

HTTP 协议基础

报文结构、状态码

HTTP 报文

起始行 (start line)

头部 (header)

空行 (CRLF)

实体 (entity/body)

```
GET / HTTP/1.1
Host: www.baidu.com
User-Agent: curl/7.79.1
Accept: */*
```

```
HTTP/1.1 200 OK
Accept-Ranges: bytes
Cache-Control: private, no-cache, no-store, proxy-revalidate, no-transform
Connection: keep-alive
Content-Length: 2381
Content-Type: text/html
Date: Mon, 18 Apr 2022 08:58:23 GMT
Etag: "588604dc-94d"
Last-Modified: Mon, 23 Jan 2017 13:27:56 GMT
Pragma: no-cache
Server: bfe/1.0.8.18
Set-Cookie: BDORZ=27315; max-age=86400; domain=.baidu.com; path=/
```

```
<!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><meta
content=always name=referrer><link rel=stylesheet type=text/css href=http://
s1.bdstatic.com/r/www/cache/bdorz/baidu.min.css><title>.....</
title></head> <body link=#0000cc> <div id=wrapper> <div id=head> <div
class=head_wrapper> <div class=s_form> <div class=s_form_wrapper> <div id=lg> <img
hidefocus=true src=//www.baidu.com/img/bd_logo1.png width=270 height=129> </div>
<form id=form name=f action=//www.baidu.com/s class=fm> <input type=hidden
name=bdorz_come value=1> <input type=hidden name=ie value=utf-8> <input type=hidden
name=f value=8> <input type=hidden name=rsv_bp value=1> <input type=hidden
name=rsv_idx value=1> <input type=hidden name=tn value=baidu><span class="bg
s_ipt_wr"><input id=kwd name=wd class=s_ipt value maxlength=255 autocomplete=off
autofocus></span><span class="bg s_btn_wr"><input type=submit id=su
value=..... class="bg s_btn"></span> </form> </div> </div> <div id=u1> <a
href=http://news.baidu.com name=tj_trnews class=mnav>.....</a> <a href=http://
www.hao123.com name=tj_trhao123 class=mnav>hao123</a> <a href=http://map.baidu.com
name=tj_trmap class=mnav>.....</a> <a href=http://v.baidu.com name=tj_trvideo
```

1 client pkt, 2 server pkts, 1 turn.

如何用 TCP 发送 HTTP 请求

```
telnet www.baidu.com 80
```

```
GET / HTTP/1.1
```

```
Host: www.baidu.com
```

```
User-Agent: curl/7.79.1
```

```
Accept: */*
```

请求行



```
GET / HTTP/1.1
Host: www.baidu.com
User-Agent: curl/7.79.1
Accept: */*

HTTP/1.1 200 OK
Accept-Ranges: bytes
Cache-Control: private, no-cache, no-store, proxy-revalidate, no-transform
Connection: keep-alive
Content-Length: 2381
Content-Type: text/html
Date: Mon, 18 Apr 2022 08:58:23 GMT
Etag: "588604dc-94d"
Last-Modified: Mon, 23 Jan 2017 13:27:56 GMT
Pragma: no-cache
Server: bfe/1.0.8.18
Set-Cookie: BDORZ=27315; max-age=86400; domain=.baidu.com; path=/
```

