Anatomy of a Web Connection: A Brief Analysis

Title: Anatomy of a Web Connection: A Brief Analysis

Author: Mário Francisco Costa Silva

Date: 25/03/2021

**Index**

[1. Context 2](#_Toc68355619)

[2. Objectives 2](#_Toc68355620)

[3. Process 2](#_Toc68355621)

[3.1 Diagnostic Tool: Traceroute 2](#_Toc68355622)

[3.2 Hop 2](#_Toc68355623)

[3.3 Request Timed Out 2](#_Toc68355624)

[3.4 Traceroute results 3](#_Toc68355625)

[3.4.1 From my Home Network 3](#_Toc68355626)

[3.4.2 From UA Network 3](#_Toc68355627)

[4. Analysis and Discussion 4](#_Toc68355628)

[4.1 Traceroute interpretation 4](#_Toc68355629)

[4.1.1 My Home Network 4](#_Toc68355630)

[4.1.2 UA Network 5](#_Toc68355631)

[4.2 Different traceroute results 6](#_Toc68355632)

[4.2.1 Ran at different times 6](#_Toc68355633)

[4.2.2 Ran at different locations 6](#_Toc68355634)

[4.3 Protocols and Mechanisms 7](#_Toc68355635)

[4.4 Entities 7](#_Toc68355636)

[4.5 Social and economic implications 8](#_Toc68355637)

[5. Conclusion 9](#_Toc68355638)

[6. References 9](#_Toc68355639)

[7. Appendix 10](#_Toc68355640)

[7.1 Packets routes 10](#_Toc68355641)

[7.2 Different traceroute results 11](#_Toc68355642)

# Context

This report was proposed by the teacher Manuel de Oliveira Duarte for the APSEI course.

The immense growth of the internet presence in the people’s lives requires **many technologies, processes, actors and operations that aren’t usually perceived by the “common user”. Such operations might have profound social and economic implications.**

# Objectives

The purpose of this report is to find out the essential steps that occur on a simple connection to a web site and understand what happens on the background of a simple web connection while identifying some aspects like the architecture, technologies, processes, actors, business models involved and then estimating possible social and economic implications associated to all of these aspects.

# Process

To find the main steps for a connection to a web site I will be using a computer with Windows OS, the Command Prompt and a diagnostic tool Traceroute to perform the command “tracert [www.cmu.edu](http://www.cmu.edu)”. This process will be repeated but in different scenarios, once from my home network, other connected to UA network through a VPN and then these two will be repeated at different times as well.

## Diagnostic Tool: Traceroute

Traceroute is a great network diagnostic tool for tracking in real-time the pathway that a packet on an IP network from the source to destination, identifying the IP addresses of all routers it pinged in between. It also records the time taken for each **hop** the packet makes during its course.

Traceroute tool uses Internet Control Message Protocol – (ICMP) messages and relies on a function called TTL – (Time to Live) in the header of this Layer 3 protocol. ICMP operates between two hosts at Layer 3 (Network) level of the OSI model.

## Hop

A hop is when a packet goes from one network to another travelling through routers, so each router along the data path constitutes a hop.

The traceroute command measures the number of router hops from one host to another. Hop counts are often useful to find faults in a network or to discover if the routing is correct.

Upon executing the command, it sets a default hop limit of 30, known as time to live (TTL) in IPv4 and hop limit in IPv6, this limits the number of hops a packet is allowed, each time a router receives a packet, it modifies the packet, decrementing the time to live (TTL). This limit is what prevents packets from being stuck in a loop forever.

## Request Timed Out

Sometimes, a traceroute has a hard time accessing a device. In these situations, it may show a message saying, “Request timed out,” along with an asterisk. This indicates that the router it reached was configured to deprioritize or automatically reject ICMP packets, which is done because ICMP is not categorized as essential traffic by many routers so for security reasons they might reject it or give it low priority.

