ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Computer Networks: Lab 5 - HTTP

Luca Bedogni

Department of Computer Science and Engineering
University of Bologna

A brief introduction

- HTTP is the Hypertext Transfer Protocol
- Listening on the 80 port
- Client/Server protocol
- Different requests can be made by the client
  - GET
  - POST
  - HEAD
  - PUT
  - DELETE
  - TRACE
  - OPTIONS
  - CONNECT
- In the reply, the first digit of the code tells the kind of reply
  - 200 OK
  - 404 Bad Request
  - 500 Internal Server Error
  - …

Basic HTTP excercise

- If you can't do a live capture, use the *http-ethereal-trace-1* file

- Start wireshark packet capture
- Put http in the display filters
- Go to http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file1.html
- Stop capture
- Try to answer the following

- Is your browser running HTTP 1.0 or 1.1?
- What languages (if any) does your browser indicate that it can accept to the server?
- What is the IP address of your computer?
- What is the status code returned to your computer?
- When was the HTML file last modified at the server?
- How many bytes of content are returned?

Example run

```
   397 12.997757    192.168.1.111   128.119.245.12   HTTP     617 GET /wireshark-labs/HTTP-wireshark-file1.html HTTP/1.1
   399 13.136494    128.119.245.12  192.168.1.111    HTTP     305 HTTP/1.1 304 Not Modified

▶ Frame 397: 617 bytes on wire (4936 bits), 617 bytes captured (4936 bits) on interface 0
▶ Ethernet II, Src: Apple_90:b9:76 (d0:e1:40:90:b9:76), Dst: Technico d7:50:66 (e0:b9:e5:d7:50:66)
▶ Internet Protocol Version 4, Src: 192.168.1.111, Dst: 128.119.245.12
▶ Transmission Control Protocol, Src Port: 56711, Dst Port: 80, Seq: 1, Ack: 1, Len: 551
▼ Hypertext Transfer Protocol
  ▼ GET /wireshark-labs/HTTP-wireshark-file1.html HTTP/1.1\r\n
    ▼ [Expert Info (Chat/Sequence): GET /wireshark-labs/HTTP-wireshark-file1.html HTTP/1.1\r\n]
        [GET /wireshark-labs/HTTP-wireshark-file1.html HTTP/1.1\r\n]
        [Severity level: Chat]
        [Group: Sequence]
      Request Method: GET
      Request URI: /wireshark-labs/HTTP-wireshark-file1.html
      Request Version: HTTP/1.1
  Host: gaia.cs.umass.edu\r\n
  Connection: keep-alive\r\n
  Cache-Control: max-age=0\r\n
  Upgrade-Insecure-Requests: 1\r\n
  User-Agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10_12_6) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/70.0.3538.77 Safari/537.36\r\n
  Accept: text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,image/apng,*/*;q=0.8\r\n
  Accept-Encoding: gzip, deflate\r\n
  Accept-Language: en-US,en;q=0.9,it;q=0.8\r\n
  If-None-Match: "80-57aae85e44920"\r\n
  If-Modified-Since: Thu, 15 Nov 2018 06:59:01 GMT\r\n
  \r\n
  [Full request URI: http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file1.html]
  [HTTP request 1/1]
  [Response in frame: 399]
```

Example run

```
    53 5.108002    192.168.1.111   128.119.245.12   HTTP     506 GET /wireshark-labs/HTTP-wireshark-file1.html HTTP/1.1
    55 5.246816    128.119.245.12  192.168.1.111   HTTP     552 HTTP/1.1 200 OK  (text/html)
```

```
▶ Frame 55: 552 bytes on wire (4416 bits), 552 bytes captured (4416 bits) on interface 0
▶ Ethernet II, Src: Technico_d7:50:66 (e0:b9:e5:d7:50:66), Dst: Apple_90:b9:76 (d0:e1:40:90:b9:76)
▶ Internet Protocol Version 4, Src: 128.119.245.12, Dst: 192.168.1.111
▶ Transmission Control Protocol, Src Port: 80, Dst Port: 57634, Seq: 1, Ack: 441, Len: 486
▼ Hypertext Transfer Protocol
  ▼ HTTP/1.1 200 OK\r\n
    ▼ [Expert Info (Chat/Sequence): HTTP/1.1 200 OK\r\n]
        [HTTP/1.1 200 OK\r\n]
        [Severity level: Chat]
        [Group: Sequence]
      Response Version: HTTP/1.1
      Status Code: 200
      [Status Code Description: OK]
      Response Phrase: OK
    Date: Thu, 15 Nov 2018 14:39:17 GMT\r\n
    Server: Apache/2.4.6 (CentOS) OpenSSL/1.0.2k-fips PHP/5.4.16 mod_perl/2.0.10 Perl/v5.16.3\r\n
    Last-Modified: Thu, 15 Nov 2018 06:59:01 GMT\r\n
    ETag: "80-57ad5e5e44020"\r\n
    Accept-Ranges: bytes\r\n
  ▶ Content-Length: 128\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
```

