# 第 2 章 应用层

P4, P7, P8, P13, P16, P17
补充 1，2，3

P4. Consider the following string of ASCII characters that were captured by Wireshark when the browser sent an HTTP GET message (i.e., this is the actual content of an HTTP GET message). The characters <cr><lf> are carriage return and line-feed characters (that is, the italized character string <cr> in the text below represents the single carriage-return character that was contained at that point in the HTTP header). Answer the following questions, indicating where in the HTTP GET message below **(i.e., which headers)** you find the answer.

```
GET /cs453/index.html HTTP/1.1<cr><lf>Host: gai
a.cs.umass.edu<cr><lf>User-Agent: Mozilla/5.0 (
Windows;U; Windows NT 5.1; en-US; rv:1.7.2) Gec
ko/20040804 Netscape/7.2 (ax) <cr><lf>Accept:ex
t/xml, application/xml, application/xhtml+xml, text
/html;q=0.9, text/plain;q=0.8,image/png,*/*;q=0.5
<cr><lf>Accept-Language: en-us,en;q=0.5<cr><lf>Accept-
Encoding: zip,deflate<cr><lf>Accept-Charset: ISO
-8859-1,utf-8;q=0.7,*;q=0.7<cr><lf>Keep-Alive: 300<cr>
<lf>Connection:keep-alive<cr><lf><cr><lf>
```

a. What is the URL of the document requested by the browser?
b. What version of HTTP is the browser running?
c. Does the browser request a non-persistent or a persistent connection?
d. What is the IP address of the host on which the browser is running?
e. What type of browser initiates this message? Why is the browser type needed in an HTTP request message?

**a. 请求文档的 URL 为：http://gaia.cs.umass.edu/cs453/index.html ，来自于 Host:头部和 GET 请求行**
**b. 浏览器运行的 HTTP 版本为 HTTP/1.1，来自于 GET 请求行**
**c. 浏览器请求持续连接，来自于 Connection：keep-alive 头部**
**d. 从 HTTP 请求中无法得到浏览器所在主机的 IP 地址，该 IP 地址可以从携带 HTTP 请求的 IP 分组（IP 头部 + TCP 段头部 + HTTP 请求）中的头部字段中得到**
**e. 浏览器类型为 Mozilla/5.0，来自于 User-Agent 头部。HTTP 服务器可以根据 User-Agent 的信息进行内容协商，为不同的浏览器提供不同的版本（比如桌面版和移动版）**

P7. Suppose within your Web browser you click on a link to obtain a Web page. The IP address for the associated URL is not cached in your local host, so a DNS lookup is necessary to obtain the IP address. Suppose that n DNS servers are visited before your host receives the IP address from DNS; the successive visits incur an RTT of RTT$_1$, . . . , RTT$_n$. Further suppose that the Web page associated with the link contains exactly one object, consisting of a small amount of HTML text. Let RTT$_0$ denote the RTT between the local host and the server containing the object. Assuming zero transmission time of the object, how much time elapses from when the client clicks on the link until the client receives the object?

$\boldsymbol{2RTT_0 + RTT_1 + RTT_2 + \cdots + RTT_n}$

P8. Referring to Problem P7, suppose the HTML file references eight very small objects on the same server. Neglecting retransmission times, how much time elapses with

(a) Non-persistent HTTP with no parallel TCP connections?
(b) Non-persistent HTTP with the browser configured for 6(or 5) parallel connections?
(c) Persistent HTTP?

a 非持续连接，不采用并行连接 $\boldsymbol{18RTT_0 + RTT_1 + RTT_2 + \cdots + RTT_n}$
b 非持续连接，6 或 5 条并行连接 $\boldsymbol{6RTT_0 + RTT_1 + RTT_2 + \cdots + RTT_n}$
c 采用持续连接，缺省不进行流水线：$\boldsymbol{10RTT_0 + RTT_1 + RTT_2 + \cdots + RTT_n}$
采用流水线方式的持续连接：$\boldsymbol{3RTT_0 + RTT_1 + RTT_2 + \cdots + RTT_n}$

P13. Consider sending over HTTP/2 a Web page that consists of one video clip, and five images. Suppose that the video clip is transported as 2000 frames, and each image has three frames.

(a) If all the video frames are sent first without interleaving, how many “frame times” are needed until all five images are sent?
(b) If frames are interleaved, how many frame times are needed until all five images are sent.

(a) 2015 帧时
(b) 18 帧时

P16. How does SMTP mark the end of a message body? How about HTTP? Can HTTP use the same method as SMTP to mark the end of a message body? Explain.

SMTP 通过某一行正好为.表示邮件体的结束。
HTTP 通过 Content-Length 头部给出消息体的长度。也可采用块编码（Chunk Encoding）或关闭 TCP 连接。
HTTP 不能采用 SMTP 的方法表示消息体的结束，HTTP 的消息体可以是二进制数据。

P17. Read RFC 5321 for SMTP. What does MTA stand for? Consider the following received spam e-mail (modified from a real spam e-mail). Assuming only the originator of this spam e-mail is malicious and all other hosts are honest, identify the malicious host that has generated this spam e-mail, **provide its IP address and hostname**.

```
From - Fri Nov 07 13:41:30 2008
Return-Path: <tennis5@pp33head.com>
Received: from barmail.cs.umass.edu (barmail.cs.umass.
edu
[128.119.240.3]) by cs.umass.edu (8.13.1/8.12.6) for
<hg@cs.umass.edu>; Fri, 7 Nov 2008 13:27:10 -0500
Received: from asusus-4b96 (localhost [127.0.0.1]) by
barmail.cs.umass.edu (Spam Firewall) for <hg@cs.umass.
edu>; Fri, 7
Nov 2008 13:27:07 -0500 (EST)
Received: from asusus-4b96 ([58.88.21.177]) by barmail.
cs.umass.edu
for <hg@cs.umass.edu>; Fri, 07 Nov 2008 13:27:07 -0500
(EST)
Received: from [58.88.21.177] by inbnd55.exchangeddd.
com; Sat, 8
Nov 2008 01:27:07 +0700
From: ”Jonny” <tennis5@pp33head.com>
To: <hg@cs.umass.edu>

Subject: How to secure your savings
```

MTA 表示 Mail Transfer Agent，即邮件传输代理

恶意主机为 asusus-4b96，IP 地址为 58.88.21.177

补充 1. 采用一个合适的 DNS 查找工具（比如 nslookup, dig, host 等），采用迭代查询方式查找 www.cs.fudan.edu.cn 的 IP 地址，在整个解析过程中，假设没有缓存。请给出途中经过的 DNS 服务器，说明每个步骤所查询的问题（资源记录）以及服务器对于该问题的答案（主要资源记录），在描述资源记录时可以采用 name type value 的形式。

**注意可能会有多个答案，关键检查是否采用迭代查询方式：**
**本地域名服务器 → 根域名服务器 → cn 域名服务器 → edu.cn 域名服务器 → fudan.edu.cn 域名服务器**

补充 2 一个二进制文件大小是 4560 字节。如果使用 Base64 编码，且在每发出 110 字节及最后结尾处插入<CRLF>。试问该文件编码后大小是多少？<CRLF>在统计时当成一个字符。

**6136 字节**

补充 3 将下述以二进制形式描述的 4 个字节进行 base64 编码，给出编码后的 ASCII 字符串。

11001100 10000001 00111000 00001100

**zIE4DA==**