## Traceroute results

### From my Home Network

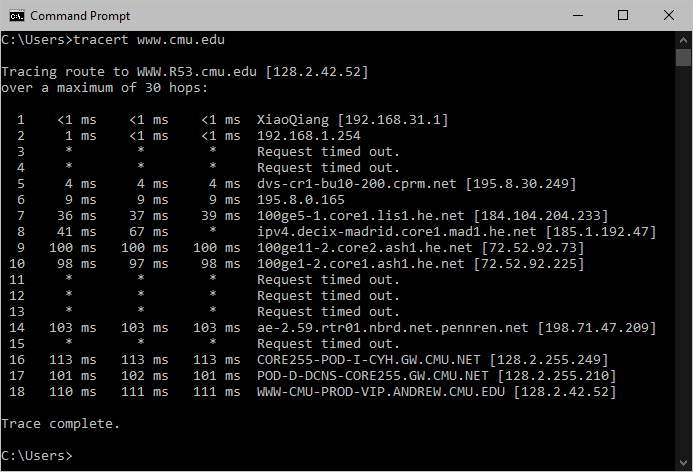


Figure 1 - Result of the traceroute command from my home network. Date: 15:30 27/03/2021

### From UA Network

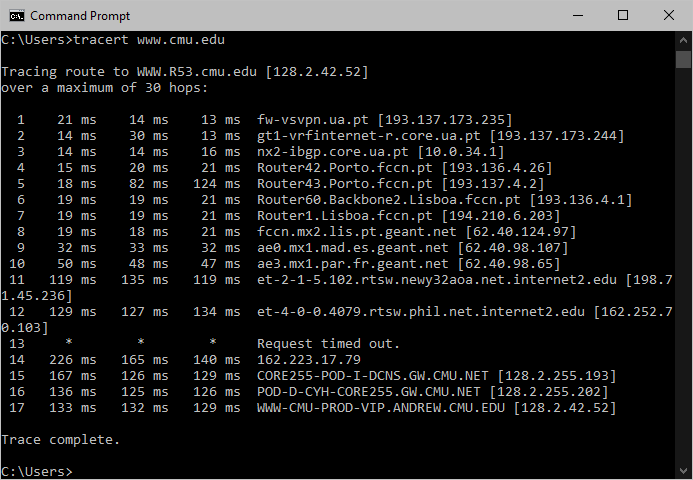


Figure 2 - Result of the traceroute command from UA network. Date: 16:00 27/03/2021

# Analysis and Discussion

## Traceroute interpretation

Here’s an interpretation of the results obtained from executing the tracert command on:

### My Home Network

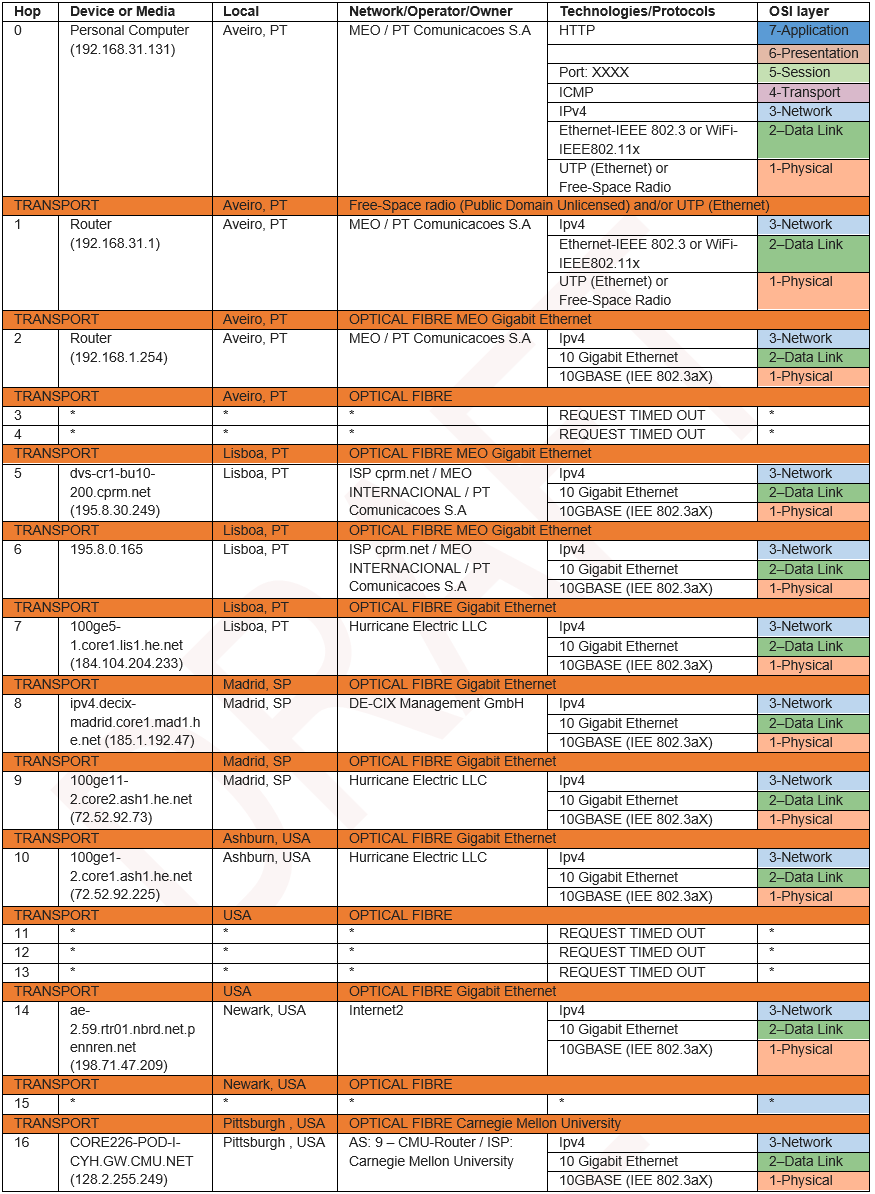




Table 1 - Analysis and interpretation of the traceroute result ran from my home network.

### UA Network



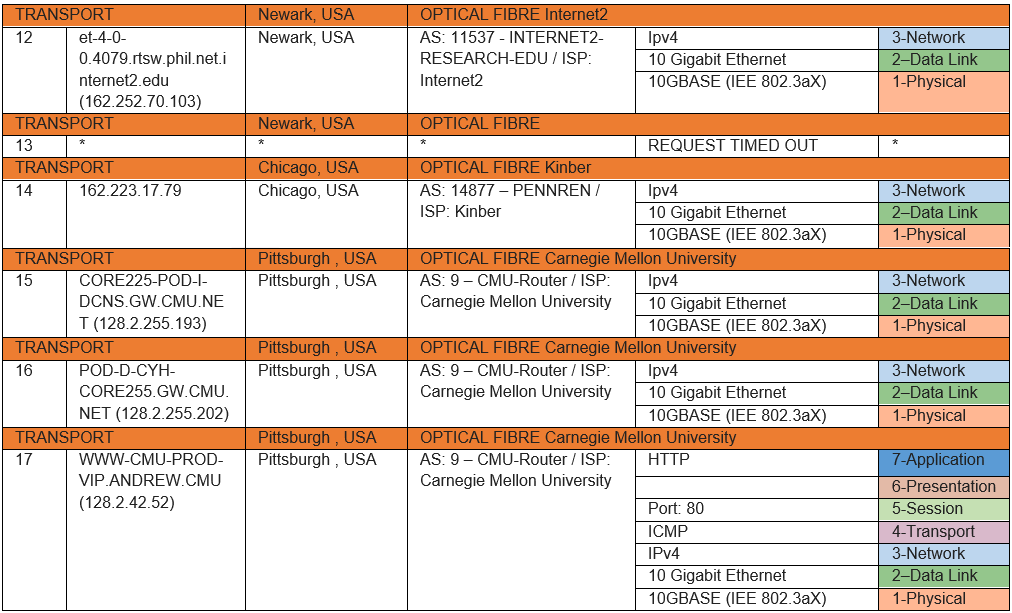


Table 2 - Analysis and interpretation of the traceroute result ran from UA network.

## Different traceroute results

From these results it is obvious that two different paths were taken when executed the same command but from different networks. Also, the traceroute command was executed more than once for each of these locations to also test different times (see **appendix - figures 5, 6**).

