**西北民族大学数学与计算机科学学院考查课程答题纸**

**（2022-2023学年第一学期 期末考试）**

**课程名称：计算机网络 课程代码： 91074B07**

**专 业：计算机科学与技术班 级： 1班**

**学 号：P201713305 姓 名： 谭秋林**

**总 分： 评阅教师（签字）**

**基于****Wireshark的IP、TCP、HTTP协议分析**

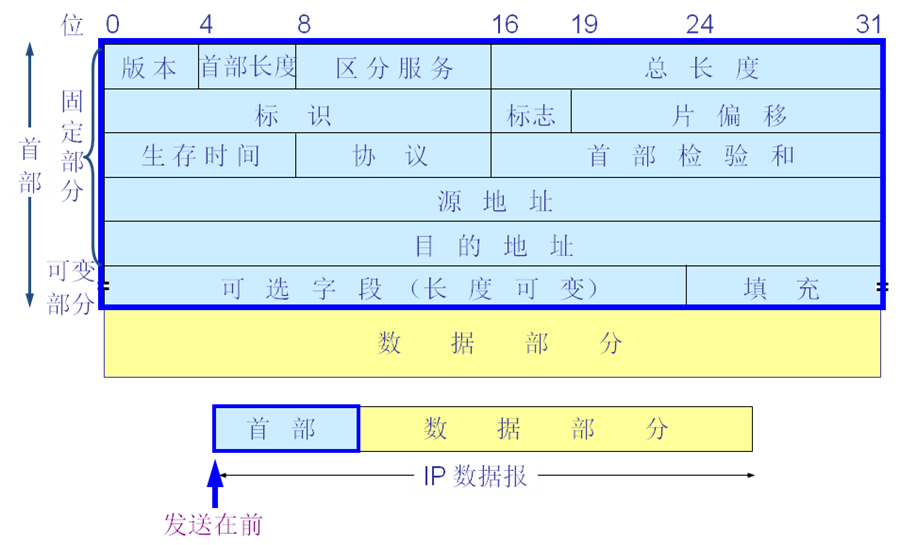
**第一章IP协议分析**

**目的：**首先本次掌握使用[Wireshark](https://so.csdn.net/so/search?q=Wireshark&spm=1001.2101.3001.7020)抓取TCP/IP协议数据包的技能，能够深入分析IP帧格式。通过抓包和分析数据包来理解TCP/IP协议，进一步提高理论联系实践的能力。

操作重点难点：1. 利用Wireshark抓IP包及IP包的分析。

2. 分析抓取到的IP包。

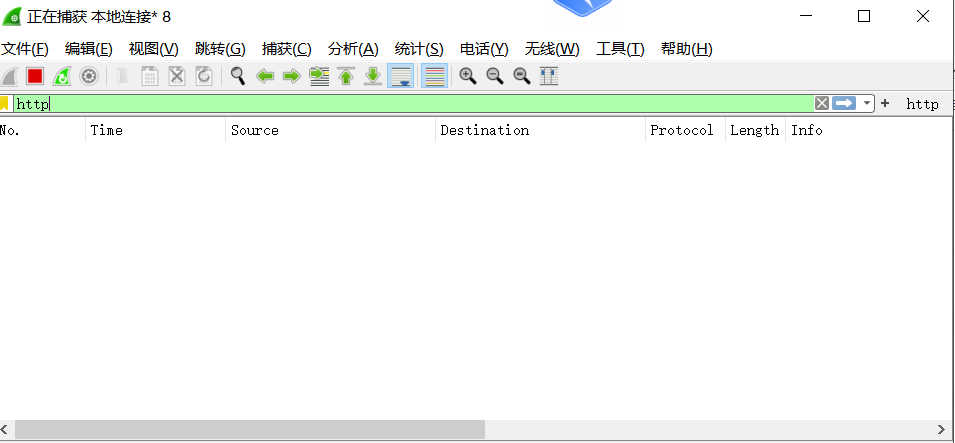
本次操作要抓的包，IP协议是因特网上的中枢。它定义了独立的网络之间以什么样的方式协同工作从而形成一个全球户联网。因特网内的每台主机都有IP地址。数据被称作数据报的分组形式从一台主机发送到另一台。每个数据报标有源IP地址和目的IP地址，然后被发送到网络中。如果源主机和目的主机不在同一个网络中，那么一个被称为路由器的中间机器将接收被传送的数据报，并且将其发送到距离目的端最近的下一个路由器。



**操作过程**

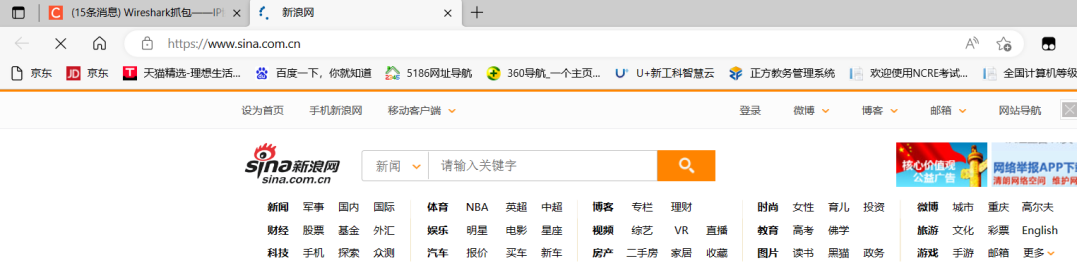
1. **第一步：协议确定，目标网站的确定**

在该步骤首先我们采取Wireshark软件并选择其中的HTTP服务，并确定了新浪网站作为抓取地址，其目标网址的URL为<http://www.sina.com.cn/>。



1. **第二步：抓取目标网址的包**

在该步骤中首先我们，启动wireshart软件，并且在第一步骤添加的过滤HTTP服务下启动抓包，并在浏览器的地址栏输入<http://www.sina.com.cn/>。

