Anatomy of a Web Connection: A Brief Analysis

Individual Work

Pedro Monteiro
97484

Aspetos Profissionais e Sociais da Engenharia Informática

March 23, 2022

Contents

1 Introduction

Nowadays, the use of the web is increasingly common, being part of almost everyone's life. It is possible to work remotely, shop and have meetings online, through the web, which is why it is important to know and understand what is behind this whole process.
This assignment will focus on understanding what is involved in a web connection, from processes, technologies, models and the social and economic impacts that may exist associated with them.

2 Summary and Objectives

According to APSEI[1] curricular plan, this assignment is the result of the first individual work, which has as main objective the analyze of a web connection and the operations behind a web connection.
The main two objectives are:

- Identify the technologies, processes, actors and business models involved in an web connection;
- Identify some of possible social and economic implications associated with the identified technologies, processes, actors and business models;

3 Framework

As mentioned before, the use of the web is constantly growing these days, and therefore many actions are carried out using this "tool". For example, the simple act of opening a browser can be done through a lot of different technologies, and there are even operations not perceived by a normal user.
These operations can have some ethical, social and economic implications, and therefore, should not go unnoticed by an informatics engineer, who should be aware of such possible implications.

4 Technologies

In the following points, very important functionalities and methods that are directly related to a web connection will be described. All these concepts were covered in courses from previous years, such as CD[2] and RS[3], but it is extremely important to understand them well so that you can understand everything that goes on in a web connection.

[1] Aspetos Profissionais e Sociais da Engenharia Informática
[2] Computação Distribuída
[3] Redes e Serviços

4.1 OSI Model

The OSI[4] Model was published in 1984 by the International Organization for Standardization (ISO), and this model consists on a conceptual framework used to describe the functions of a networking system. It characterizes computing functions into a universal set of rules.
The OSI model describes seven different abstraction layers that computer systems use to communicate over a network:

- Physical Layer (Layer 1): the lowest layer of the OSI reference model. Responsible for the actual physical connection between the devices. It is responsible for transmitting individual bits from one node to the next;
- Data Link Layer (Layer 2): responsible for the node-to-node delivery of the message. The main function is to make sure data transfer is error-free from one node to another, over the physical layer;
- Network Layer (Layer 3): works for the transmission of data from one host to the other located in different networks. Takes care of packet routing i.e. selection of the shortest path to transmit the packet;
- Transport Layer (Layer 4): provides services to the application layer and takes services from the network layer. Also provides the acknowledgement of the successful data transmission.
- Session Layer (Layer 5): responsible for the establishment of connection, maintenance of sessions, authentication, and also ensures security.
- Presentation Layer (Layer 6): the data from the application layer is extracted here and manipulated according to the required format.
- Application Layer (Layer 7): implemented by the network applications. These applications produce the data, which has to be transferred over the network.

4.2 Web Browser

A web browser is an application software that allows to access the WWW[5] or a local website. When a request of a web page is made from a website, the web browser retrieves the necessary content from a web server and then displays the page on the requester's device.
A web browser is not the same thing as a search engine. These two concepts are often confused. A search engine is a website that provides links to other websites. However, to connect to a website's server and display its web pages, a user must have a web browser installed.

Figure 1: Examples of some Web Browsers.

[4]Open Systems Interconnection Model
[5]World Wide Web

4.3 HTTP

HTTP[6] is an application-layer[7] protocol designed for communication between web browsers and web servers. It is used for for transmitting hypermedia documents, such as HTML, and follows a classical client-server model, with a client opening a connection to make a request, then waiting until it receives a response. HTTP is a stateless[8] protocol.

4.4 DNS

DNS[9] is the phonebook of the Internet. Users access information online through domain names, like ua.pt. Web browsers interact through IP[10] addresses. DNS translate these domain names to IP adresses so browser can load Internet resources. The process involves converting a hostname (www.ua.pt) into a computer-friendly IP address (192.198.1.13).