### Ran at different times

By comparing each result from the same location, but at different times, I can conclude that the paths taken are not always the same. However, the paths are very similar, having only 1 or 2 hops different.

One reason behind this variation may be the way the network configuration is made on each hop and factors like current traffic can affect how the routing is handled, so the paths may be different accordingly to the algorithms configured on those networks. This explains why there is different results for traceroutes to the same location from the same source.

Not only the path of all hops from source to destination may be different but, in each hop, each packet sent can travel different paths to the next hop and then on the way back. Latencies are affected not only by the time it took to reach a destination but also on the time to return to the source, since these two paths may be different, latencies can be heavily affected by only one of the paths.

Another possible reason is the fact that a router along the path may become unavailable and so it needs to go for another hop. It can be unavailable for several reasons, it went offline, got heavily requested or maybe its firewall blocked the address from where the packet came.

### Ran at different locations

As said before, the traceroute was executed from two different networks, one from my home network and the other from UA’s.

UA uses a network, Eduroam, that provides an easy and secured access when connecting to another institution that also uses it. For this reason, it has different possible network accesses and chooses some exclusives and probably better paths with less latencies.

## Protocols and Mechanisms

* **OSI Model**

OSI stands for Open Systems Interconnection and it’s a seven layers architecture with each layer having specific functionality to perform. All these 7 layers work collaboratively to transmit the data from one person to another across the globe and they depend from the layers bellow them. With this model it is possible to better describe a network system. The seven OSI layers are in ascending order: Physical, Data Link, Network, Transport, Session, Presentation, Application.

* **Web Browser**

**A web browser is an application used to access web pages and gaining access to a large variety of information resources (documents, images, videos, etc) located in computers throughout the world. A web browser is a piece of software capable of retrieving and presenting information resources originating in different locations of the web. A browser also has the ability to travel across these locations looking for the desired information as determined by appropriate addresses in the form of URLs (Uniform Resource Locators).**

* **TCP**

**Transmission Control Protocol (TCP) involves some of the fundamental mechanisms of the Internet protocol suite. Its main functionality is to ensure that all received bytes at one end of a communication system are identical to the bytes that are sent from the other end and are in the correct order. In approximate terms, it can be considered as being located at the level of the transport layer of the OSI model.**

* **HTTP**

Hyper Text Transfer Protocol is a communication protocol located in the Application layer of the OSI model. It’s used by information systems as a way to transfer Hypertext documents in the World Wide Web.

* **ICMP and IPv4**

The Internet Control Message Protocol is a support protocol used to send messages between network devices, it is mostly used to send error messages and operational information indicating success or failure when communicating with another IP address. **The Internet Protocol version 4 (IPv4) defines a set of rules for the Internet and assigns addresses to devices on a network.**

* **IEEE 802.11X**

**IEEE 802.11X refers to the IEEE802.11 standard, known as Wi-Fi, to define communication over a Wireless Local Area Network (WLAN), it specifies a set of MAC and PHY protocols to do so.**

* **ISP**

**An Internet Service Provider is an organization that aims to provide access and usage to the Internet.**

* **AS**

**AS stands for Autonomous System, it’s a collection of connected Internet Protocol routing prefixes controlled by one or several network operators**

## Entities

This are some of the entities that were necessary to make the connections possible:

* **MEO - PT Comunicacoes S.A**

**MEO is an operator company that belongs to Altice Portugal, it provides Mobile and residential telecommunications and Internet.**

* **Fundacao para a Ciencia e a Tecnologia, I.P**

**FCT is the Portuguese national funding agency for science, research and technology.**

* **Hurricane Electric LLC**

**Hurricane Electric is a global Internet service provider offering IPv4 and IPv6 Internet access, transit, tools, and network applications, as well as data center colocation and hosting services in San Jose, California, and in Fremont, California, where the company is based.**

* **GÉANT**

**GÉANT is a European Data Network used by the Research and Education community, connecting NRENs across Europe to facilitate the collaboration between international projects.**

* **Kinber**

**KINBER is a research and education network that provides a variety of infrastructure services to communities throughout Pennsylvania.**

