

同濟大學

TONGJI UNIVERSITY

《WEB 技术》

实验报告

实验名称

实验 1 万维网运行原理分析

姓名

2151393 程子洋

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

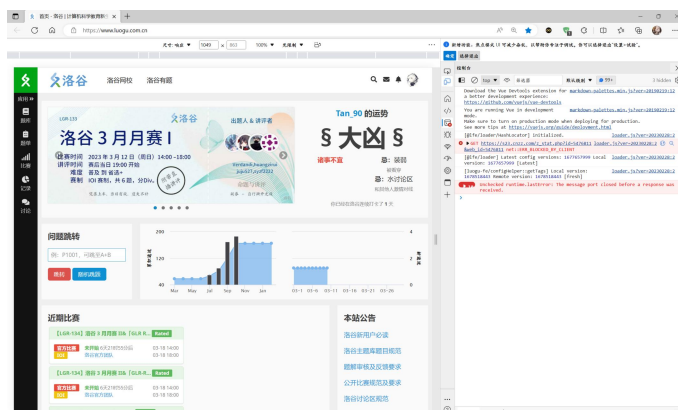
2023 年 3 月 11 日

一、实验步骤

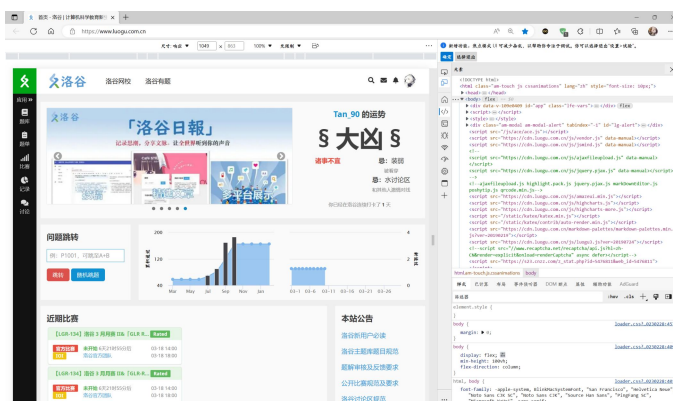
1. 选取网址：本次实验选取算法学习平台洛谷 <https://www.luogu.com.cn/>作为实验网址。
2. 进入网址：首先在浏览器中输入网站网址，进入网站页面。



3. 网站分析：进入网站后打开开发者工具，即可进行网页分析。

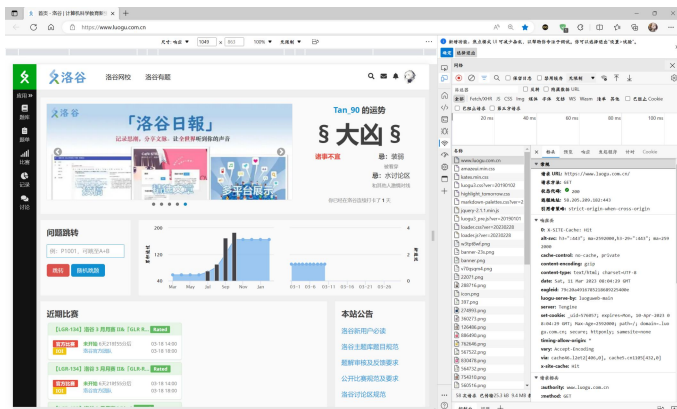


4. 分析网页组成：在开发者工具中选择“元素”一项，即可看到网页的 html 组成。



5. 分析 HTTP 协议：在开发者工具中选择“网络”一项。从“名称”一项中找到网站对应的域名(<https://www.luogu.com.cn/>), 之后在右侧导航栏中选择“标头”一项，即可对其

