Unit 28, Assignment 2 – How the Web Works

Thomas Robinson

SOLIHULL COLLEGE

Task One (P1)

Technology	Description	Examples
WWW	The World Wide Web is a collection of webpages that can be accessed over the Internet	
Web Browser	A Web Browser is a piece of software that allows one to view pages on the World Wide Web	Mozilla Firefox, Google Chrome, Microsoft Edge, Vivaldi, Brave
ISP	An Internet Service Provider facilitates a fixed-line connection to the Internet	Virgin Media, Sky, Vodafone, TalkTalk, Hyperoptic, Andrews & Arnold
Gateway	A device that facilitates the connection to devices on other networks	
Packet	An encapsulated piece of data that contains a header stating—primarily—the source and destination IP address of the data.	
Domain Name	A string of characters that can be used to identify an organisation or individual on the Internet	Google.com, Waitrose.com, Gov.uk, TfWM.org.uk
Domain Name Register	A company or organisations that allows the registration of a domain name with the domain's central registry	Porkbun, Namecheap
Web Hosting	The process of providing access to a page or service on the World Wide Web	Providers include Amazon Web Services, Google Cloud Platform, Vercel
Server	A specialized computer used for hosting applications and services	
HTTP	Hyper Text Transfer Protocol is the application-layer protocol used for transmitting web pages	
HTTPS	A version of HTTP that encrypts data while in-transit using certificates	
TCP IP	A common name for the IP Suite: a way of representing the various protocols used by the Internet	
SMTP	Simple Mail Transfer Protocol – used for sending emails	
FTP	File Transfer Protocol – used for sending files across a network or the Internet	
Database	A collection of data, usually arranged in rows and columns	
Cookies	Small pieces of information that can be saved by a specific webpage to enable interactivity and to remember user information on future visits	
Web Server	A server that is running an application designed to serve web pages.	Web server applications include

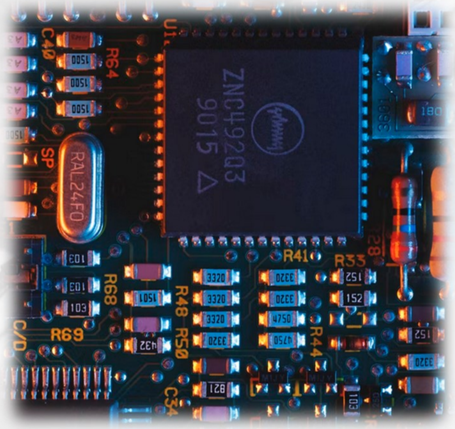
		Apache, NGINX and Microsoft IIS
Proxy server	A server that acts as an intermediary between a client and a server, often used to filter web traffic	
Programming requirements: PHP (Shopping cart), HTML, Java script etc	PHP: A server-side language used to incorporate ‘back-end’ functionality into a website, such as access to an internal database HTML: A markup language used to structure the content of a webpage JavaScript: A client-side language used to add interactivity and functionality to a webpage	
DNS (Software on the server)	A DNS server is used to resolve a domain name or a hostname (eg: google.com) to an IP address (eg: 216.58.212.142)	DNS server applications include DNSMAQ, Microsoft AD

Website Performance

Client-Side

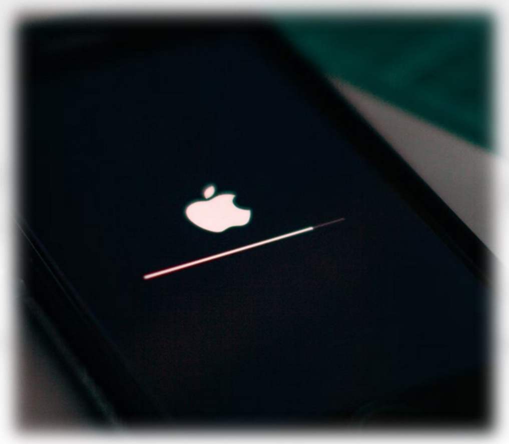
Internet Speed

The users’ internet connection speed can affect the loading time of large assets on the website since they cannot be downloaded as quickly



Hardware

A slow computer with weak hardware may struggle to render a website with a lot of heavy, interactive content



Software

An out-of-date web browser or operating system that is not compatible with modern standards can cause websites to perform sub-optimally.

