

计算机网络 课程实验报告

实验名称	HTTP 代理服务器的设计与实现					
姓名	王艺丹		院系	计算机		
班级	1903601		学号	1190201303		
任课教师	李全龙		指导教师	李全龙		
实验地点	G207		实验时间	2021.10.30		
实验课表现	出勤、表现得分(10)		实验报告		实验总分	
	操作结果得分(50)		得分(40)			
教师评语						

实验目的:

- 1. 熟悉并掌握 Socket 网络编程的过程与技术;
- 2. 深入理解 HTTP 协议,掌握 HTTP 代理服务器的基本工作原理;
- 3. 掌握 HTTP 代理服务器设计与编程实现的基本技能。

实验内容:

- 1. 设计并实现一个基本 HTTP 代理服务器。要求在指定端口(例如 8080)接收来自客户的 HTTP 请求并且根据其中的 URL 地址访问该地址 所指向的 HTTP 服务器(原服务器),接收 HTTP 服务器的响应报文,并 将响应报文转发给对应的客户进行浏览。
- 2. 设计并实现一个支持 Cache 功能的 HTTP 代理服务器。要求能缓 存原服务器响应的 对象,并能够通过修改请求报文(添加 if-modified-since 头行),向原服务器确认缓存 对象是否是最新版本。(选作内容,加分项 目,可以当堂完成或课下完成)
- 3. 扩展 HTTP 代理服务器,支持如下功能: (选作内容,加分项目, 可以当堂完成或课下完成)
 - a) 网站过滤: 允许/不允许访问某些网站:
 - b) 用户过滤: 支持/不支持某些用户访问外部网站;
 - c) 网站引导:将用户对某个网站的访问引导至一个模拟网站(钓鱼)。

实验过程:

总体思路:

首先了解一下客户端和服务器端的任务:

客户端:

- 1. 根据目标服务器IP地址与端口号创建套接字,并连接服务器(三次握手)
- 2. 发送请求报文
- 3. 接收返回报文
- 4. 关闭连接

服务器端:

- 1.创建套接字, 绑定套接字的本地IP地址和端口号, 对端口进行监听。
- 2. 从连接请求队列中取出一个连接请求,并同意连接。在TCP连接过程中进行了三次握手。
 - 3.接收客户端请求
 - 4.对请求进行响应,发送响应数据
 - 5.关闭连接
- 一. 设计并实现一个基本 HTTP 代理服务器

首先初始化套接字socket,再使用blind函数将该套接字与服务器host地址绑定,同时绑定指定端口号10240;利用while(true)循环与listen函数实现对指定端口的持续监听;使用accept 函数对请求进行接收和相应(多线程提高效率);再使用recv和send函数,接收来自客户端的 HTTP 请求,并通过代理服务器将该请求转发给服务器;同时,服务器也将获得的响应发给代理服务器,然后再由代理服务器将该响应发送给客户端;处理完成后,等待200ms关闭该线程,并清理缓存;重复循环处理下一个请求。主要功能函数具体实现见源码及注释。

二. 设计并实现一个支持 Cache 功能的 HTTP 代理服务器 主要思想为:

为 cache分配缓存,通过MakeCache函数将数据存入缓存:

当访问某网站时,首先通过url寻找本地对应缓存文件,若无对应文件,即为第一次访问

某网站,则代理服务器将该请求返回的响应数据写入缓存即相应文件中;

若匹配到本地缓存文件,利用MakeNewHTTP函数改造请求报文,即为报文增加ifmodified-since头部行,再向服务器端发送请求,解析Date字段获得最网站最新更新时间;通过服务器返回的数据码判断是否为最新的数据,若返回304,则内容并未再次更新,直接将缓存中的内容发给服务器即可;若返回200,则令服务器同时返回响应报文,将此响应报文直接发给客户端,同时更新本地缓存。

三. 扩展 HTTP 代理服务器

a) 网站过滤:允许/不允许访问某些网站;

