VIETNAM NATIONAL UNIVERSITY, HANOI
VNU UNIVERSITY OF SCIENCE
MATHEMATICS - INFORMATICS - MECHANICS

FINAL PROJECT

Interacting with the world through the Internet

Group :	**20**
Class :	**K66A4 Computer and Information Science**
Subject :	**English for Computer and Information Science**
Supervisor :	**Do Thanh Ha**

HA NOI - 2023

GROUP 20'S MEMBER

Full Name	Student ID	Major
Phan Thi Thu Trang	21001593	K66A4 Computer and Information Science
Pham Hoai Nam	21000394	K66A4 Computer and Information Science
Pham Hong Quan	21001575	K66A4 Computer and Information Science

INTRODUCTION

Nowadays, the Internet appears everywhere and distributes the number of utilities for citizens in every aspect of living. To master Computer Science and give the opportunity for future jobs, we must grasp general bits of knowledge. Comprehending the Internet is so necessary and extremely important for all developers.

The final project namely "Interacting with the world through the Internet" helps you raise awareness about how the Internet works, the structure of the Internet[1] and its applications, especially the World Wide Web. We not only provide professional theories but also apply them to generate a demo Website.

Our report includes 3 main parts:

1. The overview of Computer Networks
2. The Internet
3. The World Wide Web

Noticeably, we use the TCP/IP model instead of the OSI traditional model and apply some terms of each layer in each section. TCP/IP model marked a difference far from OSI Model and approached clearly our topic better than.

Our demo Website project that reflects some parts of how the World Wide Web works can be viewed by scanning the below QR or by clicking on the link: `https://trangphan10.github.io/Group20FinalProject/`

If you have any questions for group 20, please contact:

- Name: Phan Thi Thu Trang (Leader)
- Email: phanthithutrang_t66@hus.edu.vn
- Phone: 0967889129

Table Of Content

1 Overview of Computer Network

1.1 Definition

A computer network is a group of interconnected computers and devices that are able to communicate with each other and share resources such as data, files, printers, and applications

1.2 Outstanding features of Computer Network

Several technologies have been implemented to new problems as our computers became increasingly connected. Here are some of the key technologies that have played a critical role in the development of computer networks:

1. **Ethernet**
   Ethernet is a standard for connecting computers and other bias in a original area network(LAN). It was first developed by Xerox Corporation in the 1970s and has since become one of the most widely used LAN technologies.

2. **MAC address**
   MAC addresses are unique identifiers assigned to network devices at the factory. They are used to identify devices on a network and to ensure that data is sent to the correct device.

   But, as network traffic goes up, the probability that two computers will attempt to write data at the same time also increases. This is called a collision and we can solve this problem by using the Network switches.

3. **Network switches** [4]
   (a) **Clarification**
      - Network switches are devices used to connect multiple devices on a network.
      - They allow for more efficient data transmission and help to reduce network congestion.

   (b) **Switching**
      - **Circuit switching** : Circuit switching is a communication method where a dedicated physical circuit is established between two devices for the duration of their communication session, and the circuit remains dedicated to those two devices until the session is over

- **Message switching** : Message switching is a communication method where a message is forwarded from one device to another through a series of intermediate devices, called nodes, with each node temporarily storing the message until it can be forwarded to the next node until the message reaches its final destination.
- **Packet switching** : Packet switching is a communication met where data is broken up into small packets and sent individually over the network, with each packet containing the address of its destination and being routed independently through the network, and then reassembled into the original data at the final destination.

2 The Internet

2.1 Definition

The Internet is a large, distributed network including many other small networks. It's arranged as an ever-enlarging web of interconnected devices. In that network, Network Layer and the Transport Layer play important roles, they work together to encapsulate data and sent it all over the network.

2.2 Network Layer

Network layer is responsible for the routing of data packets between different networks, such as the Internet. It determines the optimal path for data transmission based on the destination address and network topology.

1. **IP address**

   An IP address is a unique identifier assigned to each device on a network that uses the Internet Protocol for communication. An IP address is a numeric label assigned to each device, consisting of a series of four numbers separated by dots, such as 192.168.1.1.

