

第6章

应用层

6.6 电子邮件

01

电子邮件的作用

02

电子邮件系统的组成

03

简单邮件传送协议SMTP的基本工作过程

04

电子邮件的信息格式

05

多用途因特网邮件扩展

06

常用的邮件读取协议

07

基于万维网的电子邮件

01 电子邮件的作用

■ 电子邮件E-mail是因特网上**最早流行的一种应用**，并且仍然是当今因特网上最重要、最实用的应用之一。

■ 传统的电话通信属于实时通信，存在以下两个缺点：

- ☐ 电话通信的主叫和被叫双方必须同时在场；
- ☐ 一些不是十分紧迫的电话也常常不必要地打断人们的工作或休息。

■ 而电子邮件与邮政系统的寄信相似。

- ① 发件人将邮件发送到自己使用的**邮件服务器**；
- ② 发件人的邮件服务器将收到的邮件按其目的地址转发到收件人邮件服务器中的收件人邮箱；
- ③ 收件人在方便的时候访问收件人邮件服务器中自己的邮箱，获取收到的电子邮件。

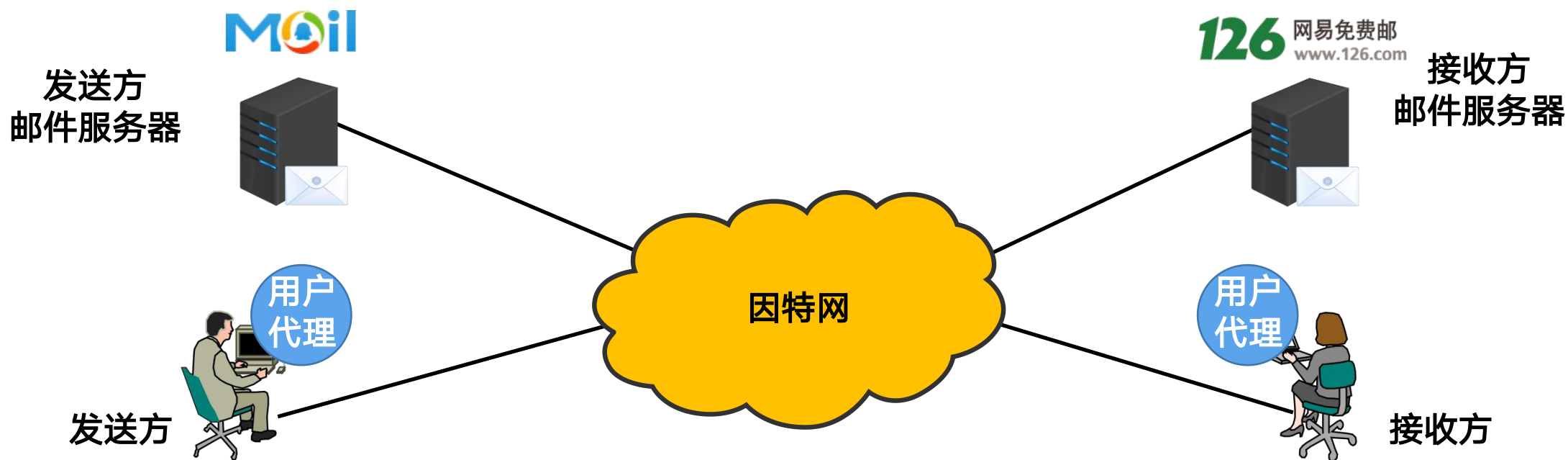
02 电子邮件系统的组成

■ 电子邮件系统采用**客户/服务器方式**。

■ 电子邮件系统的三个主要组成构件：**用户代理**，**邮件服务器**，以及**电子邮件所需的协议**。

☐ **用户代理**是用户与电子邮件系统的接口，又称为**电子邮件客户端软件**。

☐ **邮件服务器**是电子邮件系统的基础设施。因特网上所有的因特网服务提供者ISP都有邮件服务器，其功能是**发送和接收邮件**，同时还要负责维护用户的邮箱。



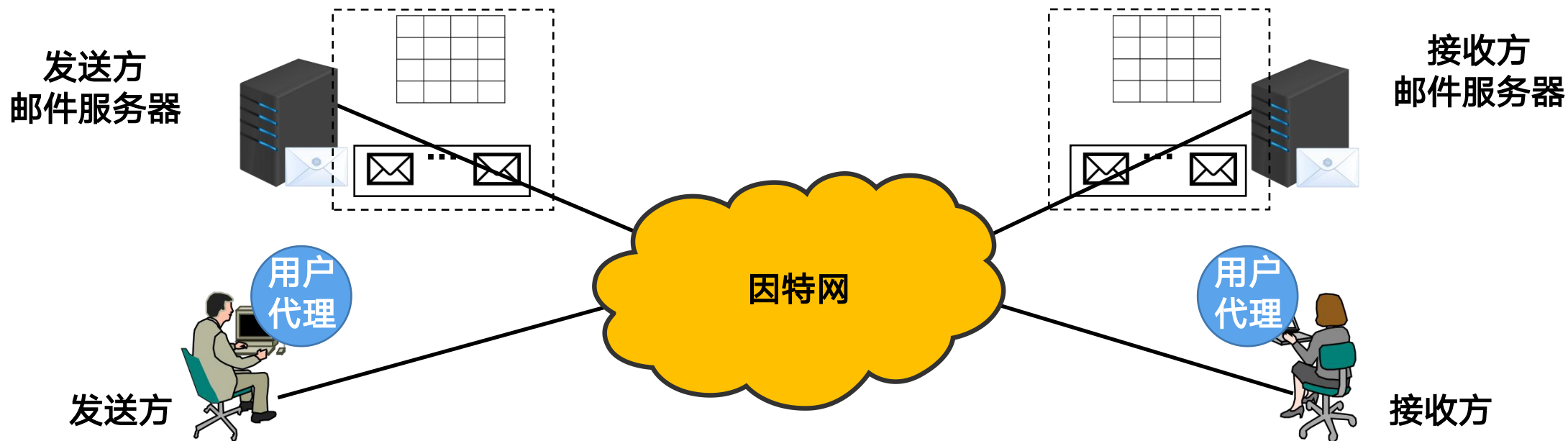
02 电子邮件系统的组成

■ 电子邮件系统采用**客户/服务器方式**。

■ 电子邮件系统的三个主要组成构件：**用户代理**，**邮件服务器**，以及**电子邮件所需的协议**。

