

同濟大學

TONGJI UNIVERSITY

《WEB 技术》

实验报告（大作业）

实验名称

万维网运行原理分析

小组成员

王灏博（2153067）

学院（系）

电子信息与工程学院

专 业

计算机科学与技术

任课教师

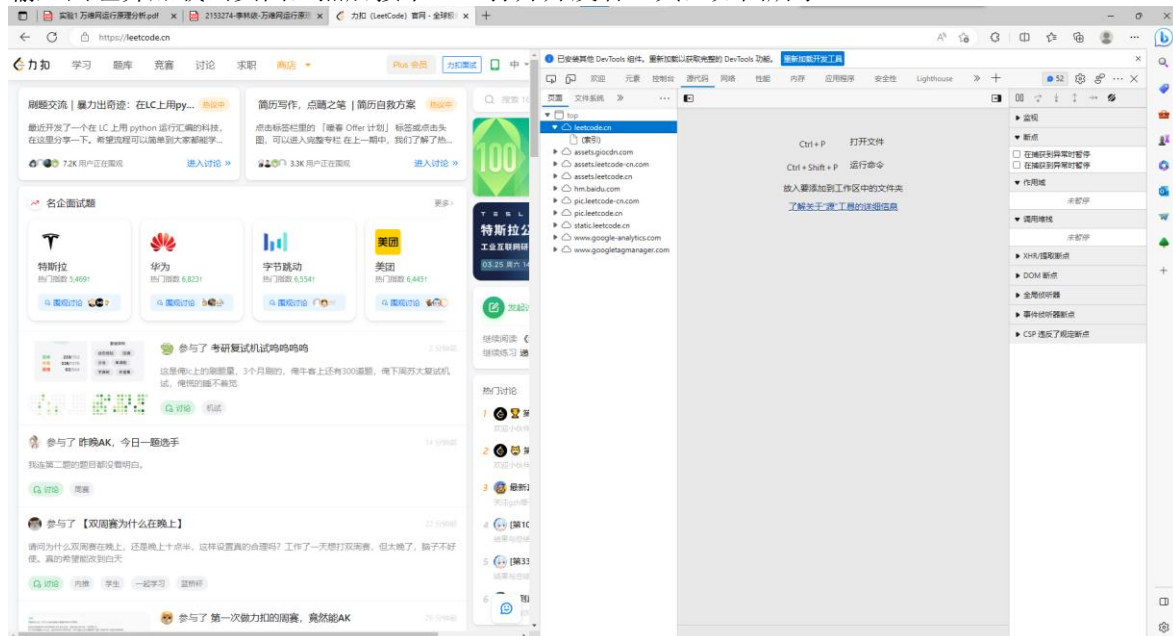
郭玉臣

日 期

2023 年 3 月 18 日

一、选择网址

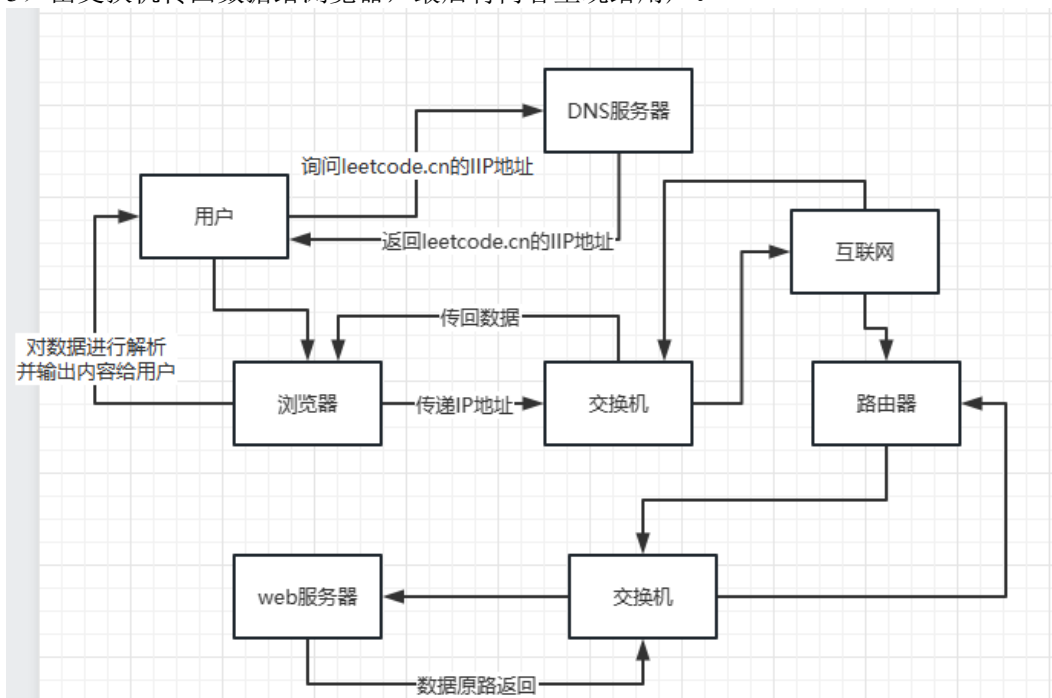
本次实验选择<https://leetcode.cn>，即力扣中文社区作为本次实验的网址。在 Google Chrome 输入网址并加载出页面，然后按下 F12 打开开发者工具，如图所示。