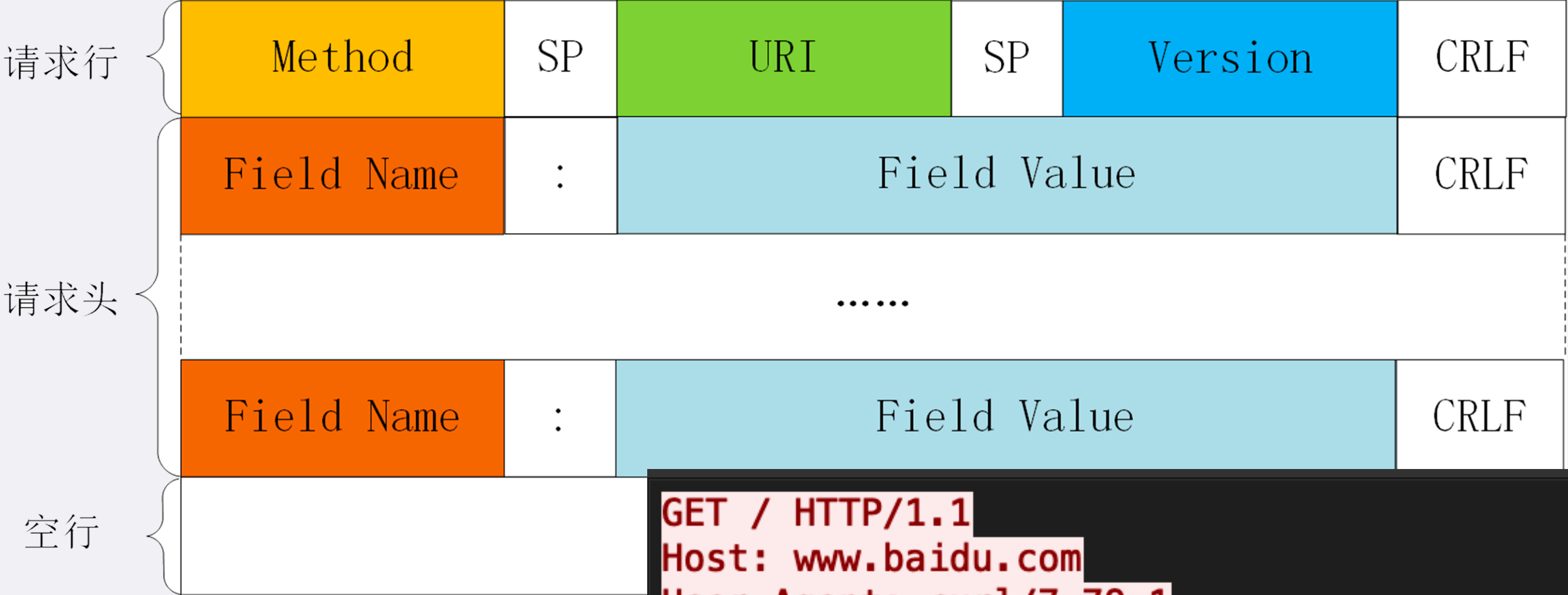

状态行



```
GET / HTTP/1.1
Host: www.baidu.com
User-Agent: curl/7.79.1
Accept: */*

HTTP/1.1 200 OK
Accept-Ranges: bytes
Cache-Control: private, no-cache, no-store, proxy-revalidate, no-transform
Connection: keep-alive
Content-Length: 2381
Content-Type: text/html
Date: Mon, 18 Apr 2022 08:58:23 GMT
Etag: "588604dc-94d"
Last-Modified: Mon, 23 Jan 2017 13:27:56 GMT
Pragma: no-cache
Server: bfe/1.0.8.18
Set-Cookie: BDORZ=27315; max-age=86400; domain=.baidu.com; path=/
```

头部字段



```
GET / HTTP/1.1
Host: www.baidu.com
User-Agent: curl/7.79.1
Accept: */*

HTTP/1.1 200 OK
Accept-Ranges: bytes
Cache-Control: private, no-cache, no-store, proxy-revalidate, no-transform
Connection: keep-alive
Content-Length: 2381
Content-Type: text/html
Date: Mon, 18 Apr 2022 08:58:23 GMT
Etag: "588604dc-94d"
Last-Modified: Mon, 23 Jan 2017 13:27:56 GMT
Pragma: no-cache
Server: bfe/1.0.8.18
Set-Cookie: BDORZ=27315; max-age=86400; domain=.baidu.com; path=/
```

请求方法

1. GET: 获取资源, 可以理解为读取或者下载数据;
2. HEAD: 获取资源的元信息;
3. POST: 向资源提交数据, 相当于写入或上传数据;
4. PUT: 类似 POST;
5. DELETE: 删除资源;
6. CONNECT: 建立特殊的连接隧道;
7. OPTIONS: 列出可对资源实行的方法;
8. TRACE: 追踪请求 - 响应的传输路径