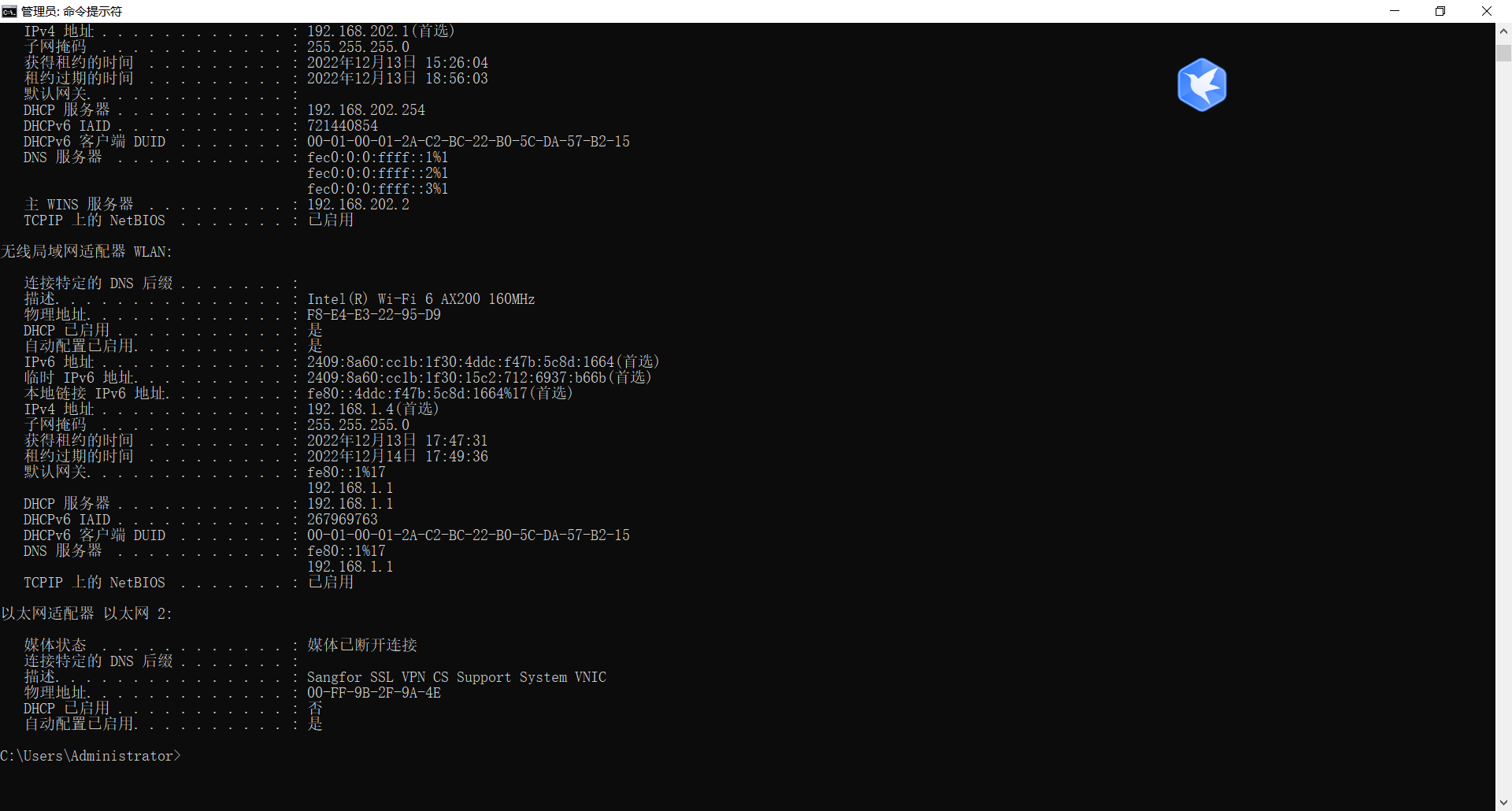


1. **第三步：保存数据包并处理**

wireshart显示过滤器得到先关数据包：通过抓包获得大量的数据包，为了对数据包分析的方便，需要使用过滤器，添加本机IP地址和IP协议过滤条件。

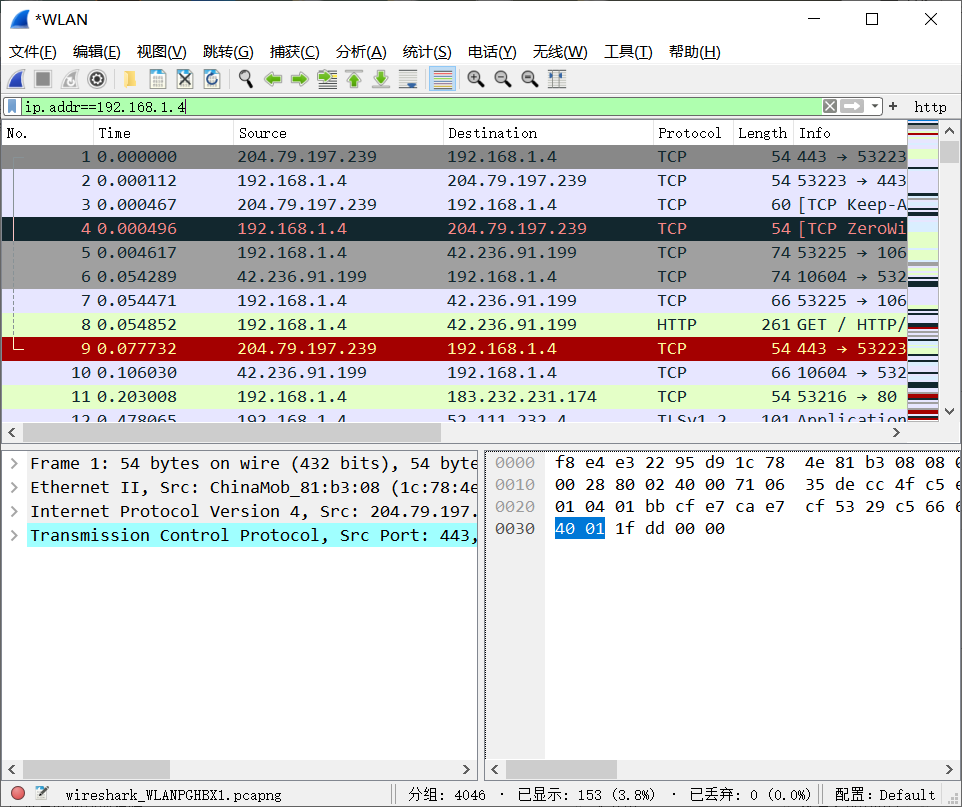
1. **获取本机IP地址**

打开本机的命令提示符号=，并通过ipconfig/all命令进行查看本机的IP地址。



1. **数据筛选**

在工具栏上的Filter对话框中填入过滤条件：ip.addr==192.168.1.4，过滤结果如下：



1. **选择报文类型**

双击点击一条tcp报文进入详细信息，那为什么不选Protocol类型为IP的协议呢。答案是没有，tcp报文正是基于ip协议的，tcp是传输层协议，而ip是它底下的网络层协议。

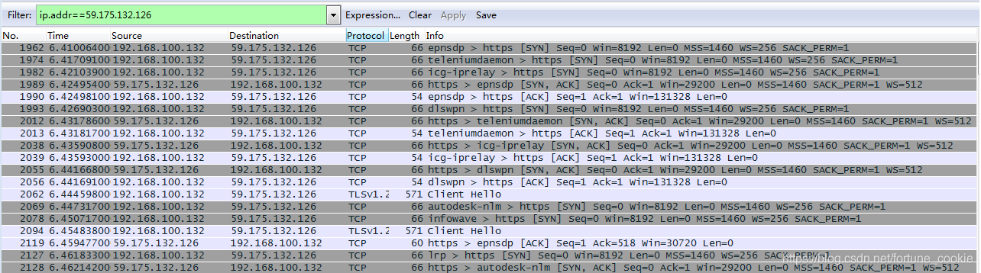
1. **TCP服务筛选**

网址地址查询：

文本

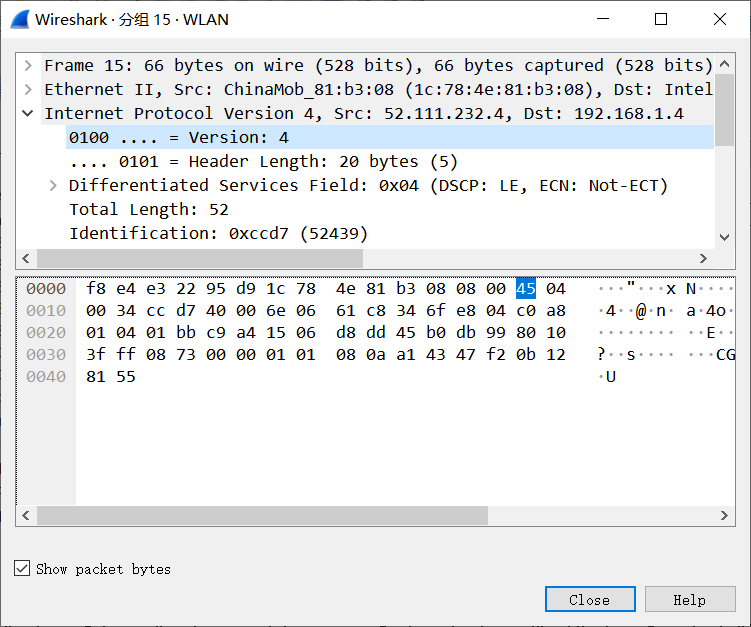
中度可信度描述已自动生成

查找到网址IP并进行筛选, ip.addr==59.175.132.126，过滤结果如下：



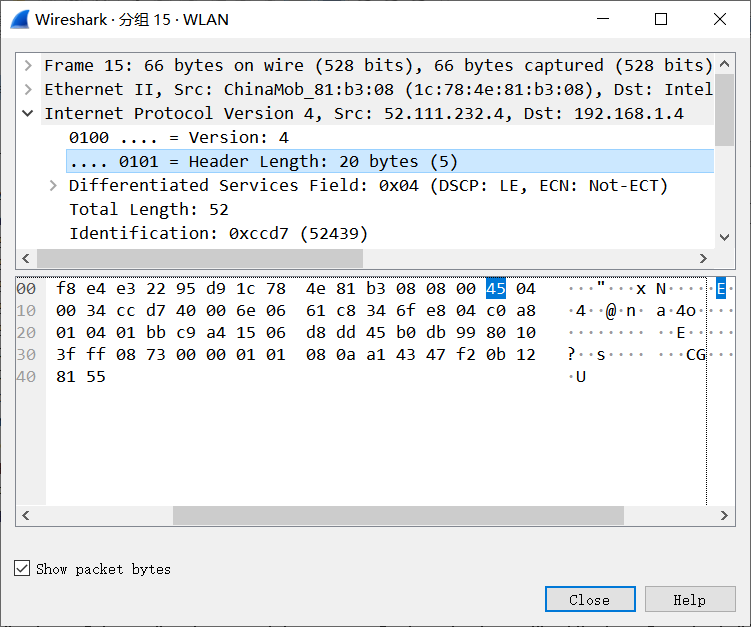
1. **第四步：协议分析**
2. **版本号部分**

版本(4bit)。IP报文中，版本占了4位，用来表示该协议采用的是那一个版本的IP，相同版本的IP才能进行通信，例如IPV4和IPV6的版本之间是不可以进行相互之间的消息传递。在进行程序编程时应该进行相关的IP协议的规定，确保版本之间的通信。



1. **首部长度字段**

首部长度(4bit)。该字段表示整个ip包头的长度，其中数的单位是4字节。即二进制数0000-1111（十进制数0-15），其中一个最小长度为0字节，最大长度为60字节。一般来说此处的值为0101，表示头长度为20字节。



1. **区分服务字段**

区分服务(8bit)。该字段用来获得更好的服务，在旧标准中叫做服务类型，但实际上一直未被使用过。1998 年这个字段改名为区分服务。只有在使用区分服（DiffServ）时，这个字段才起作用。在一般的情况下都不使用这个字段。

图形用户界面, 应用程序, Word

描述已自动生成

1. **总长度字段**

总长度(16bit)。该字段指首部和数据之和的长度，单位为字节，因此数据报的最大长度为 52439 字节。总长度必须不超过最大传送单元 MTU。

图形用户界面, 应用程序

描述已自动生成

1. **标志段区间**

标志(3bit)。标志(flag)占 3 位，目前只有前两位有意义。标志字段的最低位是 MF (More Fragment)。MF=1 表示后面“还有分片”。MF=0 表示最后一个分片。标志字段中间的一位是 DF (Don't Fragment) 。只有当 DF=0 时才允许分片。