□ **用户代理**是用户与电子邮件系统的接口，又称为**电子邮件客户端软件**。

□ **邮件服务器**是电子邮件系统的基础设施。因特网上所有的因特网服务提供者ISP都有邮件服务器，其功能是**发送和接收邮件**，同时还要负责维护用户的邮箱。



02 电子邮件系统的组成

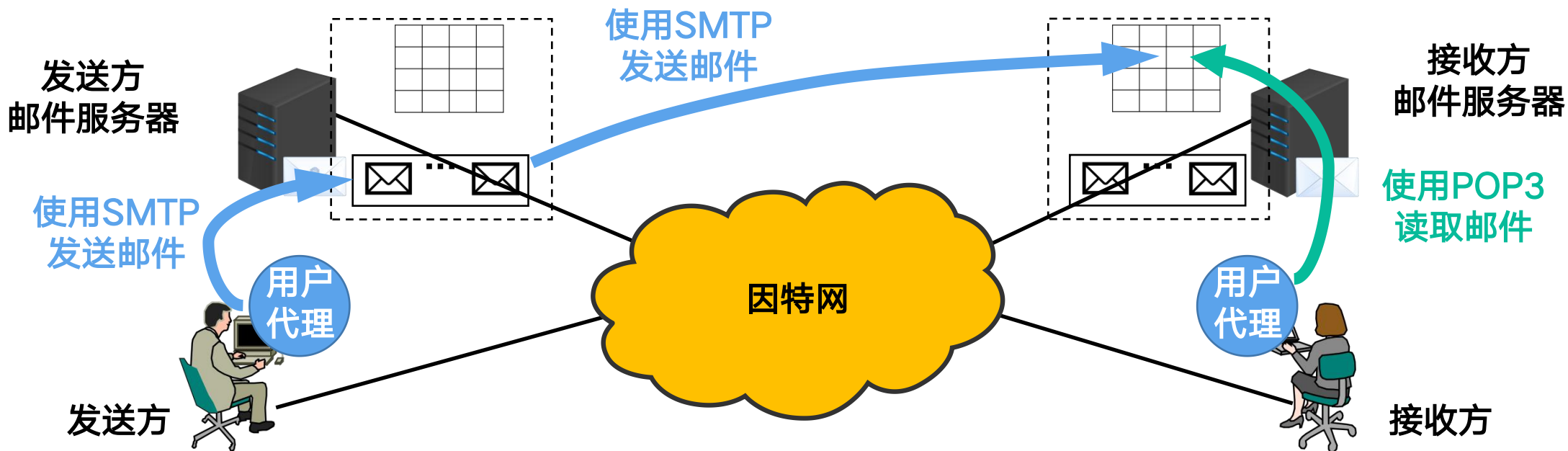
■ 电子邮件系统采用**客户/服务器方式**。

■ 电子邮件系统的三个主要组成构件：**用户代理**，**邮件服务器**，以及**电子邮件所需的协议**。

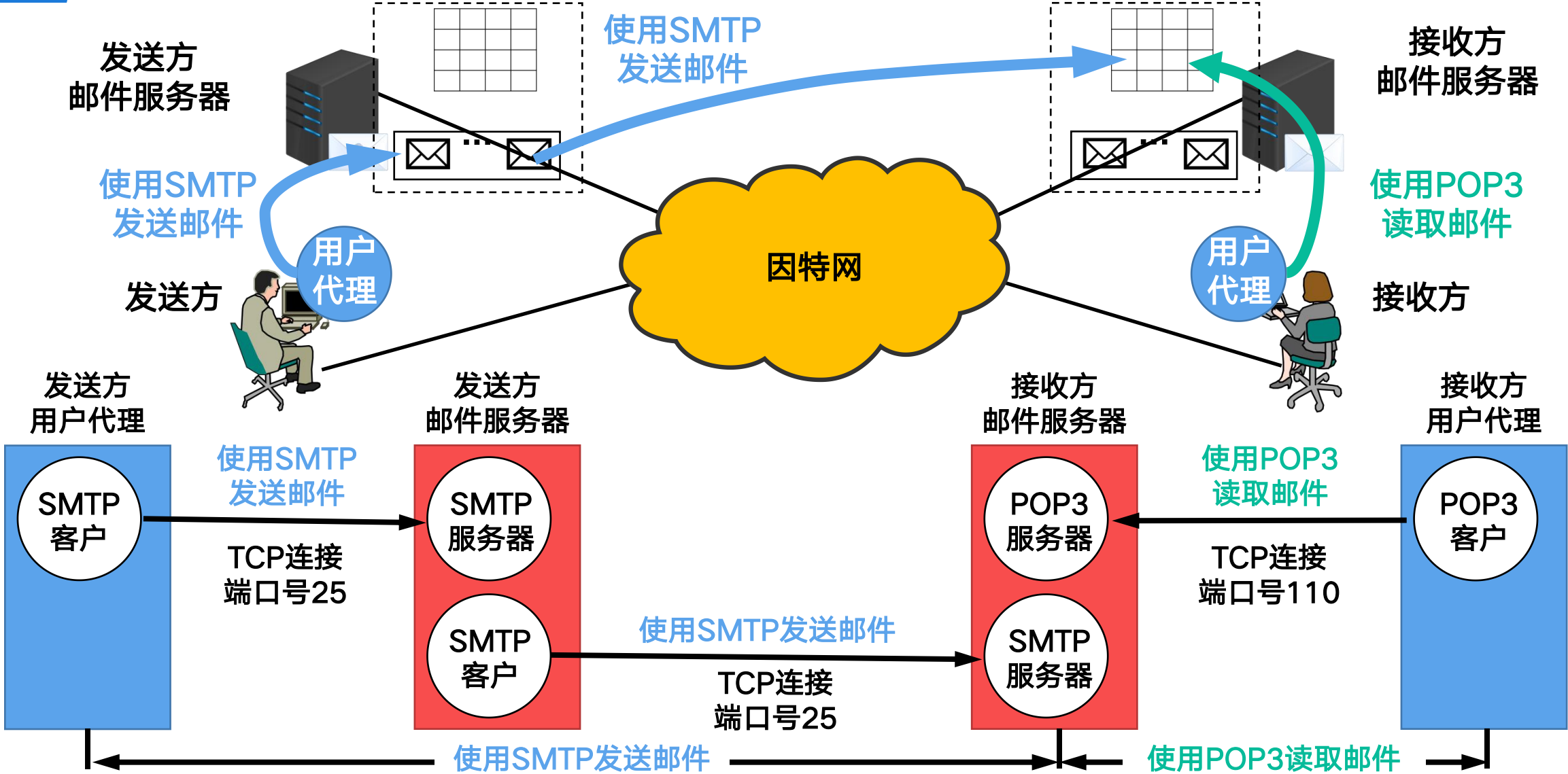
□ **用户代理**是用户与电子邮件系统的接口，又称为**电子邮件客户端软件**。

□ **邮件服务器**是电子邮件系统的基础设施。因特网上所有的因特网服务提供者ISP都有邮件服务器，其功能是**发送和接收邮件**，同时还要负责维护用户的邮箱。

□ 协议包括邮件**发送协议**（例如SMTP）和邮件**读取协议**（例如POP3，IMAP）。



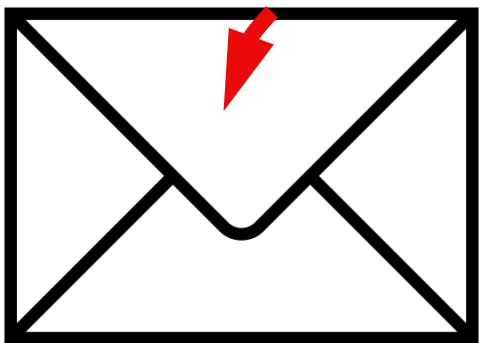
02 电子邮件系统的组成



04 电子邮件的信息格式

电子邮件的信息格式并不是由SMTP定义的，而是在[RFC 822]中单独定义的。这个RFC文档已在2008年更新为[RFC 5322]。一个电子邮件有**信封**和**内容**两部分。而内容又由**首部**和**主体**两部分构成。

邮件系统将自动提取
所需信息并写在信封上



信封

首部

From: zhengfang@mail.hzau.edu.cn
To: 计算机网络课程教学团队所有教师的邮箱
Cc: 信息学院各位领导的邮箱
Subject: 新冠肺炎疫情期间网上教学计划