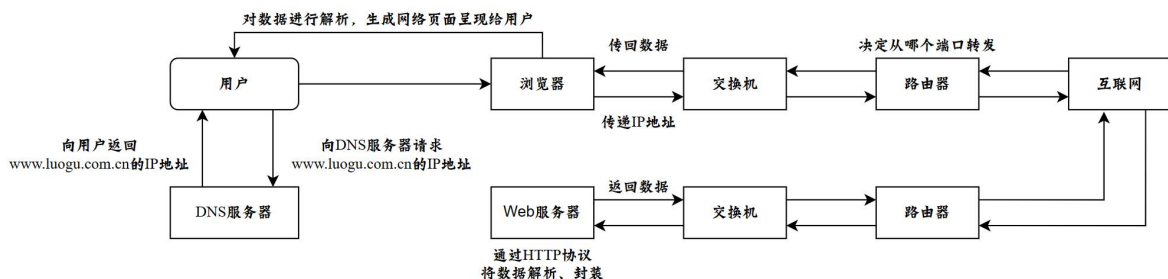
HTTP 协议进行查看。



二、实验内容

(1) 分析网络拓扑和数据流向

如图所示：



(2) 分析单个网页组成

```

<!DOCTYPE html>
<html class="am-touch js cssanimations" lang="zh" style="font-size: 10px;">
  <head> ... </head>
  <body> ... </body>
</html>
  
```

可以看到，整个网页由一个 html 标签组成，分为“head”和“body”两个组成成分，在 html 标签中，通过属性“class”指明了类名为“am-touch js cssanimations”，通过属性“lang”指明了网站的语言为简体中文，通过“style”中的“font-size”参数指明了网页的字体大小为 10px；

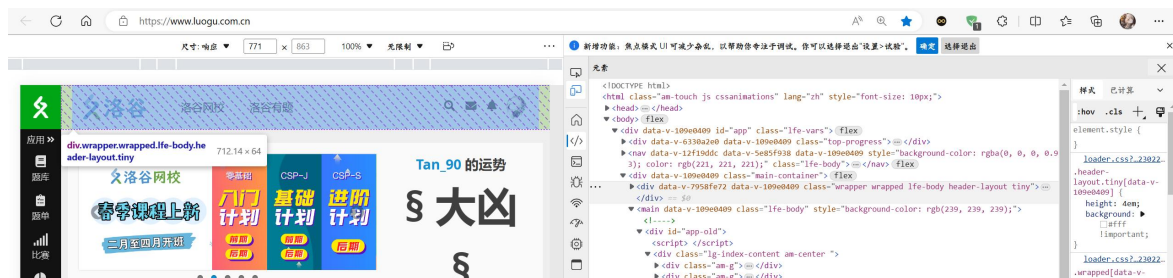
```

<head>
  <style type="text/css"> ... </style>
  <style> ... </style>
  <style id="ace-tm"> ... </style>
  <style id="ace_editor.css"> ... </style>
  <meta charset="utf-8">
  <meta http-equiv="X-UA-Compatible" content="IE=edge">
  <title>首页 - 洛谷 | 计算机科学教育新生态</title>
  <meta name="viewport" content="width=device-width, initial-scale=1, maximum-scale=1, user-scalable=no">
  <meta name="renderer" content="webkit">
  
```



```
<!DOCTYPE html>
<html class="am-touch js cssanimations" lang="zh" style="font-size: 10px;">
  <head> ... </head>
  <body> <flex>
    <div data-v-109e0409 id="app" class="lfe-vars"> <flex>
      <div data-v-6330a2e0 data-v-109e0409 class="top-progress"> ... </div>
      ...
      <nav data-v-12f19ddc data-v-5e85f938 data-v-109e0409 style="background-color: rgba(0, 0, 0, 0.93); color: rgb(221, 221, 221);" class="lfe-body"> ... </nav> <flex == $0>
        <div data-v-109e0409 class="main-container"> <flex>
          <div data-v-7958fe72 data-v-109e0409 class="wrapper wrapped lfe-body header-layout tiny"> ... </div>
          <main data-v-109e0409 class="lfe-body" style="background-color: rgb(239, 239, 239);">
            <!--
            <div id="app-old">
              <script> </script>
              <div class="lg-index-content am-center">
                <div class="am-g"> ... </div>
                <div class="am-g"> ... </div>
                <div class="am-g"> ... </div>
              </div>
              <div class="am-modal am-modal-prompt" tabindex="-1" id="report">
                <div class="am-modal-dialog"> ... </div>
              </div>
              <script> ... </script>
            </div>
            <div data-v-7958fe72 data-v-109e0409 class="wrapper wrapped lfe-body"> ... </div>
          </main>
        </div>
      <div data-v-7958fe72 data-v-109e0409 class="wrapper wrapped lfe-body"> ... </div>
    </div>
  </body>
</html>
```

