HTTP

一 实验目的

1. 掌握 HTTP 协议的作用。
2. 了解 HTTP 报文的格式。
3. 了解 HTTP 的工作原理。

二 预备知识

1. HTTP 简介

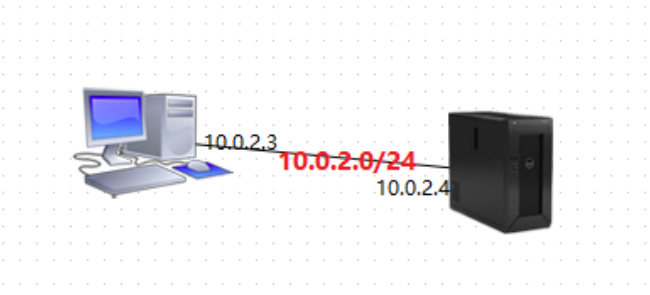
HTTP 协议是 Hyper Text Transfer Protocol（超文本传输协议）的缩写,是用于从万维网（WWW:World Wide Web ）服务器传输超文本到本地浏览器的传送协议。其基于 TCP/IP 通信协议来传递数据（HTML 文件, 图片文件, 查询结果等）。工作于客户端-服务端架构之上。浏览器作为 HTTP 客户端通过 URL 向 HTTP 服务端即 WEB 服务器发送所有请求。Web 服务器根据接收到的请求后，向客户端发送响应信息。

2. URL

HTTP 使用统一资源标识符（Uniform Resource Identifiers, URI）来传输数据和建立连接。URL（UniformResourceLocator）是一种特殊类型的 URI，包含了用于查找某个资源的足够的信息，中文称作统一资源定位符,是互联网上用来标识某一处资源的地址。

三 实验环境

在右上方的实验拓扑图中选择”**HTTP**”，点击连线配置子网网段（10.0.2.0/24），实验拓扑如下图所示：



然后点击提交实验，等待资源分配成功后，点击图标按全屏访问即可进入设备。

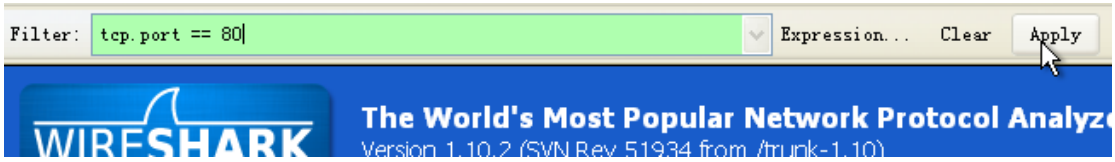
四 实验内容

1. 启动 http 服务

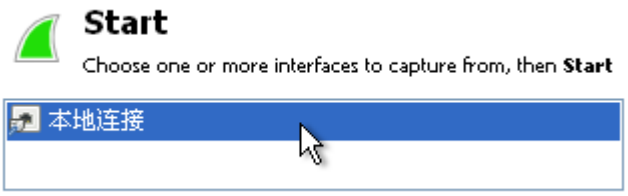
首先进入 Server，（用户名：centos 密码：centos），运行命令 **sudo /sbin/service httpd start**,启动 http 服务。

2. 抓包

我们进入 PC 系统，打开桌面的 Wireshark 软件（抓包软件），在工具栏上选择 抓包-》抓包过滤条件，输入 **tcp.port == 80**，点击 Apply：



然后再选择本地连接，点击 Start，开始抓包。界面选择如图所示：



接下来我们打开浏览器，输入 10.0.2.4 ，抓包如图所示：

14	13.215123	0 10.0.2.3	10.0.2.4	HTTP	380 GET / HTTP/1.1
15	13.215534	0 10.0.2.4	10.0.2.3	TCP	54 http > bsquare-voip [ACK] Seq=1 Ack=327 Win=27872 Len=0
16	13.216296	0 10.0.2.4	10.0.2.3	TCP	1414 [TCP segment of a reassembled PDU]
17	13.216316	0 10.0.2.4	10.0.2.3	TCP	1414 [TCP segment of a reassembled PDU]
18	13.216328	0 10.0.2.4	10.0.2.3	TCP	1414 [TCP segment of a reassembled PDU]
19	13.216340	0 10.0.2.4	10.0.2.3	HTTP	1179 HTTP/1.1 403 Forbidden (text/html)
20	13.216366	0 10.0.2.3	10.0.2.4	TCP	54 bsquare-voip > http [ACK] Seq=327 Ack=5206 Win=65535 Ler
21	13.691388	0 10.0.2.3	10.0.2.4	HTTP	443 GET /noindex/css/fonts/Light/OpenSans-Light.woff HTTP/1.
22	13.692110	0 10.0.2.4	10.0.2.3	HTTP	510 HTTP/1.1 404 Not Found (text/html)
23	13.694965	0 10.0.2.3	10.0.2.4	HTTP	441 GET /noindex/css/fonts/Bold/OpenSans-Bold.woff HTTP/1.1
24	13.695452	0 10.0.2.4	10.0.2.3	HTTP	508 HTTP/1.1 404 Not Found (text/html)
25	13.702557	0 10.0.2.3	10.0.2.4	HTTP	447 GET /noindex/css/fonts/Light/OpenSans-Light.ttf HTTP/1.1
26	13.703117	0 10.0.2.4	10.0.2.3	HTTP	509 HTTP/1.1 404 Not Found (text/html)
27	13.723606	0 10.0.2.3	10.0.2.4	HTTP	445 GET /noindex/css/fonts/Bold/OpenSans-Bold.ttf HTTP/1.1
28	13.724188	0 10.0.2.4	10.0.2.3	HTTP	507 HTTP/1.1 404 Not Found (text/html)
29	13.847300	0 10.0.2.3	10.0.2.4	HTTP	361 GET /favicon.ico HTTP/1.1

