TDTS06 Lab 1 – Wireshark Lab: HTTP

Oskar von Heideken, oskvo980, ED5
Max Wennerfeldt, maxwe987, ED5

1. The Basic HTTP GET/response interaction

This section will answer questions 1 to 7 along with Task A. The packets are included below.

```
[http GET packet: ]
Hypertext Transfer Protocol
    GET /wireshark-labs/HTTP-wireshark-file1.html HTTP/1.1\r\n [Answer to Q1 found here]
    Host: gaia.cs.umass.edu\r\n
    Connection: keep-alive\r\n
    Upgrade-Insecure-Requests: 1\r\n
    User-Agent: Mozilla/5.0 (Windows NT 10.0; WOW64) AppleWebKit/537.36 (KHTML, like Gecko)
Chrome/52.0.2743.116 Safari/537.36\r\n
    [Answer to Q2 found 3 following rows below]:
    Accept: text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,*/*;q=0.8\r\n
    Accept-Encoding: gzip, deflate, sdch\r\n
    Accept-Language: sv-SE,sv;q=0.8,en-US;q=0.6,en;q=0.4\r\n
    \r\n
    [Full request URI: http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file1.html]
    [HTTP request 1/2]
    [Response in frame: 2]
    [Next request in frame: 3]

[http RESPONSE packet: ]
Hypertext Transfer Protocol
    HTTP/1.1 200 OK\r\n [Answer to Q1, Q4 found here]
    Date: Thu, 01 Sep 2016 14:47:24 GMT\r\n
    Server: Apache/2.4.6 (CentOS) OpenSSL/1.0.1e-fips PHP/5.4.16 mod_perl/2.0.9dev Perl/v5.16.3\r\n
    Last-Modified: Thu, 01 Sep 2016 05:59:01 GMT\r\n [Answer to Q5 found here]
    ETag: "80-53b6be916836e"\r\n
    Accept-Ranges: bytes\r\n
    Content-Length: 128\r\n [Answer to Q6 found here]
    Keep-Alive: timeout=5, max=100\r\n
    Connection: Keep-Alive\r\n
    Content-Type: text/html; charset=UTF-8\r\n
    \r\n
    [HTTP response 1/2]
    [Time since request: 0.127852000 seconds]
    [Request in frame: 1]
    [Next request in frame: 3] (this is icon req.)
    [Next response in frame: 4] (this is icon resp.)
Line-based text data: text/html
```

Answers to questions 1-7:

1. *Q: "Is your browser running HTTP version 1.0 or 1.1? What version of HTTP is the server running?"*
   A: HTTP 1.1 both for the browser and server
2. *Q: "What languages (if any) does your browser indicate that it can accept to the server? In the captured session, what other information (if any) does the browser provide the server with regarding the user/browser?"*
   A: The browser indicates that is supports Swedish (sv-SE) and US English (en-US). The browser also indicated that is accepts various encodings and content formats.
3. *Q: "What is the IP address of your computer? Of the gaia.cs.umass.edu server?"*
   IP of computer: 130.236.124.143, IP of server: 128.119.245.12, this is found in the Wireshark top panel under destination and source.
4. *Q: "What is the status code returned from the server to your browser?"*
   A: 200 OK
5. *Q: "When was the HTML file that you are retrieving last modified at the server?"*
   A: Thu.01 sep 2016 05:59:01 GMT

6. *Q: "How many bytes of content are being returned to your browser?"*
   A: The length of the message returned is 128 bytes, the whole HTTP packet is 542 bytes (the GET is 416)
7. *Q: "By inspecting the raw data in the "packet bytes" pane, do you see any http headers within the data that are not displayed in the "packet details" pane? If so, name one."*
   A: No signs of raw data packets in the pane that did not appear in the packet details, this was investigated using the highlight function in the bottom Wireshark panel.

2. The HTTP CONTDITIONAL GET/response interaction

This section will cover questions 8-11 in the lab PM.

8. *Q:" Inspect the contents of the first HTTP GET request from your browser to the server. Do you see an "IF-MODIFIED-SINCE" line in the HTTP GET?"*
   A: No "IF-MODIFIED-SINCE" in the first HTTP GET was found. This is due to that no cached copy of the requested file locally, therefore a request is made for the latest copy of the file
9. *Q:" Inspect the contents of the server response. Did the server explicitly return the contents of the file? How can you tell?"*
   A: Yes! The server returned the content in plain text in the package.
10. *Q: "Now inspect the contents of the second HTTP GET request from your browser to the server. Do you see an "IF-MODIFIED-SINCE:" line in the HTTP GET? If so, what information follows the "IF-MODIFIED-SINCE:" header?"*
    A: Yes! "`If-Modified-Since: Thu, 01 Sep 2016 05:59:01 GMT\r\n`". Now the get request checks if the content/file has been updated since the latest requested version.
11. *Q: "What is the HTTP status code and phrase returned from the server in response to this second HTTP GET? Did the server explicitly return the contents of the file? Explain."*
    A: `304 not modified`, no, the file as not been modified since, therefore the locally cached copy is the latest. The server did not explicitly deliver the content since it was cached (locally).

