

**中国科学技术大学计算机学院**

## **计算机网络实验报告**

### **实验二**

## **利用 Wireshark 观察 http 报文**

学    号：PE20060014

姓    名：王晨

专    业：计算机科学与技术

指导老师：张信明

中国科学技术大学计算机学院

2020 年 10 月 26 日

## 一、 实验目的

- 1、 熟悉并掌握 Wireshark 的使用方法。
- 2、 通过抓取观察和分析 http 报文结构，理解 http 协议。

## 二、 实验原理

Wireshark 是非常流行的网络封包分析软件，功能十分强大。可以截取各种网络封包，显示网络封包的详细信息。Wireshark 使用 Npcap 作为接口，直接与网卡进行数据报文交换，监听共享网络上传送的数据包。Npcap 是替代 WinPcap 的新型 Windows 网络数据包截获软件。能够比原有的 WinPcap 数据包获得更好的抓包性能，并且稳定性更好。

## 三、 实验条件

- 1、 硬件条件：Mac
- 2、 软件条件：Mac OS Catalina

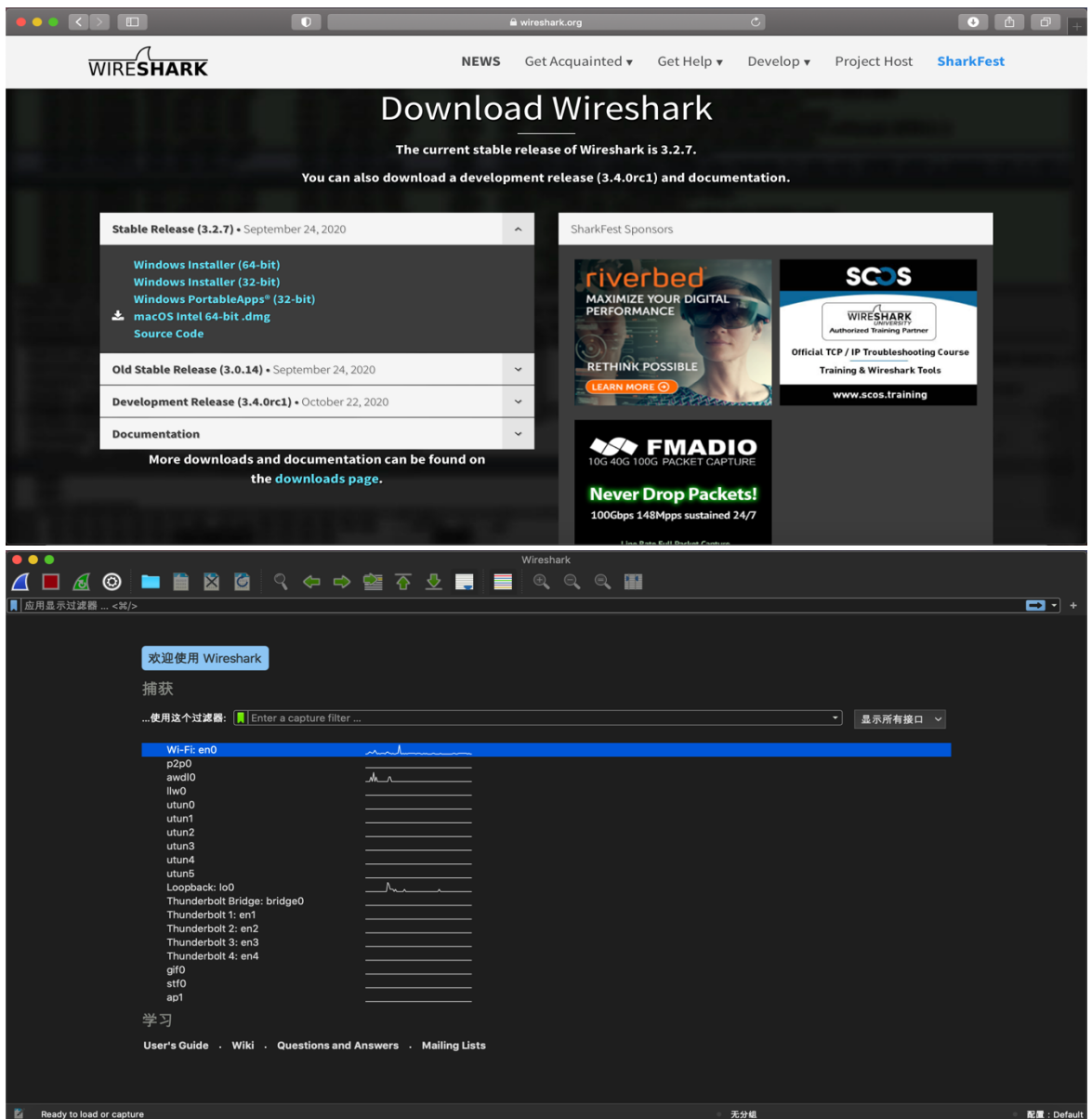
Chrome Web Browser

Wireshark 3.2.7

## 四、 实验过程

### 1、Wireshark 的安装

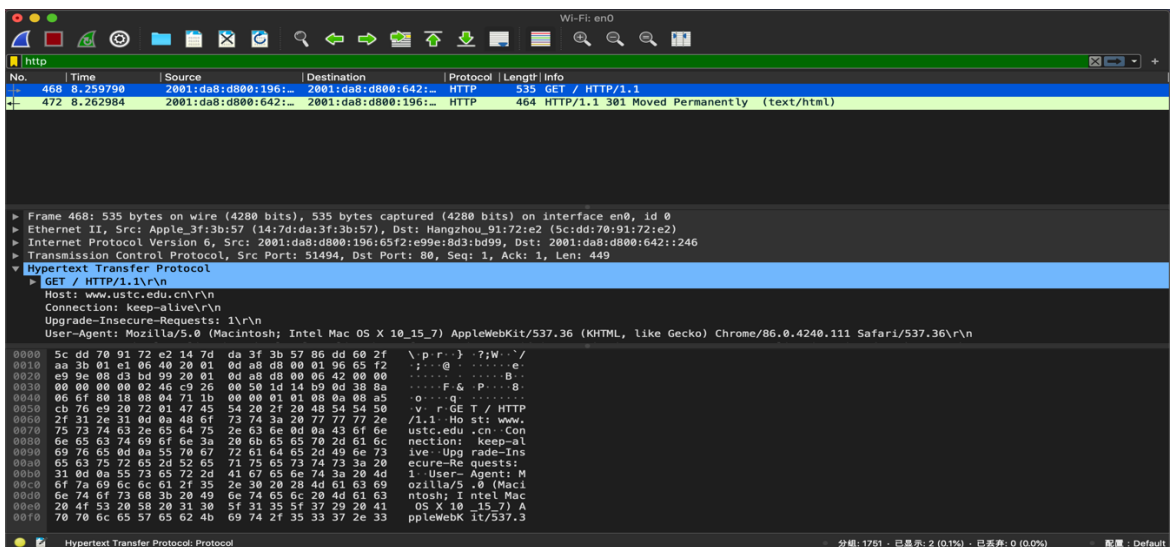
在网上下载安装最新 MacOS 版本，按照引导完成安装



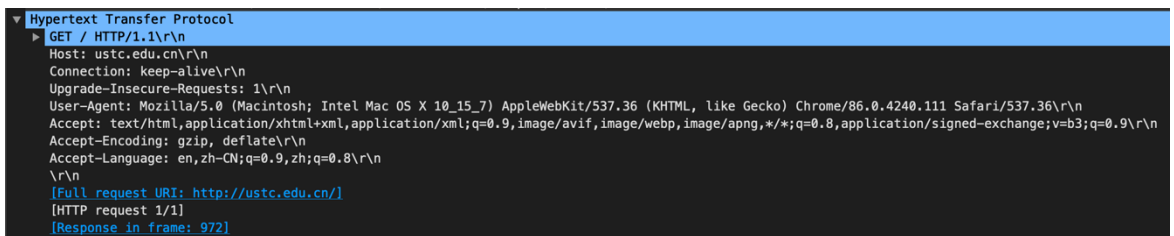
## 2、 利用 Wireshark 观察 http 报文并回答问题

### 0. 利用过滤筛选出 http 报文信息并分析

- (1)、启动浏览器（chrome，edge 等），清空浏览器缓存
- (2)、启动 wireshark，并启动捕获分组
- (3)、输入网址：ustc.edu.cn
- (4)、停止捕获，设置过滤 http 筛选出 http 条目，得到下图：



请求报文：



第一部分：请求行，用来说明请求类型,要访问的资源以及所使用的 HTTP 版本.