4.5 TCP/IP

TCP/IP[11] represents a set of protocols that allow different devices that make up a network to communicate with each other. TCP/IP architecture is also layered, but is composed by only four levels, instead of the seven layers as in the OSI model.

- Network Access Layer (Layer 1): corresponds to the combination of Data Link Layer and Physical Layer of the OSI model. The protocols in this layer provide the means for the system to deliver data to the other devices;
- Internet Layer (Layer 2): parallels the functions of OSI's Network layer. Defines the protocols which are responsible for logical transmission of data over the entire network.
- Transport Layer (Layer 3): end-to-end layer used to deliver messages to a host. Provides a point-to-point connection rather than hop-to- hop, between the source host and destination host to deliver the services reliably.
- Application Layer (Layer 4): performs the functions of top three layers of the OSI model: Application, Presentation and Session Layer. Responsible for node-to-node communication and controls user-interface specifications.

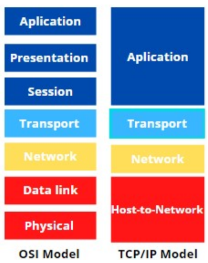

Figure 2: OSI and TCP/IP models

[6]Hypertext Transfer Protocol
[7]Layer 7 of OSI Model
[8]Server does not keep any data between two requests
[9]Domain Name System
[10]Internet Protocol
[11]Transmission Control Protocol/Internet Protocol

5 Traceroute Command

5.1 Traceroute Definition

Command line utility used to show the route that is taken by data packets as they travel across the Internet to their destination. Internet is a global network of routers that allows computers and servers the hability to communicate with each other from all over the world, and these routers communicate with each other so they can direct or route the data packets to their intended destination. The traceroute utility is a tool that is used to find out the exact path a data packet is taken from the sender to the destination. Traceroute helps finding problems like bottlenecks, such as why and where a connection to a server might be lagging. It uses ICMP's[12] Ping command to find out how many different devices are between the computer initiating the traceroute and the target, and works by manipulating[13] the packets time to live value or TTL[14].

5.2 Traceroute Workflow

1. Source sends a packet with TTL=1;
2. The Router decrements the TTL by 1, which changes the value to 0. The packet is dropped and the router sends an ICMP TTL Exceeded message;
3. The Source receives the "ICMP TTL Exceeded" message and adds the router IP to the Traceroute hops table;
4. Then the process starts over again with TTL=2;
5. And it continues like this by incrementing the TTL by 1 until it reaches its destination;

5.3 Traceroute Usage

In the Command Prompt window, just type **tracert** followed by the destination, either an IP Address or a Domain Name, and press Enter.

- `tracert www.xxx.yyy`

5.4 Traceroute vs Ping

Traceroute and ping commands are not the same. While ping is a utility that helps users to check if a particular IP address is accessible or not, traceroute, as has been said before, traces a data packet from one computer to the host, and will also show the number of steps[15] required to reach there. Also, traceroute pings not only the destination, but each router on its way to destination.

[12]Internet Control Message Protocol
[13]Every time a data packet reaches a hop, the TTL value is decreased by one
[14]Time to Tive - number of times the packet can be rebroadcast by the next host
[15]Hops

6 Traceroutes Logs

```
pedro@pedro:~$ traceroute -I www.louvre.fr
traceroute to www.louvre.fr (89.185.38.196), 30 hops max, 60 byte packets
 1  gt2-edu-alunos.core.ua.pt (192.168.63.253)  3.181 ms  3.115 ms  3.098 ms
 2  10.1.0.118 (10.1.0.118)  3.083 ms  3.073 ms  3.062 ms
 3  gt2-vrfinternet-r.core.ua.pt (193.137.173.243)  3.048 ms  3.037 ms  3.109 ms
 4  Router41.Porto.fccn.pt (193.136.4.26)  5.162 ms  5.151 ms  5.140 ms
 5  Router40.Porto.fccn.pt (194.210.7.208)  7.568 ms  7.557 ms  7.546 ms
 6  Router60.Lisboa.fccn.pt (193.136.1.10)  7.519 ms  7.073 ms  7.026 ms
 7  Router3.Lisboa.fccn.pt (194.210.6.203)  8.790 ms  8.779 ms  8.768 ms
 8  Claranet.AS8426.gigapix.pt (193.136.251.5)  15.558 ms  15.547 ms  15.536 ms
 9  sd-ar11-te-0-0-2-1.router.fr.clara.net (212.43.193.149)  46.820 ms  50.205 ms  50.161 ms
10  mail.mdlsit-pw-web.msp.fr.clara.net (89.185.58.93)  46.805 ms  46.793 ms  46.782 ms
11  89.185.38.196 (89.185.38.196)  47.308 ms  47.297 ms  48.192 ms
```