主体

计算机网络课程教学团队的各位老师：
大家好！
根据学校***

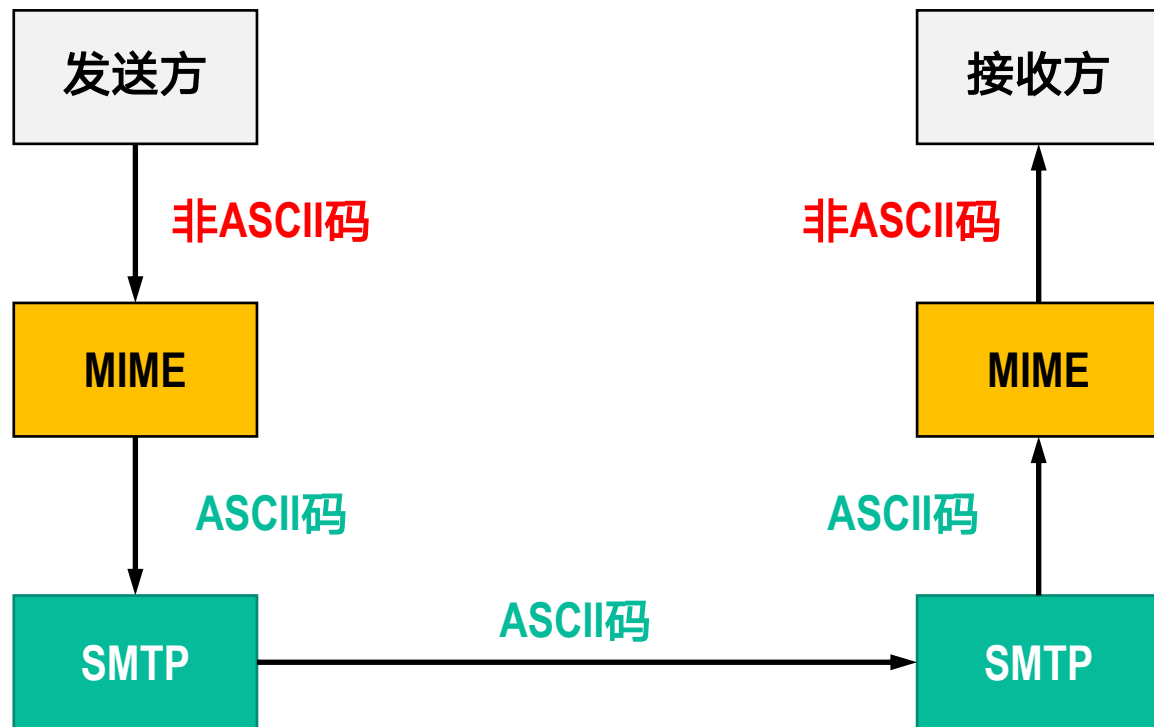
内容

05 多用途因特网邮件扩展

■ SMTP协议只能传送ASCII码文本数据，不能传送可执行文件或其他的二进制对象。

■ SMTP不能满足传送多媒体邮件（例如带有图片、音频或视频数据）的需要。并且许多其他非英语国家的文字（例如中文、俄文、甚至带有重音符号的法文或德文）也无法用SMTP传送。

■ 为解决SMTP传送非ASCII码文本的问题，提出了多用途因特网邮件扩展 (Multipurpose Internet Mail Extensions, MIME)。



05 多用途因特网邮件扩展

■ SMTP协议只能传送ASCII码文本数据，不能传送可执行文件或其他的二进制对象。

■ SMTP不能满足传送多媒体邮件（例如带有图片、音频或视频数据）的需要。并且许多其他非英语国家的文字（例如中文、俄文、甚至带有重音符号的法文或德文）也无法用SMTP传送。

■ 为解决SMTP传送非ASCII码文本的问题，提出了多用途因特网邮件扩展 (Multipurpose Internet Mail Extensions, MIME)。

☐ 增加了5个新的邮件首部字段，这些字段提供了有关邮件主体的信息。

☐ 定义了许多邮件内容的格式，对多媒体电子邮件的表示方法进行了标准化。

☐ 定义了传送编码，可对任何内容格式进行转换，而不会被邮件系统改变。

■ 实际上，MIME不仅仅用于SMTP，也用于后来的同样面向ASCII字符的HTTP。

06

常用的邮件读取协议

邮局协议 (Post Office Protocol, POP)

- POP3是其第三个版本，是因特网正式标准。
- 非常简单、功能有限的邮件读取协议。
- 用户只能以**下载并删除方式**或**下载并保留方式**从邮件服务器下载邮件到用户方计算机。
- **不允许用户在邮件服务器上管理自己的邮件。**
(例如创建文件夹，对邮件进行分类管理等)。

因特网邮件访问协议 (Internet Message Access Protocol, IMAP)

- IMAP4是其第四个版本，是因特网建议标准。
- 功能比POP3强大的邮件读取协议。
- **用户在自己的计算机上就可以操控邮件服务器中的邮箱**，就像在本地操控一样，因此IMAP是一个联机协议。