Cache

- Most of the browsers perform cache operations
- This means that they use the so-called conditional GET
- Basically, they ask whether a web page has been modified since a given date
  - In case the server sees a modified-after file, it returns it
  - Otherwise, the browser serves it from the cache

Example run - cache

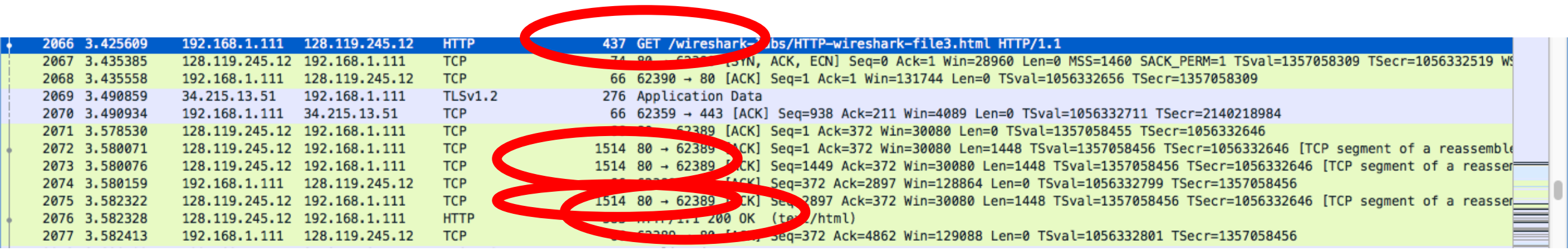
```
    397 12.997757    192.168.1.111   128.119.245.12    HTTP          617 GET /wireshark-labs/HTTP-wireshark-file1.html HTTP/1.1
    399 13.136494    128.119.245.12  192.168.1.111     HTTP          305 HTTP/1.1 304 Not Modified

▶ Frame 399: 305 bytes on wire (2440 bits), 305 bytes captured (2440 bits) on interface 0
▶ Ethernet II, Src: Technico_d7:50:66 (e0:b9:e5:d7:50:66), Dst: Apple_90:b9:76 (d0:e1:40:90:b9:76)
▶ Internet Protocol Version 4, Src: 128.119.245.12, Dst: 192.168.1.111
▶ Transmission Control Protocol, Src Port: 80, Dst Port: 56711, Seq: 1, Ack: 552, Len: 239
▼ Hypertext Transfer Protocol
  ▼ HTTP/1.1 304 Not Modified\r\n
    ▼ [Expert Info (Chat/Sequence): HTTP/1.1 304 Not Modified\r\n]
        [HTTP/1.1 304 Not Modified\r\n]
        [Severity level: Chat]
        [Group: Sequence]
      Response Version: HTTP/1.1
      Status Code: 304
      [Status Code Description: Not Modified]
      Response Phrase: Not Modified
  Date: Thu, 15 Nov 2018 14:31:55 GMT\r\n
  Server: Apache/2.4.6 (CentOS) OpenSSL/1.0.2k-fips PHP/5.4.16 mod_perl/2.0.10 Perl/v5.16.3\r\n
  Connection: Keep-Alive\r\n
  Keep-Alive: timeout=5, max=100\r\n
  ETag: "80-57aae95e44920"\r\n
  \r\n
  [HTTP response 1/1]
  [Time since request: 0.138737000 seconds]
  [Request in frame: 397]
```

HTTP – dealing with large files

- Make sure your browser cache is cleared
- Start wireshark packet sniffer
- Download http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file3.html
- Stop packet sniffing

- Answer these:
  - How many GET requests are sent?
  - What is the response status code?
  - How many packets are needed to download the whole page?