Figure 3: Traceroute Command executed to *Louvre*, on 03/20/2022, 15:35pm

6.1 Interpretation

As we can see there are several rows divided into columns on the log. Each row represents a **hop** along the route. It's like there is a check-in point where the signal gets its next set of directions. The first 2 columns tell where the hop actually landed, and then there are three numerical values known as the Round-Trip Time (RTT), which refers to the amount of time that a given data packet takes to reach its destination and route back an ICMP message to the source. By default, traceroute routes three packets of data to test each hop. Every packet routes an ICMP error message back to the source when it reaches a device on the network. This allows traceroute to determine the RTT of that packet and does not necessarily indicate an error.

6.2 Players Involved

The main players involved are:

- source computer: computer where traceroute command is executed;
- network routers: all routers/devices through which data packets pass;
- destination: destination router/website of data packets;

In our example, Figure 3, the source was my personal computer, routers from the Aveiro University network were used, routers in Porto, Lisbon and Paris, and in the end, the destination, the website of the Louvre museum.

6.3 Meaning of *** in traceroute

Sometimes traceroute logs have few lines consisting of "***". What does it means? It means that there was a **Request Timed Out**, and there are three possibilities:

- ICMP/UDP may not be configured: most likely the device that was hit was not configured to reply to ICMP/UDP traffic. Sometimes nodes are programmed not to respond to traceroute requests (for example, due to security reasons). This result does not mean that the traffic wasn't passed;
- packets were dropped due to an issue on the network. These results are usually packet timeouts, or the traffic has been blocked by a firewall;
- other nodes might give the traceroute a very low priority, if they are very busy at the time of being interrogated;

6.4 Each Hop Analysis and Operations/Technologies Used

Hop	Device or Media	Local	Network/Operator/Owner	Technologies/Protocols	OSI layer
0	Personal Computer (192.168.51.53)	DETI UA	UA Ethernet Network / STIC / Aveiro University	HTTP	7 - Application
					6 - Presentation
				Port: XXXX	5 - Session
				TCP	4 - Transport
				IPv4	3 - Network
				WiFi-IEEE802.11x	2 - Data Link
				UTP (Ethernet) or Free-Space Radio	1 - Physical
TRANSPORT		UA	Free-Space radio (Public Domain Unlicensed) and/or UTP (Ethernet)		
1	gt2-edu-alunos.core.ua.pt (192.168.63.253)	STIC UA	UA Ethernet Network / STIC / Aveiro University	IPv4	3 - Network
				Fast Ethernet (802.2; 802.3)	2 - Data Link
				100BASE-T (802.3)	1 - Physical
TRANSPORT		UA	OPTICAL FIBRE Campus Backbone (Gigabit Ethernet)		
2	10.1.0.118 (10.1.0.118)	STIC UA	UA Ethernet Network / STIC / Aveiro University	IPv4	3 - Network
				Gigabit Ethernet (IEEE 802.3-2008)	2 - Data Link
				Gigabit Ethernet (IEEE 802.3-2008)	1 - Physical
TRANSPORT		UA	OPTICAL FIBRE Campus Backbone (Gigabit Ethernet)		
3	Router gt2-vrfinternet-r.core.ua.pt (193.137.173.243)	STIC UA	UA Ethernet Network / STIC / Aveiro University	IPv4	3 - Network
				Gigabit Ethernet (IEEE 802.3-2008)	2 - Data Link
				Gigabit Ethernet (IEEE 802.3-2008)	1 - Physical
TRANSPORT		Linha do Norte	OPTICAL FIBRE FCCN Backbone (40λX40GB / DWDM)		
4	Router Router41.Porto.fccn.pt (193.136.4.26)	Estação Campanhã Porto	RCTS IP / FCCN / FCCN	IPv4	3 - Network
				10 Gicabit ethernet	2 - Data Link
				GE, OTN, SDH, SONET, etc	1 - Physical

