

به نام خدا

تهیه کننده: ابراهیم صدیقی

شماره دانشجویی: ۹۹۳۱۰۹۸

سوال (۱)

الف) بهینه نبودن پروتکل HTTP از ماهیت بدون حالت آن ناشی می‌شود، که منجر به افزایش تاخیر به دلیل تنظیمات مکرر اتصال و قطع شدن آن می‌شود. علاوه بر این، فقدان رمز گذاری داخلی و سربار سربرگ به ناکارآمدی کمک می‌کند.
ب) با وجود این اشکالات، اولین نسخه از پروتکل (HTTP/1.0) به دلیل سادگی، سازگاری با سیستم های قدیمی و سهولت پیاده سازی همچنان به طور گسترده مورد استفاده قرار می‌گیرد. بسیاری از برنامه‌ها و سرور های وب موجود به دلیل پشتیبانی و آشنایی گسترده به HTTP/1.0 متکی هستند، حتی اگر نسخه‌های جدیدتر مانند HTTP/2 و HTTP/3 پیشرفت‌های قابل توجهی در عملکرد ارائه می‌دهند.

سوال (۲)

HTTP3 که با نام HTTP/3 نیز شناخته می‌شود، آخرین نسخه پروتکل انتقال ابرمتن است. یکی از تغییرات مهم در HTTP3 پروتکل حمل و نقل اساسی است. برخلاف پیشینیان خود که از TCP استفاده می‌کردند، HTTP3 از Quick UDP Internet Connections (QUIC) به عنوان پروتکل انتقال استفاده می‌کند. هدف این تغییر بهبود عملکرد با کاهش تاخیر و افزایش امنیت است.
یکی دیگر از تفاوت های کلیدی، چندقدر سازی درخواست ها است. در HTTP3 چندین درخواست را می‌توان به طور همزمان از طریق یک اتصال ارسال کرد که باعث بهبود کارایی و سرعت بارگذاری صفحات وب می‌شود. علاوه بر این، HTTP3 دارای ویژگی هایی مانند اولویت بندی جریان و فشرده سازی هدر است که انتقال داده ها را بیشتر بهینه می‌کند.
این پیشرفت‌ها در HTTP3 منجر به تجربه مرور وب سریعتر و ایمن‌تر برای کاربران می‌شود. توانایی پروتکل برای مدیریت بهتر از دست دادن بسته ها و برقراری سریعتر اتصالات به بهبود عملکرد، به ویژه در شبکه های با تاخیر بالا کمک می‌کند. به طور کلی، HTTP3 یک پروتکل ارتباطی کارآمدتر و قابل اعتمادتر برای برنامه های کاربردی وب مدرن ارائه می‌دهد.

سوال (۳)

CRUD در پروتکل HTTP به عملیات ایجاد، خواندن، به‌روزرسانی و حذف اشاره دارد. این عملیات به ترتیب با روش های GET، PUT، POST و DELETE مطابقت دارد.
هنگام دسته بندی روش های HTTP، دو جنبه اصلی را در نظر می‌گیریم: ایمنی و عدم توانایی. روش های ایمن مانند GET منابع را تغییر نمی‌دهند، در حالی که روش های ناایمن مانند POST می‌توانند وضعیت سرور را تغییر دهند. روش های Idempotent، مانند GET و PUT، می‌توانند بدون تغییر نتیجه تکرار شوند، در حالی که روش های غیر توانمند مانند POST ممکن است با هر اجرا به نتایج متفاوتی منجر شوند.
در اینجا نحوه هماهنگی عملیات CRUD با دسته ها آمده است:

- Safe and sound: read (GET)
- Safe and non-competent: unacceptable
- Insecure and Powerless: Update (PUT)
- Insecure and non-capable: create (POST), delete (DELETE)

سوال (۴)

Flag httpOnly در هدر Set-Cookie یک ویژگی امنیتی مهم است که به کاهش حملات اسکریپت بین سایتی (XSS) کمک می‌کند. وقتی این پرچم تنظیم می‌شود، دسترسی کوکی را فقط به درخواست های HTTP محدود می‌کند و از دسترسی اسکریپت های سمت سرویس گیرنده به آن جلوگیری می‌کند. این باعث افزایش امنیت اطلاعات حساس ذخیره شده در کوکی‌ها می‌شود، مانند شناسه‌های جلسه یا نشانه‌های احراز هویت.