用 HEAD 判断 url 是否存在

`curl -I https://store-tg1.cvte.com/c341e63e-80af-477e-a29e-ccca653e2838.mp4`

```
➔ ~ curl -I https://store-tg1.cvte.com/c341e63e-80af-477e-a29e-ccca653e2838.mp4
HTTP/2 200
date: Mon, 18 Apr 2022 09:03:27 GMT
content-type: video/mp4
content-length: 17839845
server: openresty
accept-ranges: bytes
access-control-allow-origin: *
access-control-expose-headers: X-Log, X-Reqid
access-control-max-age: 2592000
cache-control: public, max-age=31536000
content-disposition: inline; filename="c341e63e-80af-477e-a29e-ccca653e2838.mp4"; fi
ca653e2838.mp4
content-md5: 16PezbYoDrDv0gWaw16g7g==
content-transfer-encoding: binary
etag: "lt0hFaLABl0ThS5Icvack37t61G"
last-modified: Mon, 18 Apr 2022 03:57:00 GMT
x-reqid: 5gYAAADYJBx54eYW
x-svr: IO
x-qiniu-zone: 0
x-log: X-Log
x-via: 1.1 PS-WUX-01Ljs29:7 (Cdn Cache Server V2.0)[64 200 0], 1.1 PS-000-01AAe53:4
x-ws-request-id: 625d295e_PS-000-01AAe53_33543-22030
access-control-allow-methods: GET,OPTIONS
access-control-allow-credentials: true
access-control-allow-headers: *
```

```
➔ ~ curl -I https://store-tg1.cvte.com/c341e63e-80af-477e-a29e-ccca653e2838.mp4a
HTTP/2 404
date: Mon, 18 Apr 2022 09:03:39 GMT
content-type: application/json
content-length: 30
server: openresty
access-control-allow-origin: *
access-control-expose-headers: X-Log, X-Reqid
access-control-max-age: 2592000
x-reqid: uy0AAAAIhzY08uYW
x-svr: IO
x-log: X-Log
x-via: 1.1 PS-TNH-01ZBu207:3 (Cdn Cache Server V2.0)[310 404 2], 1.1 PS-000-01gJe140
x-ws-request-id: 625d296b_PS-000-01AAe53_33334-43564
access-control-allow-methods: GET,OPTIONS
access-control-allow-credentials: true
access-control-allow-headers: *
```


状态码

- 1××: 提示信息，表示目前是协议处理的中间状态，还需要后续的操作；
- 2××: 成功，报文已经收到并被正确处理；
- 3××: 重定向，资源位置发生变动，需要客户端重新发送请求；
- 4××: 客户端错误，请求报文有误，服务器无法处理；
- 5××: 服务器错误，服务器在处理请求时内部发生了错误

HTTP 响应码

1XX

1xx: 请求已接收到, 需要进一步处理才能完成

100 Continue: 上传大文件前使用

- 由客户端发起请求中携带 Expect: 100-continue

101 Switch Protocols: 协议升级使用

- 如升级 websocket 或者 http/2.0

```
GET /echo HTTP/1.1
Host: 10.211.55.3:18080
Connection: Upgrade
Pragma: no-cache
Cache-Control: no-cache
User-Agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7)
like Gecko) Chrome/100.0.4896.88 Safari/537.36
Upgrade: websocket
Origin: http://10.211.55.3:18080
Sec-WebSocket-Version: 13
Accept-Encoding: gzip, deflate
Accept-Language: zh-CN,zh;q=0.9,en-US;q=0.8,en;q=0.7
Sec-WebSocket-Key: MKo/wnVQGDdDZWWhQQCPNA==
Sec-WebSocket-Extensions: permessage-deflate; client_max_wir

HTTP/1.1 101 Switching Protocols
Upgrade: websocket
Connection: Upgrade
Sec-WebSocket-Accept: 4G020hBX500pqC7QyVkXdHqYvdw=

....oG...+...(...f..Hello world!.....s...i....
```

