TÉCNICO LISBOA

Computer Networks — 2023/24, Q2		**Assignment:**	Lab 3
The Web & The HTTP Protocol		**Version:**	1.0
Issued:	2023-11-20	**Submission Due:**	2023-11-26

> **Submission note.** You will need to submit all the answers to this lab's questions in Moodle, under "Lab 3: The Web & HTTP".

1 Web/HTTP

The goals of this lab are:

1. Analyze waterfall charts and understand resource dependencies in a webpage.
2. Understand the difference between a *persistent* and a *non-persistent* HTTP connection.
3. Understand the concept of virtual hosting.
4. Show why unencrypted authentication requests are unsecure.

> **Important.** Using the provided virtual machine execute the command `git pull origin master` in `/home/rc/lab-files`. Next, execute `/home/rc/lab-files/lab-http/setup.sh <student #>`, replacing `<student #>` with your own IST student ID number (it should be a 5-6 digit number). Leave the command running until you have completed the lab, at which point you can stop it with `Ctrl-C`.

Inside the VM, open the browser (the VM comes with `Firefox`) and visit `cats.rc.com`. Take a look at its HTML (if you are using `Firefox` you can right-click the page and select *View Page Source*, or press `Ctrl+U`).

Q1 Considering the HTML you just analyzed, how many HTTP requests does the browser need to make in order to fully render the page?

Still in `cats.rc.com`, now open the developer tools of your browser and access the *Network* tab (you can access the developer tools by pressing `Ctrl+Shift+C`). Now refresh the page. You can now see every HTTP request made by the browser regarding `cats.rc.com`. Take special attention to the waterfall chart drawn on the right side.

Q2a Analyze the waterfall chart. Are all requests made at the same time, or is a specific request made first?

Q2b Why does this happen?

Open `wireshark` and watch the *Loopback* interface. Add the filter `http.host == "cats.rc.com"`, go back to the browser and refresh the page. Now come back again to `wireshark`. You should see the same requests that were presented in the waterfall chart. Right-click on the *Get /* request, and select *Conversation filter > TCP*. This gives you all the TCP and HTTP packets associated with that specific request.

Q3a Is the connection *persistent* or *non-persistent*?

Q3b Justify your answer.

Now visit `dogs.rc.com`. The source page is very similar to the one retrieved by visiting `cats.rc.com`: it should request the same number of resources (but now instead of images of cats, we receive images of dogs). Open `wireshark`, apply the filter `http.host == "dogs.rc.com"`, and refresh the page. Again, select the packet corresponding to the HTTP request *Get /*, and apply the TCP conversion filter.

Q4a Is the connection *persistent* or *non-persistent*?

Q4b Justify your answer.

Q5 Compare the destination IP of the requests *Get /* to `dogs.rc.com` and `cats.rc.com`. How does the server disambiguate the content it should present (in other words: why dogs now, and not cats)? *Hint: look carefully at the HTTP request headers, and investigate the concept of virtual hosting.*

Using `wireshark`, apply the filter `http.host == "dogs.rc.com"`. Even without accessing the webpage, you should see requests coming in. It seems like someone else is accessing `dogs.rc.com`...

Q6 What is the source IP of the misterious request? How about its type? Which URL is being accessed?

Visit the misterious URL. It seems that you need to authenticate yourself to visit this webpage...

Q7a Analyze the HTTP headers of the misterious request. Specifically, retrieve the authentication information, and use it to access the protected webpage. Explain which header you used, and provide the authentication information retrieved.

Q7b This authentication mechanism uses base64 to encode its information. How secure is this mechanism?