图形用户界面, 应用程序, Word

描述已自动生成

1. **片偏移段区间**

片偏移(13 bit)。该字段指出较长的分组在分片后某片在原分组中的相对位置。片偏移以 8 个字节为偏移单位。

图形用户界面, 应用程序

描述已自动生成

1. **生存时间段区间**

生存时间(8 bit)。记为 TTL (Time To Live)数据报在网络中可通过的路由器数的最大值。

图形用户界面, 应用程序, Word

描述已自动生成

1. **协议区字段**

协议(8 bit)。该字段指出此数据报携带的数据使用何种协议以便目的主机的 IP 层将数据部分上交给哪个处理过程。

图形用户界面, 应用程序, Word

描述已自动生成

1. **首部检验区字段**

首部检验和(16 bit)。该字段只检验数据报的首部不检验数据部分

图形用户界面, 应用程序

描述已自动生成

1. **源地址/目的地址(32bit)**

图形用户界面, 应用程序, Word

描述已自动生成

图形用户界面, 应用程序, Word

描述已自动生成

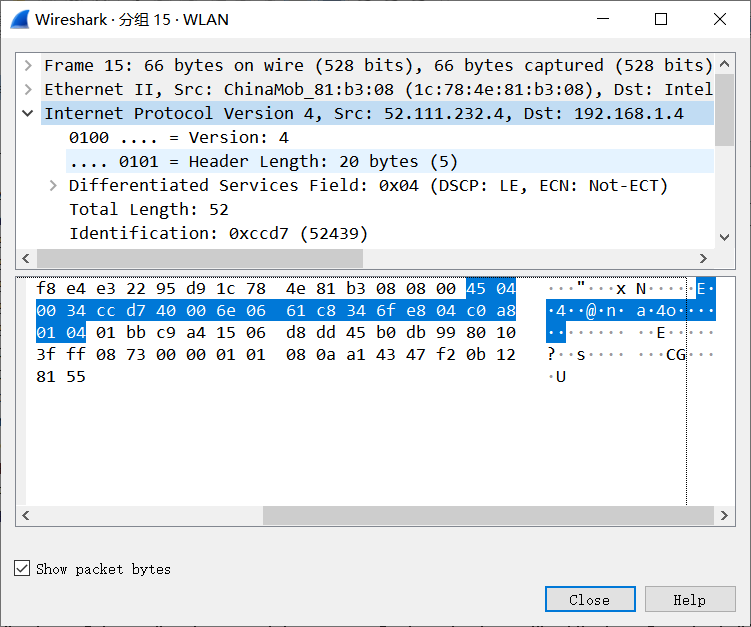
1. **可选字段**

可选字段，一般一些特殊的要求会加在这个部分。

图形用户界面, 应用程序

描述已自动生成

1. **数据字段**

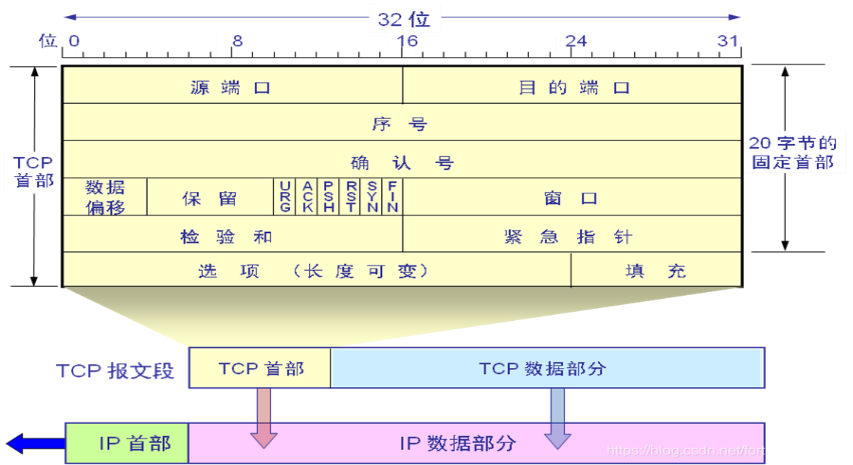


**第二章 TCP协议分析**

**目的：**掌握使用[Wireshark](https://so.csdn.net/so/search?q=Wireshark&spm=1001.2101.3001.7020)抓取TCP/IP协议数据包的技能，能够深入分析TCP帧格式及“TCP三次握手”。通过抓包和分析数据包来理解TCP/IP协议，进一步提高理论联系实践的能力。

操作重点难点：1. 利用Wireshark抓TCP包及TCP包的分析。 2. 分析抓取到的TCP包。

TCP协议是在计算机网络中使用最广泛的协议，很多的应用服务如FTP,HTTP,SMTP等在传输层都采用TCP协议，因此，如果要抓取TCP协议的数据包，可以在抓取相应的网络服务的数据包后，分析TCP协议数据包，深入理解协议封装，协议控制过程以及数据承载过程。两幅图分别是TCP帧格式及TCP三次握手。

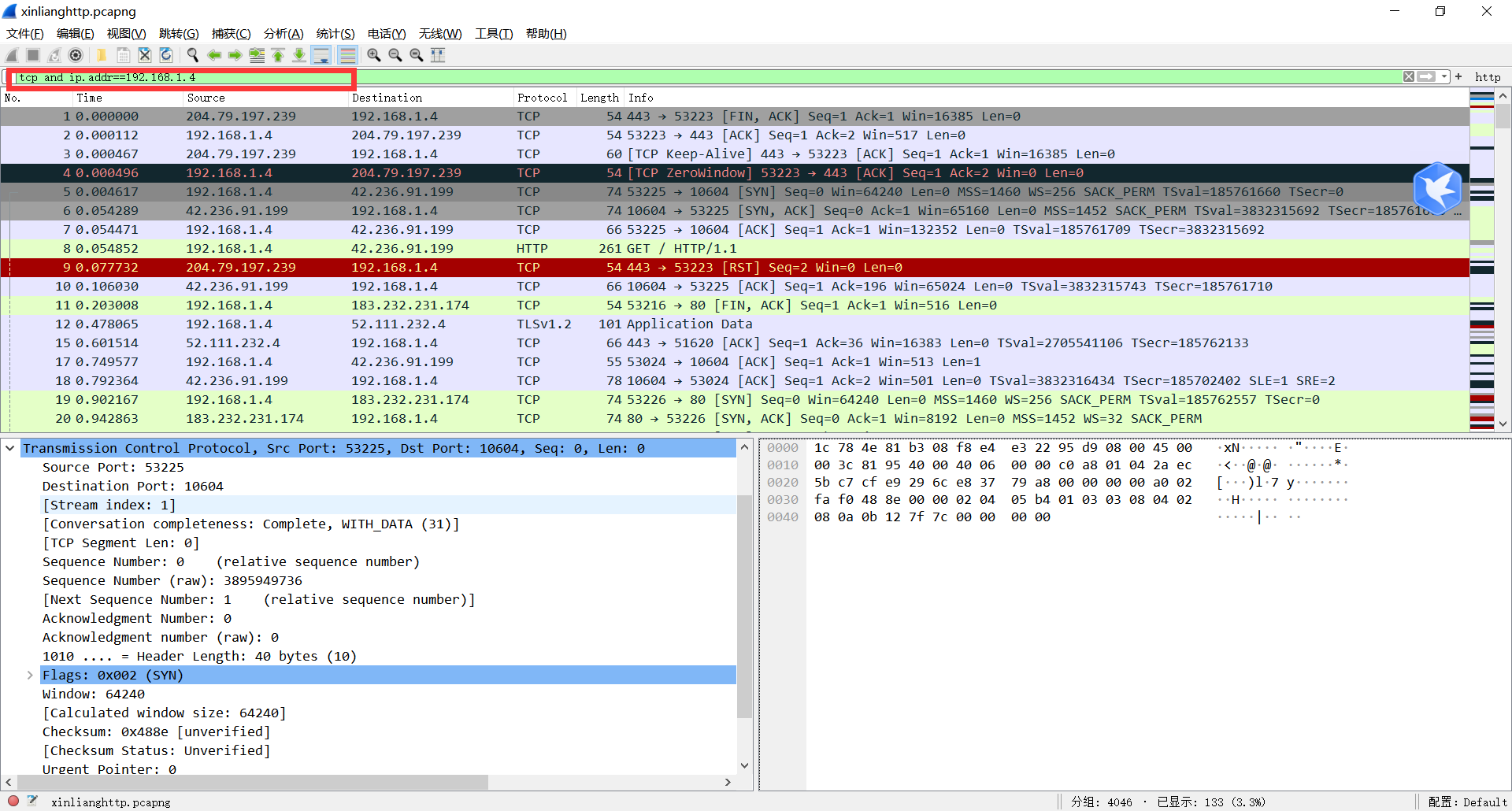


**操作过程：**

1. **第一步：条件筛选**

在IP协议抓包的数据中进行TCP协议的相关筛选。在上面

tcp and ip.addr==196.168.1.4，进行过滤筛选得到。



1. **第二步：分析三次握手**

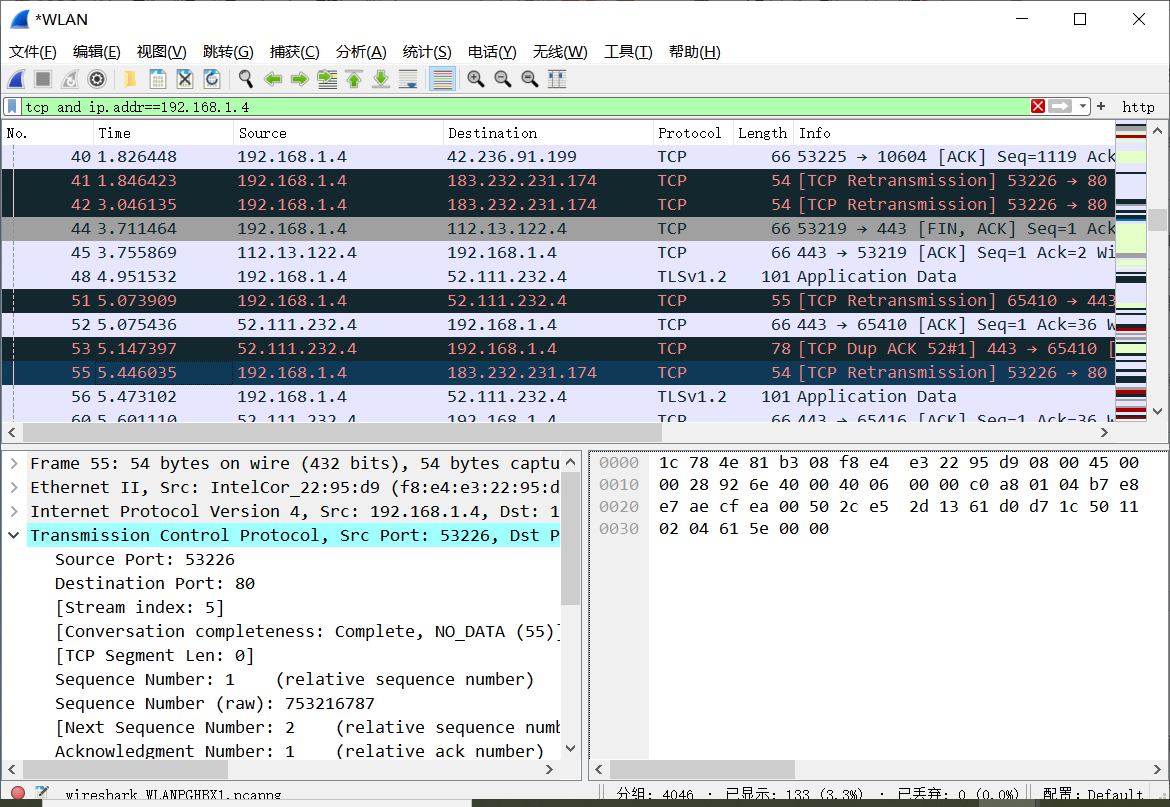
图形用户界面, 应用程序

描述已自动生成

1. **第三步：分析TCP协议**
2. **原端口/目的端口区段**

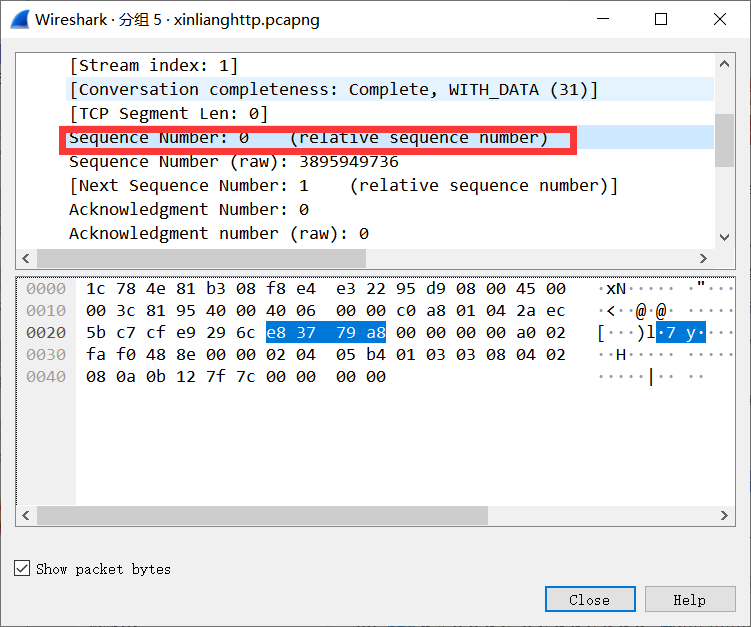
原端口/目的端口(16bit)。如下图所示，源端口为53226，标识了发送进程；目的端口为80，标识了接收方进程。