   IP identify the host or network interface, allowing devices to communicate with each other over a network and provide location information, enabling data to be routed to and from the correct destination. The network layer has two versions of the IP: IPv4 and IPv6. [5]

IPv4	IPv6
IPv4 has a 32-bit address length	IPv6 has a 128-bit address length
It Supports Manual and DHCP address configuration	It supports Auto and renumbering address configuration
In IPv4 end to end, connection integrity is Unachievable	In IPv6 end to end, connection integrity is Achievable
Fragmentation is performed by Sender and forwarding routers	In IPv6 fragmentation is performed only by the sender
The Security feature is dependent on application	IPSEC is an inbuilt security feature in the IPv6 protocol
IPv4 are divided into five different classes A, B, C, D, E.	IPv6 does not have any classes of IP address.

2. **Routing protocol**

   Routing in the network layer refers to the process of selecting the optimal path for sending data packets from the source to the destination across multiple interconnected networks.

   When a packet is sent from a source device to a destination device, it may need to traverse multiple networks, such as local area networks (LANs), wide area networks (WANs), or the internet. The routers in each network use routing protocols to exchange information with each other about the best path for the packet to reach its destination. The routers then use this information to make forwarding decisions and send the packet on its way.

   Routing is a crucial function in the network layer because it ensures that data packets are delivered to their intended destination in a timely and efficient manner. Without routing, data packets could get lost or delayed, causing disruptions to network communication.

2.3 Transport Layer

The Transport Layer is responsible for providing end-to-end communication between devices on a network. It is responsible for the segmentation, reassembly, and reliable delivery of data between the source and destination devices.

The Transport Layer adds a header to the data received from the Session Layer and passes it to the Network Layer for delivery to the destination device. The header contains information such as source and destination port numbers, sequence and acknowledgment numbers, and control flags. The header information is used by the Transport Layer to ensure the proper delivery of data to the destination device.

The Transport Layer defines two protocols: TCP (Transmission Control Protocol) and UDP (User Datagram Protocol).

1. **UDP**

   UDP (User Datagram Protocol) is a connectionless protocol, which means that it does not establish a dedicated end-to-end connection before transmitting data. In UDP, data is sent in the form of datagrams or packets, without verifying whether the destination device is ready to receive them or not. Therefore, UDP is often referred to as an unreliable protocol.

   To send data using UDP, the sender application first creates a UDP packet that includes the source and destination port numbers, the length of the packet, and the payload (actual data to be transmitted). The packet is then sent to the network layer for transmission. At the receiving end, the UDP packet is received and processed by the transport layer protocol on the receiving device.

   UDP also includes a checksum field that is used to detect errors in the datagram during transmission. The checksum is calculated using a complex algorithm that takes into account the source and destination IP addresses, the protocol type (UDP), and the length of the datagram. The checksum is calculated at the source device and added to the UDP header.

   However, UDP does not have any flow control or error recovery mechanisms, making it a lightweight protocol with low overhead. Because of its simplicity, UDP is often used for applications that require fast data transmission and do not need reliable data delivery, such as video streaming, online gaming, and real-time communication.

2. **TCP**

   TCP (Transmission Control Protocol) is a transport layer protocol used to establish reliable, connection-oriented communication between applications running on different devices on a network. Here is a basic explanation of how TCP works:

   - Establishing a connection: Before any data can be transmitted, the two devices must first establish a connection. This is done through a three-way handshake process, where the two devices exchange packets to establish and synchronize the sequence and acknowledgment numbers.

   - Data transfer: Once the connection is established, the data transfer can begin. The data is broken down into segments, each with its own sequence number, and sent to the destination device. The destination device acknowledges each segment it receives, and the sender retransmits any segments that are not acknowledged within a certain time frame.

- Connection termination: When the data transfer is complete, the two devices terminate the connection. This is done through a four-way handshake process, where the two devices exchange packets to ensure that all data has been received and that the connection can be safely closed.

TCP also includes mechanisms for flow control, congestion control, and error recovery to ensure that the data is reliably transmitted between the devices.

3 World Wide Web

3.1 Definition

The World Wide Web (or the Web) is the most popular online information library built on the Internet or as a client-server network of the Internet. Client-server which means an operating model between clients and servers in a response-request messaging template, adheres to TCP/IP protocol.