HTTP – dealing with large files



2066	3.425609	192.168.1.111	128.119.245.12	HTTP	437	GET /wireshark-labs/HTTP-wireshark-file3.html HTTP/1.1
2067	3.435385	128.119.245.12	192.168.1.111	TCP		[SYN, ACK, ECN] Seq=0 Ack=1 Win=28960 Len=0 MSS=1460 SACK_PERM=1 TSval=1357058309 TSecr=1056332519 WS
2068	3.435558	192.168.1.111	128.119.245.12	TCP	66	62390 → 80 [ACK] Seq=1 Ack=1 Win=131744 Len=0 TSval=1056332656 TSecr=1357058309
2069	3.490859	34.215.13.51	192.168.1.111	TLSv1.2	276	Application Data
2070	3.490934	192.168.1.111	34.215.13.51	TCP	66	62359 → 443 [ACK] Seq=938 Ack=211 Win=4089 Len=0 TSval=1056332711 TSecr=2140218984
2071	3.578530	128.119.245.12	192.168.1.111	TCP		62389 [ACK] Seq=1 Ack=372 Win=30080 Len=0 TSval=1357058455 TSecr=1056332646
2072	3.580071	128.119.245.12	192.168.1.111	TCP	1514	80 → 62389 [ACK] Seq=1 Ack=372 Win=30080 Len=1448 TSval=1357058456 TSecr=1056332646 [TCP segment of a reassemble
2073	3.580076	128.119.245.12	192.168.1.111	TCP	1514	80 → 62389 [ACK] Seq=1449 Ack=372 Win=30080 Len=1448 TSval=1357058456 TSecr=1056332646 [TCP segment of a reasser
2074	3.580159	192.168.1.111	128.119.245.12	TCP		[ACK] Seq=372 Ack=2897 Win=128864 Len=0 TSval=1056332799 TSecr=1357058456
2075	3.582322	128.119.245.12	192.168.1.111	TCP	1514	80 → 62389 [ACK] Seq=2897 Ack=372 Win=30080 Len=1448 TSval=1357058456 TSecr=1056332646 [TCP segment of a reasser
2076	3.582328	128.119.245.12	192.168.1.111	HTTP		HTTP/1.1 200 OK (text/html)
2077	3.582413	192.168.1.111	128.119.245.12	TCP		Seq=372 Ack=4862 Win=129088 Len=0 TSval=1056332801 TSecr=1357058456

HTTP – dealing with embedded objects

- Start wireshark packet capture
- Download http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file4.html
- Stop wireshark packet capture

- Answer these
  - How many HTTP GET requests are issued?
  - Were the two images requests in serial or parallel?

HTTP – dealing with embedded objects

1398	2.254098	192.168.1.111	128.119.245.12	HTTP	506	GET /wireshark-labs/HTTP-wireshark-file4.html HTTP/1.1
1404	2.394507	128.119.245.12	192.168.1.111	HTTP	1139	HTTP/1.1 200 OK (text/html)
1406	2.567983	192.168.1.111	128.119.245.12	HTTP	477	GET /pearson.png HTTP/1.1
1417	2.708508	128.119.245.12	192.168.1.111	HTTP	781	HTTP/1.1 200 OK (PNG)
1421	2.762795	192.168.1.111	128.119.245.12	HTTP	491	GET /~kurose/cover_5th_ed.jpg HTTP/1.1
1562	3.250975	128.119.245.12	192.168.1.111	HTTP	1472	HTTP/1.1 200 OK (JPEG JFIF image)

- 3 HTTP GET requests
- Images were requested in serial

HTTP – dealing with authentication procedures

- Start wireshark capture
- Go to http://gaia.cs.umass.edu/wireshark-labs/protected_pages/HTTP-wireshark- file5.html
- Username is **wireshark-students** and password is **network**
- Stop wireshark

- Answer these
  - What is the HTTP response to the first HTTP GET?
  - When you send the authentication, which new field is contained in the HTTP message?