Sequence Number：发送序列号  
Acknowledgment Number：确认序列号  
Flags: SYN-同步序列号  
Window size value：窗口大小  
Checksum：检验和  
Urgent potiner：紧急指针



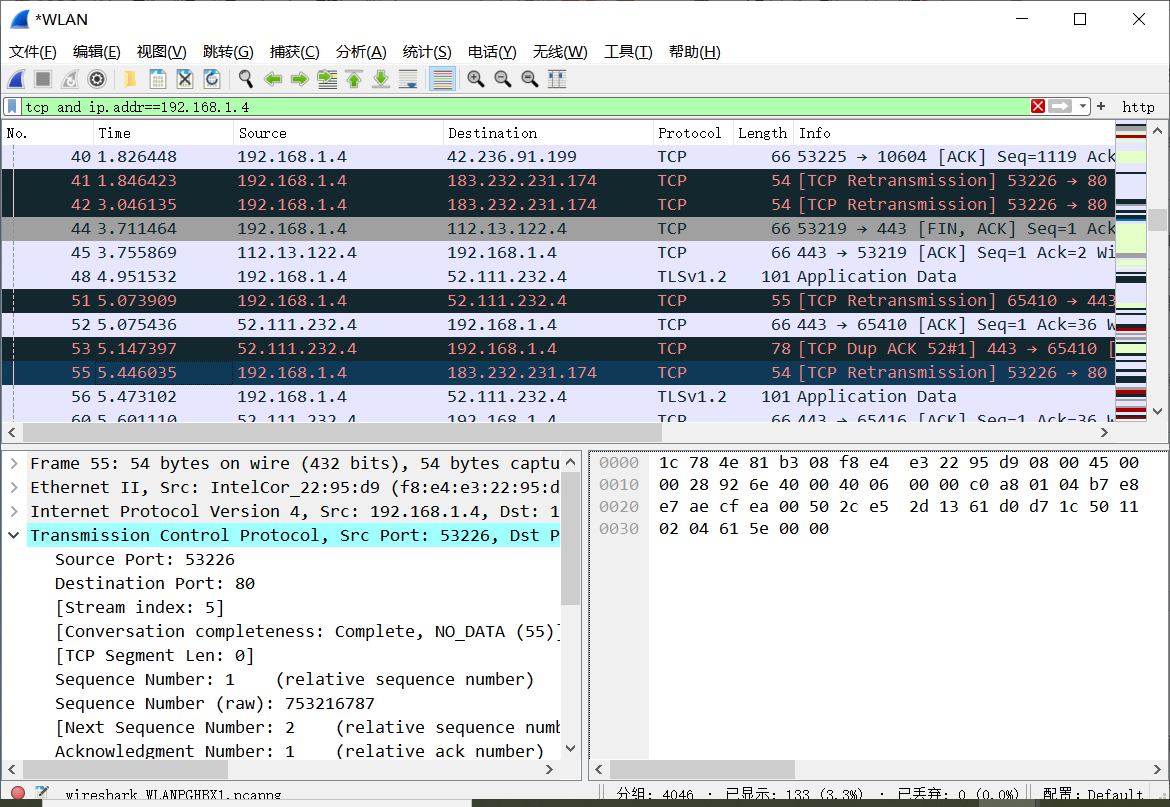
1. **序列号区段**

序列号(32bit)。如下图所示，发送序列号Sequence Number为0，标识从源端向目的端发送的数据字节流，它表示在这个报文端中的第一个数据字节的顺序号，序列号是32位的无符号类型，序号表达达到2^32 - 1后又从0开始， 当建立一个新的连接时，SYN标志为1，系列号将由主机随机选择一个顺序号ISN(Initial Sequence Number)



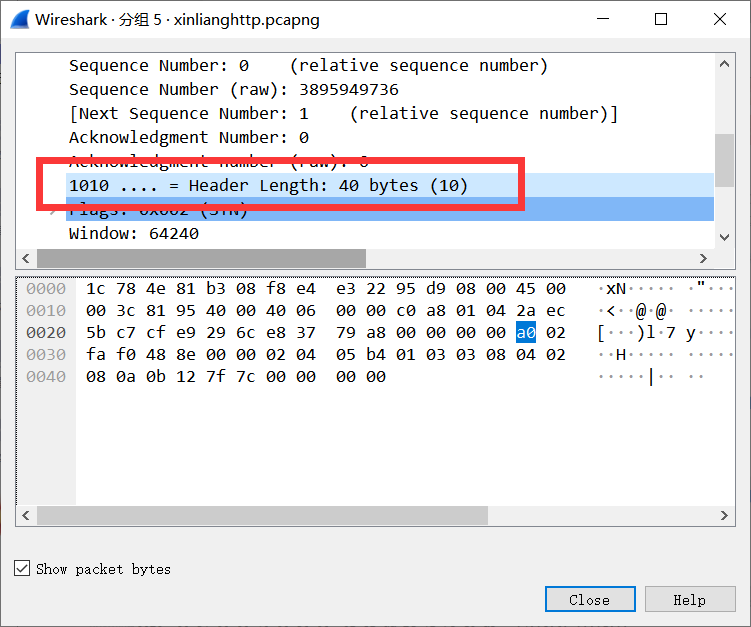
1. **确认号区段**

确认号(32bit)。如下图所示，确认号Acknowledgment Number为1，包涵了发送确认一端所期望收到的下一个顺序号。因此确认序列号应当是上次成功接收到数据的顺序号加1。只有ACK标志为1时确认序号字段才有效。TCP为应用层提供全双工服务，这意味着数据能在两个方向上独立的进行传输，因此连接的两断必须要保证每个方向上的传输数据顺序。



1. **偏移区段**

偏移(4bit)。如下图所示，偏移32bytes，这里的偏移实际指的是TCP首部的长度Header length，它用来表明TCP首部中32bit字的数目，通过它可以知道一个TCP包它的用户数据从哪里开始。



1. **保留位区段**

保留位(6bit)。如下图所示，保留位Reserved未设置。

图形用户界面, 应用程序

描述已自动生成

1. **标志位区段**

图形用户界面, 应用程序

描述已自动生成

URG(Urgent Pointer Field Significant):紧急指针标志，用来保证TCP连接不被中断，并且督促中间设备尽快处理这些数据，图中其值为1。

ACK(Acknowledgement Field Signigicant)：确认号字段，该字段为1时表示应答字段有效，即TCP应答号将包含在TCP报文中，图中其值为1。

PSH(Push Function): 推送功能，所谓推送功能指的是接收端在接收到数据后立即推送给应用程序，而不是在缓冲区中排队，图中其值为0。

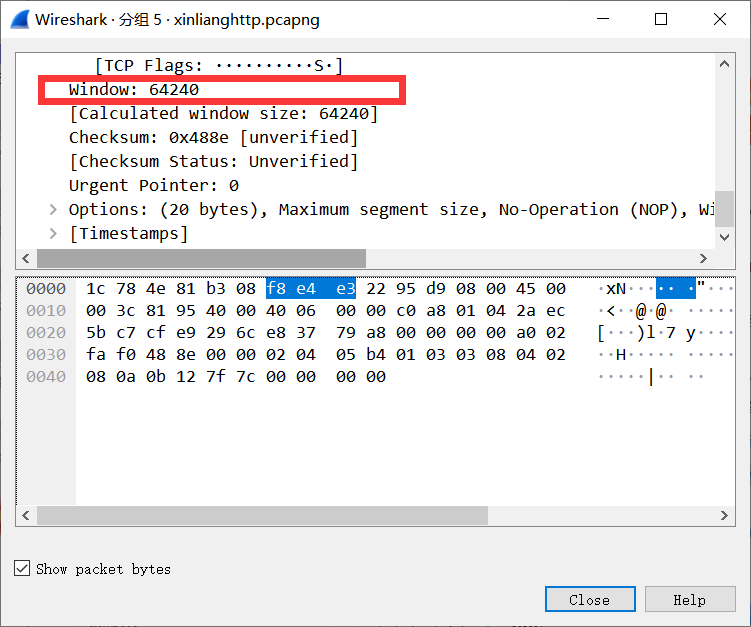
RST(Reset the connection): 重置连接，不过一搬表示断开一个连接，图中其值为0。

SYN(Synchronize sequence numbers):同步序列号，用来发起一个连接请求，图中其值为1。

FIN(No more data from sender)表示发送端发送任务已经完成（既断开连接）。

1. **窗口大小区段**

窗口大小(16bit)。 如下图所示，窗口大小Windows size value为64240，表示源主机最大能接收64240字节。



1. **校验和区段**

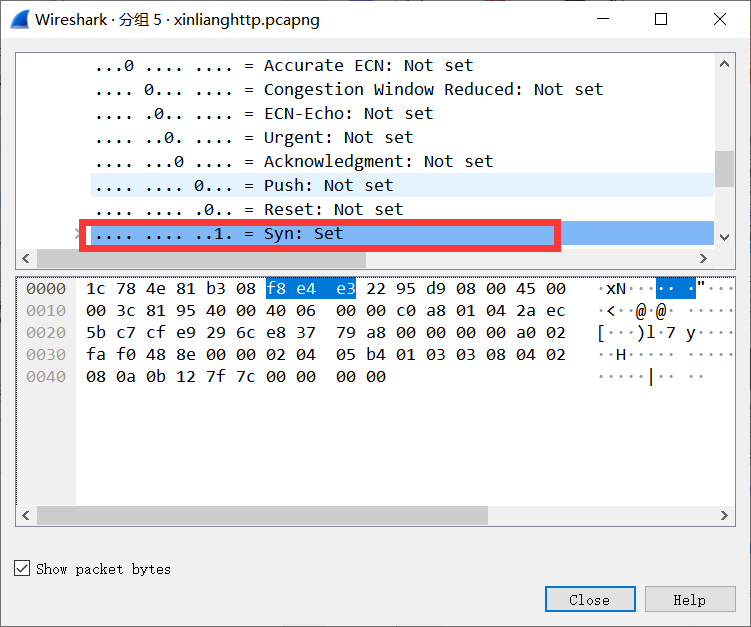
校验和(16bit)。如下图所示，校验和Checksum为0xc24f，包含TCP首部和TCP数据段，这是一个强制性的字段，一定是由发送端计算和存储，由接收端进行验证。

