

Network Programming



北京大学

March 27, 2021

Connection

- Connection
- Server

Echo Server

- Overview
- Socket
- Socket Address
- Client Side
- Server Side

Web

- Web Basics
- HTTP
- Proxy

Exam review

- ▶ 点对点: 连接一对进程

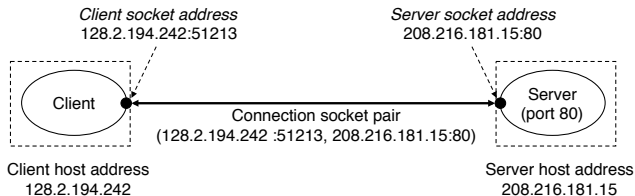
- ▶ 点对点: 连接一对进程
- ▶ 全双工: 数据同时可以在两个方向流动

- ▶ 点对点: 连接一对进程
- ▶ 全双工: 数据同时可以在两个方向流动
- ▶ 可靠: 字节的发送的顺序和收到的一致

- ▶ 点对点: 连接一对进程
- ▶ 全双工: 数据同时可以在两个方向流动
- ▶ 可靠: 字节的发送的顺序和收到的一致
- ▶ 套接字 (socket): 连接的端点, 地址为 IP addr:port

- ▶ 点对点: 连接一对进程
- ▶ 全双工: 数据同时可以在两个方向流动
- ▶ 可靠: 字节的发送的顺序和收到的一致
- ▶ 套接字 (socket): 连接的端点, 地址为 IP addr:port
- ▶ 端口 (port): 16 位无符号数 ($2^{16} = 65536$), 分为临时端口 (ephemeral port) 和知名端口 (well-known port), 用于标识进程.

- ▶ 点对点: 连接一对进程
- ▶ 全双工: 数据同时可以在两个方向流动
- ▶ 可靠: 字节的发送的顺序和收到的一致
- ▶ 套接字 (socket): 连接的端点, 地址为 IP addr:port
- ▶ 端口 (port): 16 位无符号数 ($2^{16} = 65536$), 分为临时端口 (ephemeral port) 和知名端口 (well-known port), 用于标识进程.



Definition (服务器 (server))

长时运行 (守护) 进程, 识别请求并处理

Definition (服务器 (server))

长时运行 (守护) 进程, 识别请求并处理

常见知名端口

Web(HTTP)(80), FTP(20, 21), SSH(22), Telnet(23),
Mail(SMTP)(25), HTTPS(443)

Connection

Connection

Server

Echo Server

Overview

Socket

Socket Address

Client Side

Server Side

Web

Web Basics

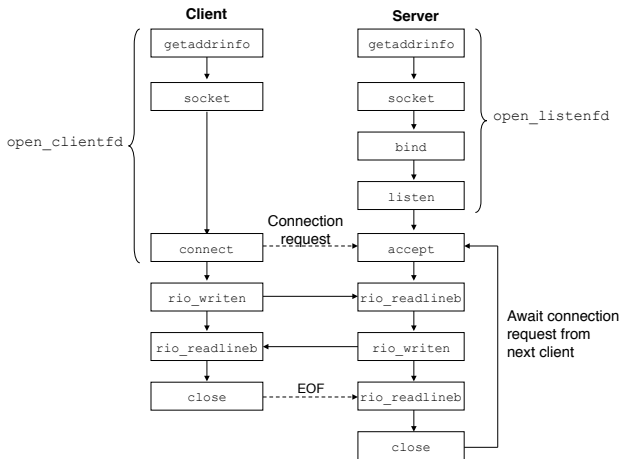
HTTP

Proxy

Exam review

将客户端发来的内容原样发回
如何实现？

1. 开启服务器
2. 开启客户端
3. 交换数据
4. 关闭客户端
5. 断开客户端



可能考默写 (18. 六)

IPv4 sockaddr_in: sin_family(AF_INET), sin_port, sin_addr,
sin_zero (大端法)
通用结构 sockaddr: sa_family, sa_data

string → sockaddr list

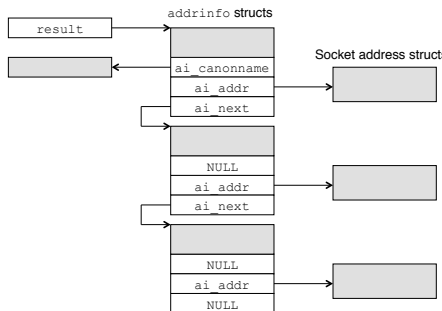
返回 result, 其指向套接字地址链表
(addrinfo), 供客户端按需访问
协议无关

(e.g. 不必区分 IPv4/IPv6)

调用后需用 freeaddrinfo 清理垃圾

getnameinfo 与之相反,

用于解析连接信息



open_clientfd: getaddrinfo, socket(创建 socket 描述符),
connect(建立与服务器的连接)

RIOs: robust I/O for applications such as network programs that
are subject to short counts

close: 关闭连接

open_listenfd: getaddrinfo, socket, bind, listen(主动套接字 → 监听套接字)

accept: 等待并处理来自客户端的连接请求

echo: 持续读写直至收到 EOF

监听描述符 (listenfd) 已连接描述符 (connfd)