برای پیاده سازی `flag httpOnly` در یک کوکی، به سادگی آن را در هدر `Set-Cookie` به این صورت قرار دهید:

```
Set-Cookie: sessionId=12345; Secure; HttpOnly
```

با افزودن `HttpOnly` به تنظیمات کوکی خود، سیستم دفاعی برنامه وب خود را در برابر اسکریپت های مخربی که سعی در سرقت داده های حساس دارند، تقویت می کنید. بهترین روش استفاده از `flag httpOnly` برای کوکی های حاوی اطلاعات حساس است.

(سوال ۵)

در برنامه های کاربردی وب، یک پروکسی فرورود از طرف کلاینت ها عمل می کند و درخواست ها را رهگیری و به سرورها ارسال می کند. این به عنوان یک واسطه برای درخواست های مشتری برای دسترسی به منابع از اینترنت عمل می کند. از سوی دیگر، یک پروکسی معکوس در مقابل سرورها قرار می گیرد و درخواست های مشتری ورودی را مدیریت می کند و آنها را در سرورهای باطن مناسب توزیع می کند. با بارگذاری وظایف از سرورها، امنیت، عملکرد و مقیاس پذیری را افزایش می دهد. پروکسی های فرورود معمولاً برای فیلتر کردن محتوا، کنترل دسترسی و ناشناس بودن استفاده می شوند. ابزارهایی مانند `Squid` و `Apache HTTP Server` و `Proxy` می توانند به عنوان پروکسی های فرورود عمل کنند. پروکسی های معکوس برای متعادل سازی بار، خاتمه `SSL` و ذخیره سازی استفاده می شود. نمونه هایی از ابزارهای پروکسی معکوس شامل `Nginx` و `HAProxy` هستند که به طور موثر ترافیک ورودی به سرورهای باطن را مدیریت می کنند.

(سوال ۶)

`gRPC`، یک چارچوب `RPC` با کارایی بالا و منبع باز (`Remote Procedure Call`) که توسط گوگل توسعه یافته است، چندین مزیت را نسبت به ارتباطات سنتی `HTTP` ارائه می دهد. یکی از مزیت های کلیدی کارایی آن در سریال سازی داده ها با استفاده از بافرهای پروتکل است که نسبت به `JSON` که در `API` های `HTTP` استفاده می شود فشرده تر و سریعتر پردازش می شوند. `gRPC` همچنین از جریان دو طرفه پشتیبانی می کند و امکان ارتباط بلادرنگ بین مشتریان و سرورها را فراهم می کند. علاوه بر این، `gRPC` پشتیبانی داخلی برای متعادل کردن بار، احراز هویت و تولید کد خودکار در چندین زبان ارائه می کند و بهروری توسعه دهنده را افزایش می دهد. به طور کلی، ویژگی های `gRPC` آن را به انتخابی قانع کننده برای ساخت برنامه های کاربردی وب سریع، قابل اعتماد و کارآمد در مقایسه با `HTTP` تبدیل می کند.

(سوال ۷)

سرویس `Shekan` که به دور زدن تحریم ها معروف است، با تغییر مسیر ترافیک شبکه از طریق سرورهای پروکسی واقع در مناطق نامحدود عمل می کند. با پنهان کردن مبدأ ترافیک، از شناسایی فرار می کند و به کاربران اجازه می دهد به محتوای محدود دسترسی داشته باشند. این فرآیند شامل رمزگذاری داده ها برای جلوگیری از نظارت و استفاده از پروتکل های تونل سازی برای ایجاد مسیرهای امن است. درک مفاهیمی مانند سرورهای پروکسی، رمزگذاری و تونل سازی برای درک اینکه چگونه سرویس `Shekan` کاربران را قادر می سازد تا به طور موثر تحریم ها را دور بزنند، بسیار مهم است.

(سوال ۸)

شبکه تحویل محتوا (`CDN`) شبکه ای از سرورهای توزیع شده است که به طور استراتژیک در مکان های مختلف قرار گرفته اند تا محتوای وب را به طور کارآمدتر به کاربران ارائه دهند. `CDN` ها با ذخیره محتوا در سرورهای نزدیک به کاربر، تأخیر را کاهش می دهند و تحویل محتوا را سرعت می بخشند. این منجر به زمان بارگذاری سریع تر، بهبود عملکرد وب سایت و بهبود تجربه کاربری می شود.