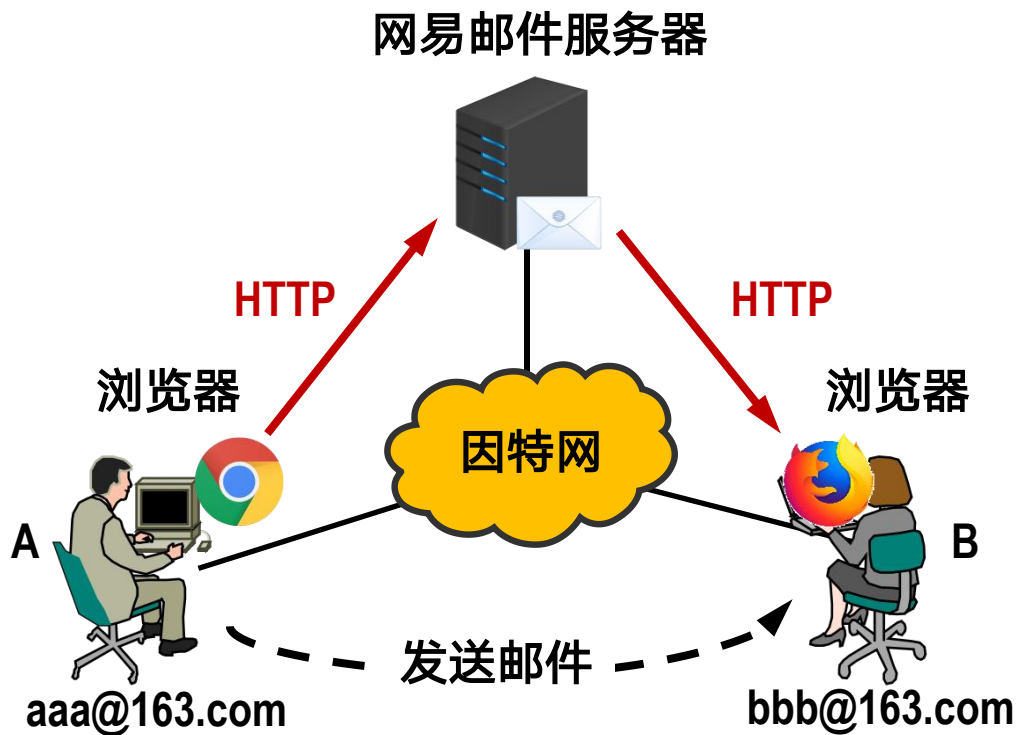
端口号为110

POP3和IMAP4都采用基于TCP连接的客户/服务器方式

端口号为143

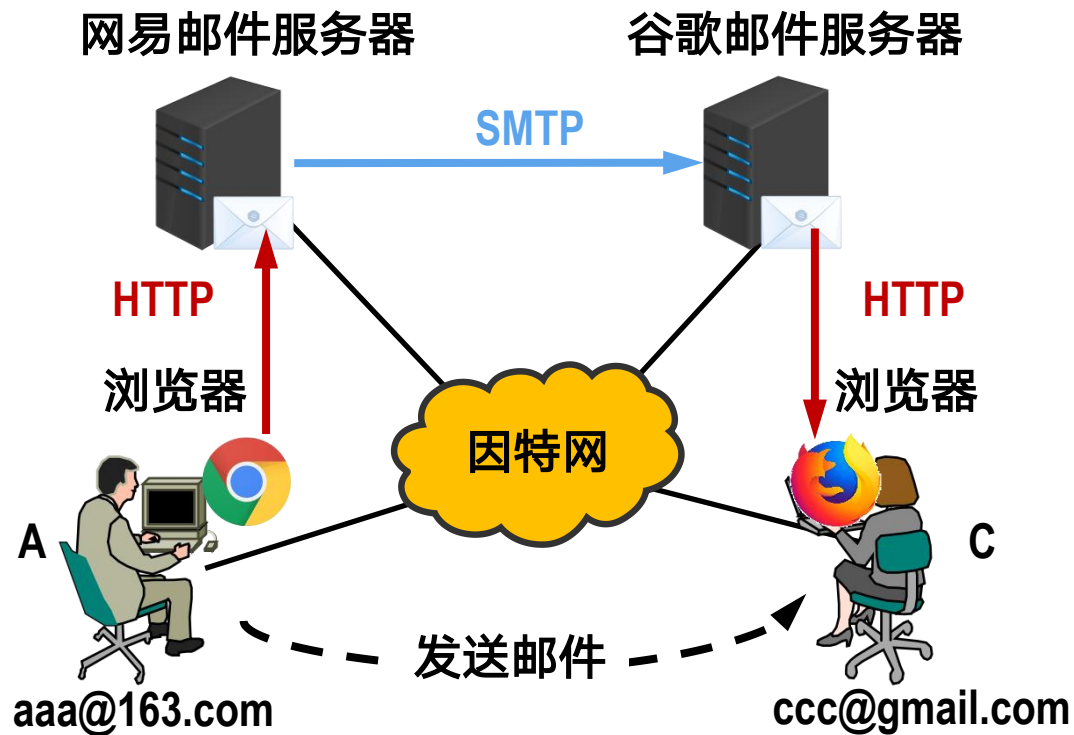
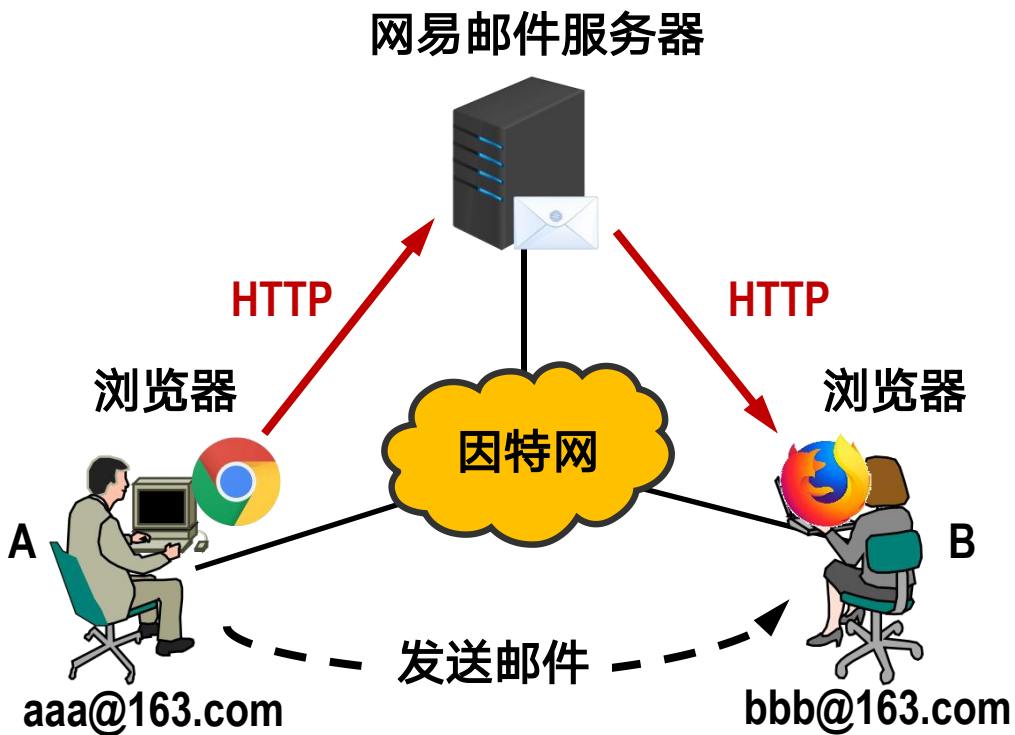
07 基于万维网的电子邮件

- 通过**浏览器**登录（提供用户名和口令）**邮件服务器万维网网站**就可以撰写、收发、阅读和管理电子邮件。这种工作模式**与IMAP很类似**，不同的是用户计算机**无需安装专门的用户代理程序**，只需要使用**通用的万维网浏览器**。
- 邮件服务器网站通常都提供非常强大和方便的邮件管理功能，用户可以在邮件服务器网站上管理和处理自己的邮件，而不需要将邮件下载到本地进行管理。



07 基于万维网的电子邮件

- 通过**浏览器**登录（提供用户名和口令）**邮件服务器万维网网站**就可以撰写、收发、阅读和管理电子邮件。这种工作模式**与IMAP很类似**，不同的是用户计算机**无需安装专门的用户代理程序**，只需要使用**通用的万维网浏览器**。
- 邮件服务器网站通常都提供非常强大和方便的邮件管理功能，用户可以在邮件服务器网站上管理和处理自己的邮件，而不需要将邮件下载到本地进行管理。



电子邮件相关练习题

【2012年 题40】若用户1与用户2之间发送和接收电子邮件的过程如下图所示，则图中①、②、③阶段分别使用的应用层协议可以是 **D**)。

A. SMTP、SMTP、SMTP

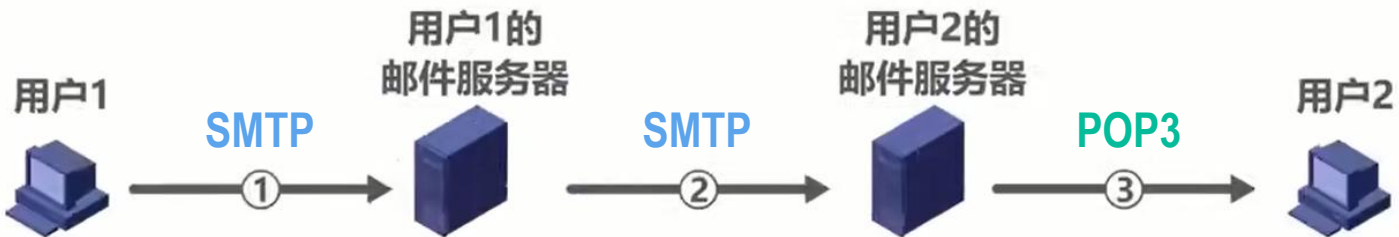
B. POP3、SMTP、POP3

C. POP3、SMTP、SMTP

D. SMTP、SMTP、POP3



解析



电子邮件相关练习题

【2013年 题40】下列关于SMTP协议的叙述中，正确的是（ A ）。

- I. 只支持传输7比特ASCII码内容
- II. 支持在邮件服务器之间发送邮件
- III. 支持从用户代理向邮件服务器发送邮件
- IV. 支持从邮件服务器向用户代理发送邮件

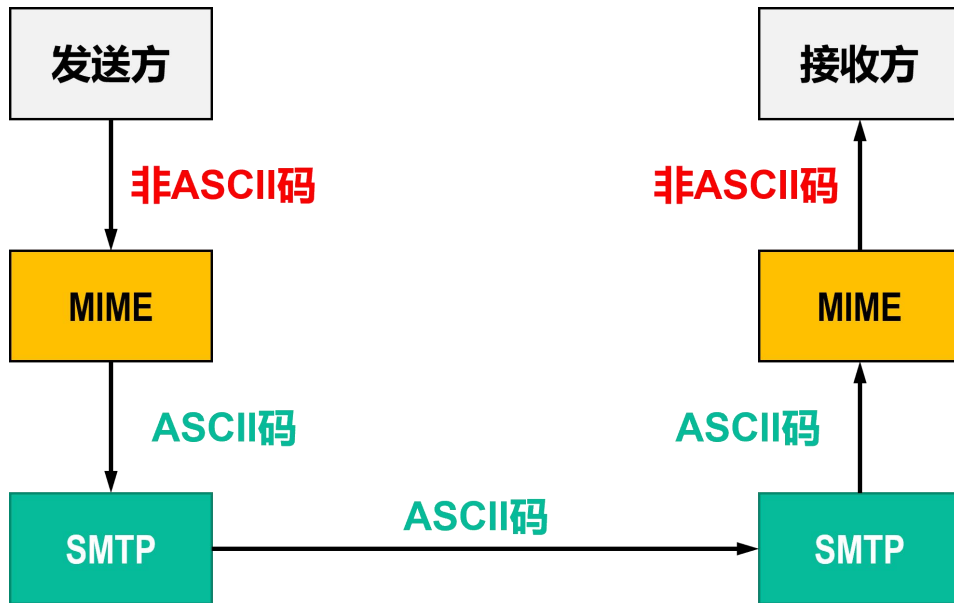
A. 仅I、II和III

B. 仅I、II和IV

C. 仅I、III和IV

D. 仅II、III和IV

解析



电子邮件相关练习题

【2013年 题40】下列关于SMTP协议的叙述中，正确的是（ A ）。

- I. 只支持传输7比特ASCII码内容 ✓
- II. 支持在邮件服务器之间发送邮件 ✓
- III. 支持从用户代理向邮件服务器发送邮件 ✓
- IV. 支持从邮件服务器向用户代理发送邮件 ✗