二、分析网络拓扑和数据流向

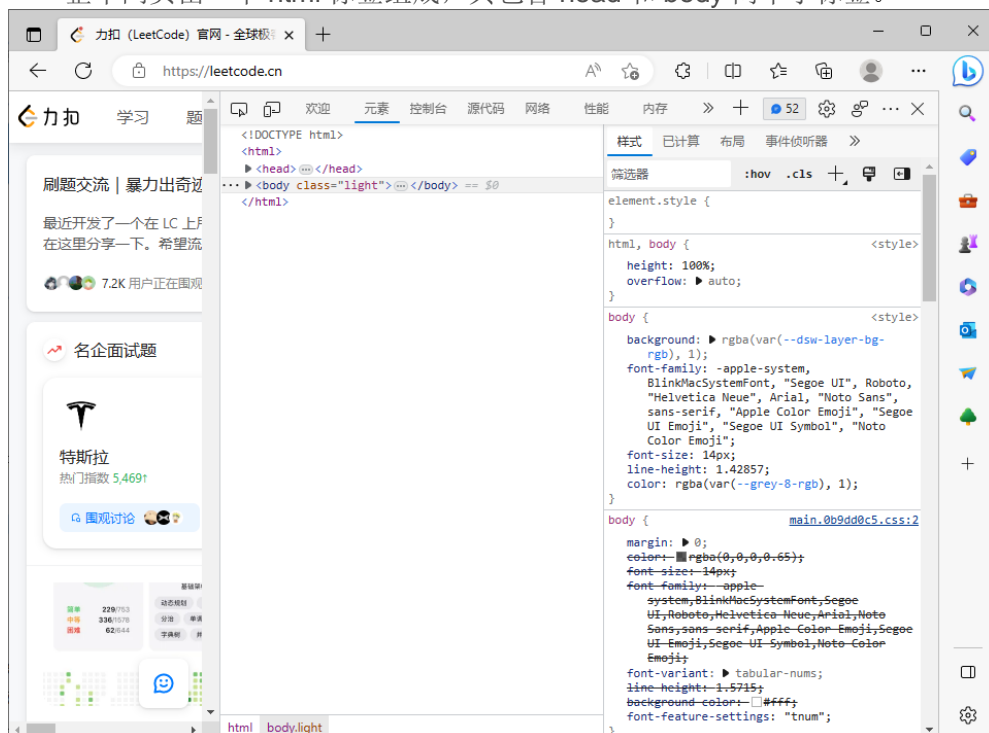
从在浏览器网址框输入网址到看到页面呈现，其间发生如下过程：

- 1) 浏览器向域名服务器（DNS）询问<https://leetcode.cn>的IP地址；
- 2) DNS 将<https://leetcode.cn>的IP地址返回给浏览器；
- 3) 浏览器经过交换机、路由器和互联网等向 web 传递 IP 地址；
- 4) web 通过 HTTP 协议解析、封装数据，并将该数据原路返回；
- 5) 由交换机传回数据给浏览器，最后将内容呈现给用户。