图形用户界面, 应用程序

描述已自动生成

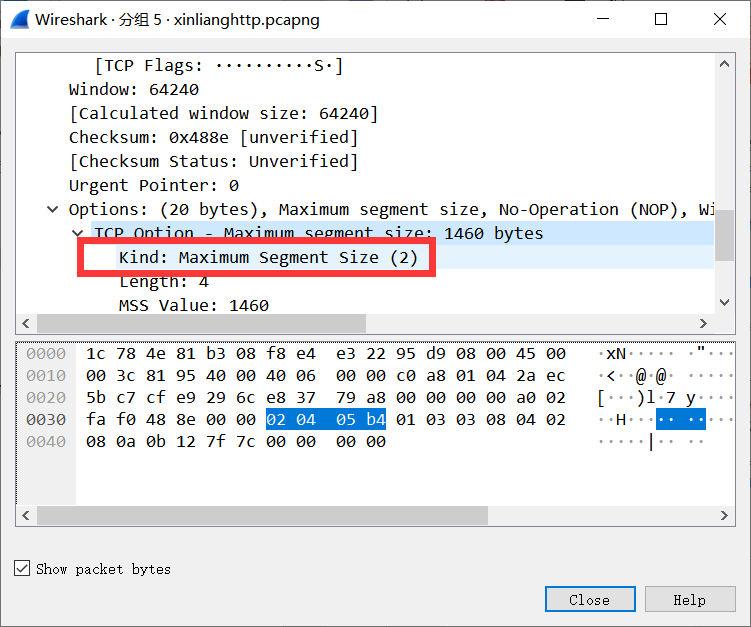
1. **紧急指针区段**

紧急指针(16bit)。如下图所示，URG标志为1，只有当URG标志置为1时该字段才有效，紧急指针是一个正的偏移量，和序号字段中的值相加表示紧急数据最后一个字节的序号。TCP的紧急方式是发送端向另一段发送紧急数据的一种方式。



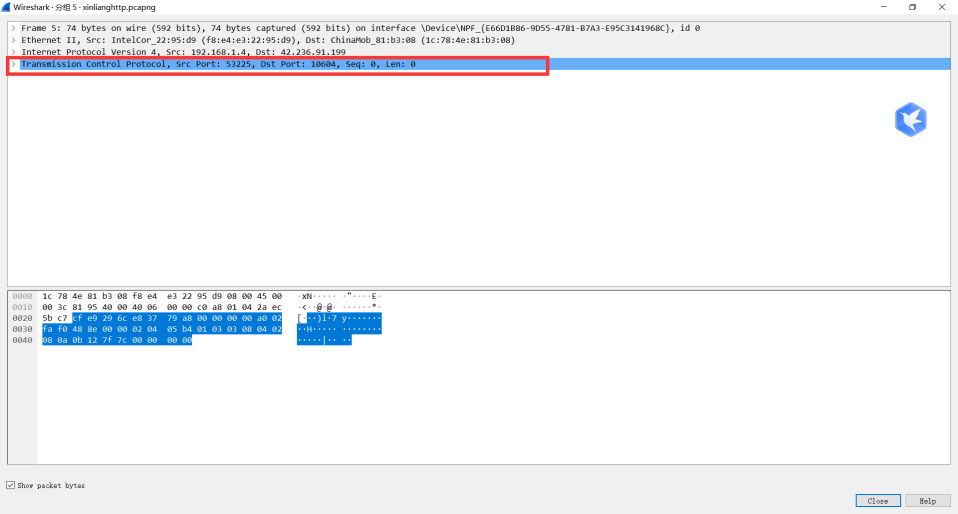
1. **TCP选项区段**

TCP选项。至少1个字节的可变长字段，标识哪个选项有效。Kind=0:选项表结束， Kind=1:无操作， Kind=2：最大报文段长度，Kind=3:窗口扩大因子， Kind=8:时间戳。如下图所示，Kind为2，代表最大报文长度MSS size。