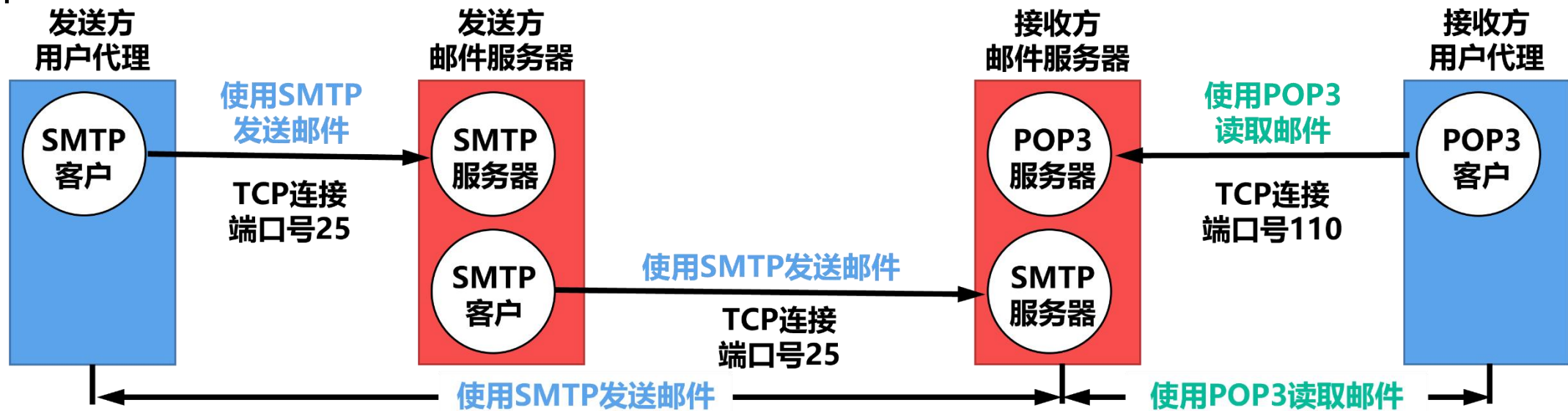
A. 仅I、II和III

B. 仅I、II和IV

C. 仅I、III和IV

D. 仅II、III和IV

解析



电子邮件相关练习题

【2018年 题40】无需转换即可由SMTP协议直接传输的内容是（ D ）。

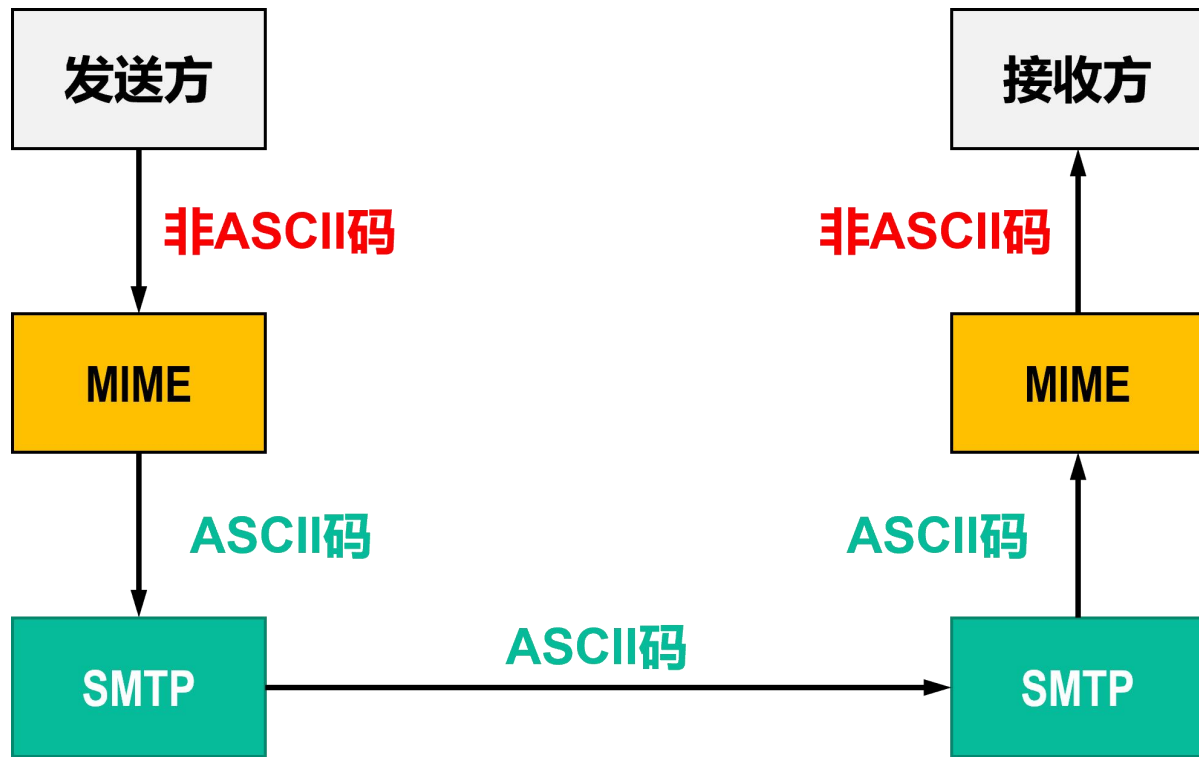
A. JPEG图形

B. MPEG视频

C. EXE文件

D. ASCII文本

解析



6.7 万维网

01

万维网概述

02

统一资源定位符

03

万维网文档

04

超文本传输协议

05

使用Cookie在服务器上记录信息

06

万维网缓存与代理服务器

01 万维网概述

- 万维网（**World Wide Web, WWW**）并非某种特殊的计算机网络。它是一个大规模的、联机式的信息储藏所，**是运行在因特网上的一个分布式应用**。
- 万维网利用网页之间的**超链接**将不同网站的网页链接成一张逻辑上的信息网。
- 万维网是欧洲粒子物理实验室的**Tim Berners-Lee**最初于1989年3月提出的。



01 万维网概述

目前比较流行的浏览器如下：



Chrome
Blink



FireFox
Gecko



Safari
WebKit

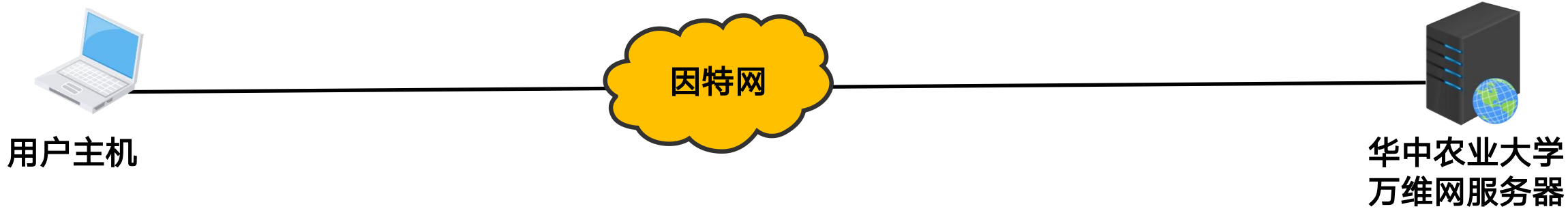
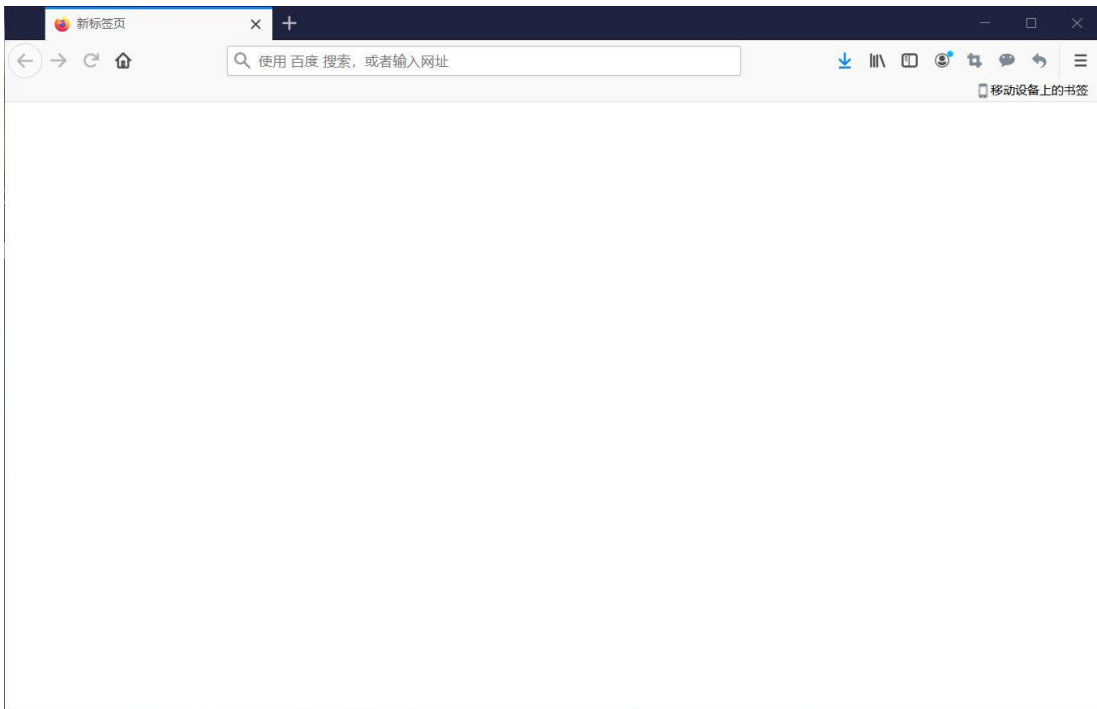