第二部分：请求头部，紧接着请求行（即第一行）之后的部分，用来说明服务器要使用的附加信息。例如：HOST 将指出请求的目的地。Connection 是连接维持方式。

第三部分：空行，请求头部后面的空行是必须的。即使第四部分的请求数据为空，也必须有空行。

第四部分：请求数据也叫主体，可以添加任意的其他数据。

响应报文：

```
▼ Hypertext Transfer Protocol
  HTTP/1.1 301 Moved Permanently\r\n
  Server: nginx\r\n
  Date: Tue, 27 Oct 2020 02:29:12 GMT\r\n
  Content-Type: text/html\r\n
  Content-Length: 162\r\n
  Connection: keep-alive\r\n
  Keep-Alive: timeout=20\r\n
  Location: https://ustc.edu.cn/\r\n
  \r\n
  [HTTP response 1/1]
  [Time since request: 0.003057000 seconds]
  [Request in frame: 969]
  [Request URI: http://ustc.edu.cn/]
  File Data: 162 bytes
▼ Line-based text data: text/html (7 lines)
  <html>\r\n
  <head><title>301 Moved Permanently</title></head>\r\n
  <body>\r\n
```

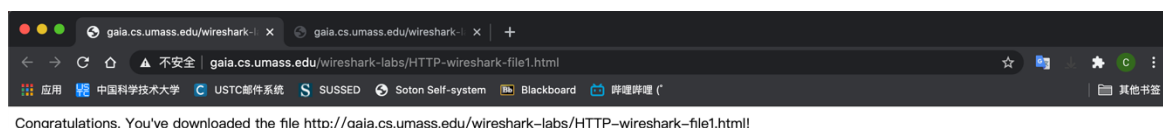
HTTP 响应也由三个部分组成，分别是：状态行、响应头部、响应包体。状态行（status line）通过提供一个状态码来说明所请求的资源情况。其中，HTTP-Version 表示服务器 HTTP 协议的版本；Status-Code 表示服务器发回的响应状态代码；Reason-Phrase 表示状态代码的文本描述。例如途中 301 就表示此地址被永久移动了。

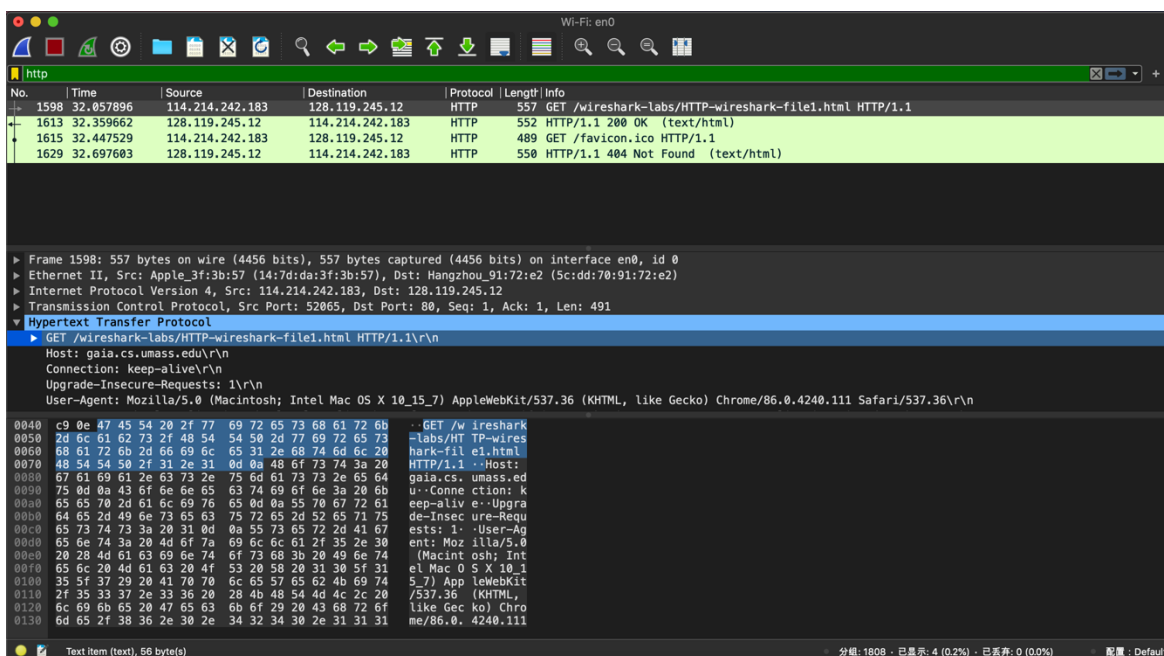
响应头部包含了一系列的附加信息，例如 Content-Type,Content-Length 等.最后部分为响应包体，包含了服务器向浏览器传输的报文实体。