Server-Side

Hardware & Processor Speed

An underperforming server may not be able to serve requests at the required rate. Factors that particularly effect this include storage access latency and processor speed.

Connection Speed

If the server is using a slower connection than the clients accessing the site, it will create a bottleneck. A website that wishes to support a large number of visitors will require a faster connection.

Programming

Poor server-side programming can result in a slower website and an overworked server, especially if there were to be a large volume of traffic

Excess Bandwidth

A large volume of traffic at one time can overwhelm the bandwidth capacity of the connection and cause a denial of service to users

Risks, Prevention & Laws

Unit 28, Assignment 2

Thomas Robinson - Solihull College - April 2022

Risks

VIRUSES

- Like a real-life virus, replicates itself in many places on a computer
- Often attaches to a vulnerable piece of software
- Can cause damage (acting as part of a botnet to perform a DDOS attack) for fun or for profit (ransomware that blackmails users/companies)
- On servers or on employee devices

TROJANS

- A trojan is a virus that is disguised as a legitimate program

ON-PAGE VIRUSES

- Viruses on a webpage will exploit vulnerabilities in a user’s web browser

Hacking & DDOS Attacks

Hacking:

- Refers to gaining unauthorised access to a computer system
- Is unlawful
- Often done for financial gain, or simply to show off one’s ability

A DDOS Attack:

- Aims to take down a website or render it inoperable (deny service).
- Overwhelms capabilities of webserver/network (can only handle so many req/s)
- May last extended periods of time – perpetrators may want ransom to stop
- Can simply be caused by too many people innocently visiting a website at once (hug of death)

Data Sniffing

- Sniffing refers to monitoring all data being sent and recieved by a device or over a network
- Used legitimatly to troubleshoot issues
- Used maliciously to read data such as passwords and other sensitive information

Identity Theft

- Obtaining enough information about a party to be able to impersonate them.
- **Examples**:
  - With passport + driving license could potentially open bank accounts
  - With social media account information could trick followers (who trust the account) into doing unlawful things

Prevention

ANTIVIRUS

- An antivirus solution can detect known malicious programs and prevent them from being executed

FIREWALLS

- A firewall can be configured to block suspicious traffic that could be attempting to gain access to parts of a network

Prevention

UPDATES

- Keeping devices up to date ensures the software is patched against known exploits and vulnerabilities

PASSWORDS & PERMISSIONS

- Applying the correct permissions to administrative areas on websites can prevent unauthorised Access
- Strong passwords are much harder to brute-force

Prevention

ENCRYPTION

- Encrypting data in-transit can prevent sniffed data from being read
- Ensures that the website content has not been modified before it arrives at a user’s browser

Laws

DATA PROTECTION ACT

- Governs handling of personal information
- Affects both data held in paper and digital forms
- Passed ’84, updated ’02, still current
- Upheld by ICO
- Healthcare Records
- Criminal Justice
- Financial Institutions
- Biological
- All businesses must register with ICO and state what information they hold and why.

The Data Protection Act governs the handling of personal information by organisations. It applies to both data held in physical and digital forms. It requires that all businesses register with the ICO (Information Commissioners Office) and state what information they hold and why.

There are eight principles of the data protection act – they are to make sure that data is:

- used fairly, lawfully, and transparently
- used for specific, explicit purposes
- used in a way that is adequate, relevant and limited to only what is necessary
- accurate and, when necessary, kept up to date
- kept for no longer than necessary
- handled in a way that ensures appropriate security, including protection against unlawful or unauthorised processing, access, loss, destruction or damage.

Laws

COPYRIGHT DESIGNS & PATENTS ACT

- The C, D & P Act protects intellectual property, such as brand names, stories and designs.

- There are different protections that can apply to works: copyright, patents, licensing, and trademarks.

- **Copyright**

  Copyright defines rights the creator has over their own work. It also applies to digital work, such as music, videos, code, etc. Does not need to be applied for – it is automatically assumed when you create a piece of work. Copyright makes it illegal to copy work without permission; the owner can take legal action.

- **Patents**

  Patents give the inventor the right to decide if/how an invention can be used by others.