In the above sniffing results we can see some of the information in the header (e.g., Q1-6) regarding HTTP versions, languages, IP addresses etc.

Then we observed the GET request when no locally cached version was present (Q8-9) versus when there was a local cached copy (Q10-11) and difference in status response from the server (code 200 vs code 304).

3. Retrieving Long Documents

The questions 12-15 in the lab PM will be answered in this section.

12. *Q: "How many HTTP GET request messages were sent by your browser?"*
    A: Only one GET request, since only one file was requested.
13. *Q: "How many data-containing TCP segments were needed to carry the single HTTP response?"*
    A: 4 TCP segments were needed to carry the response. The server split the content of the file into 4 segments because the file size being too large to send in a single packet.
14. *Q:" What is the status code and phrase associated with the response to the HTTP GET request?"*
    A: "`200 OK`" since the request/response was successful.
15. *Q:" Is there any HTTP header information in the transmitted data associated with TCP segmentation? For this question you may want to think about at what layer each protocol operates, and how the protocols at the different layers interoperate."*
    A: No, HTTP is in the application layer, and don't know that the underlying layers restrictions are. TCP is in the transport layer.

In the above observation, it is shown that only one get and one response even for large files, but several TCP transfers since the file size was large for a single TCP package. This is concluded in the answer to question 15.

4. HTML Documents with Embedded Objects

This section will cover the answers to questions 16 and 17 in the lab PM.

