

同濟大學

TONGJI UNIVERSITY

《WEB 技术》

实验报告（大作业）

实验名称 实验一 万维网运行原理分析

小组成员 黃保翔 (2351753)

学院（系） 计算机科学与技术学院

专业 计算机科学与技术

任课教师 郭玉臣

日期 2025 年 3 月 14 日

实验步骤

以www.baidu.com为例

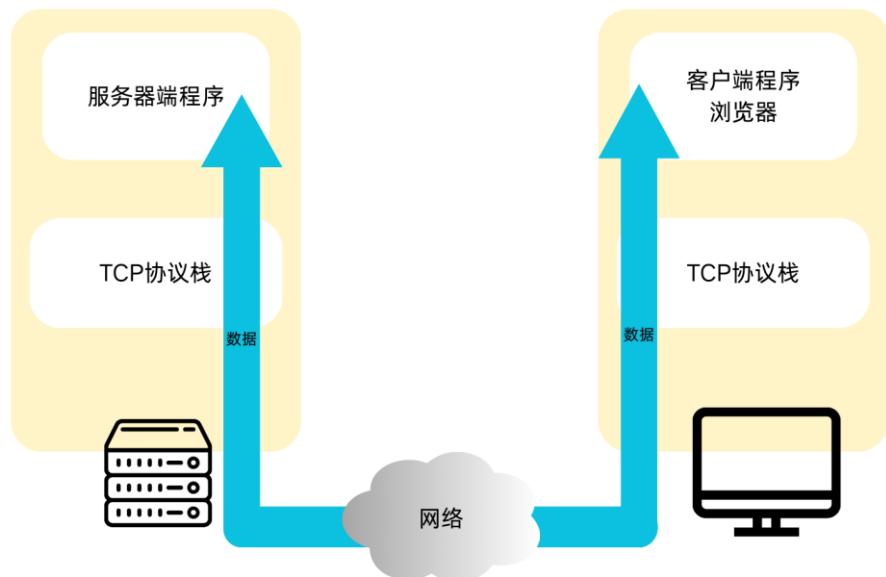
右键打开检查

在元素中可以看到页面的组成元素以及对应的原代码

在网络中点击文档，刷新页面可以看到返回的报文，点击原始，即可看到原始的报文格式

1、绘制分析网络拓扑和数据流向，按照浏览器、协议、服务器、数据等结构；

装
订
线



2、分析所选单个网页组成，按 HTML 组成要素；

选择网页www.baidu.com

2.1 导航栏为 div 盒元素



2.1.1 左侧小盒



由无序列表构成，每个项 li 内为超链接指向对应的 url

```
<a href="http://news.baidu.com" target="_blank" class="mnav c-font-normal c-color-t">
    新闻</a>
```

2.1.2 右侧小盒



内部由 3 个 div 小盒组成，分别为天气，设置，个人中心

```
▼<div id="u1" class="s-top-right s-isindex-wrap">
    ▼<div class="s-weather-wrapper">
        ▼<div id="s_mod_weather" class="s-mod-weather s-isindex-wrap hide-unknow-city">
            ...
            ▶<div class="weather-mod">...</div> == $0
            ▶<div class="unknown-city">...</div>
            ▶<div id="s_mod_setweather" class="s-mod-setweather show-content hide-warning hide">...</div>
        </div>
    </div>
    <span class="s-top-right-text c-font-normal c-color-t s-top-right-new" id="s-usersetting-top" name="tj_settingicon">设置</span>
    ▶<a id="s-top-username" class="s-top-username s-top-right-new" data-tid="2004" href="https://www.baidu.com/my/index" target="_blank">...</a>
    ▶<div id="s-user-name-menu" class="s-top-userset-menu c-floating-box c-font-normal" style="display: none; right: 33.1125px;">...</div>
    ▶<div id="s-user-setting-menu" class="s-top-userset-menu c-floating-box c-font-normal" style="display: none; right: 129px;">...</div>
    ▶<div class="guide-info guide-info-login">...</div>
</div>
```

装
订
线

2.2 中间大盒的主要元素为

Logo 图片

`img#s_lg_img.s_lg_img_gold_shown` 270 × 129



表单元素的搜索框

`form#form.fm.has-soutu` 654 × 44



2.3 热搜 div 盒

`div#s_mancard_main.s-mancard-d-main` 654 × 235



内部由 ul 无序列表填充，列表内部为超链接指向热搜

```
<ul class="s-hotsearch-content" id="hotsearch-content-wrapper">
  <li class="hotsearch-item odd" data-index="0">...</li>
  <li class="hotsearch-item even" data-index="5">...</li>
  <li class="hotsearch-item odd" data-index="1">...</li>
  <li class="hotsearch-item even" data-index="6">...</li>
  <li class="hotsearch-item odd" data-index="2">...</li>
  <li class="hotsearch-item even" data-index="7">...</li>
  <li class="hotsearch-item odd" data-index="3">...</li>
  <li class="hotsearch-item even" data-index="8">...</li>
  <li class="hotsearch-item odd" data-index="4">...</li>
  <li class="hotsearch-item even" data-index="9">...</li>
</ul>
```