3.2 Several components of World Wide Web

The World Wide Web or "The Web" contains 4 several components:

1. **URL**

   URL or Uniform Resource Locator interpret a unique address that web resource can be found on the internet. URL is honestly a zip code that links to a website or an exact page on a website when we enter into the address bar.

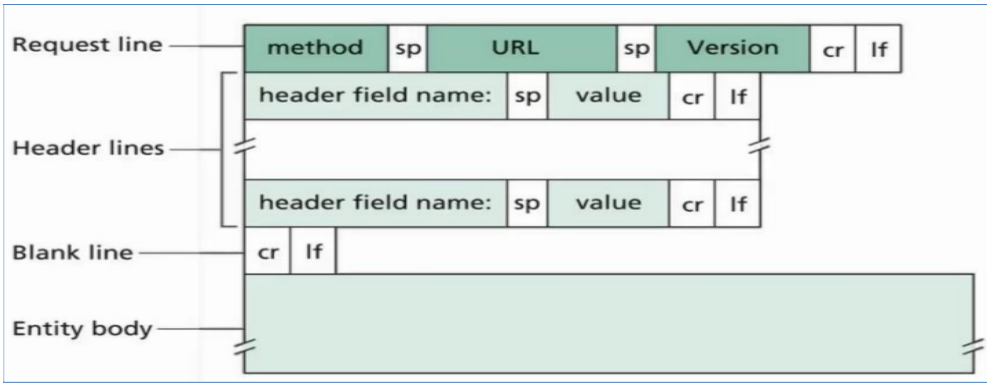
   A URL typically includes three main parts:

   - The protocol of the Internet(e.g. HTTP or HTTPS)
   - The domain name(DNS)
   - The path to the resource (e.g./index.html).

2. **HTTP**

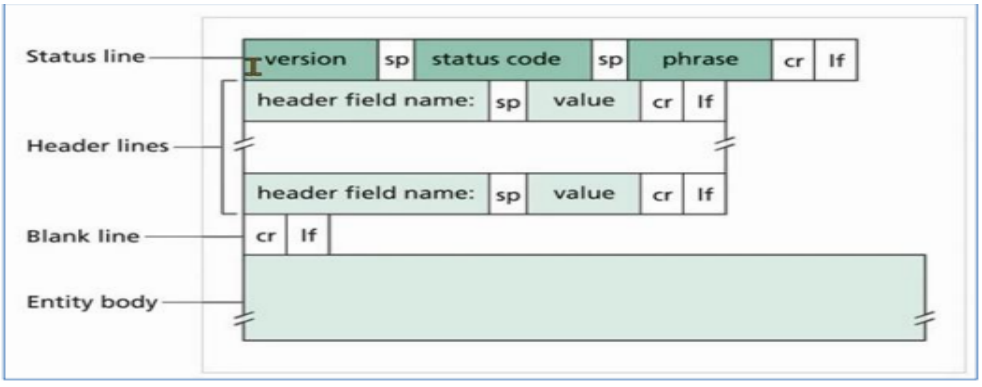
   HTTP, which stands for HyperText Transfer Protocol, is the protocol used to transfer data over the World Wide Web and an application layer protocol. It's the foundation of the web and enables the creation and sharing of information on a global scale.

   When we enter a URL on web browser, we can directly hit up the server computer for the website that we want to view. The browser will send an HTTP request to the web server and they send us all the files and data for our browser to be able to render through HTTP



An HTTP request is initiated by a client, typically a web browser, to request a resource from a web server. The request consists of several parts:

- **Request line**: It contains the HTTP method (such as GET, POST, PUT, DELETE) indicating the action to be performed, the target URL or URI (Uniform Resource Identifier) of the resource, and the HTTP version.
- **Headers**: These are key-value pairs that provide additional information about the request. They can include headers like Accept, Content-Type, Authorization, and more. Headers convey details such as the format the client expects the response in, authentication credentials, caching preferences, and more.
- **Body (optional)**: Certain types of requests, such as POST or PUT, may include a body containing data to be sent to the server. For example, when submitting a form, the form data is included in the request body.