TRANSPORT		Porto	OPTICAL FIBRE Campus Backbone (Gigabit Ethernet)		
5	Router Router40.Porto.fccn.pt (194.210.7.208)	Porto	RCTS IP / FCCN / FCCN	IPv4	3 - Network
				10 Gicabit ethernet	2 - Data Link
				GE, OTN, SDH, SONET, etc	1 - Physical
TRANSPORT		Lisbon	OPTICAL		
6	Router Router60.Lisboa.fccn.pt (193.136.1.10)	Lisbon	RCTS IP / FCCN / FCCN	IPv4	3 - Network
				10 Gicabit ethernet	2 - Data Link
				GE, OTN, SDH, SONET, etc	1 - Physical
TRANSPORT		Lisbon	OPTICAL FIBRE Campus Backbone (Gigabit Ethernet)		
7	Router Router3.Lisboa.fccn.pt (194.210.6.103)	Lisbon	RCTS IP / FCCN / FCCN	IPv4	3 - Network
				10 Gicabit ethernet	2 - Data Link
				GE, OTN, SDH, SONET, etc	1 - Physical
TRANSPORT		Lisbon	OPTICAL FIBRE Campus Backbone (Gigabit Ethernet)		
8	Router Claranet.AS8426.gigapix.pt (193.136.251.5)	Lisbon	RCTS IP / FCCN / FCCN	IPv4	3 - Network
				10 Gicabit ethernet	2 - Data Link
				GE, OTN, SDH, SONET, etc	1 - Physical
TRANSPORT		Paris	OPTICAL FIBRE Campus Backbone (Gigabit Ethernet)		
9	Router sd-ar11-te-0-0-2-1.router.fr.clara.net (212.43.193.149)	Paris	GEANT	IPv4	3 - Network
				10 Gicabit ethernet	2 - Data Link
				GE, OTN, SDH, SONET, etc	1 - Physical
TRANSPORT		Paris	OPTICAL FIBRE Campus Backbone (Gigabit Ethernet)		
10	Router mail.mdlsit-pw-web.msp.fr.clara.net (89.185.58.93)	Paris	GEANT	IPv4	3 - Network
				10 Gicabit ethernet	2 - Data Link
				GE, OTN, SDH, SONET, etc	1 - Physical
TRANSPORT		Paris	OPTICAL FIBRE Campus Backbone (Gigabit Ethernet)		
11	Router www.louvre.fr (89.185.38.196)	Paris	GEANT	HTTP	7 - Application
					6 - Presentation
				Port: XXXX	5 - Session
				TCP	4 - Transport
				IPv4	3 - Network
				Ethernet-IEEE802.3 or WiFI-IEEE802.11x	2 - Data Link
				UTP (Ethernet) or Free-Space Radio	1 - Physical

As we can see the packages traveled from the Aveiro university to Porto, then to Lisbon and finally to Paris, where the website of the Louvre museum is hosted.
We can verify that the packets arrived at the destination by verifying that the IP address obtained corresponds to the destination website, Figure 4.