3、分析所选网页的 HTTP 协议，包括请求报文、应答报文及报文关键点分析

请求报文

GET / HTTP/1.1 Accept:

text/html,application/xhtml+xml,application/xml;q=0.9,image/avif,image/webp,image/apng,*/*;q=0.8,application/signed-exchange;v=b3;q=0.7 Accept-Encoding: gzip, deflate, br, zstd Accept-Language:

zh-CN,zh;q=0.9,en;q=0.8,en-GB;q=0.7,en-US;q=0.6 Cache-Control: no-cache

Connection: keep-alive Cookie:

BAIDUID BFESS=DDDBB1B3BA4750AF769FEE348778BA3D:FG=1;

bid n=18c5d2e340f80487a25b02; MCITY=-%3A;

BIDUPSID=DDDBB1B3BA4750AF769FEE348778BA3D; PSTM=1734877042;

ZFY=ImayxxGh2kx3rrylsd|OHaUda9ij3COB1yL0KHROquM:C;

BDUSS=ViUnNtMDFUQWxFMndBdENnMkVmcjdXa3dofnMwaXVTWWp1SDRXWk5
VM1pnNzFuRVFBQUFBJCQAAAAAAQAAAAEAAADZImCDAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAA
AAAAAAAAAAAAANn2lWfZ9pVneU;

BDUSS_BFESS=ViUnNtMDFUQWxFMndBdENnMkVmcjdXa3dofnMwaXVTWWp1SD
RXWk5VM1pnNzFuRVFBQUFBJCQAAAAAAQAAAAEAAADZlmcDAAAAAAAAAAAAA
AAA
AAAAAAAAAAAAAAANn2lwfZ9pVneU; BD UPN=12314753;

RT="z=1&dm=baidu.com&si=c284d9ce-3f67-46c6-85f0-e4cda51b58fe&ss=m816

装
订
线

omij&sl=3&tt=7jz&bcn=https%3A%2F%2Ffclog.baidu.com%2Flog%2Fweirwood%3Ftype%3Dperf&ld=d3r&ul=1izhh&hd=1izhs";

H_WISE_SIDS=110085_1992049_626068_628198_632163_632959_633614_637652_632300_639930_637860_641078_641220_641425_641465_641765_642204_642323_642414_642215_642537_639679_642955_642987_642975_643252_643096_643579_641261_643826_641423_643921_643981_644027_643205_644084_644161_641767_644320_644324_644343_644368_644372_644383_644317_644236_644407_644431_644457_644495_644552_643218_644114_644585_644649_644665_644678_644715_644717_644641_644731_627285_644842_644948_644936_644976_645007_644893_645086_643282_643296_645024_645226_645258_645214_643934_645302_644899_645346_645361_645379_645313_645426_645485_637756_645427_645169_645161_645433_645565_645655_643287;

rsv_i=228a0g6w8GGHzGYYoDJNaDftRwoSCH0J884vyvj+Y1fiGiPPtEikb7n70cAhbuLjhVIX6ZhtJ47NagIgqALv5BDewj+alGg;

H_WISE_SIDS_BFESS=110085_1992049_626068_628198_632163_632959_633614_637652_632300_639930_637860_641078_641220_641425_641465_641765_642204_642323_642414_642215_642537_639679_642955_642987_642975_643252_643096_643579_641261_643826_641423_643921_643981_644027_643205_644084_644161_641767_644320_644324_644343_644368_644372_644383_644317_644236_644407_644431_644457_644495_644552_643218_644114_644585_644649_644665_644678_644715_644717_644641_644731_627285_644842_644948_644936_644976_645007_644893_645086_643282_643296_645024_645226_645258_645214_643934_645302_644899_645346_645361_645379_645313_645426_645485_637756_645427_645169_645161_645433_645565_645655_643287;

BA_HECTOR=8ga4a185208k010024018g05ahchlt1jt7oh822;

BDORZ=AE84CDB3A529C0F8A2B9DCDD1D18B695;

H_PS_PSSID=60277_61027_61668_62336_62347_62329_62370_62251_62421_62422_62480_62485_62492_62522_62456_62454_62453_62450_62587 Host:

www.baidu.com Pragma: no-cache Referer: https://cn.bing.com/ Sec-Fetch-Dest:

document Sec-Fetch-Mode: navigate Sec-Fetch-Site: cross-site Sec-Fetch-User: ?1
Upgrade-Insecure-Requests: 1 User-Agent: Mozilla/5.0

