

上海交通大学

SHANGHAI JIAO TONG UNIVERSITY

实验报告

LAB REPORT



数据通信

Winsock ex3

姓 名: 袁炜程

学 号: 516030910287

班 级: F1603602

一、实验要求

Write a client program to execute a single HTTP GET to a Web server.

- The response from "www.sjtu.edu.cn" on port 80 to the message "GET / HTTP/1.0\n\n".
- The response from "www.baidu.com" on port 80 to the message "GET / HTTP/1.0\n\n".
- The response from "www.sina.com" on port 80 to the message "GET / HTTP/1.0\n\n".

二、实验原理

1、WebClient 基本步骤

- (1) 初始化 Winsock: `WSAStartup()`
- (2) 创建 socket: `socket()`
- (3) 接受地址信息, `result` 以链表储存该信息, 每一个节点存放一个 ip 及其信息
- (4) 根据 `result` 连接 server: `connect()`
- (5) 发送接收数据: `send()`、`receive()`
- (6) 关闭连接: `shutdown()`, `closesocket()`, `WSACleanup()`

2、http 请求消息结构

- two types of HTTP messages: *request, response*
- HTTP request message:
 - ASCII (human-readable format)

```
request line  
(GET, POST,  
HEAD commands) → GET /somedir/page.html HTTP/1.1  
header lines → Host: www.someschool.edu  
User-agent: Mozilla/4.0  
Connection: close  
Accept-language: fr  
Carriage return,  
line feed → (extra carriage return, line feed)  
indicates end  
of message
```

1、请求 www.baidu.com (正常访问)

1、请求 www.baidu.com (正常访问)

```
C:\Users\hp\Desktop\数据通信\winsock_ex3\WebClient\Debug\WebClient www.baidu.com
Bytes Sent: 16
Bytes received: 512
HTTP/1.1 200 OK
Date: Tue, 07 May 2019 04:48:06 GMT
Content-Type: text/html
Content-Length: 14615
Last-Modified: Thu, 25 Apr 2019 09:47:17 GMT
Connection: Close
Vary: Accept-Encoding
Set-Cookie: BAIDUID=AC3D30DE4709080969042F6ED6A9F669:FC=1; expires=Thu, 31-Dec-37 23:55:55 GMT; max-age=2147483647; path=/; domain=.baidu.com
Set-Cookie: BIDIPSID=AC3D30DE4709080969042F6ED6A9F669; expires=Thu, 31-Dec-37 23:55:55 GMT; max-age=2147483647; path=/; domain=.baidu.com
Set-Cookie: PSTM=1557204486; expires=Thu 烫烫烫烫8浜
Bytes received: 512
, 31-Dec-37 23:55:55 GMT; max-age=2147483647; path=/; domain=.baidu.com
P3P: CP=" OTI DSP COR IVA OUR IND COM "
ETag: "5cc18225-3917"
Server: BWS/1.1
X-UA-Compatible: IE=Edge,chrome=1
Pragma: no-cache
Cache-control: no-cache
Accept-Ranges: bytes

<!DOCTYPE html><!--STATUS OK-->
<html>
<head>
  <meta http-equiv="content-type" content="text/html; charset=utf-8">
  <meta http-equiv="X-UA-Compatible" content="IE=Edge">
  <link rel="dns-prefetch" href="//s1.bdstatic.com"/>
  <link rel="dns-prefetch" href="//s1.bdstatic.com"/>
Bytes received: 512
ef="//t1.baidu.com"/>
  <link rel="dns-prefetch" href="//t2.baidu.com"/>
  <link rel="dns-prefetch" href="//t3.baidu.com"/>
  <link rel="dns-prefetch" href="//t10.baidu.com"/>
  <link rel="dns-prefetch" href="//t11.baidu.com"/>
```

2、请求 www.sjtu.edu.cn (无法连接)

```
C:\Users\hp\Desktop\数据通信\winsock_ex3\WebClient\Debug>WebClient www.sjtu.edu.cn
Unable to connect to server!
```

3、请求 www.sina.com (收到的应答是 404 Not Found)

```
C:\Users\hp\Desktop>数据通信\winsock_ex3\WebClient\Debug>WebClient www.sina.com
Bytes Sent: 16
Bytes received: 512
HTTP/1.1 404 Not Found on Accelerator
Server: nginx
Date: Tue, 07 May 2019 04:50:56 GMT
Content-Type: text/html
Content-Length: 297
Connection: close
Via: http/1.1 cern.beijing.ha2ts4.205 (ApacheTrafficServer/6.2.1 [c s f])
Cache-Control: no-store
Content-Language: en
X-Cache: MISS, 205
X-Via-CDN: f=edge,s=cern.beijing.ha2ts4.210.nb.sinaedge.com,c=2403:d400:1001:20::2807;f=Edge,s=cern.beijing.ha2ts4.205,c=58.205.212.210
X-Via-Edge: 155720465606040000000410010026391a44d

<HTML>
<HEAD>
<TITLE>Not 资源资源8派
Bytes received: 273
Found on Accelerator</TITLE>
</HEAD>

<BODY BGCOLOR="white" FGCOLOR="black">
<H1>Not Found on Accelerator</H1>
<HR>

<FONT FACE="Helvetica,Arial"><B>
Description: Your request on the specified host was not found.
Check the location and try again.
</B></FONT>
<HR>
</BODY>
: en
X-Cache: MISS, 205
X-Via-CDN: f=edge,s=cern.beijing.ha2ts4.210.nb.sinaedge.com,c=2403:d400:1001:20::2807;f=Edge,s=cern.beijing.ha2ts4.205,c=58.205.212.210
X-Via-Edge: 155720465606040000000410010026391a44d
```

四、实验思考

在请求不同服务器时出现了不同的结果：请求百度时返回了正常的页面，请求新浪

时返回“404 Not Found”，请求交大时没有返回任何信息。我觉得新浪返回 404 是因为页面需要被重定向，因为从以下抓包的结果可以看出：

No.	Time	Source	Destination	Protocol	Length	Info
253	11.245085	2403:d400:1001:3:c0cf:eee8...	2400:89c0:1132:160::210:1	HTTP	438	GET / HTTP/1.1
408	11.279092	2400:89c0:1132:160::210:1	2403:d400:1001:3:c0cf:ee...	HTTP	927	HTTP/1.1 301 Moved Permanently (text/html)
455	11.290019	2403:d400:1001:3:c0cf:eee8...	2400:89c0:1132:160::207:1	HTTP	719	GET / HTTP/1.1
619	11.318891	2400:89c0:1132:160::207:1	2403:d400:1001:3:c0cf:ee...	HTTP	570	HTTP/1.1 302 Moved Temporarily (text/html)
2359	12.748762	10.163.232.8	222.22.29.98	HTTP	283	GET /MFewTzBNMEswSTA7BgUrDgMCGgUABBSXi0cw5bD2Wlrmnasv
2477	12.779518	222.22.29.98	10.163.232.8	OCSP	1143	Response
3207	12.983963	10.163.232.8	222.22.29.98	HTTP	285	GET /MFewTzBNMEswSTA7BgUrDgMCGgUABBSXi0cw5bD2Wlrmnasv
3315	13.010802	222.22.29.98	10.163.232.8	OCSP	1174	Response

新浪会返回 302 资源已被永久移到其他地方，因此简单的”GET / HTTP/1.0”无法请求页面，我试过将该 get 请求补齐，或者是其他的 get 请求都没法访问到新浪。而交大服务器没有返回信息，可能是其防火墙直接阻止了简单的 get 访问。