* **Internet2**

**Internet2 is a community providing a secure high-speed network, cloud solutions, research support, and services tailored for research and education.**

* **GIGAPIX**

**Gigapix is FCCN’s Internet Exchange Point with the mission to allow several networks to connect each other in a more efficient way in Portugal and avoid the use of international resources.**

## Social and economic implications

Considering the result of the “traceroute” command it is clear that the paths are distinct when it is executed from my home network and from UA’s. As explained above this is because UA uses a network that is developed for research and educational purposes. For some people this can seem a bit unfair since they don’t have access to the same “network opportunities” as an institution like UA has. In my perspective I don’t see an issue with it, because this service (eduroam) is made of a collaboration between hundreds of institutions that even handle themselves some of the infrastructures needed, also it’s reserved for educational and research purposes which usually contribute for the development and evolution of all communities, so everyone benefits from it.

After reconstructing both paths with the help of google earth (see **appendix – figures 3, 4**), I measured the distance travelled by the packets between each hop from the source to its destination. The approximated results were 7700 km from my home network and 9500 km from UA’s. I did this with the intent of self realization and visualization of the network topology that needs to be implemented on a global scale in order to make a simple web connection possible for the entire web world. The packets travelled through enumerous fibre optic cables, organizations, countries and even went to the other side of the ocean. It is incredible the amount of collaboration between organizations and countries it’s necessary to reach a website. But this routes required infrasctrutures that are really expensive since they must be enourmous and very durable, some of them are owned by privated companies, which is worth mentioning that some of them probably focus a lot on making money, and some other infrastructures are owned by the governments (they also don’t guarantee the user’s priority over money).

The imense growth of the internet usage caused this complex networks to be created to allow and expand it’s usage for more areas around the world. I believe this has huge social and economic impact since the internet is literally speeding up the globalisation process by enabling communication without physical barriers for those using it. However there is also a downside to this, it builds up pressure to create these network infrastructures for all countries and some countries don’t have the same amount of resources to be able to compete with the other ones.

Another very important aspect affected by the rise of the internet are the privacy and protection of data. This is a topic that most likely will always be present when talking about internet. A branch that focus on some of these concerns is the cyber security, it tries to alert people to be aware while browsing the internet and also focus on protecting and increasing the level of confidence of the end users aswell. Even with largely developed security systems there is always going to exist online frauds, thefts, destruction and invasion of data, etc… But this situations are illegal, however when talking about privacy of data there are several legal ways of handling it. As I meantioned above, enterprises and governments don’t always prioritize the user and focus on other aspects such as the money. One example of this is the gathering of user’s information to generate adverts or suggestions (like on Youtube), going to a website and having ads all the time can be frustating but it is something the community in general as accepted as part of the internet, it doesn’t seem that bad since it can be helpful getting suggested things we might actually have some interest but it can become a problem if the information gathered to generate this suggestions are private: emails, voice data, personal notes… and sensitive: race, health, religion… In economic aspects, advertisment has a huge impact, the companies that host the adverts are usually very crowded and just the fact that people look at it may influence people in some ways, so it’s benificial for the hosting platform as well as the company/product that is being adverted.

# Conclusion

From what has been presented on this report I can conclude that the objectives were completed, I was able to identify the route of data packets when connecting to a website and learn about the entities, technologies, protocols, mechanisms and business involved and also reflect and discuss about the socio-economics impacts of all these aspects.