1.1 请求行

GET / HTTP/1.1

GET: HTTP 方法, 表示请求资源 (即 www.baidu.com 的首页)。

/: 请求的路径, 表示根目录 (首页)。

HTTP/1.1: 使用的 HTTP 版本。

1.2 请求头

Host: www.baidu.com 请求的目标主机, 表明要访问 www.baidu.com。

Connection: keep-alive 保持 TCP 连接, 避免每次请求都重新建立连接, 提高效率。

Upgrade-Insecure-Requests: 表明希望服务器将 HTTP 连接升级到 HTTPS (如果支持)。

Accept:

text/html,application/xhtml+xml,application/xml;q=0.9,image/avif,image/webp,image/apng,*/*;q=0.8,application/signed-exchange;v=b3;q=0.7

Accept-Encoding: gzip, deflate, br, zstd 客户端支持的压缩格式

Accept-Language: zh-CN,zh;q=0.9,en;q=0.8,en-GB;q=0.7,en-US;q=0.6 客户端期望的语言

Cache-Control: no-cache 表示不使用缓存, 每次都请求最新的资源

Cookie: BAIDUID_BFESS=DDDBB1B3BA4750AF769FEE348778BA3D:FG=1; ... 存储用户的会话信息

Referer: <https://cn.bing.com/> 表明请求来源于 cn.bing.com

User-Agent: Mozilla/5.0 标识客户端的类型 (浏览器、操作系统等)

响应报文

HTTP/1.1 200 OK Connection: keep-alive Content-Encoding: gzip

Content-Security-Policy: frame-ancestors 'self' https://chat.baidu.com

http://mirror-chat.baidu.com https://fj-chat.baidu.com https://hba-chat.baidu.com

https://hbe-chat.baidu.com https://njjs-chat.baidu.com https://nj-chat.baidu.com

https://hna-chat.baidu.com https://hnb-chat.baidu.com

http://debug.baidu-int.com; Content-Type: text/html; charset=utf-8 Date: Fri, 14

Mar 2025 08:41:24 GMT Server: BWS/1.1 Set-Cookie:

H_PS_PSSID=60277_61027_61668_62336_62347_62329_62370_62251_62421_62422
_62480_62485_62492_62522_62456_62454_62453_62450_62587; path=/;

expires=Sat, 14-Mar-26 08:41:24 GMT; domain=.baidu.com Traceid:

1741941684263958631411320554750166404468 X-Ua-Compatible:

IE=Edge,chrome=1 X-Xss-Protection: 1;mode=block Transfer-Encoding: chunked

2.1 响应行

HTTP/1.1 200 OK

HTTP/1.1: 服务器使用的 HTTP 版本。

200 OK: 表示请求成功，服务器正常返回数据。

2.2 响应头

Connection: keep-alive 保持 TCP 连接

Content-Encoding: gzip 服务器返回的内容经过 **gzip 压缩**

Content-Security-Policy: frame-ancestors 'self' https://chat.baidu.com

http://mirror-chat.baidu.com ... 指定哪些网站可以通过 <iframe> 嵌套该页面

Content-Type: text/html; charset=utf-8

text/html: 表示返回的内容是 HTML 网页。 **charset=utf-8**: 采用 UTF-8 编码，支持多语言字符显示。

Date: Fri, 14 Mar 2025 08:41:24 GMT 服务器的时间戳

Server: BWS/1.1 **BWS/1.1**: 表示服务器使用的是 **百度 Web Server (BWS)**

Set-Cookie:

H_PS_PSSID=60277_61027_61668_62336_62347_62329_62370_62251_62421_62422
_62480_62485_62492_62522_62456_62454_62453_62450_62587; path=/;

expires=Sat, 14-Mar-26 08:41:24 GMT; domain=.baidu.com

服务器返回 Cookie **expires=Sat, 14-Mar-26 08:41:24 GMT**: Cookie **有效期为 1 年**，到 2026 年 3 月 14 日过期。

4、心得体会

1 前端 html 只是负责形成页面的骨架，其余的样式由 css 负责，动态响应由 javascript 负责

在进行前端编程时，一般将各个元素嵌套进 div 盒内部，方便进行样式的选择

2 报文包括请求报文和响应报文

HTTP 报文由 **请求行/响应行、请求头/响应头、主体数据** 组成，每个部分都承载着不同的信息：

- **请求报文** 主要描述 **客户端的需求**，比如：
 - GET / HTTP/1.1 表示获取首页内容。
 - User-Agent 说明了用户的设备信息。
 - Cookie 带有用户的身份标识，可能影响服务器返回的内容。
- **响应报文** 主要描述 **服务器的反馈**，比如：
 - 200 OK 说明请求成功。
 - Content-Encoding: gzip 说明数据被压缩，提升传输效率。
 - Set-Cookie 可能影响后续请求，比如记录登录状态。

装
订
线