- **Trademarks**

  Trademarks cover brand names, logos and other “service marks.”

Laws

EQUALITY ACT

- Protects those with disabilities and certain protected characteristics from discrimination and allows them to challenge any they do face
- Public-sector websites providing services are requried to publish an accesibility statement under the *Public Sector Bodies (Websites and Mobile Applications) Accessibility* Regulations

Requires that organisations:

- take reasonable steps to ensure equal experience and avoid the disadvantage for persons with disabilities in their practice,
- take reasonable steps to avoid the disadvantage for persons with disabilities against persons who are not disabled by fixing physical features that could cause such,
- take reasonable steps to provide an auxiliary aid if, without it, persons with disabilities would be at a disadvantage in comparison with persons who are not disabled.

Web Architecture & Communication

Unit 28, Assignment 2

Thomas Robinson

April 2022
Solihull College

Internet Service Providers

An internet service provider is a company that facilitates a connection and access to the Internet. On a subscription basis, you gain access to their network, sometimes through the use of their own hardware.

There are many ways they can provide access, including through a telephone dial-up connection, ADSL, Fibre, DOCISS, Mobile Broadband or Satellite.

Examples include Virgin Media, Sky, Hyperoptic, A&A among many others.

A Fibre ONT (Optical Network Termination), installed by Openreach in the UK.

Domain Structure

A domain is a string of characters that identify an individual or a company on the internet. The structure of a domain includes the TLD (example.com) as well as subdomains (www.example.com or mail.example.com). Each of these strings can point to a different website or service

A row of server racks in a Google Datacentre

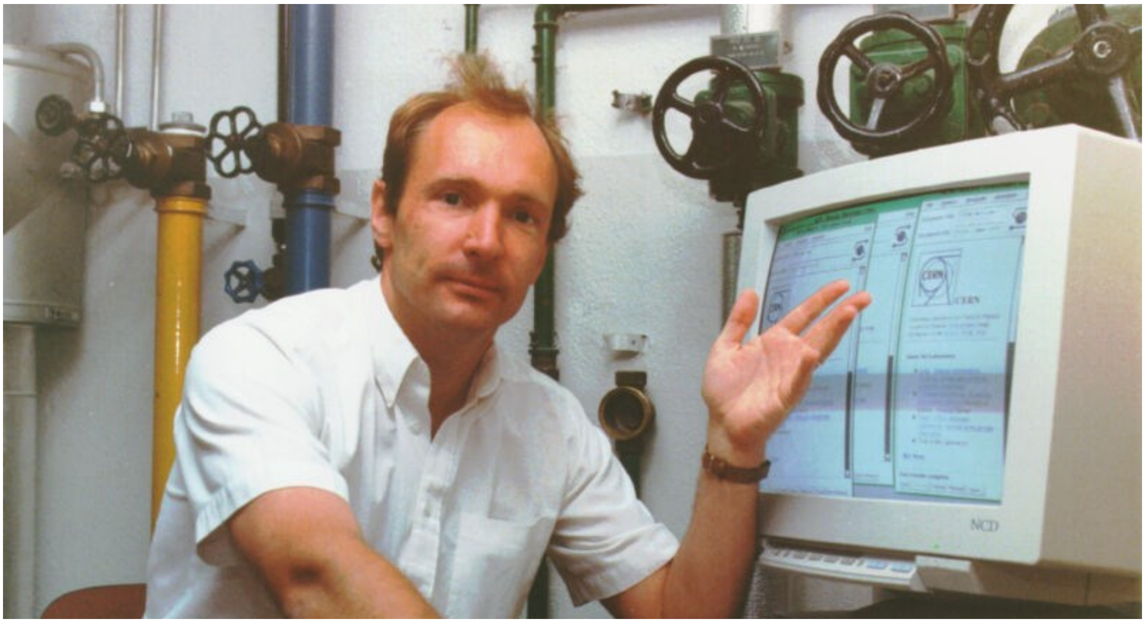
Domain Name Registrars

A domain name register is an organization from which one can purchase a domain name. This usually endeavors a yearly fee.

While domains are registered with individual registers, the extensions (such as .co.uk) are controlled by central registries (Nominet, in the case of .uk).