HTTP 响应码

2XX

2xx: 成功处理请求

200 OK: 成功返回响应。

202 Accepted: 服务器接收并开始处理请求，但请求未处理完成。

204 No Content: 成功执行了请求且不携带响应包体

206 Partial Content: 分块下载或断点续传的基础

206 状态码

<https://store-tg1.cvte.com/c341e63e-80af-477e-a29e-ccca653e2838.mp4>

The screenshot displays the developer tools of a web browser, specifically the 'Headers' section. The 'General' tab is active, showing the 'Status Code' as 206. The 'Request Headers' tab is also visible, showing the 'range' header as 'bytes=4227072-'. The 'Response Headers' tab is partially visible, showing the 'Content-Range' header as 'bytes 4227072-17839844/17839845'.

General

Request URL: <https://store-tg1.cvte.com/c341e63e-80af-477e-a29e-ccca653e2838.mp4>

Request Method: GET

Status Code: 206

Remote Address: 127.0.0.1:7890

Referrer Policy: strict-origin-when-cross-origin

Response Headers

accept-ranges: bytes

access-control-allow-credentials: true

access-control-allow-headers: *

access-control-allow-methods: GET,OPTIONS

access-control-allow-origin: *

access-control-expose-headers: X-Log, X-Rec

access-control-max-age: 2592000

age: 57

cache-control: public, max-age=31536000

content-disposition: inline; filename="c341e63e-80af-477e-a29e-ccca653e2838.mp4"

Content-Length: 13612773

content-md5: 16PezbYoDrDv0gWaw16g7g==

Content-Range: bytes 4227072-17839844/17839845

content-transfer-encoding: binary

content-type: video/mp4

Request Headers

:authority: store-tg1.cvte.com

:method: GET

:path: /c341e63e-80af-477e-a29e-ccca653e2838.mp4

:scheme: https

accept: */*

accept-encoding: identity;q=1, *,q=0

accept-language: zh-CN,zh;q=0.9,en-US;q=0.8,en;q=0.7

cache-control: no-cache

cookie: _ga=GA1.2.1554522476.1632792630

pragma: no-cache

range: bytes=4227072-

referer: <https://store-tg1.cvte.com/c341e63e-80af-477e-a29e-ccca653e2838.mp4>

sec-ch-ua: "Not A;Brand";v="99", "Chromium";v="100", "Google Chrome";v="100"

sec-ch-ua-mobile: ?0

sec-ch-ua-platform: "macOS"

sec-fetch-dest: video

sec-fetch-mode: no-cors

sec-fetch-site: same-origin

user-agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/100.0.0.0 Safari/100.0

HTTP 响应码

3XX

<https://sourl.cn/rFu6Na>

3xx: 重定向使用 Location 指向的资源或者缓存中的资源

- 301 Moved Permanently: 资源永久性的重定向
- 302 Found: 资源临时的重定向
- 304 Not Modified: 304 告诉客户端可以复用缓存。