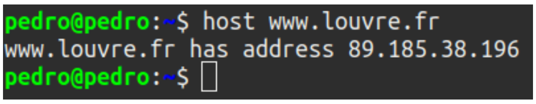

Figure 4: Louvre Museum Website IP Adress

6.5 Are the *traceroute* logs always the same for different times of the day and for different locations?

Not always, but after running multiple traceroutes commands for the same location I was able to see that there are some differences among the logs. This can be caused by the overload of the network. Sometimes the path requested first can not be used, due to a large number of requests or due to the failure of a network router. This can lead at some impacts, such as latency, or network delay. A request that normally takes 2ms may take longer to respond. What is done to avoid latency is sending the packets through another path that is as efficient as possible.

6.6 What happens during a typical web session?

6.6.1 Possible Social and Economic Implications

When we type a web address into browser, the browser finds the real address of the server through the DNS resolution and sends an HTTP request message to the server, asking it to send a copy of the website. If the server approves the request, the server sends the client a message, "200 OK", and then sends the website's files and the browser displays the page. Sometimes there are some advertisements and recommendations that are displayed too and have not been asked by the user. This is possible using, for example, data such as the location and browsing history.

This is one of the biggest impacts of a web connection. Although the ads were not created with the idea of disturbing users, but rather to help in the search for a product/item, they can be used by people with bad intentions. As discussed in the SIO[16] course last semester, there may be some attacks that aim to steal the user's credentials, without him realizing it, such as XSS[17].

7 Conclusion

This report allowed me to apply the knowledge acquired throughout the course, from algorithms more related to computer science and also to mathematics, such as graph theory. Through the research carried out and the observation of several videos on the subject, it was also possible to delve into what really happens in a web connection, from the technologies used and even the impacts associated with a simple Google search.

Furthermore, I could see that there is a network path from Aveiro to Paris, passing through Porto and Lisbon.

[16] Segurança Informática e nas Organizações
[17] Cross-site scripting

8 References

[1] "Andrew S. Tanenbaum, "Computer Networks", Prentice-Hall, 2002, ISBN 0-13-066102-3"

[2] "The OSI Model's Seven Layers" http://www.inetdaemon.com/tutorials/basic_concepts/network_models/osi_model/osi_model_seven_layers.shtml

[3] "How OSI Works" http://computer.howstuffworks.com/osi.htm

[4] "Understanding the Ping and Traceroute Commands" http://www.cisco.com/c/en/us/support/docs/ios-nx-os-software/ios-software-releases-121-mainline/12778-ping-traceroute.html

[5] "George Reynolds, "Ethics in Information Technology", Cengage Learning, 2015, ISBN-13: 9781285197159 / ISBN-10: 1285197151"

[6] "Understanding the Ping and Traceroute Commands" http://www.cisco.com/c/en/us/support/docs/ios-nx-os-software/ios-software-releases-121-mainline/12778-ping-traceroute.html

[7] "The OSI Model Defined, Explained, and Explored" https://www.forcepoint.com/cyber-edu/osi-model

[8] "Layers of OSI Model" https://www.geeksforgeeks.org/layers-of-osi-model/

[9] "Web Browser" https://en.wikipedia.org/wiki/Web_browser

[10] "HTTP" https://developer.mozilla.org/en-US/docs/Web/HTTP

[11] "What is DNS? | How DNS works" https://www.cloudflare.com/en-gb/learning/dns/what-is-dns/

[12] "TCP/IP Model" https://www.geeksforgeeks.org/tcp-ip-model/

[13] "Traceroute Command" https://www.sciencedirect.com/topics/computer-science/traceroute-command

[14] "What Is Traceroute and How Does It Work?" https://www.n-able.com/blog/what-is-traceroute-how-does-it-work

[15] "What is a Traceroute and How Do Traceroutes Work?" https://obkio.com/blog/traceroutes-what-are-they-and-how-do-they-work/

[16] "Traceroute (tracert) Explained - Network Troubleshooting" https://www.youtube.com/watch?v=up3bcBLZS74

[17] "Traceroute: Finding meaning among the stars" https://www.redhat.com/sysadmin/traceroute-finding-meaning

[18] "Difference between Ping and Traceroute" https://www.geeksforgeeks.org/difference-between-ping-and-traceroute/