# References

|  |  |
| --- | --- |
| [1] | Information Systems, Open Libraries [Online]. Available:  https://open.lib.umn.edu/informationsystems/chapter/12-3-getting-where-youre-going/ [Accessed in 28/3/2021]. |
| [2] | IP Adress Details, IP Location [Online]. Available:  https://www.iplocation.net/ip-lookup [Accessed in 28/3/2021]. |
| [3] | "What is Traceroute and How Does It Work?", Fortinet [Online]. Available:  https://www.fortinet.com/resources/cyberglossary/traceroutes [Accessed in 28/3/2021]. |
| [4]  [5] | Hop, Wikipedia, [Online]. Available:  https://en.wikipedia.org/wiki/Hop\_(networking) [Accessed in 28/3/2021].  “Online Traceroute Using MTR” HackerTarget, [Online]. Available:  https://hackertarget.com/online-traceroute/ [Accessed in 1/4/2021]. |
| [6] | IP geolocation, DBIP, [Online]. Available:  https://db-ip.com/ [Accessed in 28/3/2021]. |
| [7]  [8]  [9]  [10]  [11] | Privacy & Terms, Google, [Online]. Available: https://policies.google.com/technologies/partner-sites?hl=en-US [Accessed in 1/4/2021].  Hurricane Electric, Wikipedia, [Online]. Available:  https://en.wikipedia.org/wiki/Hurricane\_Electric [Accessed in 28/3/2021].  Gigapix, FCCN, Pedro Lorga [Online]. Available:  https://www.fccn.pt/gigapix/ [Accessed in 28/3/2021].  Internet2, [Online]. Available:  https://internet2.edu/community/about-us/ [Accessed in 28/3/2021].  Manuel de Oliveira Duarte, Richard Soares, "Anatomy of a Network Connection – Web and OTT cases", 19/2/.2018. |

# Appendix

## Packets routes

Route of the packets from my home network:

Aveiro -> Lisboa -> Madrid -> Ashburn -> Newark -> Pittsburgh ≈ 7,672.86 km

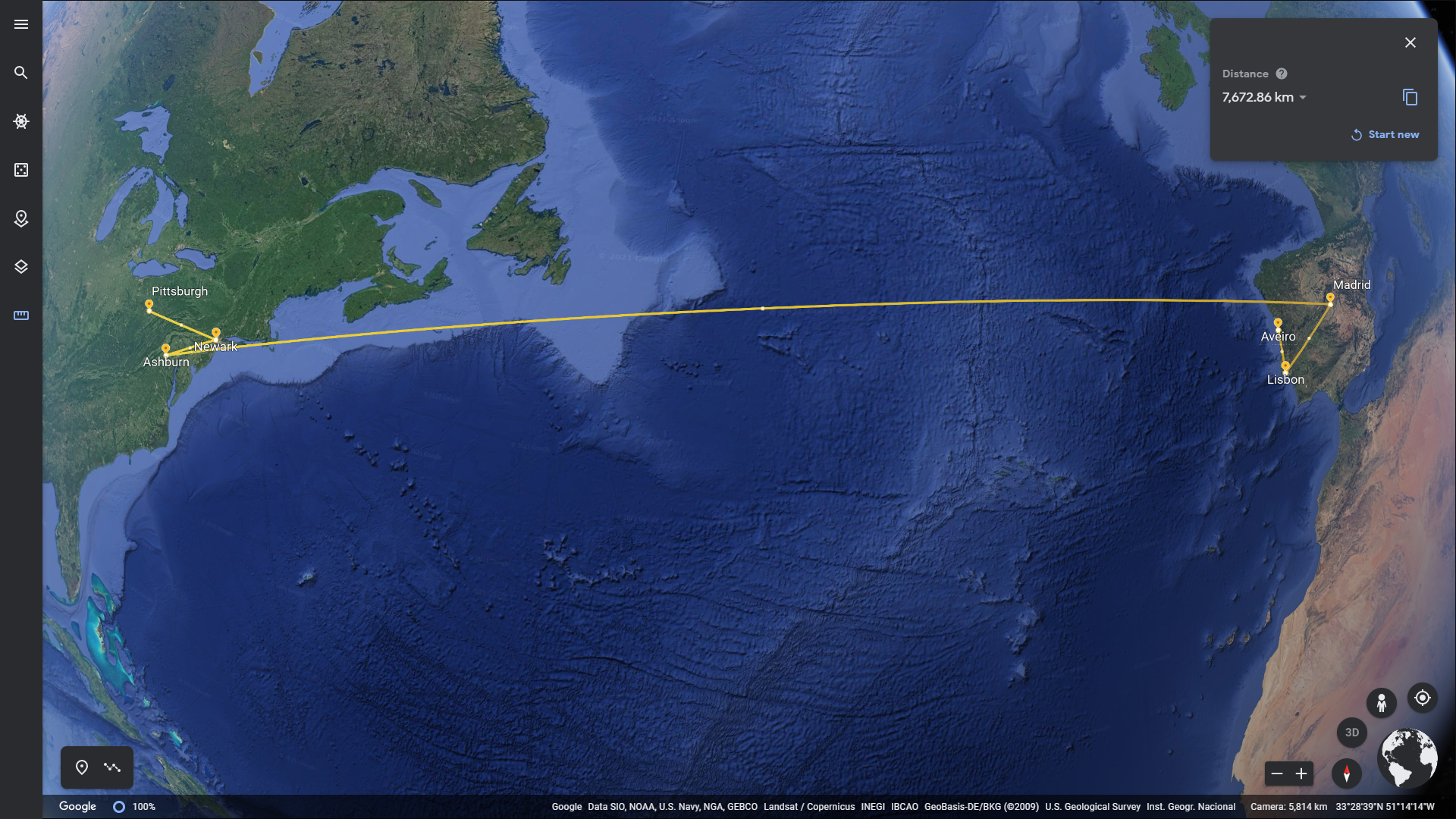


Figure 3 - Route of the data packets sent from my home network.

Route of the packets from UA network:

Aveiro -> Porto -> Lisboa -> Madrid -> Paris -> Newark -> Chicago -> Pitsburgh ≈ 9,523.14 km