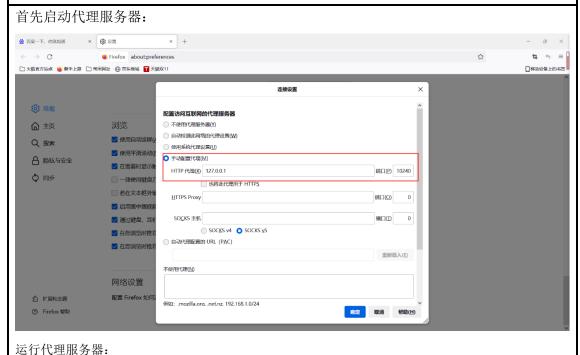
设置字符串数组存储屏蔽网站地址,对请求的 HTTP 报文头部进行解析,提取其中的访问地址 url ,并与屏蔽网站地址进行匹配,若匹配成功,则代码跳转至error 部分,打印相关提示信息,立即关闭套接字,断开连接。

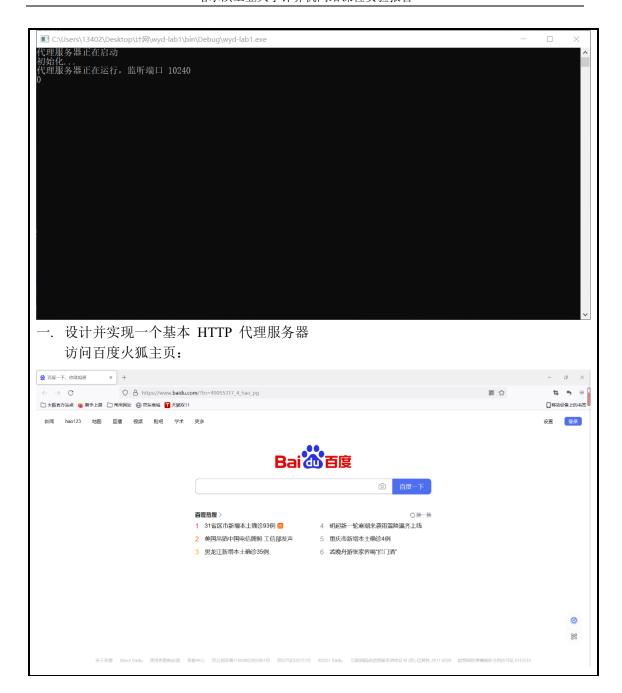
b) 用户过滤: 支持/不支持某些用户访问外部网站;

更改套接字绑定的主机地址, ProxyServerAddr.sin_addr.S_un.S_addr = inet_addr("127.0.0.1"); //仅本机用户可访问服务器; 也可以通过数组存储禁止访问的人或者资源,实现不同的用户屏蔽。

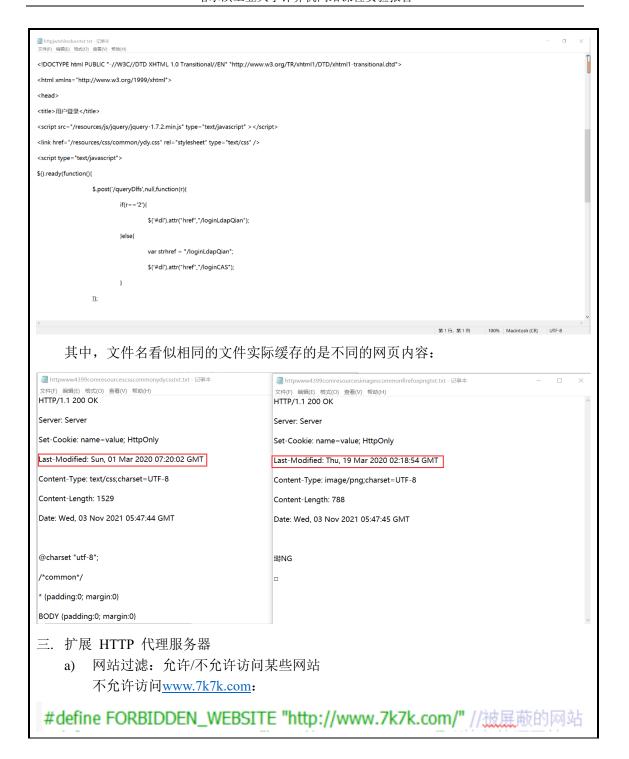
c) 网站引导:将用户对某个网站的访问引导至一个模拟网站(钓鱼) 设置字符串数组存储钓鱼网站地址,设置引导目的网站地址,同样解析匹配url,若 匹配成功,则更改HTTP头部字段的url(访问网址)与host主机地址,实现网页跳转,同时 打印输出相关提示信息。