World-Wide Web

The world-wide web is a name for the collection of web pages that are available to access over the Internet.

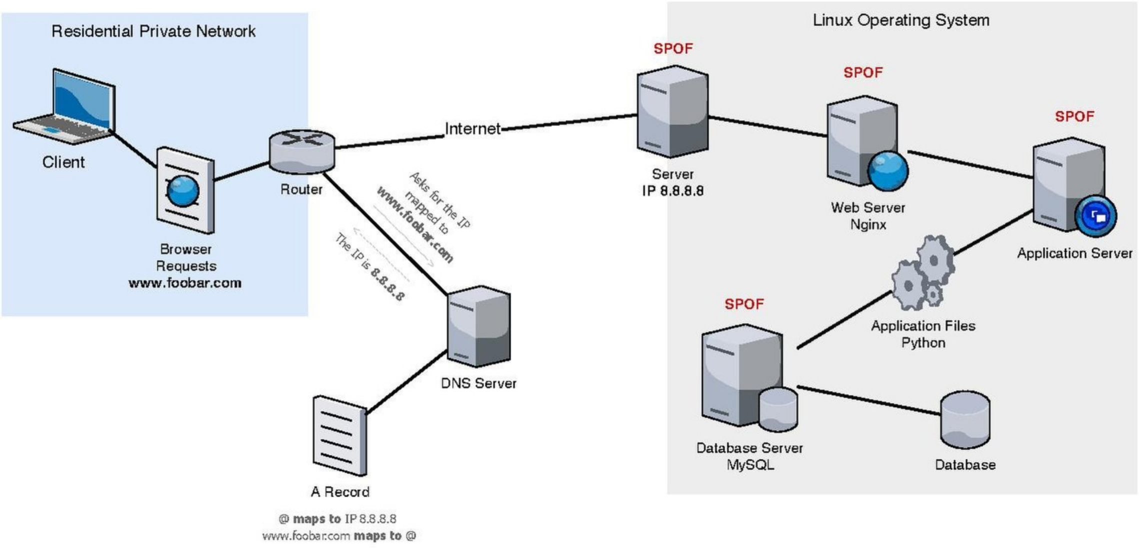


Tim Berners-Lee, the creator of the World-Wide Web at the European Organization for Nuclear Research

Web 2.0

Web 2.0 refers to a general change in the design and content of webpages. It signaled a shift to more interoperability through the use of APIs and websites more oriented to user-generated content.

Web Architecture Diagram

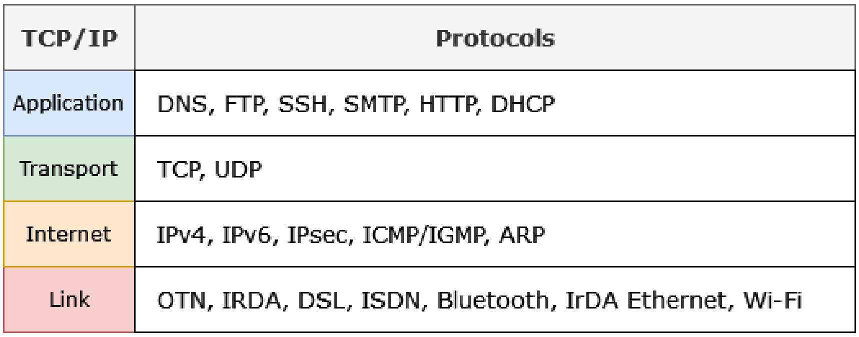


Depending on the type of website, many of the depictions on the right-hand-side of the diagram could be different.

For example:
- The server could be running Microsoft’s IIS software on a Windows Server operating system.
- A static website will not make use of a database and will instead serve unchanging pages with no ‘backend’ interaction.
- There are many that can be used server-side rather than Python, including PHP and Go.
- There are many other database server options besides MySQL such as Redis and PostgreSQL

TCP/IP

The Transmission Control Protocol/Internet Protocol stack is a collection of protocols that govern how data should be split into packets, routed, sent and received over the Internet.