“body”部分主要包含了大量 div 标签，将网页布局划分成了许多方块，同时使用 css 分别设置这些方块的样式，从而构成了整个页面。



在开发者工具中，我们可以轻触每一个标签，来观察其对应网页中的哪一部分。如图中轻触到的 div 标签对应网页中的上方导航栏。

据此，在分析整个页面的 div 标签后，我们可以得到洛谷网页的划分：



(3) 分析 HTTP 协议

▼ 请求头

```
:authority: www.luogu.com.cn
:method: GET
:path: /
:scheme: https
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
cookie: __client_id=8b5093640f50e49c93a448a1c453a4aa44dbda2f; _uid=576057; acw_tc=79c20a4b16785270000061296ecf83279181967017efecc88d41997740; cdn_sec_tc=79c20a4b16785270000061296ecf83279181967017efecc88d41997740
sec-ch-ua: "Chromium";v="110", "Not A(Brand";v="24", "Microsoft Edge";v="110"
sec-ch-ua-mobile: ?1
sec-ch-ua-platform: "Android"
sec-fetch-dest: document
sec-fetch-mode: navigate
sec-fetch-site: same-origin
sec-fetch-user: ?1
upgrade-insecure-requests: 1
user-agent: Mozilla/5.0 (Linux; Android 6.0; Nexus 5 Build/MRA58N) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/110.0.0.0 Mobile Safari/537.36 Edg/110.0.1587.69
```

请求报文信息如上图所示。

请求方法：GET 方法

请求 URL：www.luogu.com.cn

客户端接受的响应类型：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

客户端接受的数据压缩格式：gzip, deflate, br

客户端的 Cookie 信息：

```
__client_id=8b5093640f50e49c93a448a1c453a4aa44dbda2f;
_uid=576057;
acw_tc=79c20a4b16785270000061296ecf83279181967017efecc88d41997740;
cdn_sec_tc=79c20a4b16785270000061296ecf83279181967017efecc88d41997740
```

客户端缓存时间：最大时间为 0，内容接收即认为陈旧失效，不断进行重新获取

▼ 响应头

```
0: X-SITE-Cache: Hit
alt-svc: h3=":443"; ma=2592000,h3-29=":443"; ma=2592000
cache-control: no-cache, private
content-encoding: gzip
content-type: text/html; charset=UTF-8
date: Sat, 11 Mar 2023 10:28:32 GMT
eagleid: 79c20a4716785305119426861e
luogu-serve-by: luoguweb-ess198
server: Tengine
set-cookie: _uid=576057; expires=Mon, 10-Apr-2023 10:28:32 GMT; Max-Age=2592000; path=/; domain=.luogu.com.cn; secure; httponly; samesite=none
timing-allow-origin: *
vary: Accept-Encoding
via: cache36.l2et2-2[874,0], cache3.cn1105[902,0]
x-site-cache: Hit
```

响应报文信息如上图所示。

▼ 常规

请求 URL: <https://www.luogu.com.cn/>

请求方法: GET

状态代码: 200

远程地址: 58.205.209.179:443

引用策略: strict-origin-when-cross-origin

响应状态: 请求成功（该项显示在“常规”中，见上图）

客户端缓存控制: 可以缓存，需要使用对比缓存来验证缓存数据

Cookie 设置:

_uid=576057;

expires=Mon, 10-Apr-2023 10:28:32 GMT;

Max-Age=2592000; （Cookie 有效时间为 30 天）

path=/; domain=.luogu.com.cn; secure;

httponly;

samesite=none

数据压缩格式: gzip

内容类型: text/html

字符集: UTF-8

三、心得体会

在此次实验中，我从网络拓扑和数据流向、网页组成、HTTP 协议三个角度对网页进行了深入的探究。通过此次实验，我了解到了一个美观、系统、规范的网站的架构及其背后的技术原理。想要做好一个网站，不仅需要 UI 设计师对网页进行设计美化，也需要前端工程师通过熟练的 html、css 和 JS 知识对网页进行搭建，更需要后端设计师对数据库、服务器进行构建与维护。一个网站，是团队合作的结果，也是集体智慧的结晶。代码的隐蔽性、规范性，协议的完备性，都直接影响着最终的成果。

此外，在 HTTP 协议的分析中，我才真正了解到，一次请求、一次应答，其中蕴含了如此多的信息，有如此多的属性参与到了报文的传输过程中，我也不得不感叹，目前自己接触到的网页知识仅是冰山一角，未来还有更多的知识等着我们去探索。