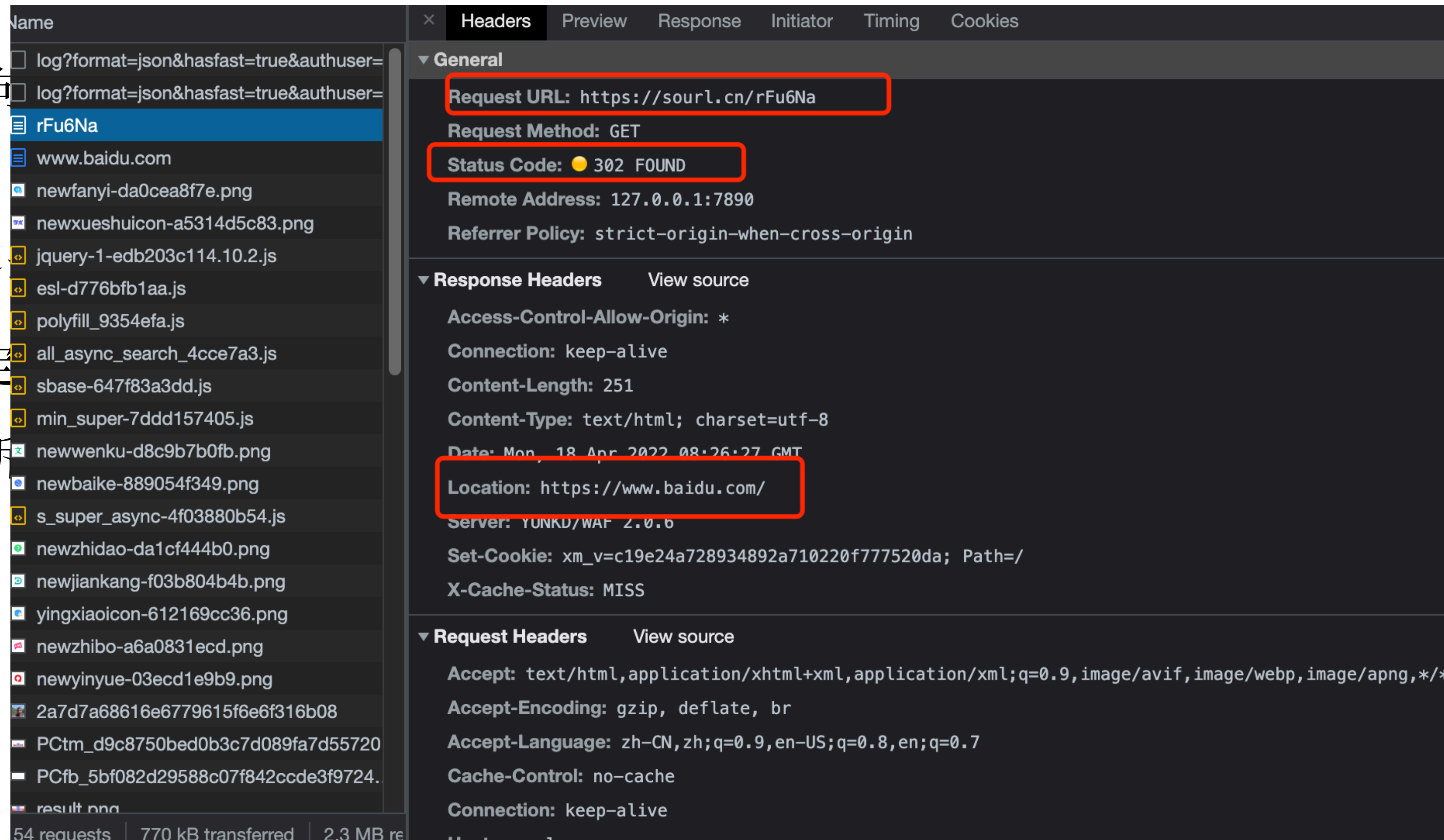
HTTP 响应码

3XX

<https://sourl.cn/rFu6Na>

3xx: 重定向使用 Location 指向

- 301 Moved Permanently: 资源永久移动
- 302 Found: 资源临时的重定向
- 304 Not Modified: 304 告诉客户端资源未修改



HTTP 响应码

4XX

- 400 Bad Request: 客户端出现了错误
- 401 Unauthorized: 用户认证信息缺失或者不正确
- 403 Forbidden: 无权限
- 404 Not Found: 服务器没有找到对应的资源

其它常见的 4XX 状态码

405 Method Not Allowed: 不允许使用某些方法操作资源，例如不允许 POST 只能 GET；

406 Not Acceptable: 资源无法满足客户端请求的条件，例如请求中文但只有英文；

408 Request Timeout: 请求超时，服务器等待了过长的时间；

409 Conflict: 多个请求发生了冲突，可以理解为多线程并发时的竞态；

413 Request Entity Too Large: 请求报文里的 body 太大；

414 Request-URI Too Long: 请求行里的 URI 太大；

429 Too Many Requests: 客户端发送了太多的请求，通常是由于服务器的限连策略；

431 Request Header Fields Too Large: 请求头某个字段或总体太大；

HTTP 响应码

5XX 服务器错误

5xx: 服务器错误

- 500 Internal Server Error: 服务器内部错误
- 502 Bad Gateway: 代理服务器无法获取到合法响应
- 503 Service Unavailable: 服务器当前忙, 暂时无法响应服务
- 504 Gateway Timeout: 超时