`CDN` ها با ذخیره منابع ثابت مانند تصاویر، فایل های `CSS` و اسکریپت ها در سرورهای سراسر جهان کار می کنند. هنگامی که کاربر یک صفحه وب را درخواست می کند، `CDN` به جای سرور مبدأ، این دارایی های ثابت را از نزدیک ترین سرور ارائه می

دهد و فاصله ای که داده ها برای طی کردن نیاز دارند را کاهش می دهد. این نزدیکی زمان بارگذاری را کاهش می دهد و تجربه مرور روان تری را برای کاربران، به ویژه برای مخاطبان جهانی تضمین می کند.

در اصل، CDN ها با استفاده از سرورهای توزیع شده جغرافیایی، تحویل محتوا را بهینه می کنند و در نهایت سرعت وب و رضایت کاربر را افزایش می دهند.

سوال عملی)

این سوال را با استفاده از BeautifulSoup با زبان python پیاده سازی میکنیم. در ابتدا باید یک صفحه که تمام تعطیلات رسمی را نشان میدهد پیدا کنیم ([این لینک](#)). حال این سایت را get میکنیم و html سازنده سایت را استخراج میکنیم. با استفاده از BeautifulSoup این html را فیلتر می کنیم و داده های مورد نظر را استخراج و به قالب قابل فهم به کاربر نشان می دهیم.

عکس کد:


```
1 import json
2
3 import requests
4 from bs4 import BeautifulSoup
5
6 url = "https://www.time.ir/fa/eventyear-%d8%aa%d9%82%d9%88%db%8c%d9%85-%d8%b3%d8%a7%d9%84%db%8c%d8%a7%d9%86%d9%87"
7
8 page = requests.get(url)
9
10 soup = BeautifulSoup(page.content, 'html.parser')
11
12 year = []
13
14 for month in soup.find_all('div', class_='eventsCurrentMonthWrapper'):
15     month_data = {
16         "month": {
17             "name": "",
18             "holidays": []
19         }
20     }
21     month_data["month"]["name"] = month.find('div', class_='eventsCurrentMonthTitle').text.strip()
22     event_holidays = month.find_all('li', class_='eventHoliday')
23     for event in event_holidays:
24         holiday_data = {
25             "date": event.find('span').text.strip(),
26             "ocation": event.contents[2].strip()
27         }
28         month_data["month"]["holidays"].append(holiday_data)
29
30     year.append(month_data)
31
32 json_output = json.dumps(year, ensure_ascii=False, indent=4)
33 print(json_output)
```

خروجی:

```
/usr/bin/python3 /home/lwall/Documents/uni/1402-1403-second-term/web/HW1/crawler/crawler.py
```

```
[  
  {  
    "month": {  
      "name": "مناسبت های ماه فروردین",  
      "holidays": [  
        {  
          "date": "۱ فروردین",  
          "ocation": "جشن نوروز/جشن سال نو"  
        },  
        {  
          "date": "۲ فروردین",  
          "ocation": "عید نوروز"  
        },  
        {  
          "date": "۳ فروردین",  
          "ocation": "عید نوروز"  
        },  
        {  
          "date": "۴ فروردین",  
          "ocation": "عید نوروز"  
        },  
        {  
          "date": "۱۲ فروردین",  
          "ocation": "روز جمهوری اسلامی"  
        },  
        {  
          "date": "۱۳ فروردین",  
          "ocation": "جشن سیزده به در"  
        },  
        {  
          "date": "۱۳ فروردین",  
          "ocation": "شهادت حضرت علی علیه السلام"  
        },  
        {  
          "date": "۲۲ فروردین",  
          "ocation": "عید سعید فطر"  
        },  
        {  

```

```
>  crawler.py
```

```

    },
    {
      "date": "۲۲ فروردین",
      "ocation": "عید سعید فطر"
    },
    {
      "date": "۲۳ فروردین",
      "ocation": "تعطیل به مناسبت عید سعید فطر"
    }
  ]
},
{
  "month": {
    "name": "مناسبت های ماه اردیبهشت",
    "holidays": [
      {
        "date": "۱۵ اردیبهشت",
        "ocation": "شهادت امام جعفر صادق علیه السلام"
      }
    ]
  }
},
{
  "month": {
    "name": "مناسبت های ماه خرداد",
    "holidays": [
      {
        "date": "۱۴ خرداد",
        "ocation": "رحلت حضرت امام خمینی"
      },
      {
        "date": "۱۵ خرداد",
        "ocation": "قیام 15 خرداد"
      },
      {
        "date": "۲۸ خرداد",
        "ocation": "عید سعید قربان"
      }
    ]
  }
}

```