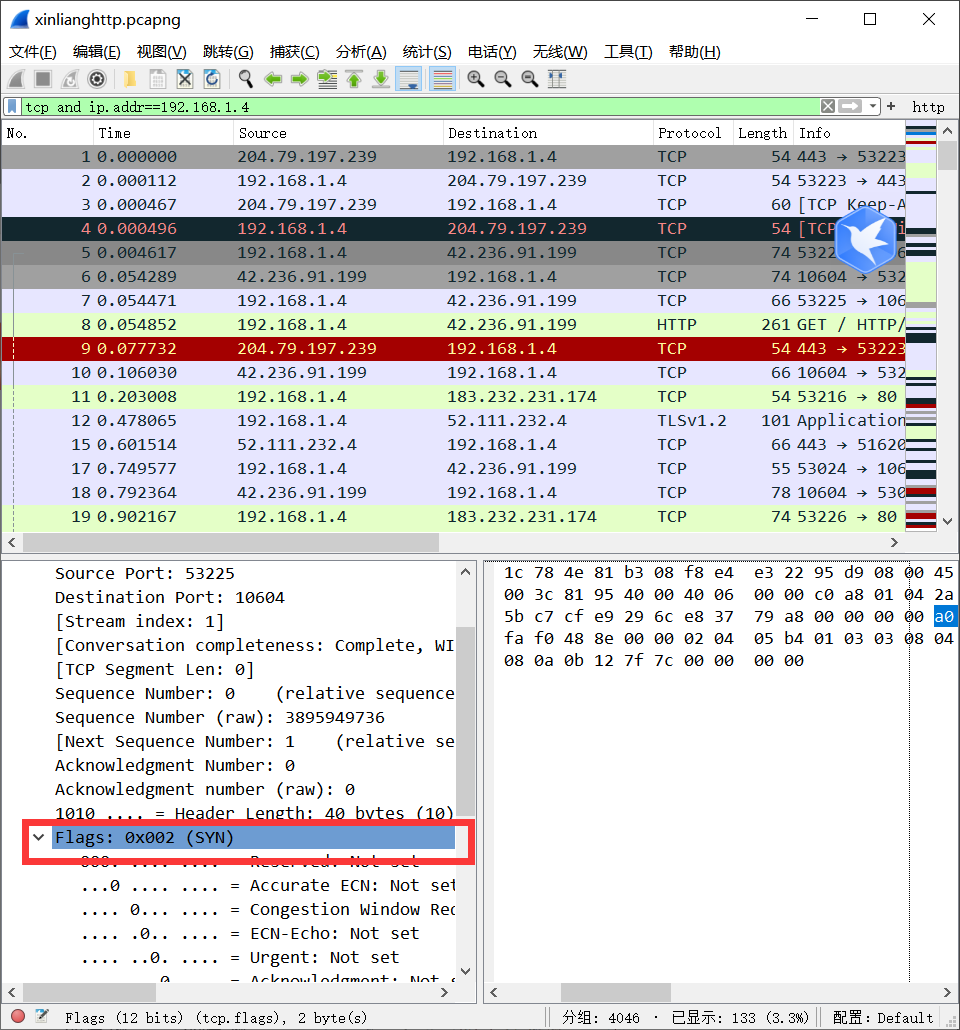
1. **数据部分区段**

数据部分。当前数据包的数据部分，如下图所示



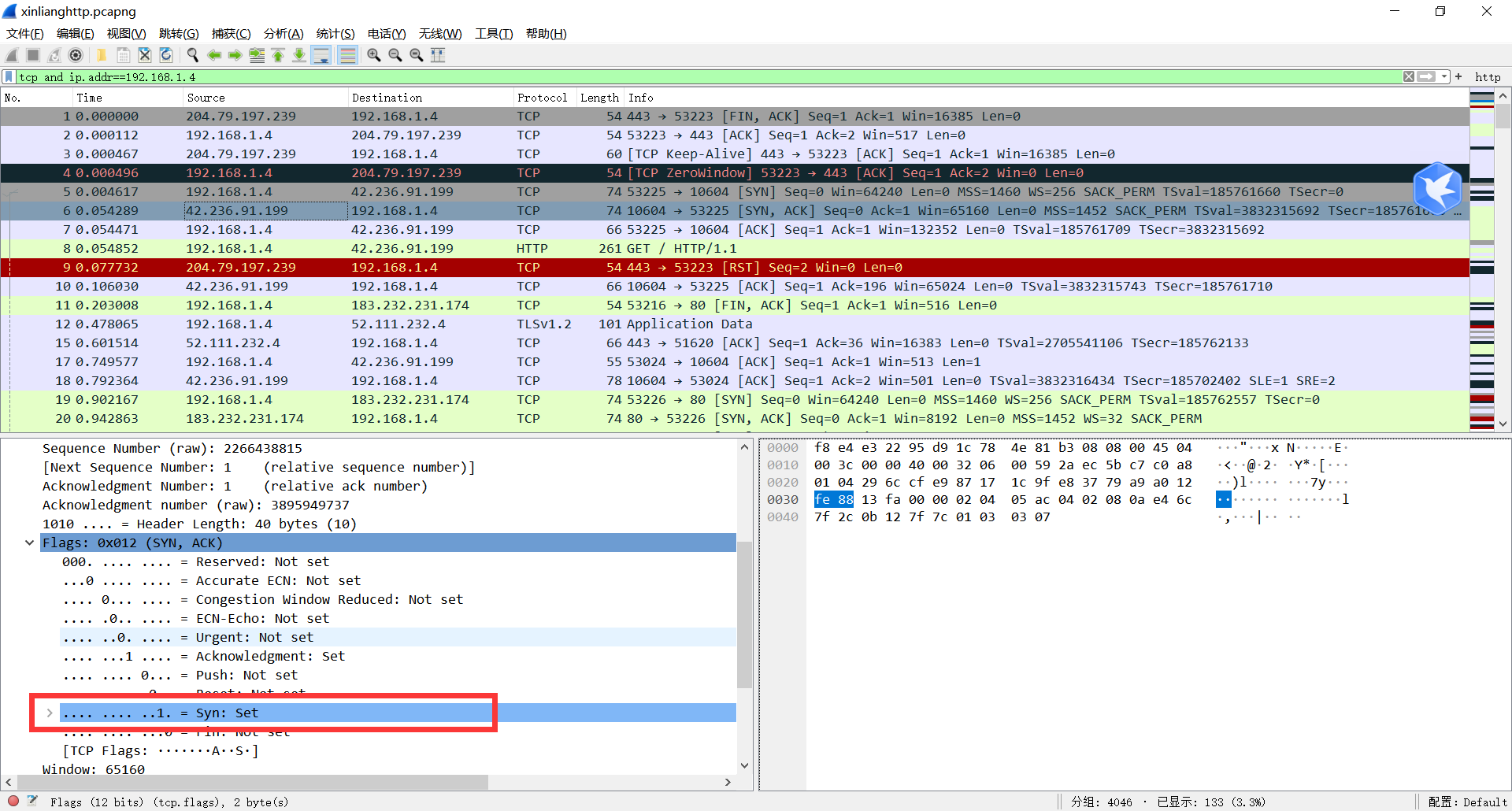
1. **第四步：TCP三次握手具体分析**
2. **第一次握手交互**

第一次握手数据包：客户端发送一个TCP，标志位为SYN，序列号为0， 代表客户端请求建立连接，如下图所示（第一条）：



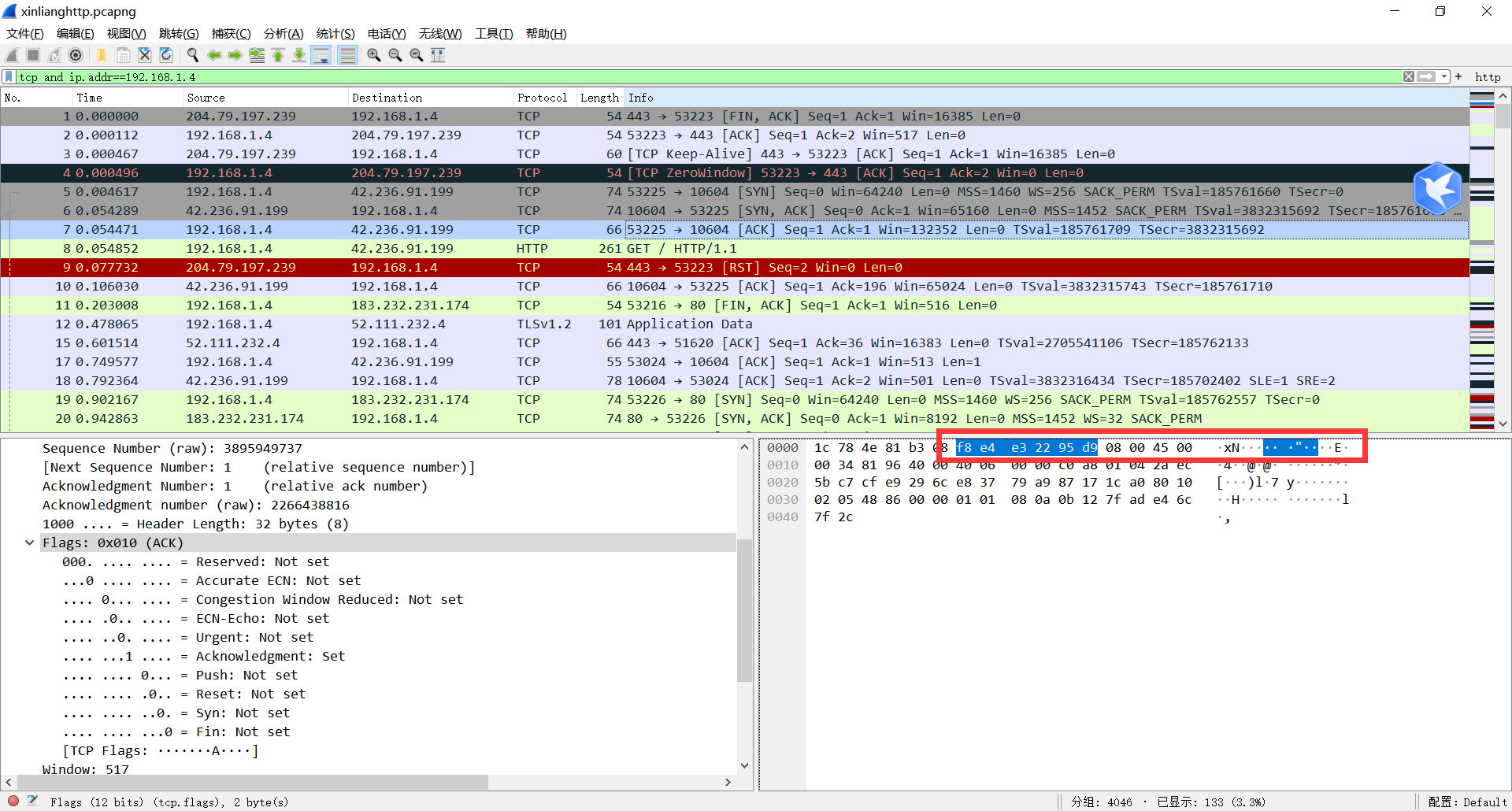
1. **第二次握手数据包**

第二次握手的数据包：服务器发回确认包, 标志位为 SYN,ACK. 将确认序号(Acknowledgement Number)设置为客户的I S N加1以.即0+1=1，如下图所示（第二条）：



1. **第三次握手数据包**

第三次握手的数据包：客户端再次发送确认包(ACK) SYN标志位为0,ACK标志位为1.并且把服务器发来ACK的序号字段+1,放在确定字段中发送给对方。在进过三次握手后和服务器建立了TCP连接，如下图所示（第三条）：



**第三章HTTP协议分析**

**目的：**熟悉并掌握Wireshark的基本操作，了解网络协议实体间的交互以及报文交换。

操作重点难点：1. 利用Wireshark抓http包及http包的分析。 2. 分析抓取到的Http包。

**基础理论**

HTTP（HyperText Transfer Protocol，超文本传输协议）是Web系统最核心的内容，是Web服务器和客户端直接进行数据传输的规则。HTTP协议是用于从WWW服务器传输超文本到本地浏览器的传送协议。可以使浏览器更加高效，使网络传输减少。不仅保证计算机正确快速地传输超文本文档，还确定传输文档中的哪一部分，以及首先显示(如文本先于图形)等。HTTP是一个应用层协议，有请求和响应构成，是一个标准的客户端服务器模型。HTTP具有以下几个特点：

**支持客户/服务器模式：**支持基本认证和安全认证；

**简单快速：**客户端向服务器请求服务时，只需传送请求方法和路径。请求方法常用的有GET、HEAD、POST。。由于HTTP协议简单，使得HTTP服务器的程序规模小，因而通信速度很快；

**灵活**：HTTP允许传输任意类型的数据对象。正在传输的类型由Content-Type加以标记；

**HTTP 0.9和1.0使用非持续连接：**限制每次连接只处理一个请求，服务器处理完客户的请求，并收到客户的应答后，即断开连接。采用这种方式可以节省传输时间。HTTP 1.1使用持续连接：不必为每个web对象创建一个新的连接，一个连接可以传送多个对象；

**无状态：**HTTP协议是无状态协议。无状态是指协议对于事务处理没有记忆能力。缺少状态意味着如果后续处理需要前面的信息，则它必须重传，这样可能导致每次连接传送的数据量增大。

**HTTP的请求格式**  
**(1)请求方法URI协议/版本**

请求方法URI协议/版本：请求的第一行是"方法URI协议/版本"  
例如:GET/sample.jsp HTTP/1.1 其中"GET"代表请求方法，"/sample.jsp"表示URI，"HTTP/1.1代表协议和协议的版本。

请求方法的8种方法

OPTIONS：返回服务器针对特定资源所支持的HTTP请求方法，也可以利用向Web服务器发送"\*"的请求来测试服务器的功能性。

HEAD：向服务器索要与GET请求相一致的响应，只不过响应体将不会被返回。这一方法可以在不必传输整个响应内容的情况下，就可以获取包含在响应消息头中的元信息。

GET：向特定的资源发出请求。注意：get方法不应当被用于产生“副作用”的操作中。例如在Web APP中，其中一个原因是GET可能会被网站蜘蛛等随意访问。

POST：向指定资源提交数据进行处理请求（比如提交表单或者上传文件）。数据被包含在请求体中。POST请求可能会导致新的资源的建立或已有资源的修改。

PUT：向指定资源位置上传其最新内容。

DELETE：请求服务器删除Request-URI所标识的资源。

TRACE：回显服务器收到的请求，主要用于测试或者诊断。

CONNECT：HTTP/1.1协议中预留给能够将连接改为管道方式的代理服务器。  
**(2) 请求头(Request Header)**

- Accept:[表示浏览器可以接受文本，网页图片等]

- Accept-Charaset: [表示接受字符编码]

- Accept-Encoding:[可以接受格式压缩后数据]

- Accept-Language:[浏览器支持的语言为中文]

- Host:[浏览器要找的主机]

- IF-MODIFIED-Since:[文件在缓存中且指明文件时间]

- Referer:[指明来源]

- User-Agent:[本机浏览器内核]

- Connection:close/Keep-Alive [保持链接,即发完数据后不关闭链接]

- Date:[浏览器发送数据的请求时间]  
**(3) 请求正文**

 请求头和请求正文之间是一个空行，这个行非常重要，它表示请求头已经结束，接下来的是请求正文。请求正文中包含客户提交的查询字符串信息:

**HTTP响应格式**  
**(1)响应首行**

包含协议版本，响应状态码，对状态码的解释。  
例如：分析”HTTP/1.1 200 OK”  
其中HTTP/1.1：为协议版本。200 ：为响应状态码，此表示响应成功。OK ：为解释响应状态码200是响应成功。