Opera
Blink

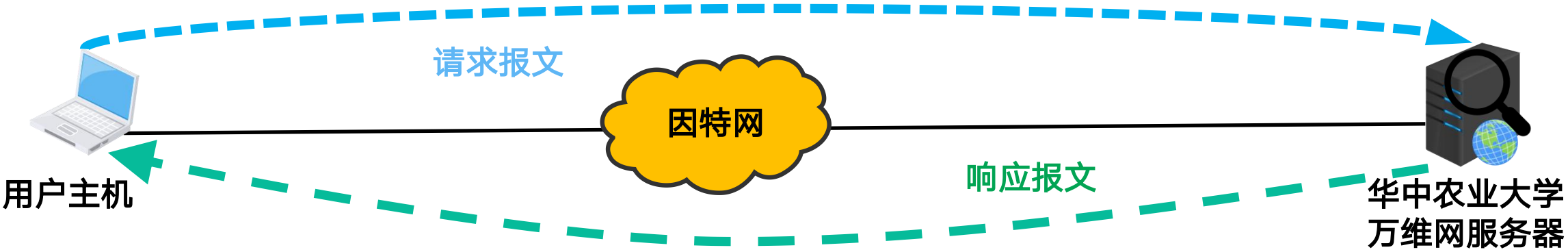
浏览器最重要的部分是**渲染引擎**，也就是**浏览器内核**。负责对网页内容进行解析和显示。

- ☐ 不同的浏览器内核对网页内容的解析也有不同，因此同一网页在不同内核的浏览器里的显示效果可能不同；
- ☐ 网页编写者需要在不同内核的浏览器中测试网页显示效果。

01 万维网概述



01 万维网概述



02 统一资源定位符

为了方便地访问在世界范围的文档，万维网使用**统一资源定位符URL**来指明因特网上任何种类“资源”的位置。

URL的一般形式由以下四个部分组成：

< 协议 > : // < 主机 > : < 端口 > / < 路径 >



02 统一资源定位符

为了方便地访问在世界范围的文档，万维网使用**统一资源定位符URL**来指明因特网上任何种类“资源”的位置。

URL的一般形式由以下四个部分组成：

< 协议 > : // < 主机 > : < 端口 > / < 路径 >



<http://www.hzau.edu.cn:80/index.htm>

02 统一资源定位符

为了方便地访问在世界范围的文档，万维网使用**统一资源定位符URL**来指明因特网上任何种类“资源”的位置。

URL的一般形式由以下四个部分组成：

< 协议 > : // < 主机 > : < 端口 > / < 路径 >



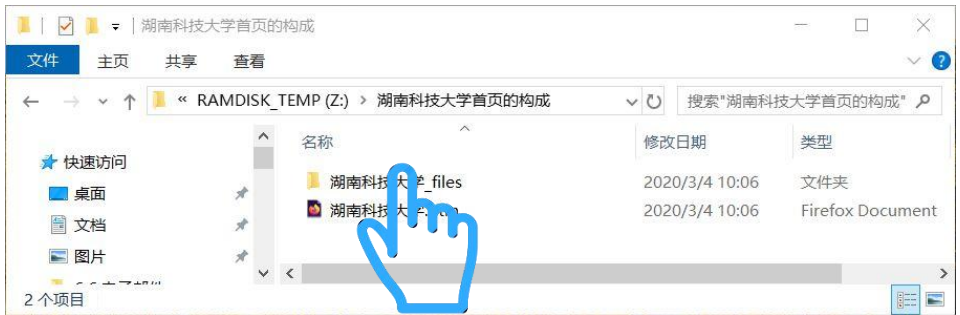
<http://www.hzau.edu.cn:80/jyxx/bksjy.htm>



03 万维网文档



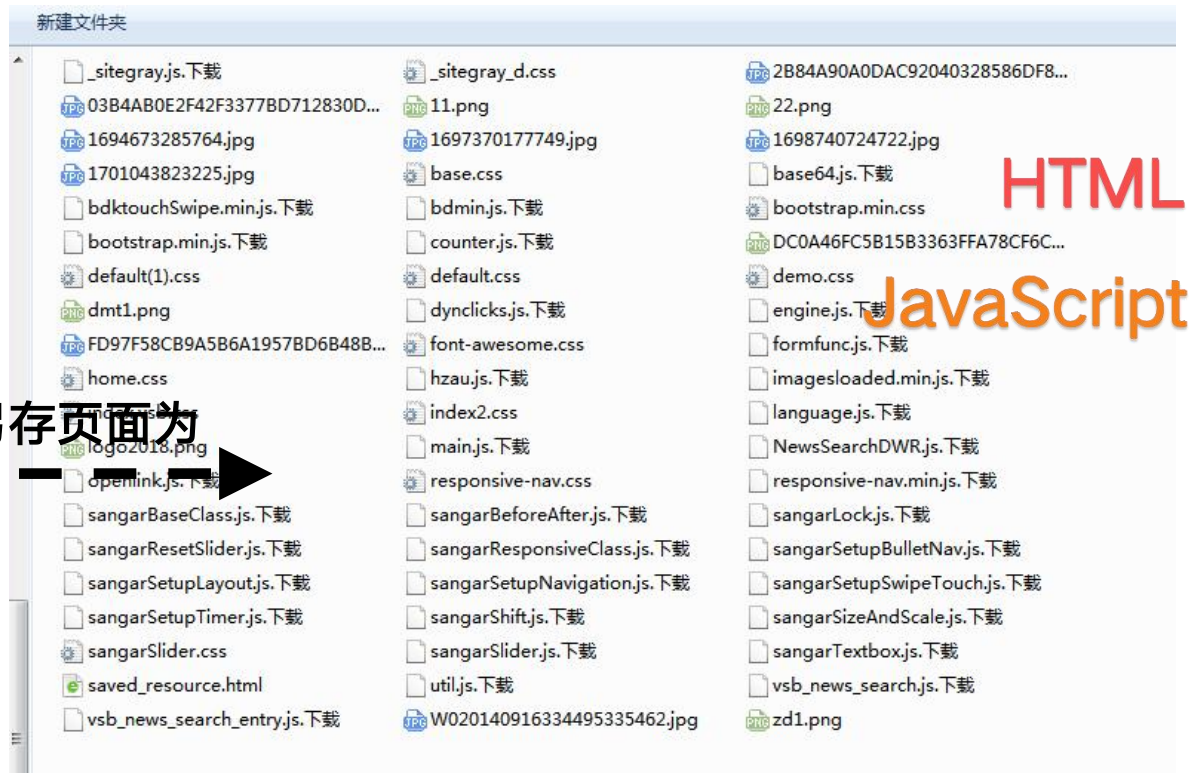
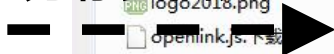
另存页面为



03 万维网文档



另存页面为



03 万维网文档

The logo for HTML, consisting of the letters "HTML" in a white, bold, sans-serif font, centered on a solid red rectangular background.