Once the server receives the HTTP request, it processes the request and generates an HTTP response. The response contains the requested

resource, along with additional information about the response itself. An HTTP response consists of the following components:

- **Status line**: It contains the HTTP version, a three-digit status code, and a brief textual status message. The status code indicates whether the request was successful or encountered an error. Common status codes include 200 (OK), 404 (Not Found), 500 (Internal Server Error), and many others.
- **Headers**: Similar to request headers, response headers provide additional information about the response. They can include headers like Content-Type, Content-Length, Cache-Control, and more. These headers convey details such as the type of content being sent, the size of the response, caching instructions for the client, and more.
- **Body**: The response body contains the actual content requested by the client, such as an HTML page, an image, or any other type of data. The format and structure of the body depend on the requested resource and the headers provided in the request.

Noticeably, developers should understand what HTTP response error codes mean. These error codes are usually displayed by three-digit number, show the status of HTTP request and returned by Web browser:

(a) 1xx Informational [6]

100 Continue: The server has received the initial part of the request and is willing to process the client's further actions.

(b) 2xx Success [6]

200 OK: The request has been successful, and the server has returned the requested resource as the response.

201 Created: The request has been fulfilled, resulting in the creation of a new resource.

204 No Content: The request has been successfully processed, but the server is not returning any content in the response.

(c) 3xx Redirection [6]

301 Moved Permanently: The requested resource has been permanently moved to a new URL. Clients should update their references.

302 Found: The requested resource has been temporarily moved to a different URL. Clients should continue using the original URL for future requests.

304 Not Modified: The client's cached copy of the resource is still valid, and the server is indicating that there is no need to retransmit the resource.

(d) 4xx Client Errors [6]

400 Bad Request: The server could not understand the client's request, typically due to malformed syntax or invalid parameters.

401 Unauthorized: The request requires user authentication. The client must provide valid credentials to access the requested resource.

403 Forbidden: The server understood the request but refuses to fulfill it. The client does not have the necessary permissions to access the resource.

404 Not Found: The requested resource could not be found on the server.

(e) 5xx Server Errors [**?**]

500 Internal Server Error: An unexpected error occurred on the server while processing the request.

502 Bad Gateway: The server acting as a gateway or proxy received an invalid response from an upstream server.

503 Service Unavailable: The server is temporarily unable to handle the request due to maintenance, overload, or other reasons.

3. **The combination of HTML, CSS and Javascripts files**

   HyperText Markup Language files that are called HTML files define the content of Websites (like the text content and images) and the structure of the web.

   Cascading Style Sheets or CSS files are responsible for styling websites. CSS files determine how the website will look

   Javascript files are the programs that allow websites to actually have functionality. It executes whatever we want

4. **Linking**

   Linking, or connecting coffers through hyperlinks, is a defining conception of the Web, abetting its identity as a collection of connected documents.

4 Conclusion

Through the above three parts, we can come to a conclusion about the role of each layer in the TCP/IP model

1. **Network Interface Layer** is responsible for providing communication services between hosts on the same network. It handles the physical addressing of data packets, data framing, and error detection and correction. Exemplifications of link layer protocols include Ethernet, Wi-Fi, Bluetooth,.. etc.

2. **Network Layer** is responsible for routing data packets between hosts on different networks. It handles IP (Internet Protocol), which is the primary protocol used for Internet communication. The Internet layer is also responsible for addressing and fragmentation of data packets.

3. **Transport Layer** provides end-to-end communication services between applications running on different hosts. It is responsible for ensuring reliable transmission of data between hosts and handling flow control, error correction, and congestion control. The most commonly used transport protocols are TCP (Transmission Control Protocol) and UDP (User Datagram Protocol).

4. **Application Layer** is responsible for providing communication services to end-users or applications. It provides a means for applications to exchange data with each other, regardless of the underlying network. Examples of protocols that operate at the application layer include HTTP, FTP, SMTP, and DNS.

5 Reference

References

[1] https://m.youtube.com/playlistlist=PL8dPuuaLjXtNlUrzyH5r6jN9ulIgZBpdo

[2] https://www.webvidyalayam.com/data-communication-and-networking/ethernet-frame-format/

[3] https://www.geeksforgeeks.org/ethernet-frame-format/

[4] https://www.gatevidyalay.com/computer-networks/

[5] https://www.geeksforgeeks.org/differences-between-ipv4-and-ipv6/

[6] https://loadfocus.com/blog/2014/07/errors-and-response-codes-in-load-testing-on-l