The HTTP plain text export is shown in below.

```
Hypertext Transfer Protocol
    GET /wireshark-labs/HTTP-wireshark-file4.html HTTP/1.1\r\n
    Host: gaia.cs.umass.edu\r\n
    Connection: keep-alive\r\n
    Upgrade-Insecure-Requests: 1\r\n
    User-Agent: Mozilla/5.0 (Windows NT 10.0; WOW64) AppleWebKit/537.36 (KHTML, like Gecko)
Chrome/52.0.2743.116 Safari/537.36\r\n
    Accept: text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,*/*;q=0.8\r\n
    Accept-Encoding: gzip, deflate, sdch\r\n
    Accept-Language: sv-SE,sv;q=0.8,en-US;q=0.6,en;q=0.4\r\n
    \r\n
    [Full request URI: http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file4.html]
    [HTTP request 1/2]
    [Response in frame: 24]
    [Next request in frame: 25]
------------------------------------------------------------
Hypertext Transfer Protocol
    HTTP/1.1 200 OK\r\n
    Date: Thu, 01 Sep 2016 16:01:00 GMT\r\n
    Server: Apache/2.4.6 (CentOS) OpenSSL/1.0.1e-fips PHP/5.4.16 mod_perl/2.0.9dev Perl/v5.16.3\r\n
    Last-Modified: Thu, 01 Sep 2016 05:59:01 GMT\r\n
    ETag: "2ca-53b6be9166fe6"\r\n
    Accept-Ranges: bytes\r\n
    Content-Length: 714\r\n
    Keep-Alive: timeout=5, max=100\r\n
    Connection: Keep-Alive\r\n
    Content-Type: text/html; charset=UTF-8\r\n
    \r\n
    [HTTP response 1/2]
    [Time since request: 0.124692000 seconds]
    [Request in frame: 22]
    [Next request in frame: 25]
    [Next response in frame: 31]
Line-based text data: text/html
------------------------------------------------------------
Hypertext Transfer Protocol
    GET /pearson.png HTTP/1.1\r\n
    Host: gaia.cs.umass.edu\r\n
    Connection: keep-alive\r\n
    User-Agent: Mozilla/5.0 (Windows NT 10.0; WOW64) AppleWebKit/537.36 (KHTML, like Gecko)
Chrome/52.0.2743.116 Safari/537.36\r\n
    Accept: image/webp,image/*,*/*;q=0.8\r\n
    Referer: http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file4.html\r\n
    Accept-Encoding: gzip, deflate, sdch\r\n
    Accept-Language: sv-SE,sv;q=0.8,en-US;q=0.6,en;q=0.4\r\n
    \r\n
    [Full request URI: http://gaia.cs.umass.edu/pearson.png]
    [HTTP request 2/2]
    [Prev request in frame: 22]
    [Response in frame: 31]
------------------------------------------------------------
Hypertext Transfer Protocol
    HTTP/1.1 200 OK\r\n
    Date: Thu, 01 Sep 2016 16:01:00 GMT\r\n
    Server: Apache/2.4.6 (CentOS) OpenSSL/1.0.1e-fips PHP/5.4.16 mod_perl/2.0.9dev Perl/v5.16.3\r\n
    Last-Modified: Sat, 06 Aug 2016 10:08:14 GMT\r\n
    ETag: "cc3-539645c7f1ee7"\r\n
    Accept-Ranges: bytes\r\n
    Content-Length: 3267\r\n
    Keep-Alive: timeout=5, max=99\r\n
    Connection: Keep-Alive\r\n
    Content-Type: image/png\r\n
    \r\n
    [HTTP response 2/2]
    [Time since request: 0.124616000 seconds]
    [Prev request in frame: 22]
    [Prev response in frame: 24]
    [Request in frame: 25]
Portable Network Graphics
```

```
---------------------------------------------------------
Hypertext Transfer Protocol
    GET /~kurose/cover_5th_ed.jpg HTTP/1.1\r\n
    Host: manic.cs.umass.edu\r\n
    Connection: keep-alive\r\n
    User-Agent: Mozilla/5.0 (Windows NT 10.0; WOW64) AppleWebKit/537.36 (KHTML, like Gecko)
Chrome/52.0.2743.116 Safari/537.36\r\n
    Accept: image/webp,image/*,*/*;q=0.8\r\n
    Referer: http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file4.html\r\n
    Accept-Encoding: gzip, deflate, sdch\r\n
    Accept-Language: sv-SE,sv;q=0.8,en-US;q=0.6,en;q=0.4\r\n
    \r\n
    [Full request URI: http://manic.cs.umass.edu/~kurose/cover_5th_ed.jpg]
    [HTTP request 1/1]
    [Response in frame: 37]
---------------------------------------------------------
Hypertext Transfer Protocol
    HTTP/1.1 302 Found\r\n
    Date: Thu, 01 Sep 2016 16:00:52 GMT\r\n
    Server: Apache\r\n
    Location: http://caite.cs.umass.edu/~kurose/cover_5th_ed.jpg\r\n
    Content-Length: 234\r\n
    Connection: close\r\n
    Content-Type: text/html; charset=iso-8859-1\r\n
    \r\n
    [HTTP response 1/1]
    [Time since request: 0.128406000 seconds]
    [Request in frame: 35]
Line-based text data: text/html
---------------------------------------------------------
Hypertext Transfer Protocol
    GET /~kurose/cover_5th_ed.jpg HTTP/1.1\r\n
    Host: caite.cs.umass.edu\r\n
    Connection: keep-alive\r\n
    User-Agent: Mozilla/5.0 (Windows NT 10.0; WOW64) AppleWebKit/537.36 (KHTML, like Gecko)
Chrome/52.0.2743.116 Safari/537.36\r\n
    Accept: image/webp,image/*,*/*;q=0.8\r\n
    Referer: http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file4.html\r\n
    Accept-Encoding: gzip, deflate, sdch\r\n
    Accept-Language: sv-SE,sv;q=0.8,en-US;q=0.6,en;q=0.4\r\n
    \r\n
    [Full request URI: http://caite.cs.umass.edu/~kurose/cover_5th_ed.jpg]
    [HTTP request 1/1]
    [Response in frame: 126]
---------------------------------------------------------
Hypertext Transfer Protocol
    HTTP/1.1 200 OK\r\n
    Date: Thu, 01 Sep 2016 16:00:52 GMT\r\n
    Server: Apache\r\n
    Last-Modified: Tue, 15 Sep 2009 18:23:27 GMT\r\n
    ETag: "78004-18a68-473a1e0e6e5c0"\r\n
    Accept-Ranges: bytes\r\n
    Content-Length: 100968\r\n
    Connection: close\r\n
    Content-Type: image/jpeg\r\n
    \r\n
    [HTTP response 1/1]
    [Time since request: 0.748807000 seconds]
    [Request in frame: 47]
JPEG File Interchange Format
```

16. *Q: "How many HTTP GET request messages were sent by your browser? To which Internet addresses were these GET requests sent?"*
    A: Four "GET" commands were sent, 2 of them was to "gaia", one was to "caite" and one to "manic". The client requested the HTML file, parsed the HTML content and found that there was 2 embedded images. The pearson logo was hosted on the server where the HTML file were stored and the cover file was hosted by another server. The cover photo link points to manic, redirecting to caite resulting in the response from manic where to find the cover photo content in a total of 2 requests.
17. *Q: "Can you tell whether your browser downloaded the two images serially, or whether they were downloaded from the two web sites in parallel? Explain."*
    The images are downloaded serially, since the get request for the cover photo is sent after the logo picture has arrived, the data is downloaded in series.

To get pictures embedded in the content, a GET command is needed. The GET commands for the picture took some investigating to understand, since the link to the first picture redirected to another site, which was not visible in the browser.

5. HTTP Authentication

This section will cover questions 18 and 19, regarding HTTP authentication.

18. *Q: "What is the server's response (status code and phrase) in response to the initial HTTP GET message from your browser?"*
    A: `401 Unautorized`
19. *Q:" When your browser sends the HTTP GET message for the second time, what new field is included in the HTTP GET message?"*
    A: The new field is the "`Autorization:`" field which indicates base64 encoding.