超文本标记语言 (HyperText Markup Language, **HTML**)

使用多种“标签”来描述网页的结构和内容

The logo for CSS, consisting of the letters "CSS" in a white, bold, sans-serif font, centered on a solid blue rectangular background.

层叠样式表 (Cascading Style Sheets, **CSS**)

从审美的角度来描述网页的样式

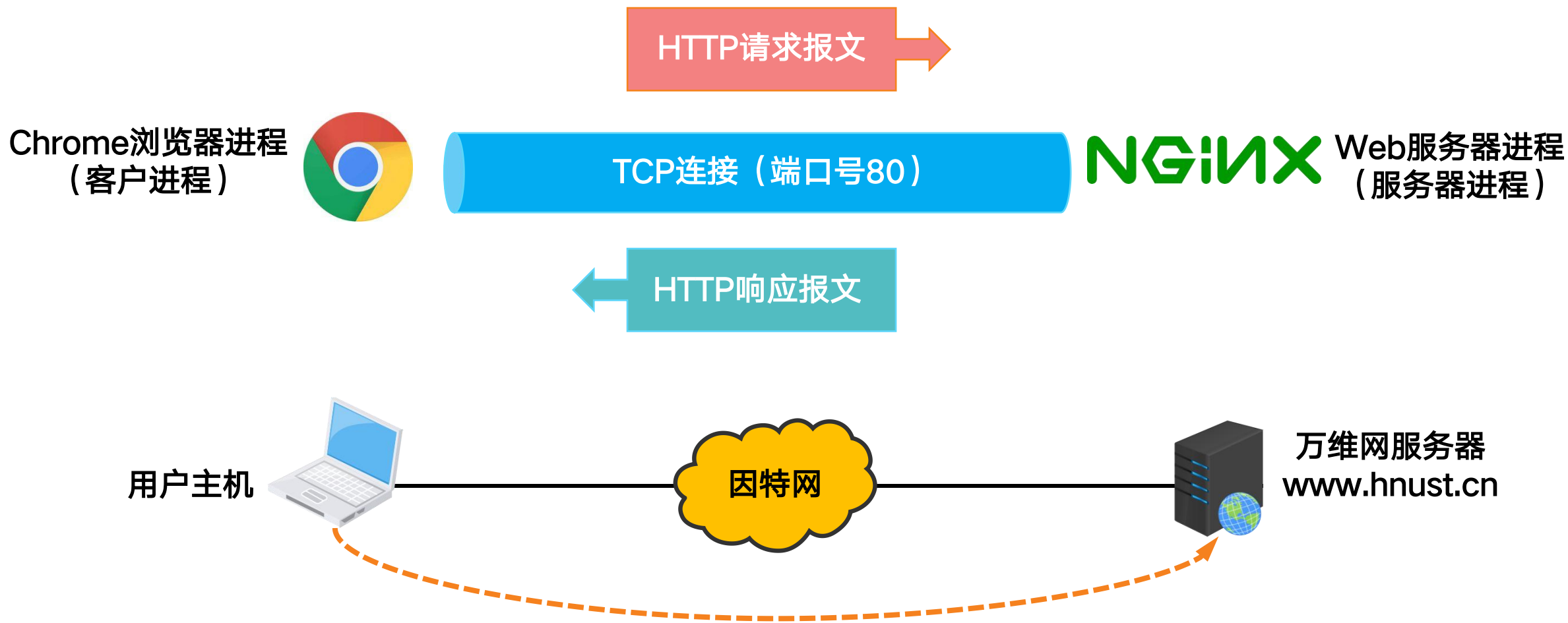
The logo for JavaScript, consisting of the text "JavaScript" in a white, bold, sans-serif font, centered on a solid orange rectangular background.

一种脚本语言（和Java没有任何关系）

控制网页的行为

04 超文本传输协议HTTP(HyperText Transfer Protocol)

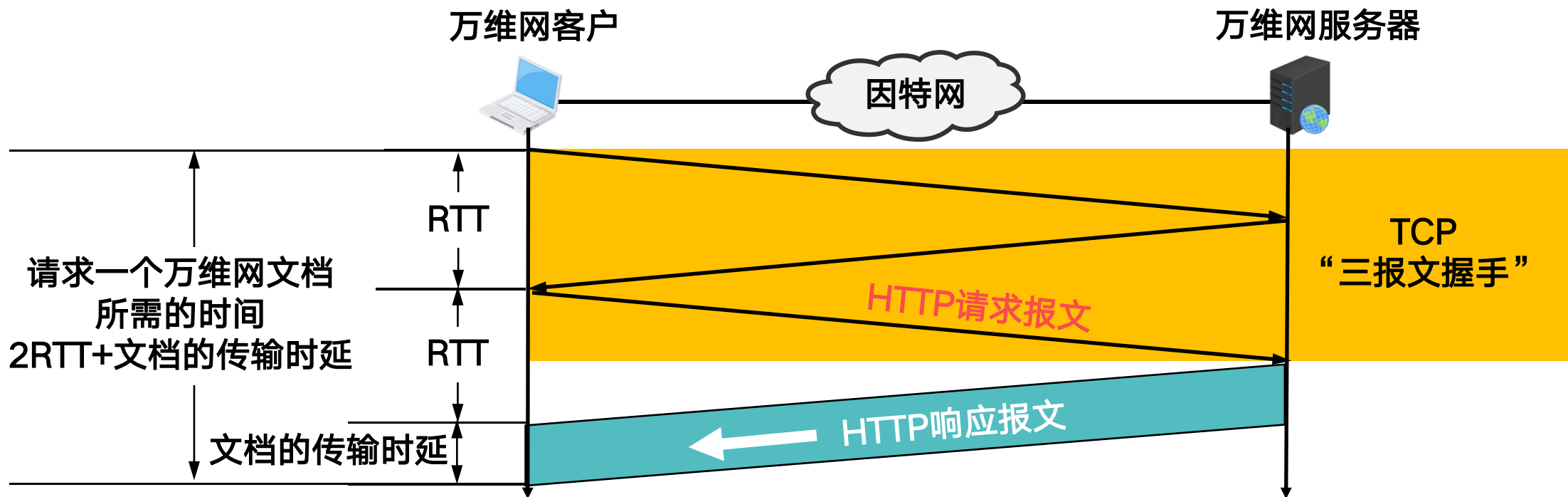
■ HTTP定义了浏览器（即万维网客户进程）怎样向万维网服务器请求万维网文档，以及万维网服务器怎样把万维网文档传送给浏览器。



04 超文本传输协议HTTP(HyperText Transfer Protocol)

■ HTTP/1.0采用**非持续连接**方式。在该方式下，每次浏览器要请求一个文件都要与服务器建立TCP连接，当收到响应后就立即关闭连接。

- **每请求一个文档就要有两倍的RTT的开销**。若一个网页上有很多引用对象（例如图片等），那么请求每一个对象都需要花费2RTT的时间。
- 为了减小时延，**浏览器通常会建立多个并行的TCP连接同时请求多个对象**。但是，这会大量占用万维网服务器的资源，特别是万维网服务器往往要同时服务于大量客户的请求，这会使其负担很重。



04 超文本传输协议HTTP(HyperText Transfer Protocol)