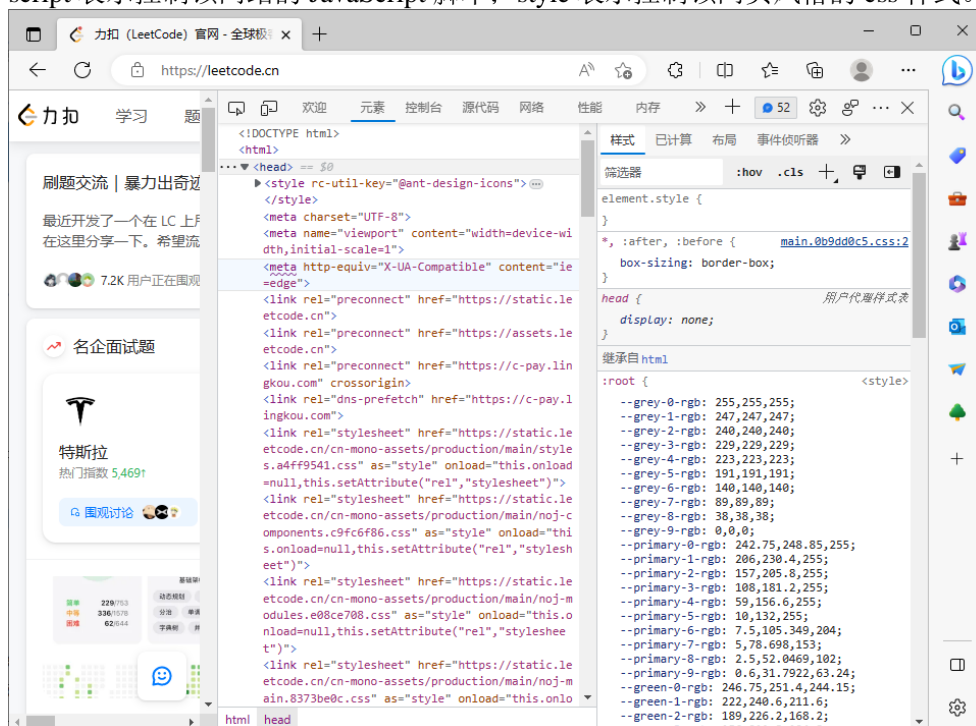


三、分析网页组成

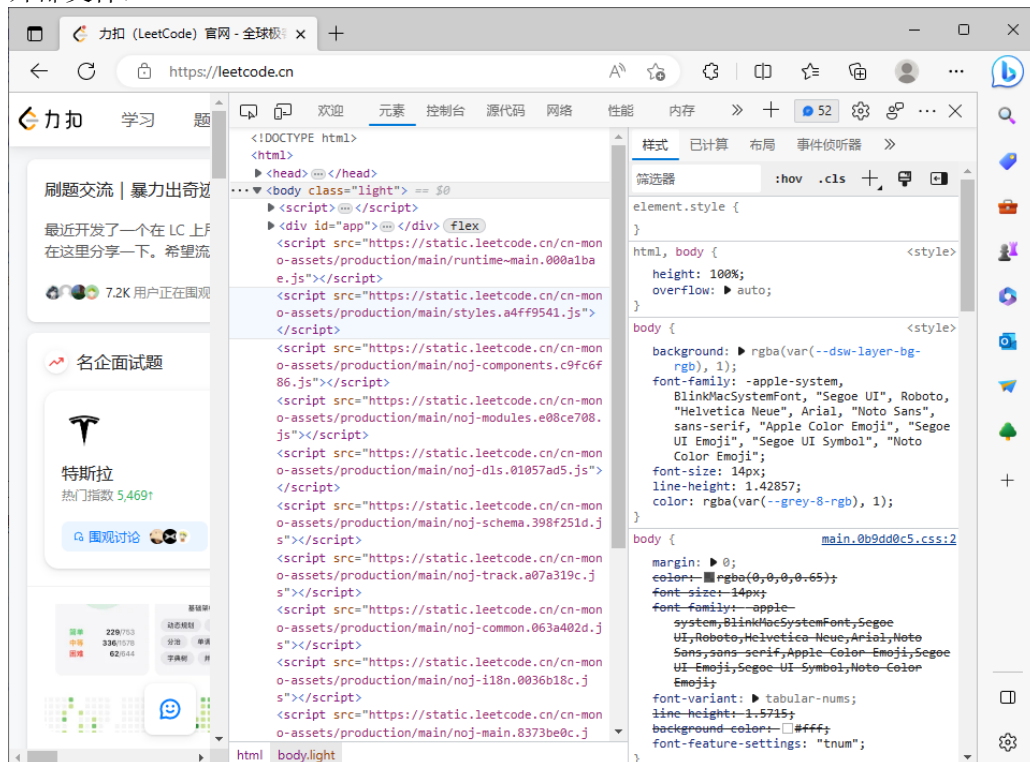
整个网页由一个 html 标签组成，其包含 head 和 body 两个子标签。



对于子标签 head，它含有 title、meta、link、script、style 等子标签。在这些子标签中，title 只有一个，表示网页的标题；meta 提供一些网页的基本信息；link 表示链接到外部文件的信息；script 表示控制该网站的 JavaScript 脚本；style 表示控制该网页风格的 css 样式。