常见状态响应码6个：

200 请求成功，浏览器把响应回来的信息显示在浏览器端。

404 客户端出错，在浏览器端请求一个不存在的资源时会出现404状态码。

405 客户端错误的一种，表示当前的请求方式不支持。如：服务器端只对GET请求做了处理，而客户端的请求是post方式的，这个时候会出现405状态码。

500 服务器端错误，如：服务器端代码出现空指针等异常，浏览器就会收到服务器发送的500状态码。

302 表示重定向。如：浏览器访问一个资源，服务器响应给浏览器一个302的状态码，并且通过响应头Location发送了一个新的url，告诉浏览器去请求这个url。这就是重定向。

304 第一次访问一个资源后，浏览器会将该资源缓存到本地，第二次再访问该资源时，若该资源没有发生改变，则服务器响应给浏览器304状态码，告诉浏览器使用本地缓存的资源。  
**(2)响应头信息**

-Server：服务器告诉浏览器，当前响应的服务类型和版本。

-Content-Type：服务器告诉浏览器响应内容的类型和字符编码。如：值为text/html;charset=utf-8。说明响应信息的类型是文本类型中的html，使用的字符编码是utf-8。

-Content-Length：服务器告诉浏览器，响应内容的长度字节数。

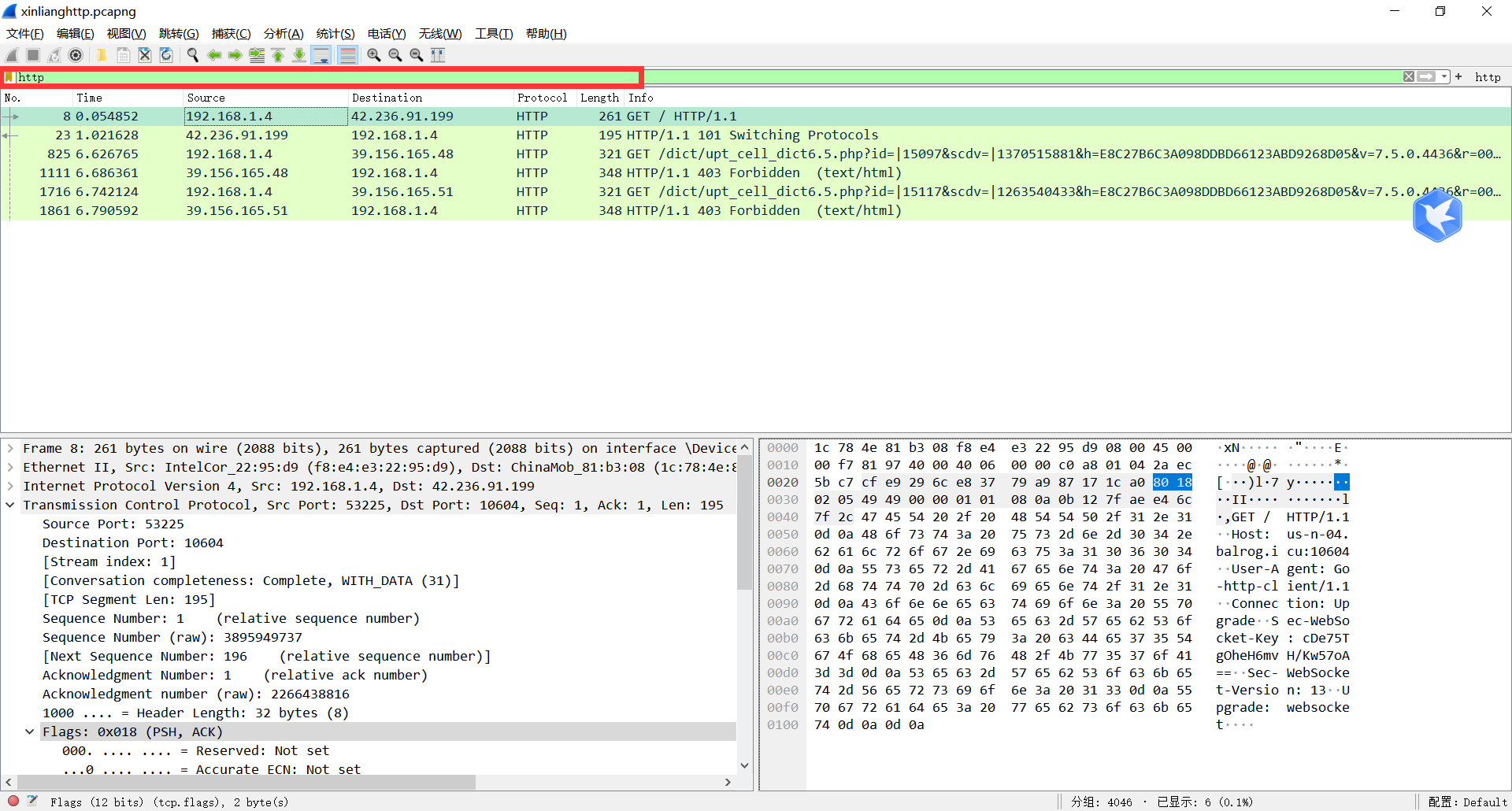
-Date：表示是服务器是响应回浏览器的时间，注意该时间是按照格林尼治标准时间(GMT)来计算。  
**(3)响应正文**

空行连接响应头和响应体。响应正文即为返回资源内容。浏览器可以直接识别响应正文html文件。

**操作过程：**

1. **第一步：筛选过滤**

在IP协议抓包的数据中进行HTTP协议的相关筛选。在上面http进行过滤筛选得到。



1. **第二步：HTTP协议分析**
2. **对请求报文分析**

图形用户界面

描述已自动生成

GET”：向特定的资源发出请求。

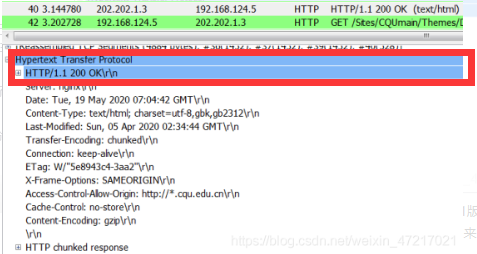
”HOST”首部在HTTP1.1版本中是必须的，描述了URL中的主机，在这里是http:// www.cqu.edu.cn/。

”USER-AGENT”首部显示了web服务器浏览器和本机windows系统。

Accept包括：Accept、 Accept-Language 、Accept-Encoding。

Connection首部描述了有关TCP连接的信息,通过此连接发送HTTP请求和响应,表明在发送请求之后连接是否保持活动状态及保持多久大多数HTTP1.1连接是持久的,意思是在每次请求后不关闭TCP连接，而保持该连接以接受从同一台服务器发来的多个请求

1. **对响应报文分析**



响应发送”HTTP/1.1 200 ok" ,状态码200表示响应成功，用HTTP1.1版本发送网页。

Cache-control首部，用于描述是否将数据的副本存储或高速缓存起来，此处为no-store 表示所有内容不会存在缓存 或internet临时文件中。

Content-length首部描述了数据长度为1676字节

Date首部表示了数据发送时间为：2020.5.19 07:04:42 其中GMT表示格林尼治标准时间。

**获取长文件**

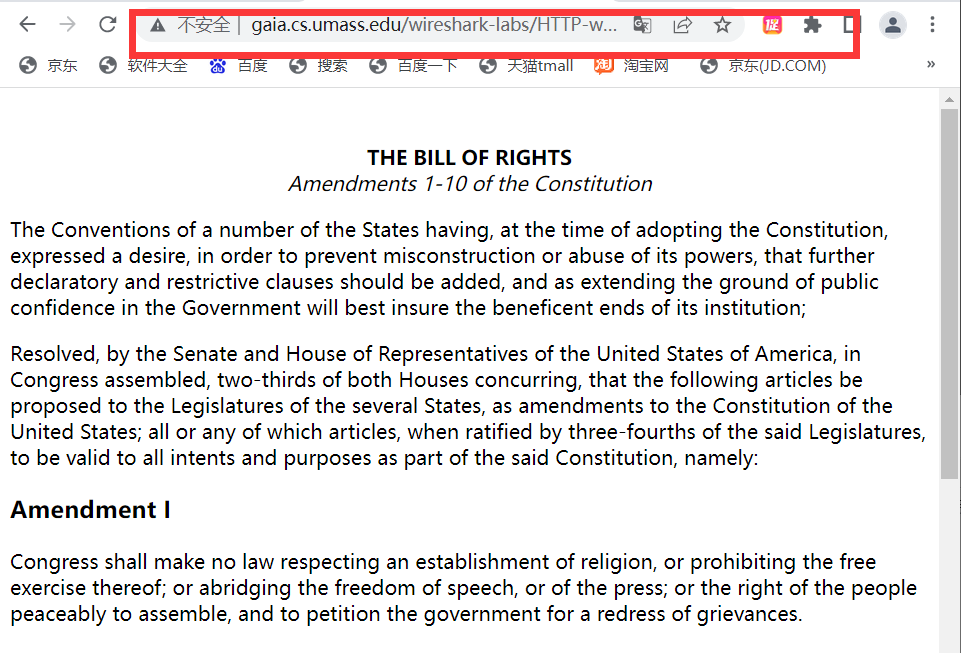
（1）启动浏览器，将浏览器的缓存清空。

（2）启动Wireshark，开始Wireshark分组捕获。

（3）在浏览器的地址栏中输入以下URL:

http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file3.html

浏览器将显示一个相当长的美国权力法案文本。



（4）停止Wireshark分组捕获，在显示过滤筛选编辑框中输入“http”,分组列表子窗口中将只显示所捕获到的HTTP消息。

图形用户界面, 文本, 应用程序

描述已自动生成

**嵌有对象的HTML文档**

（1）启动浏览器，将浏览器的缓存清空。

（2）启动Wireshark。开始Wireshark分组捕获。

（3）在浏览器的地址栏中输入以下URL:

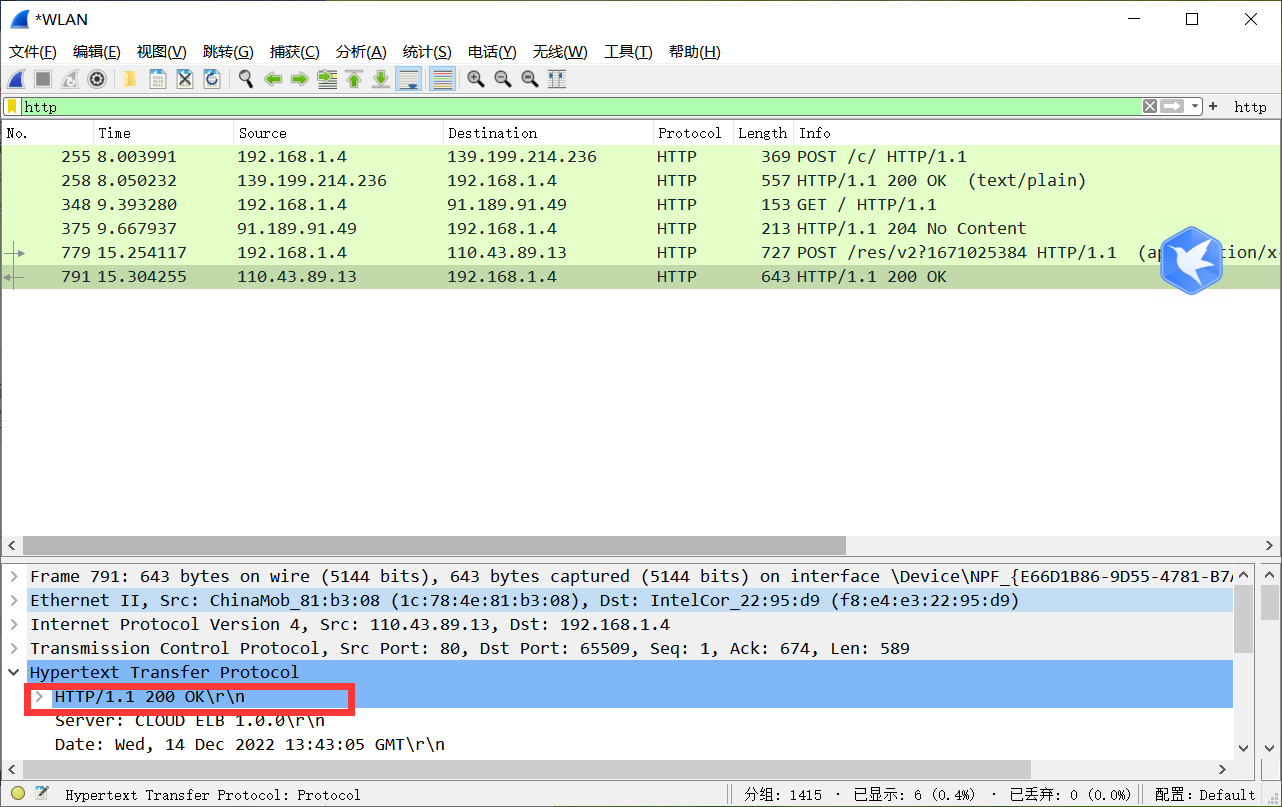
http://gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file4.html

（4）停止Wireshark分组捕获，在显示过滤筛选说明处输入“http”,分组列表子窗口中将只显示所捕获到的HTTP消息。

**HTTP具体分析过程**

**1.你的浏览器使用的是HTTP1.0，还是HTTP1.1？你所访问的Web服务器所使用HTTP协议的版本号是多少？**

使用的HTTP1.1 HTTP1.1

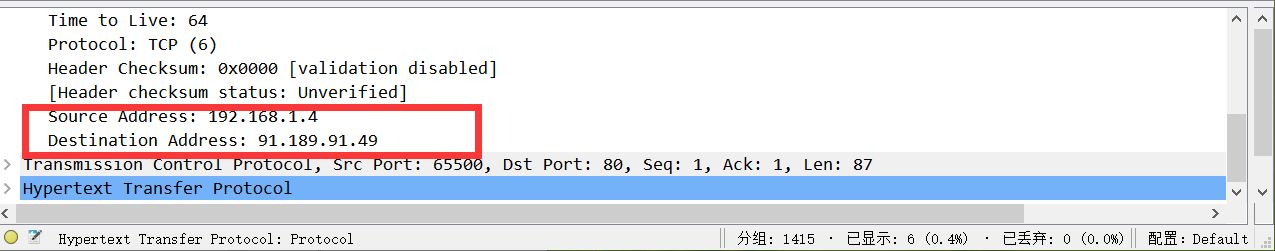


**2.你的浏览器向服务器指出它能接收何种语言版本的对象？**



**3.你的计算机的IP地址是多少？服务器gaia.cs.umass.edu的IP地址是多少？**

我的计算机的IP是192.168.1.4 服务器的IP 是91.189.91.49



**4.从服务器向你的浏览器返回response消息的状态代码是多少？**

200 OK

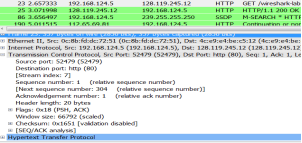
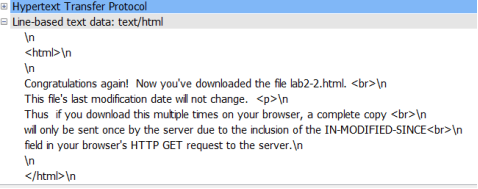
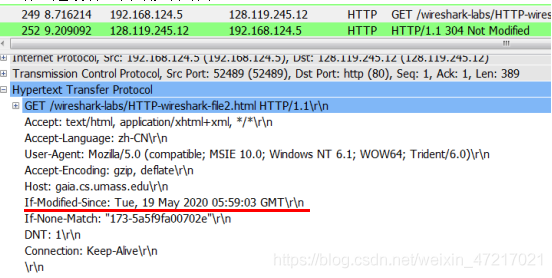
图形用户界面, 应用程序, Word

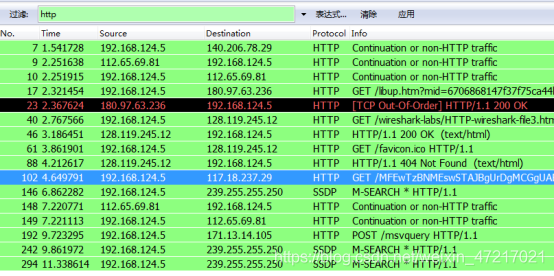
描述已自动生成

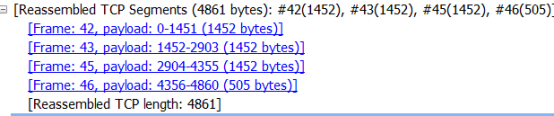
**5.你从服务器上所获取的HTML文件的最后修改时间是多少?**在2020年4月5日 02：34：44

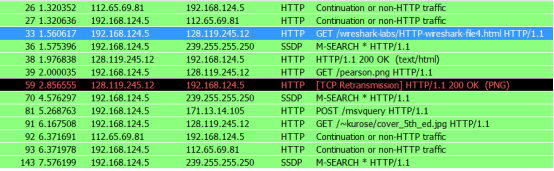
在这里插入图片描述

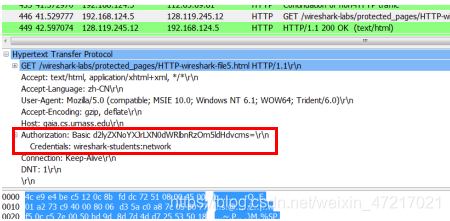
**6.返回到你的浏览器的内容一共多少字节?**  
4498字节

**7.分析你的浏览器向服务器发出的第一个HTTP GET请求的内容，在该请求报文中，是否有一行是: IF-MODIFIED-SINCE?**没有  
  
**8.分析服务器响应报文的内容，服务器是否明确返回了文件的内容?如何获知?**明确返回了  
  
**9.分析你的浏览器向服务器发出的第二个“HTTP GET”请求，在该请求报文中是否有一行是:IF -MODIFIED- SINCE?如果有，在该首部行后面跟着的信息是什么?**有，是最后一次的修改时间。  
  
**10. 服务器对第二个HTTP GET请求的响应中的HTTP状态代码是多少?服务器是否明确返回了文件的内容?请解释。**  
在这里插入图片描述  
 304状态码表示：服务器没有明确返回文件内容。原因:浏览器端缓存页面最后修改时间与服务器端时间一致，返回304状态码，客户端接到之后，就直接把本地缓存文件显示到浏览器中。

**11 . 你的浏览器一共发出了多少个HTTP GET请求?** 总共四个,分别为序号17,46,61,102  


1. **承载这一个HTTP响应报文一共需要多少个data-containing TCP报文段?**需要四个data-containing TCP报文段。  
   **13.与这个HTTPGET请求相对应的响应损文的状态代码和状态短语是什么?**在这里插入图片描述**14.在被传送的数据中一共有多少个HTTP状态行与TCP-induced”continuation"有关?**一个

**15.你的浏览器一共发出了多少个和打开的网址相关的HTTP GET请求?这些请求被发送到的目的地的IP地址是多少?** 发送了三个和打开的网址相关的HTTP GET请求，目的IP 为：128.119.245.12  
  
**16.浏览器在下载这两个图片时，是串行下载还是并行下载?请解释**。  
 两张图片是串行下载的，因为两张图片是连续请求，并由时间看出等第一张图片请求得到回复后才继续第二张图片的请求

**17.对于浏览器发出的最初的HTTP GET请求，服务器的响应是什么(状态代码和状态短语)?**  
 响应是：401 Unauthorized  
在这里插入图片描述18.**当浏览器发出第二个HTTP GET请求时，在HTTP GET报文中包含了哪些新的字段?**

**第四章 总结**

在期末计算机网络的大作业让我学到了不少东西。随着社会发展，计算机网络的应用已经非常广泛，渗透到智能仪表、工业控制、家用电器、计算机网络和通信网络、医用电器、汽车等领域。学好单片机真的很重要，这次期末大作业过程中，现实搜集了网上些资料和查了许多相关书籍，抓包分析时遇到了许多问题，后请教同学和查阅相关资料这些问题得以解决

本次的计算机网络的大作业中，我们对计算机网络中一些协议有了更加深层次的理解，其中包括IP协议，tcp协议和HTTP协议，IP协议又作为tcp协议的基础进行实现，而在我们所分析中并不能直接的找到IP协议的相关东西。在这个互联网蓬勃发展的时代，作为我们计算机专业的学生，我们对其互联网的基础应该有着比其他专业同学有更深刻的理解而计算机网络的知识是我们必不可缺少的专业技能。

这让我懂得了做任何事情都要有耐心，并且要细心，做任何事情都不可能一帆风顺，遇到挫折要积极面对，要有信心去解决，虚心向他人请教，这样也会让自己少走许多弯路。

学以致用，通过这次期末大作业把学到的计算机网络的理论知识与实际结合起来，更加巩固了以前的学习，并且对计算机网络认识更加深刻。最后，特别感谢老师在本学期中的指导和帮助。