18. —.11

11. 下面有关套接字接口(Socket API)的叙述中, 错误的是
- A. 套接字接口常常被用来创建网络应用
 - B. Windows 10 系统没有实现套接字接口
 - C. `getaddrinfo()` 和 `getnameinfo()` 可以被用于编写独立于特定版本的 IP 协议的程序
 - D. `socket()` 函数返回的描述符, 可以使用标准 Unix I/O 函数进行读写

18. —.11

11. 下面有关套接字接口(Socket API)的叙述中, 错误的是

- A. 套接字接口常常被用来创建网络应用
- B. Windows 10 系统没有实现套接字接口
- C. `getaddrinfo()` 和 `getnameinfo()` 可以被用于编写独立于特定版本的 IP 协议的程序
- D. `socket()` 函数返回的描述符, 可以使用标准 Unix I/O 函数进行读写

B

Connection

- Connection
- Server

Echo Server

- Overview
- Socket
- Socket Address
- Client Side
- Server Side

Web

- Web Basics
- HTTP
- Proxy

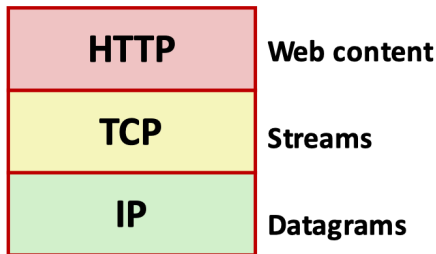
Exam review

HTTP: 超文本传输协议

MIME: html, plain, image, ...

静态内容: 返回服务器上文件

动态内容: 返回程序运行结果



Static:

`http://info.cern.ch:80/hypertext/WWW/TheProject.html`

Dynamic:

`http://nbhbdm.cn:80/?s=%e7%99%be%e5%ba%a6`

客户端利用前缀推断连接协议, 服务器地址和端口

服务器利用后缀判断请求类型并找到相应文件, 最小后缀为"/"

URI

HTTP 请求 = 请求行 $\times 1$ + 请求报头 $\times k$ + 空行 $\times 1$
(换行使用 $\backslash r \backslash n$)

请求行 = method URI version

method: GET, POST, ...

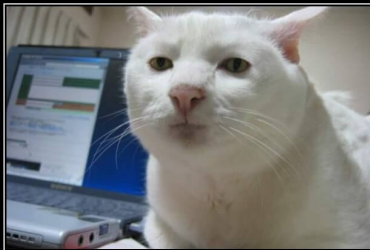
请求报头 = header-name:header-data (e.g. User-Agent, Referer)

HOST 报头 (HTTP/1.1): 供代理缓存使用

HTTP 响应 = 响应行 $\times 1$ + 响应报头 $\times k$ + 空行 $\times 1$ + 响应主体
响应行 = version status-code status-message



404
Not Found



200
OK

<https://http.cat>

服务器利用 `fork+execve` 执行子进程以应对动态请求
服务器向子进程的参数传递遵循 CGI(通用网关接口)
重定向标准输出到已连接描述符

子进程要负责输出 `Content-type`, `Content-length` 响应报头及空行.

```
bash:makoshark> telnet whaleshark.ics.cs.cmu.edu 15213
Trying 128.2.210.175...
Connected to whaleshark.ics.cs.cmu.edu (128.2.210.175).
Escape character is '^]'.
GET /cgi-bin/adder?15213&18213 HTTP/1.0
```

HTTP request sent by client

```
HTTP/1.0 200 OK
Server: Tiny Web Server
Connection: close
Content-length: 117
Content-type: text/html
```

*HTTP response generated
by the server*

```
Welcome to add.com: THE Internet addition portal.
<p>The answer is: 15213 + 18213 = 33426
<p>Thanks for visiting!
Connection closed by foreign host.
bash:makoshark>
```

*HTTP response generated
by the CGI program*

客户端 \Rightarrow 代理服务器 \Rightarrow 原始服务器

优点: 提高访问速度 (利用缓存), 控制对内部资源的访问, 过滤内容, 隐藏真实 IP (e.g. Tor), 突破内容过滤机制限制 (GFW), 转码

客户端 \Rightarrow 代理服务器 \Rightarrow 原始服务器

优点: 提高访问速度 (利用缓存), 控制对内部资源的访问, 过滤内容, 隐藏真实 IP (e.g. Tor), 突破内容过滤机制限制 (GFW), 转码

显式代理

客户端显式配置使用代理, 此时客户端知道正向代理的存在
请求需给出完整 URL

客户端 \Rightarrow 代理服务器 \Rightarrow 原始服务器

优点: 提高访问速度 (利用缓存), 控制对内部资源的访问, 过滤内容, 隐藏真实 IP (e.g. Tor), 突破内容过滤机制限制 (GFW), 转码

显式代理

客户端显式配置使用代理, 此时客户端知道正向代理的存在
请求需给出完整 URL

透明代理

客户端没有进行任何配置, 不知道代理的存在, 代理在中途拦截并完成请求

Connection

- Connection
- Server

Echo Server

- Overview
- Socket
- Socket Address
- Client Side
- Server Side

Web

- Web Basics
- HTTP
- Proxy

Exam review

数连接最大值, 端口 (13. 七, 14. 八)

socket 连接过程 (13. 七, 14. 八, 18. 六)

connection socket pairs 的合法性判断/书写 (13. 一.18)
写代码 (16. 七, 实现基于进程的并发 echo server, 补全
open_listenfd)

静态/动态内容的概念 (14. 八, 15. 一.17)
与并发编程的联系

谢谢!