## 1. The Basic HTTP GET/response interaction

按照上述方法打开 <http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file1.html>

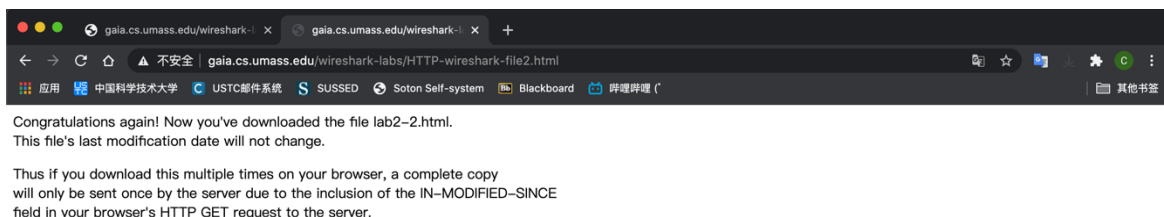
得到如下窗口：

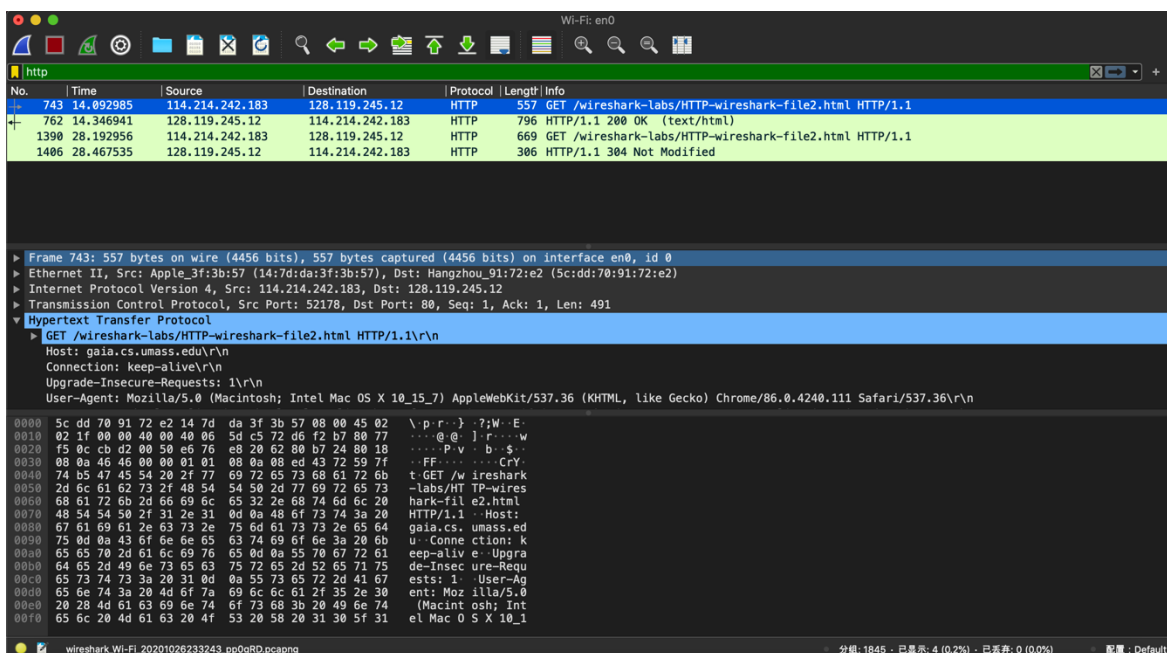




## 2. The HTTP CONDITIONAL GET/response interaction

清除浏览器缓存，重新开始捕获，重新打开浏览器，打开网页 <http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file2.html> 并快速刷新一次，停止捕获。得到如下窗口：



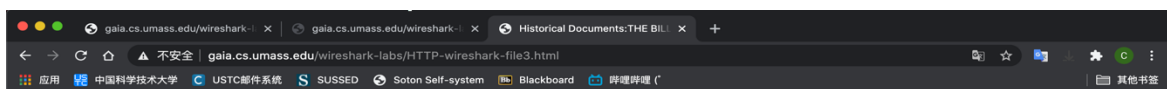


### 3. Retrieving Long Documents

清除浏览器缓存，重新开始捕获，重新打开浏览器，打开网页

<http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file3.html>，停止捕获。得到如下窗口：

□：



#### THE BILL OF RIGHTS Amendments 1-10 of the Constitution

The Conventions of a number of the States having, at the time of adopting the Constitution, expressed a desire, in order to prevent misconstruction or abuse of its powers, that further declaratory and restrictive clauses should be added, and as extending the ground of public confidence in the Government will best insure the beneficent ends of its institution;

Resolved, by the Senate and House of Representatives of the United States of America, in Congress assembled, two-thirds of both Houses concurring, that the following articles be proposed to the Legislatures of the several States, as amendments to the Constitution of the United States; all or any of which articles, when ratified by three-fourths of the said Legislatures, to be valid to all intents and purposes as part of the said Constitution, namely:

##### Amendment I

Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the government for a redress of grievances.