3. 分析

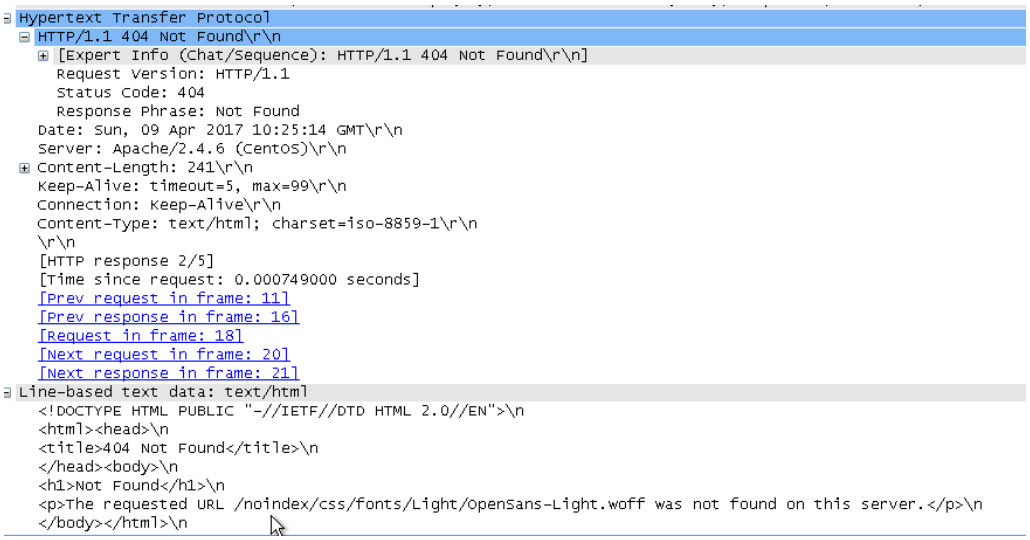
分析上面抓取的数据包可以得出通过 http 协议从服务器端获取数据流程如下：

1. 浏览器（10.0.2.3）向服务器（10.0.2.4）发出连接请求。此为 TCP 三次握手第一步，此时从图中可以看出，为 SYN，seq：x（x=0）
2. 服务器（10.0.2.4）回应了浏览器（10.0.2.3）的请求，并要求确认，此时为：SYN，ACK，此时 seq：y（y 为 0），ACK：x+1（为 1）。此为三次握手的第二步；
3. 浏览器（10.0.2.3）回应了服务器（10.0.2.4）的确认，连接成功。为：ACK，此时 seq：x+1（为 1），ACK：y+1（为 1）。此为三次握手的第三步；
4. 浏览器（10.0.2.3）发出一个页面 HTTP 的 GET 请求；
5. 服务器（10.0.2.4）确认；
6. 服务器（10.0.2.4）发送类型数据及发送状态响应码 200 OK；
7. 剩下就是 TCP 四次挥手释放连接的过程

分析 HTTP 报文，每个头域由一个域名，冒号（:）和域值三部分组成。域名是大小写无关的，域值前可以添加任何数量的空格符，头域可以被扩展为多行，在每行开始处，使用至少一个空格或制表符。HTTP 请求报文如下图所示。

```
⊟ Hypertext Transfer Protocol
  ⊟ GET /noindex/css/fonts/Light/OpenSans-Light.woff HTTP/1.1\r\n
    ⊞ [Expert Info (Chat/Sequence): GET /noindex/css/fonts/Light/OpenSans-Light.woff HTTP/1.1\r\n]
      Request Method: GET
      Request URI: /noindex/css/fonts/Light/OpenSans-Light.woff
      Request Version: HTTP/1.1
  Host: 10.0.2.4\r\n
  User-Agent: Mozilla/5.0 (Windows NT 5.1; rv:52.0) Gecko/20100101 Firefox/52.0\r\n
  Accept: application/font-woff2;q=1.0,application/font-woff;q=0.9,*/*;q=0.8\r\n
  Accept-Language: zh-CN,zh;q=0.8,en-US;q=0.5,en;q=0.3\r\n
  Accept-Encoding: identity\r\n
  Referer: http://10.0.2.4/noindex/css/open-sans.css\r\n
  Connection: keep-alive\r\n
  \r\n
  [Full request URI: http://10.0.2.4/noindex/css/fonts/Light/OpenSans-Light.woff]
  [HTTP request 2/5]
  [Prev request in frame: 11]
  [Response in frame: 19]
  [Next request in frame: 20]
```

HTTP 回应的消息如下图所示:



其中各个字含义如下：

1. Accept：浏览器端可以接受的媒体类型
2. Accept-Encoding：浏览器申明自己接收的编码方法
3. Accept-Language：浏览器申明自己接收的语言
4. User-Agent：告诉 HTTP 服务器，客户端使用的操作系统和浏览器的名称和版本
5. Host：请求报头域主要用于指定被请求资源的 Internet 主机和端口号
6. Connection：是否保持 TCP 连接
7. 而在回应报文中：
8. Date：生成消息的具体时间和日期
9. Server：指明 HTTP 服务器的软件信息
10. Last-Modified：用于指示资源的最后修改日期和时间
11. Content-Length：指明实体正文的长度，以字节方式存储的十进制数字来表示
12. Connection：是否保持 TCP 连接
13. Content-Type：WEB 服务器告诉浏览器自己响应的对象的类型和字符集
14. Line-based text data：响应数据

请根据各个字段的含义结合 HTTP 请求报文与响应报文深入理解 HTTP 协议的工作原理。