实验结果:



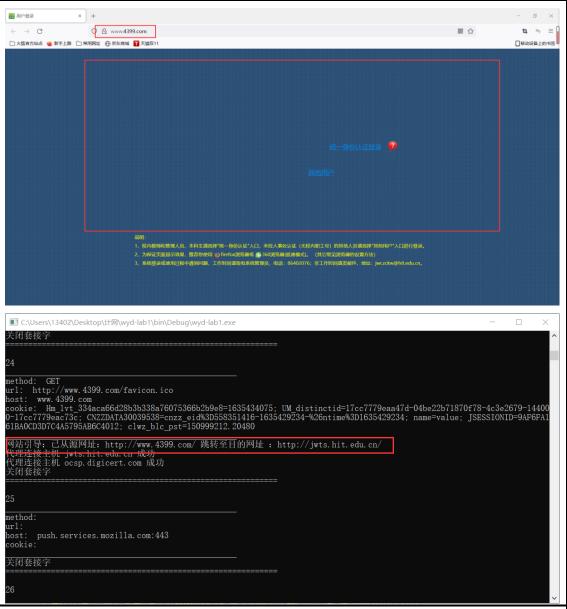






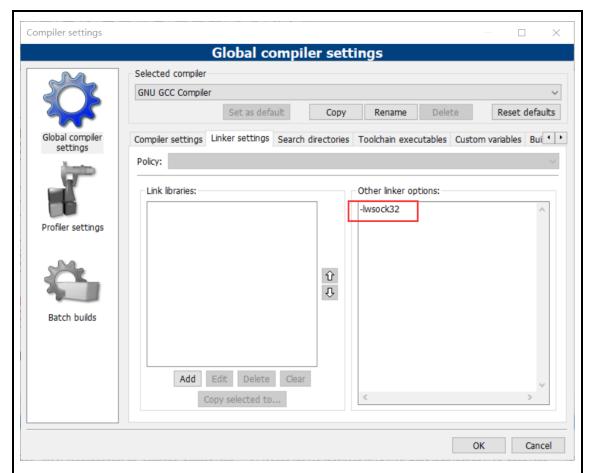


#define FISHING_WEB_DEST "http://jwts.hit.edu.cn/" //钓鱼的目的网址



问题讨论:

1) 实验指导书的参考代码 #pragma comment(lib,"Ws2_32.lib") 编译失败,查询资料得知: CodeBlocks 使用MingGW 编译,而MingGW不支持#pragma comment(lib,"Ws2_32.lib") 的写法,该命令是静态链接Ws2_32.lib 库,CB需要在设置里加上 -lws2_32 或 -lwsock32,如下图所示:



- 2) 不支持 strtok s() 函数,这里需要切换成 strtok() 函数
- 3) 不可以在goto语句后边定义新变量

心得体会:

结合实验过程和结果给出实验的体会和收获:

- ◆ 本次实验让我对 socket 编程有了初步的了解,掌握了 HTTP 代理服务器的基本原理, 对 HTTP 请求和响应原理有了更深的认识;
- ◇ 对TCP/IP协议有了更深的理解;
- ◆ 对网站钓鱼、网站屏蔽等有了深刻的理解;
- ◆ 让我对 socket 编程产生了浓厚的兴趣,对计算机网络的学习有了更好的理解与热情。