##### Amendment II

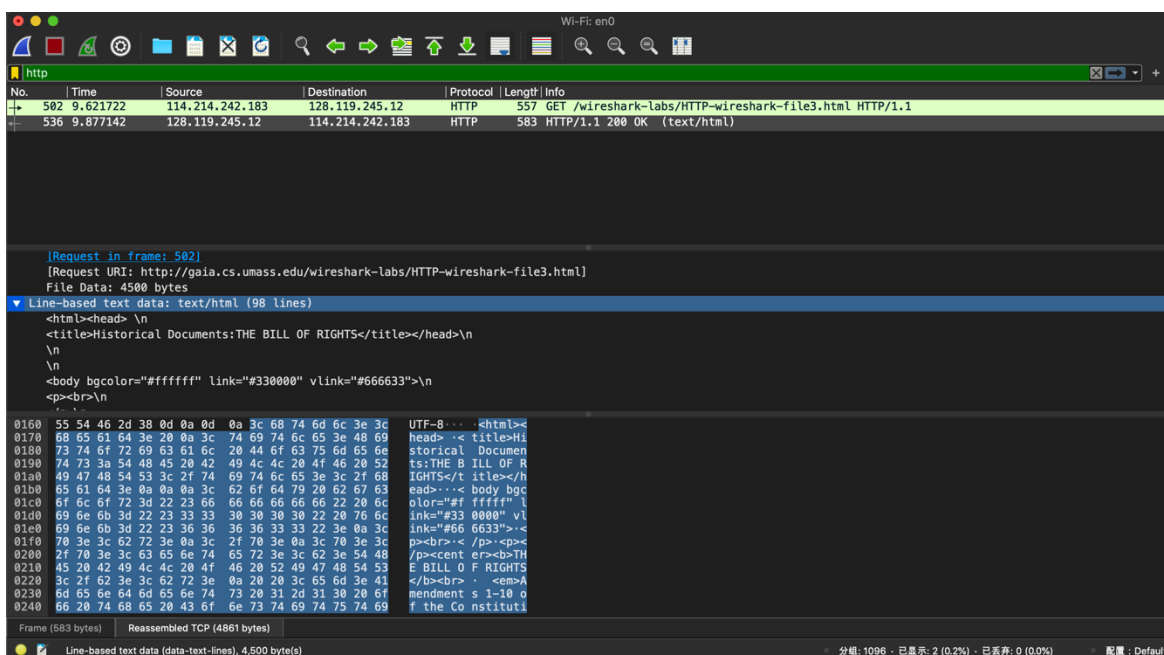
A well regulated militia, being necessary to the security of a free state, the right of the people to keep and bear arms, shall not be infringed.

##### Amendment III

No soldier shall, in time of peace be quartered in any house, without the consent of the owner, nor in time of war, but in a manner to be prescribed by law.

##### Amendment IV

The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no warrants shall issue, but upon probable cause, supported by oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.