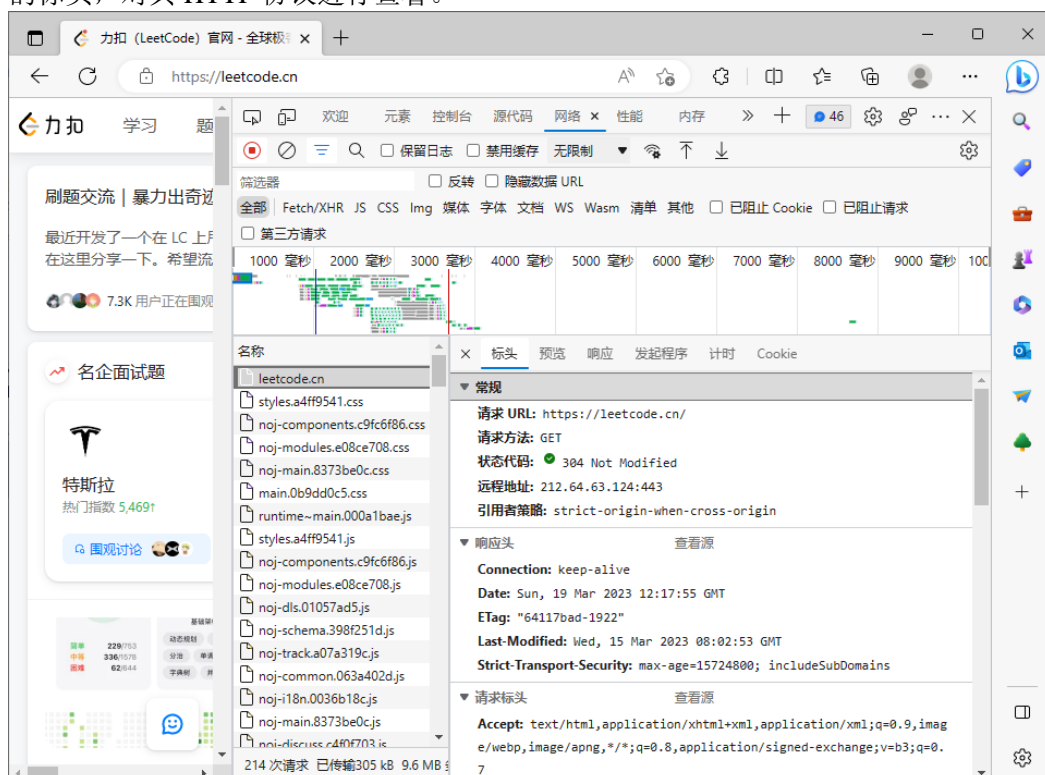


对于子标签 `body`，它表示网页的主体部分。Leetcode 中文社区有很多小方块，由 `div` 标签控制，每个小方块有各自的属性和风格；同时有很多 `script` 标签来进行脚本的编写（很多都链接到外部文件）。



四、分析 HTTP 协议

在开发者工具中点击网络标签，并刷新当前页面，在下方的名称中找到 `leetcode.cn`，电吉他的标头，对其 HTTP 协议进行查看。



×

标头

预览

响应

发起程序

计时

Cookie

▼ 常规

请求 URL: https://leetcode.cn/

请求方法: GET

状态代码: 304 Not Modified

远程地址: 212.64.63.124:443

引用者策略: strict-origin-when-cross-origin

▼ 响应头

查看源

Connection: keep-alive

Date: Sun, 19 Mar 2023 12:17:55 GMT

ETag: "64117bad-1922"

Last-Modified: Wed, 15 Mar 2023 08:02:53 GMT

Strict-Transport-Security: max-age=15724800; includeSubDomains

▼ 请求标头

查看源

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

Accept-Encoding: gzip, deflate, br

Accept-Language: zh-CN,zh;q=0.9,en;q=0.8,en-GB;q=0.7,en-US;q=0.6

Cache-Control: max-age=0

Host: leetcode.cn

If-Modified-Since: Wed, 15 Mar 2023 08:02:53 GMT