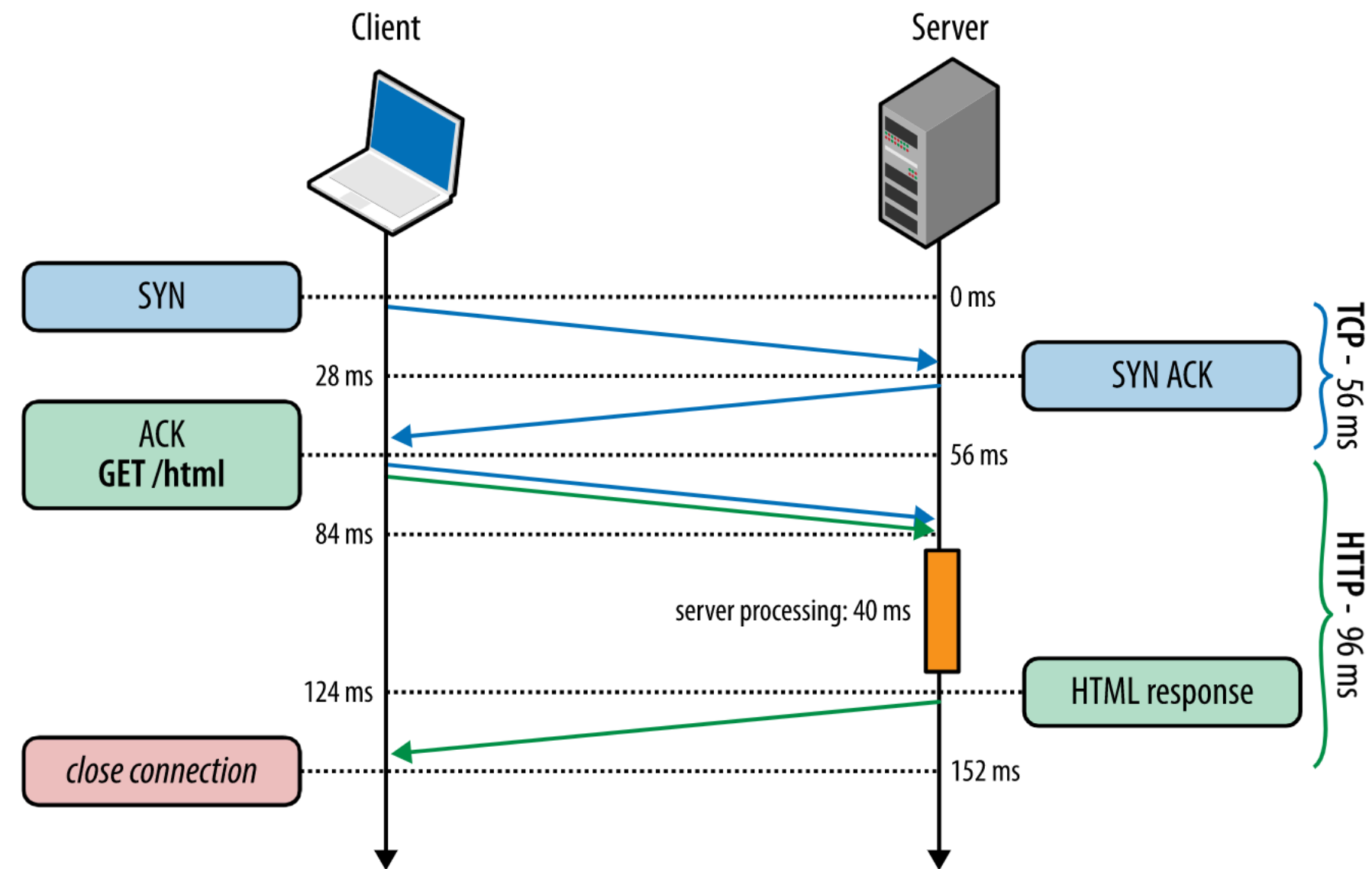
Connection 头部

短连接与长连接

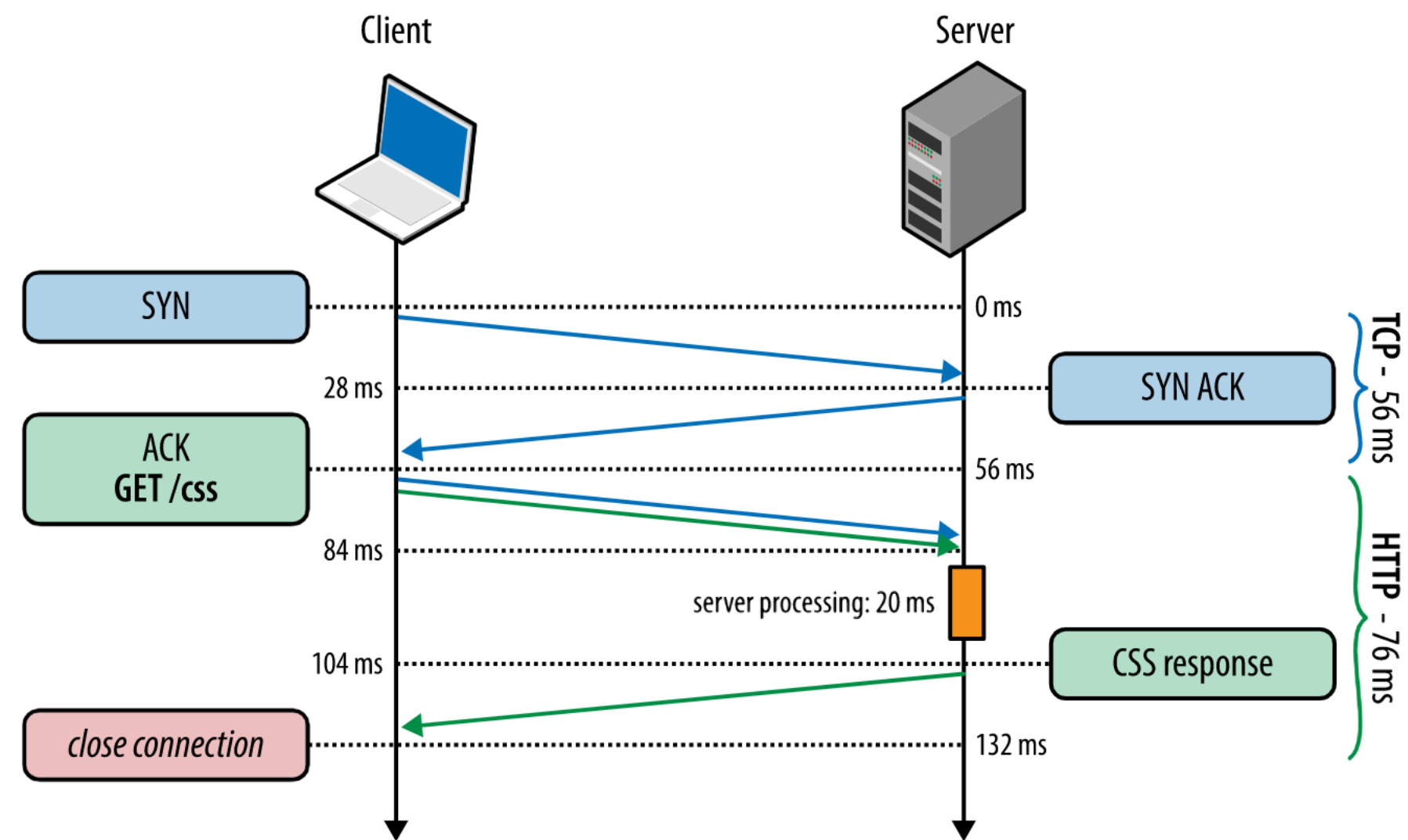
Keep-Alive: 长连接

Close: 短连接

TCP connection #1, Request #1: HTML request



TCP connection #2, Request #2: CSS request



TCP connection #1, Request #1-2: HTML + CSS