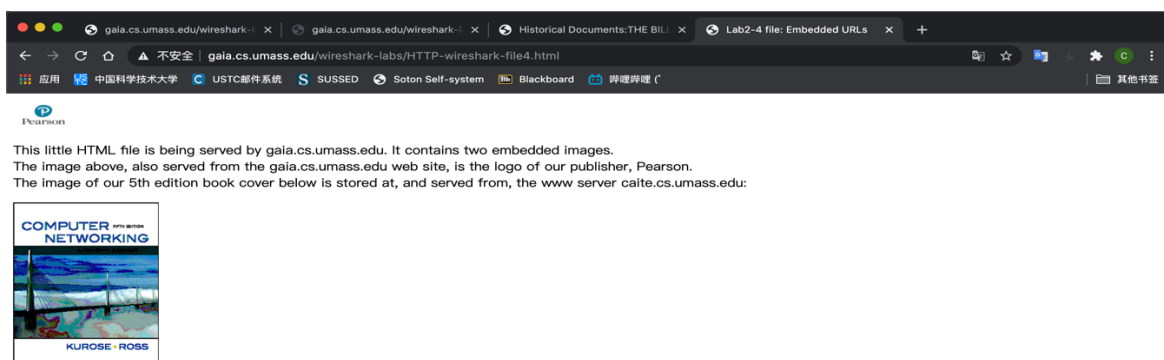


## 4. HTML Documents with Embedded Objects

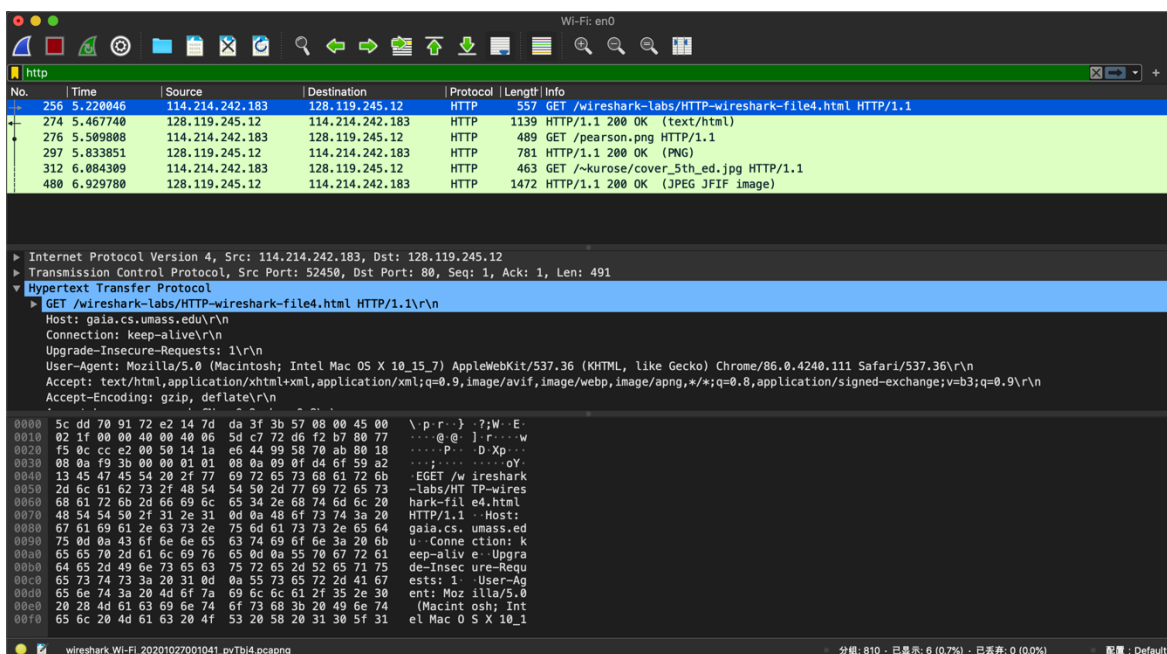
清除浏览器缓存，重新开始捕获，重新打开浏览器，打开网页

<http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file4.html> ,停止捕获。得到如下窗

口:



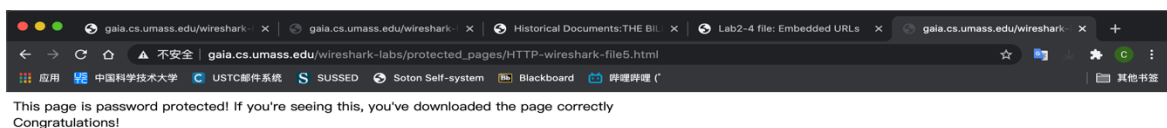


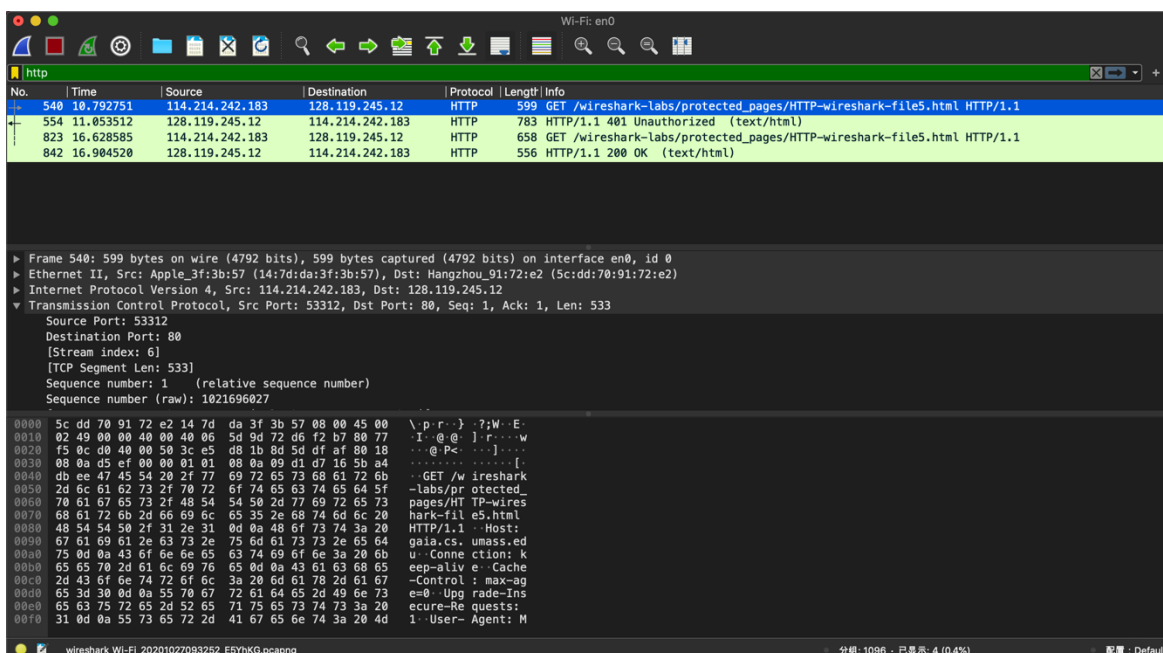


## 5 HTTP Authentication

清除浏览器缓存，重新开始捕获，重新打开浏览器，打开网页

[http://gaia.cs.umass.edu/wireshark-labs/protected\\_pages/HTTP-wireshark-file5.html](http://gaia.cs.umass.edu/wireshark-labs/protected_pages/HTTP-wireshark-file5.html)，输入用户名和密码，停止捕获。得到如下窗口：





## 五、 结果分析

# Questions in Wireshark Lab: HTTP

1. Is your browser running HTTP version 1.0 or 1.1? What version of HTTP is the server running?

都是 http-1.1

2. What languages (if any) does your browser indicate that it can accept to the server?

英文和中文:Accept-Language: en,zh-CN;q=0.9,zh;q=0.8\r\n

3. What is the IP address of your computer? Of the gaia.cs.umass.edu server?

我的 IP 是 114.214.242.183; Server 的 IP 是 128.119.245.12

4. What is the status code returned from the server to your browser?

Status code:200

5. When was the HTML file that you are retrieving last modified at the server?

Last-Modified: Mon, 26 Oct 2020 05:59:02 GMT

6. How many bytes of content are being returned to your browser?

128 个字节: [Content length: 128]

7. By inspecting the raw data in the packet content window, do you see any headers within the data that are not displayed in the packet-listing window? If so, name one.

For example, Server: Apache/2.4.6 (CentOS) OpenSSL/1.0.2k-fips PHP/7.4.11 mod\_perl/2.0.11 Perl/v5.16.3\r\n

8. 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?

第一条没有，但是第二条有。

9. Inspect the contents of the server response. Did the server explicitly return the contents of the file? How can you tell?

是的，Content-Length: 371\r\n; Line-based text data: text/html (10 lines)下面有具体的 10 行 text。

10. 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?

是的，If-Modified-Since: Mon, 26 Oct 2020 05:59:02 GMT\r\n,

这个时间和上一次相应报文的时: Last-Modified: Mon, 26 Oct 2020 05:59:02 GMT\r\n 是相同的。

11. 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.

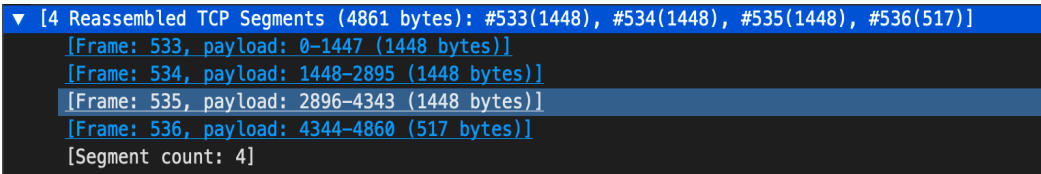
返回的 Status code 为 304 Not modified。服务器没用返回实际内容，这是由于浏览器对服务器的请求中包含了 If-Modified-Since 字段，因此服务器不会再次发送在这之前的同样文本。

12. How many HTTP GET request messages were sent by your browser?

一个。

13. How many data-containing TCP segments were needed to carry the single HTTP response?

4 个。



```
▼ [4 Reassembled TCP Segments (4861 bytes): #533(1448), #534(1448), #535(1448), #536(517)]
  [Frame: 533, payload: 0-1447 (1448 bytes)]
  [Frame: 534, payload: 1448-2895 (1448 bytes)]
  [Frame: 535, payload: 2896-4343 (1448 bytes)]
  [Frame: 536, payload: 4344-4860 (517 bytes)]
  [Segment count: 4]
```

14. What is the status code and phrase associated with the response to the HTTP GET request?

200 OK。

15. Are there any HTTP status lines in the transmitted data associated with a TCP-induced “Continuation”?

没有。HTTP 中没有这种消息。

16. How many HTTP GET request messages were sent by your browser? To which Internet addresses were these GET requests sent?

3 条 get 都向 128.119.245.12。

17. Can you tell whether your browser downloaded the two images serially, or whether they were downloaded from the two web sites in parallel? Explain.

似乎是 serially 的。因为两者的 get 请求相差了 0.35s，还是一个比较大的时间差。

18. What is the server's response (status code and phrase) in response to the initial HTTP GET message from your browser?