HTTP – dealing with authentication procedures

```
    39 2.714979    192.168.1.111   128.119.245.12  HTTP      512 GET /wireshark-labs/protected_pages/HTTP-wireshark- HTTP/1.1
    41 2.902874    128.119.245.12  192.168.1.111   HTTP      783 HTTP/1.1 401 Unauthorized  (text/html)
   440 12.089870   192.168.1.111   128.119.245.12  HTTP      571 GET /wireshark-labs/protected_pages/HTTP-wireshark- HTTP/1.1
   444 12.324196   128.119.245.12  192.168.1.111   HTTP      586 HTTP/1.1 404 Not Found  (text/html)
   690 22.240954   192.168.1.111   128.119.245.12  HTTP      581 GET /wireshark-labs/protected_pages/HTTP-wireshark-file5.html HTTP/1.1
   692 22.380628   128.119.245.12  192.168.1.111   HTTP      556 HTTP/1.1 200 OK  (text/html)
```

```
▶ [Timestamps]
  TCP payload (717 bytes)
▼ Hypertext Transfer Protocol
  ▼ HTTP/1.1 401 Unauthorized\r\n
    ▼ [Expert Info (Chat/Sequence): HTTP/1.1 401 Unauthorized\r\n]
        [HTTP/1.1 401 Unauthorized\r\n]
        [Severity level: Chat]
        [Group: Sequence]
      Response Version: HTTP/1.1
      Status Code: 401
      [Status Code Description: Unauthorized]
      Response Phrase: Unauthorized
  Date: Thu, 15 Nov 2018 15:31:43 GMT\r\n
  Server: Apache/2.4.6 (CentOS) OpenSSL/1.0.2k-fips PHP/5.4.16 mod_perl/2.0.10 Perl/v5.16.3\r\n
  WWW-Authenticate: Basic realm="wireshark-students only"\r\n
▶ Content-Length: 381\r\n
  Keep-Alive: timeout=5, max=100\r\n
  Connection: Keep-Alive\r\n
```

HTTP – dealing with authentication procedures

```
    39  2.714979    192.168.1.111   128.119.245.12   HTTP     512  GET /wireshark-labs/protected_pages/HTTP-wireshark- HTTP/1.1
    41  2.902874    128.119.245.12  192.168.1.111    HTTP     783  HTTP/1.1 401 Unauthorized  (text/html)
   440  12.089870   192.168.1.111   128.119.245.12   HTTP     571  GET /wireshark-labs/protected_pages/HTTP-wireshark- HTTP/1.1
   444  12.324196   128.119.245.12  192.168.1.111    HTTP     586  HTTP/1.1 404 Not Found  (text/html)
   690  22.240954   192.168.1.111   128.119.245.12   HTTP     581  GET /wireshark-labs/protected_pages/HTTP-wireshark-file5.html HTTP/1.1
   692  22.380628   128.119.245.12  192.168.1.111    HTTP     556  HTTP/1.1 200 OK  (text/html)
```

```
▶ [Timestamps]
  TCP payload (505 bytes)
▼ Hypertext Transfer Protocol
  ▼ GET /wireshark-labs/protected_pages/HTTP-wireshark- HTTP/1.1\r\n
    ▼ [Expert Info (Chat/Sequence): GET /wireshark-labs/protected_pages/HTTP-wireshark- HTTP/1.1\r\n]
        [GET /wireshark-labs/protected_pages/HTTP-wireshark- HTTP/1.1\r\n]
        [Severity level: Chat]
        [Group: Sequence]
      Request Method: GET
      Request URI: /wireshark-labs/protected_pages/HTTP-wireshark-
      Request Version: HTTP/1.1
  Host: gaia.cs.umass.edu\r\n
  Connection: keep-alive\r\n
  Authorization: Basic d2lyZXNoYXJrLXN0dWRlbnRzOm5ldHdvcms=\r\n
    Credentials: wireshark-students:network
  Upgrade-Insecure-Requests: 1\r\n
  User-Agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10_12_6) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/70.0.3538.77 Safari/537.36\r\n
  Accept: text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,image/apng,*/*;q=0.8\r\n
```

HTTP Username and passwords

- Are sent in clear
  - Do you think that **d2lyZXNoYXJrLXN0dWRlbnRzOm5ldHdvcms=** is encrypted?
  - It is just a Base64 encoding
    - Just base64 decode **d2lyZXNoYXJrLXN0dWRlbnRz** and you get the username
    - Then base64 decode **Om5ldHdvcms=** and you get the password
    - Try it yourself at http://www.motobit.com/util/base64-decoder-encoder.asp