If-None-Match: W/"64117bad-1922"

sec-ch-ua: "Microsoft Edge";v="111", "Not(A:Brand";v="8", "Chromium";v="111"

sec-ch-ua-mobile: ?0

sec-ch-ua-platform: "Windows"

Sec-Fetch-Dest: document

Sec-Fetch-Mode: navigate

Sec-Fetch-Site: none

Sec-Fetch-User: ?1

Upgrade-Insecure-Requests: 1

User-Agent: Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/111.0.0.0 Safari/537.36 Edg/111.0.1661.44

1、请求报文

从请求报文中，我们可以得到如下信息：

- 1)Accept 表示客户机支持的数据类型
- 2)Accept-Encoding 表示客户及支持的数据压缩格式
- 3)Accept-Language 表示客户及支持的语言
- 4)Cache-Control 表示对资源缓存进行控制
- 5)Cookie 表示存放数据，将客户端的 Cookie 传送给服务端
- 6)User-Agent 表示用户的软件环境

2、响应报文

- 1)Content-Type 表示网页内容类型和编码方式
- 2)Date 表示当前时间
- 3)ETag 表示服务器资源是否变化

五、心得体会

通过此次实验，我大致明白了网页的运行框架，对于一个完善的网站，特别是这种大型知名网站，要兼顾美观、安全、规范等要求。在编写网站的内容部分时，尽量将每一条信息分好类、通过不同的标签分别管理，不仅可以提高效率，也可以降低出错的概率，同时便于维护，值得学习。与此同时，也感慨互联网发展速度如此迅猛，认识到我们正在学习的是一个时刻都在进步的学科。