401 Unauthorized.

19. When your browser's sends the HTTP GET message for the second time, what new field is included in the HTTP GET message?

Authorization: Basic d2lyZXNoYXJrLXN0dWR1bnRzOm5ldHdvcm9m=\r\n

Credentials: wireshark-students:network

## Questions in Wireshark Lab: Getting Started

1. List up to 10 different protocols that appear in the protocol column in the unfiltered packet-listing window in step 7 above.

ARP, TCP, UDP, TLSv1, SSL, HTTP, DNS

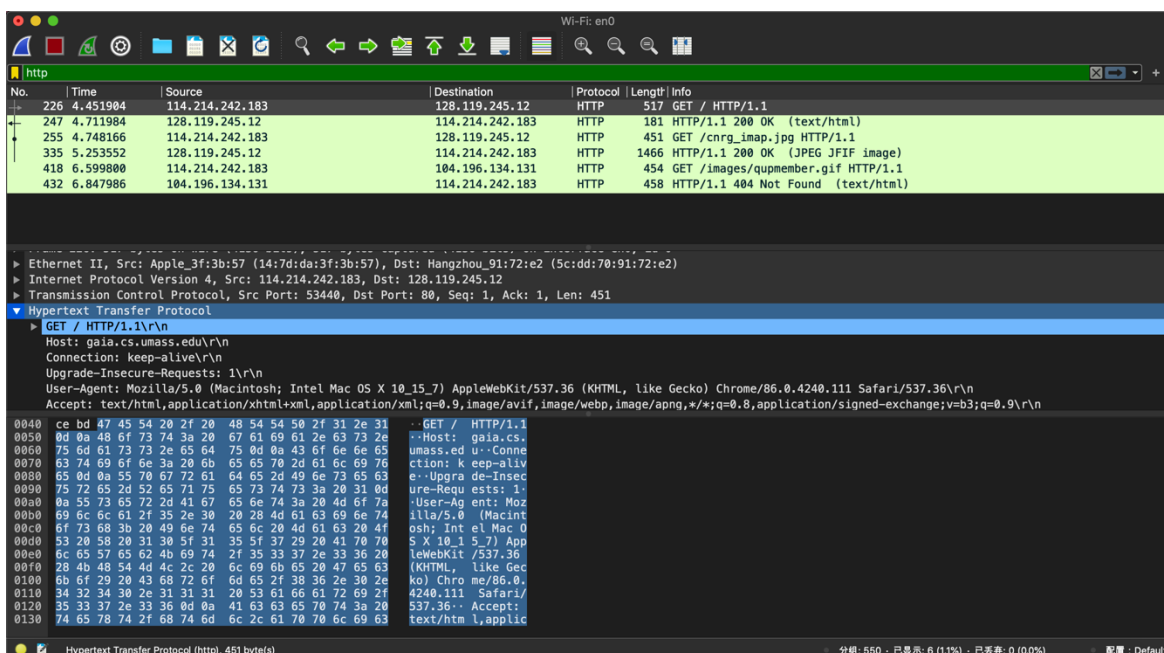
2. How long did it take from when the HTTP GET message was sent until the HTTP OK reply was received?

约 0.35s

3. What is the Internet address of the gaia.cs.umass.edu (also known as www-net.cs.umass.edu)? What is the Internet address of your computer?

我的 IP 是 114.214.242.183; Server 的 IP 是 128.119.245.12

4. Print the two HTTP messages displayed in step 9 above. To do so, select *Print* from the Wireshark *File* command menu, and select "*Selected Packet Only*" and "*Print as displayed*" and then click OK.



## 分析 HTTP 中 get 和 post 请求方式的区别。

1. get 一般用于从服务器上获取数据，post 一般用于向服务器请求提交数据，要提交的数据位于信息头后面的实体中。
2. get 是把参数数据队列加到提交表单的 ACTION 属性所指的 URL 中，值和表单内各个字段一一对应，在 URL 中可以看到。post 是通过 HTTPpost 机制，将表单内各个字段与其内容放置在 HTML HEADER 内一起传送到 ACTION 属性所指的 URL 地址。用户看不到这个过程。
3. 对于 get 方式，服务器端用 Request.QueryString 获取变量的值，对于 post 方式，服务器端用 Request.Form 获取提交的数据。
4. get 传送的数据量较小，不能大于 2KB。post 传送的数据量较大，一般被默认为不受限制。但理论上，IIS4 中最大量为 80KB，IIS5 中为 100KB。（get 和 post 的传送数据大小跟各个浏览器、操作系统以及服务器的限制也有关）。
5. get 安全性非常低，当通过 get 方法提交数据时，用户名和密码将出现在 URL 上。post 安全性较高。