TCP/IP	Protocols
Application	DNS, FTP, SSH, SMTP, HTTP, DHCP
Transport	TCP, UDP
Internet	IPv4, IPv6, IPsec, ICMP/IGMP, ARP
Link	OTN, IRDA, DSL, ISDN, Bluetooth, IrDA Ethernet, Wi-Fi

At the **application layer**, HTTP (Hyper-Text Transfer Protocol) is the protocol used to send and receive pages on the WWW. HTTPS is a secure variant which provides encryption for data in-transit.

DNS (Domain Name System) acts as the phone book of the Internet. It translates human-readable domain names (such as example.com) to routable Internet Protocol addresses (such as 93.184.216.34).

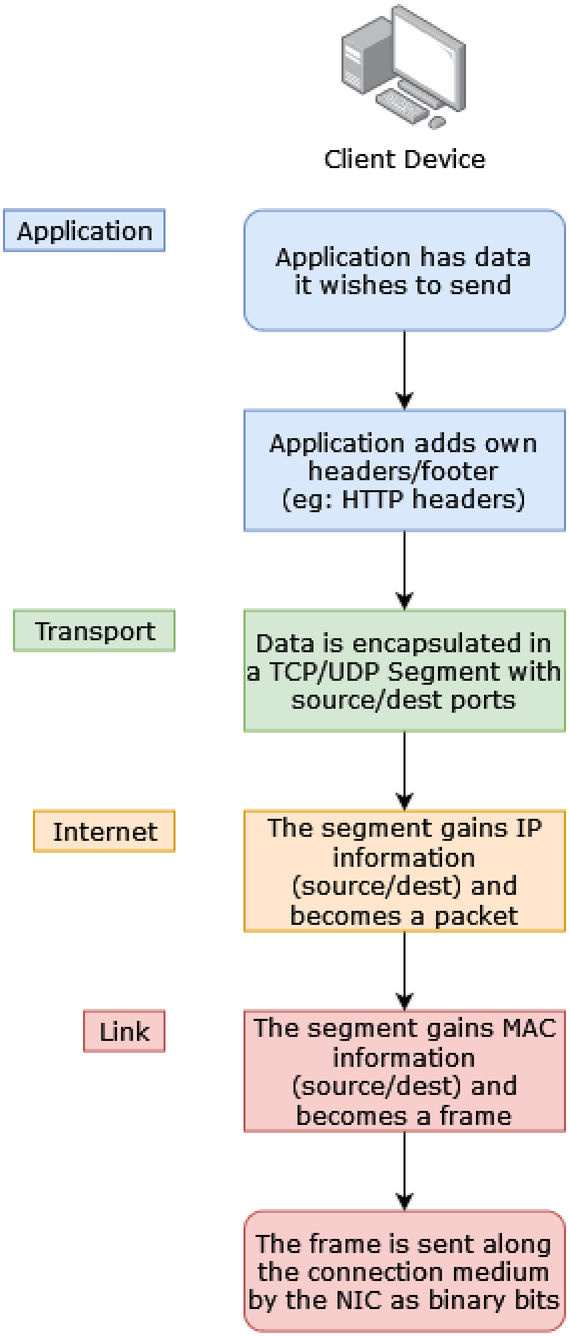
FTP (File Transfer Protocol) is an older standard used for transferring files between computers, though it is often still used today.

At the **transport layer**, data is split into segments. In the TCP/IP model, they will either be TCP or UDP in type. These will contain a source and destination port which will allow the data to be sent to the correct program on the computer. For example, port 443 is the port used by web servers wishing to receive HTTPS traffic.

At the **internet layer**, the segments are further encapsulated into packets. These will contain a source and destination IP address to allow the data to traverse the many networks between the source and the destination.

At the **link layer**, the packets are again encapsulated into a frame. At this layer, MAC addresses are used to send the packet to the correct device within the network of an ISP, organization or home.

