliefs
> - health
> - sexual health
> - criminal records

MI: Web 2.0, Blogs, Cloud Computing and Online Applications

Web 2.0

Web 1.0 was what was considered the "static web". All the information only changed when the website was updated, and nothing was reactive. Website were more like books, in that they could only be read and nothing could be done with them other than that. This was good for the time, as web pages didn't really need to do much at the time. The World Wide Web was not as easily accessible or widely used as it is now, so nothing extra was needed. After a while the Internet grew and more people needed to use the service for different things. Because of this, new programming languages such as Javascript and PHP were developed to repond to the user's inputs. These webpages are "reactive" web pages and are considered a standard in Web 2.0. Some popular examples of Web 2.0 pages are social media websites such as `Facebook`, `Twitter` and `Reddit`, but Web 2.0 also covers any website with a forum or commenting system, or just any general way for users to upload their content. This means that websites such as `BBC News`, `YouTube` and `Wikipedia` are all technically Web 2.0, even if most would not consider any of them, especially the latter, to have any sort of user interactivity.

Blogs

Blogs are a way for people to communicate with the world. They allow people to spread their thoughts and opinions with one another and just let people get things off their chest. Blogs are normally written in an informal format and normally contain articles that newspapers refer to "opinion pieces", meaning that they may be biased in some way. There are many different types of blogs that can be found on the World Wide Web, but the most common are those that are owned by private individuals. These blogs can talk about anything at all, but they usually stick to one subject so they they can pander to their target demographic. For instance, a beauty and makeup blog would not be expected to write articles on quantum mathematics. Because of this, many people turn to blogs for information that relates to their personal opinion. This generates what is called an "echo chamber", meaning that an audience only hears what they want to hear. This is a dangerous precident as it means that those who only get their information from blogs may be misinformed. Some people use blogs to share things that they have done. For example, the popular Internet artist XKCD uses blogs as a form of allowing people to share their work and promote their content. This shows that there are many different ways to use a blog, and the only real requirement is that the posts are in an informal format, and are only posted by a one person or a small group of individuals.

Cloud Computing

Cloud computing is the practice of generating, saving, storing and loading information from an external server. This allows companies to rent server space in order to do all of their computing and file saving without having to host the physical servers themselves. This means that they can save lots of money as they do not have to purchase the servers, install them, keep using up space to store them, cool them, repair them or upgrade them. All of the hard work of keeping the servers maintained is down to a subcontracted company, meaning that businesses can spend more money on R&D or expanding the company. Another thing that most hosting companies have is a Service Level Agreement, which means that the host agrees to having a specific uptime percentage, and if they do not then their customers either get some free store credit or money off their next payment. Good hosting compies that have a good infrastructre and a very high uptime normally set their Service Level Agreement to be between 98.00% and 99.99% uptime, depending on how confident they are. Service Level Agreements only apply to faults that are caused by the hosting company. Anything that the customer's do that causes their service to go down in their own fault.

Online Applications

Online Applications, also known as `SaaS` or `Software as a Service`, is a form of cloud computing that performs

all of the server based actions on the cloud, and allows the client to perform actions on the client that are then sent to the servers and performed. An upside of this is that client doesn't have to install anything, meaning that the service can be used on any device. Another upside is that sharing files is incredibly easy, as they are already saved to the cloud. A downside to SaaS is that it requires a constant Internet connection in order to edit the files. Another downside is that these services normally have a monthly fee rather than a one-time payment.

DI: TCP/IP

Role of TCP/IP

TCP/IP is a model that has four layers. It is used to standardize how the TCP/IP stack works. The OSI model is different from the TCP/IP model, and should not be confused with the TCP/IP model. I shall provide an example for each of the layers in the TCP/IP model:

Layer I: Network Access

The lowest layer of the TCP/IP stack is the network access layer. It handles individual pulses on the wire. It also words with wireless networking by being on the antennas. This layer only knows about the signals on the wire, which it interprets into 0s and 1s, and is therefore not aware of the packet structure of the higher up levels. Because it is so low and pimative, it is considered as a part of the kernel level driver package.

Layer II: Internet Layer

The second layer on the TCP/IP model is the internet layer. It works with the packets that are found within the TCP part of the system. It recieves them from the levels tha are higher up, forming data that can be sent over wires and cables. Within these packets is framentations of the data, along with the source and destination IP addresses. The packet is also set when in the Internet layer when it is on it's initial route, meaning that further nodes will be assisted by this.

Layer III: Transport Layer

The third layer within the TCP/IP stack is the transport layer, and this is liable for ensuring that the connection between the two nodes is kept constant and open or else the packets will be either not sent or lost when either transmitting or recieving the data, as the connection is closed during the transmission. This layer also performs cyclic checks to ensure that the other node is kept alive, by also sending a keepAlive signal.

Layer IV: Application Layer

The final layer within the TCP/IP model is the application layer. This final layer is in charge of transporting specific data such as `HTTP`, `HTTPS`, `FTP`, `SSH` and `TELNET`. As the application layer is the highest layer on the TCP/IP model, it is used as an interface with regular APIs. This is used by developers for programming.

How TCP/IP Links To The Application Layer

Within TCP/IP, the application layer deals with all communication protocols that are used within the various process-to-process programs on the system. Along with this, the application layer also contains the interfaces needed for the communication systems. This is used to standardize communication and ensure that all systems have the same protocols, in order to make communcation between the systems easy and quick. This means that

The following is a flow diagram of how the TCP/IP stack works.