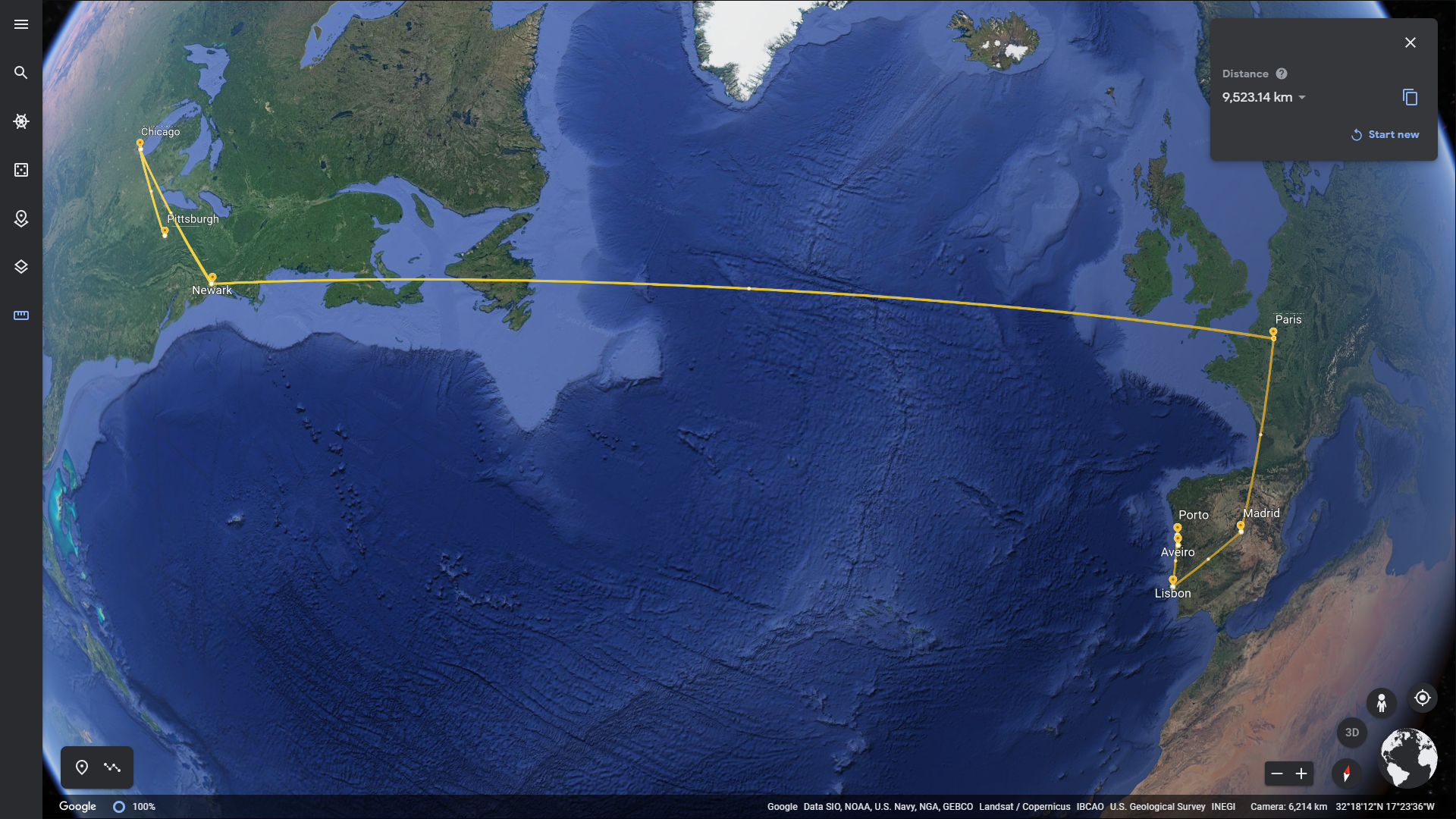


Figure 4 - Route of the data packets sent from UA network.

## Different traceroute results

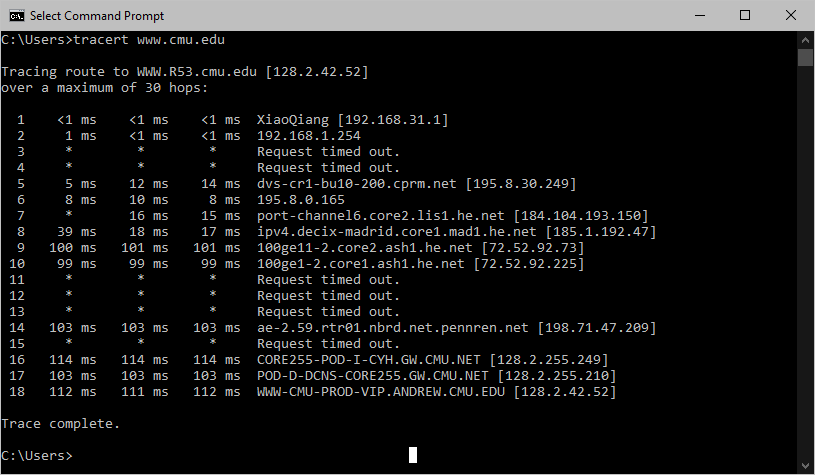


Figure 5 - Result of the traceroute command from my home network. Date: 14:00 30/03/2021

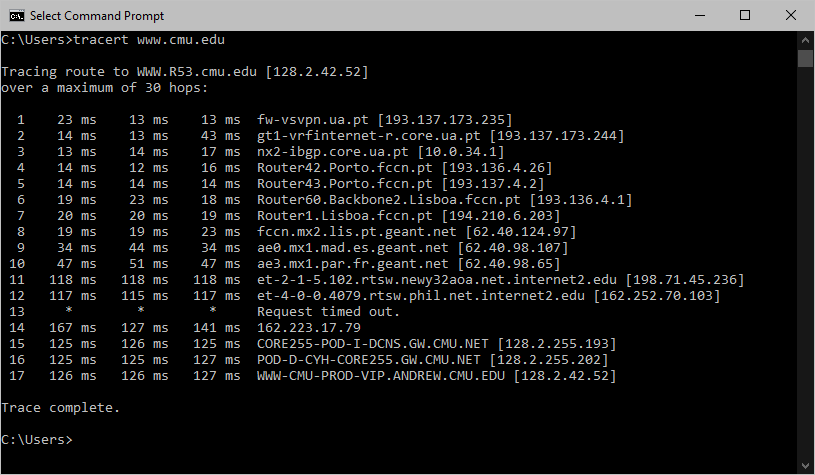


Figure 6 - Result of the traceroute command from UA network. Date: 14:05 30/03/2021