TCP/IP Client Sending



Client Device

Application

Application has data it wishes to send

Application adds own headers/footer (eg: HTTP headers)

Transport

Data is encapsulated in a TCP/UDP Segment with source/dest ports

Internet

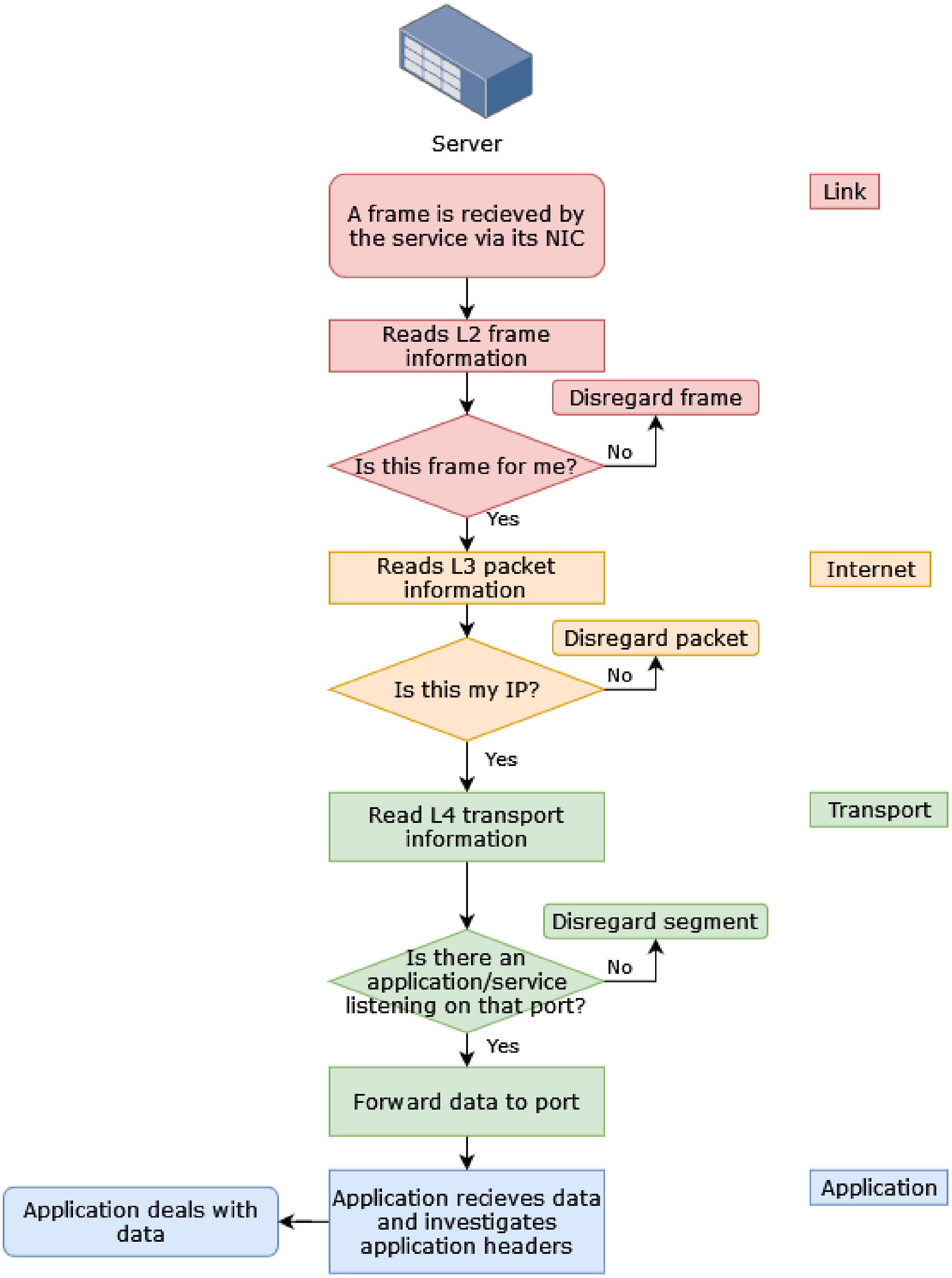
The segment gains IP information (source/dest) and becomes a packet

Link

The segment gains MAC information (source/dest) and becomes a frame

The frame is sent along the connection medium by the NIC as binary bits

TCP/IP Server Receiving


Server
A frame is recieved by the service via its NIC
Link
Reads L2 frame information
Disregard frame
No
Is this frame for me?
Yes
Reads L3 packet information
Internet
Disregard packet
No
Is this my IP?
Yes
Read L4 transport information
Transport
Disregard segment
No
Is there an application/service listening on that port?
Yes
Forward data to port
Application deals with data
Application recieves data and investigates application headers
Application