■ HTTP/1.0采用**非持续连接**方式。在该方式下，每次浏览器要请求一个文件都要与服务器建立TCP连接，当收到响应后就立即关闭连接。

☐ **每请求一个文档就要有两倍的RTT的开销**。若一个网页上有很多引用对象（例如图片等），那么请求每一个对象都需要花费2RTT的时间。

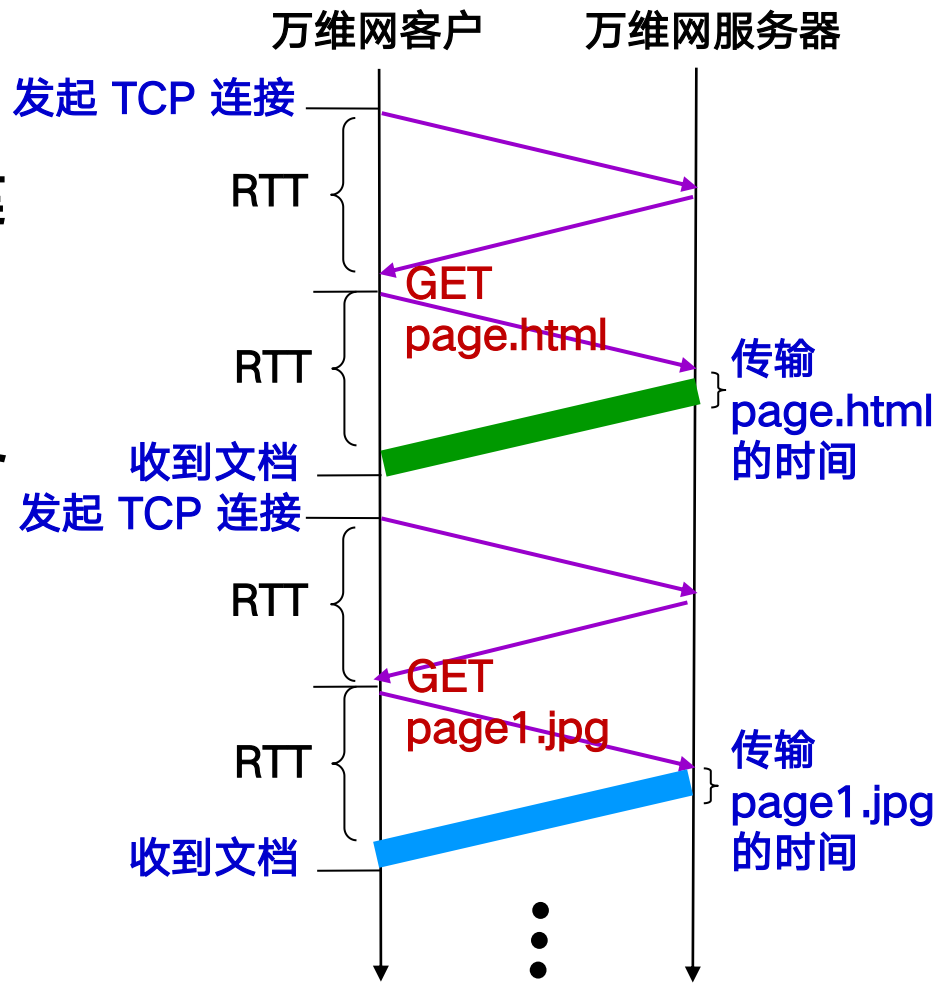
☐ 为了减小时延，**浏览器通常会建立多个并行的TCP连接同时请求多个对象**。但是，这会大量占用万维网服务器的资源，特别是万维网服务器往往要同时服务于大量客户的请求，这会使其负担很重。

■ HTTP/1.1采用**持续连接**方式。在该方式下，万维网服务器在发送响应后仍然保持这条连接，使同一个客户（浏览器）和该服务器可以继续在这条连接上传送后续的HTTP请求报文和响应报文。这并不局限于传送同一个页面上引用的对象，而是只要这些文档都在同一个服务器上就行。

☐ 为了进一步提高效率，HTTP/1.1的持续连接还可以使用**流水线**方式工作，即浏览器在收到HTTP的响应报文之前就能够连续发送多个请求报文。这样的一个个的请求报文到达服务器后，服务器就发回一个接一个的响应报文。这样就节省了很多个RTT时间，使TCP连接中的空闲时间减少，提高了下载文档的效率。

协议 HTTP/1.0 的主要缺点

- 每请求一个文档就要有
两倍 RTT 的开销。
- 客户和服务器的每一次建立新的 TCP 连接都要
分配缓存和变量。
- 这种**非持续连接**使服务器的负担很重。



假设一个主页
page.html 上有 10
个链接的图片:
page1.jpg,
page2.jpg,
page3.jpg,
...

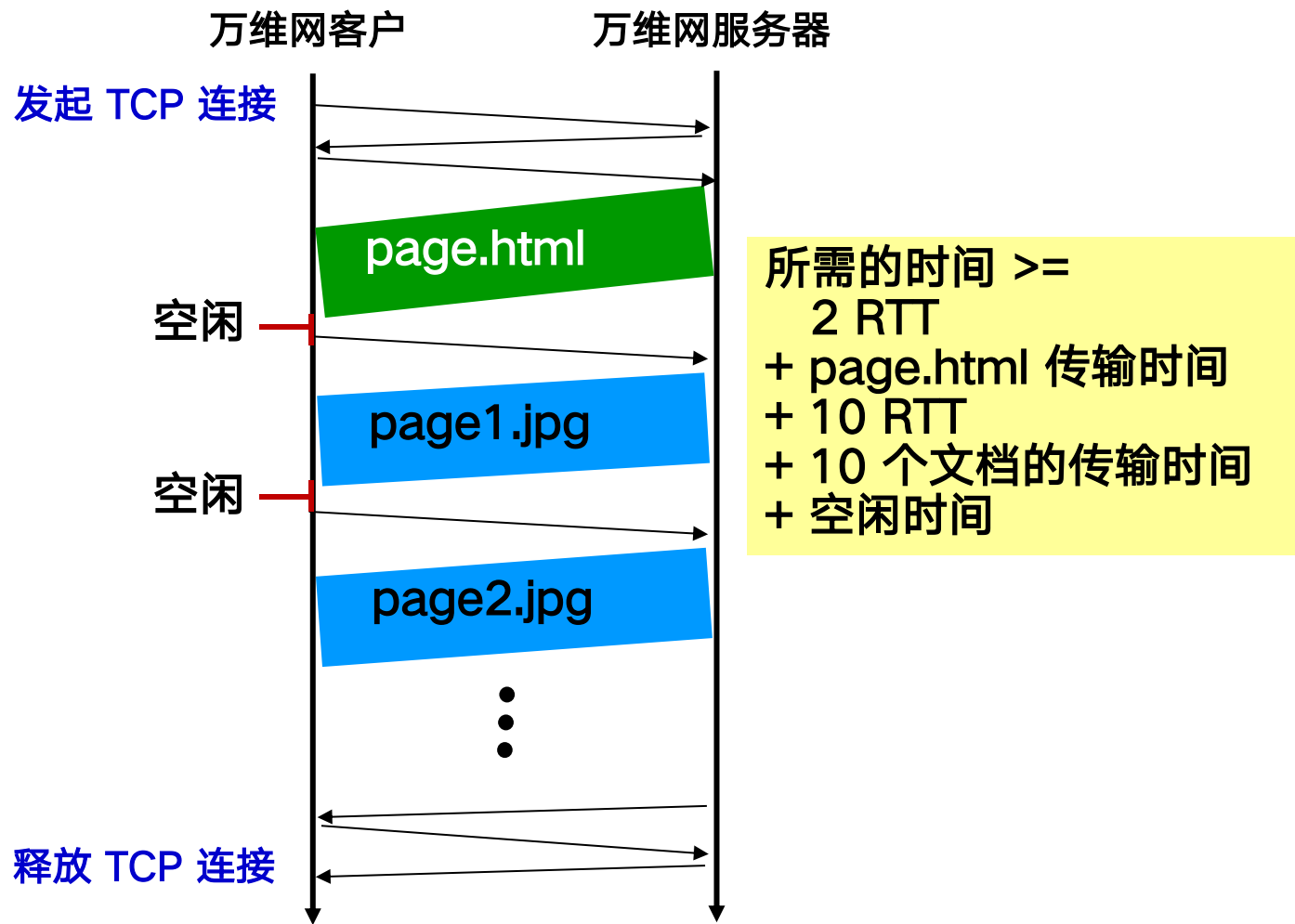
所需的时间 \geq
22 RTT
+ 11 个文档的传输时间

协议 HTTP/1.1 使用持续连接

- **持续连接**（persistent connection）：服务器在发送响应后仍然在一段时间内**保持**这条连接（不释放），使同一个客户（浏览器）和该服务器可以继续在这条连接上传送后续的 HTTP 请求报文和响应报文。
- 只要文档都在同一个服务器上，就可以继续使用该 TCP 连接。
- **两种工作方式：**
 - ◆ 非流水线方式 (without pipelining)
 - ◆ 流水线方式 (with pipelining)。

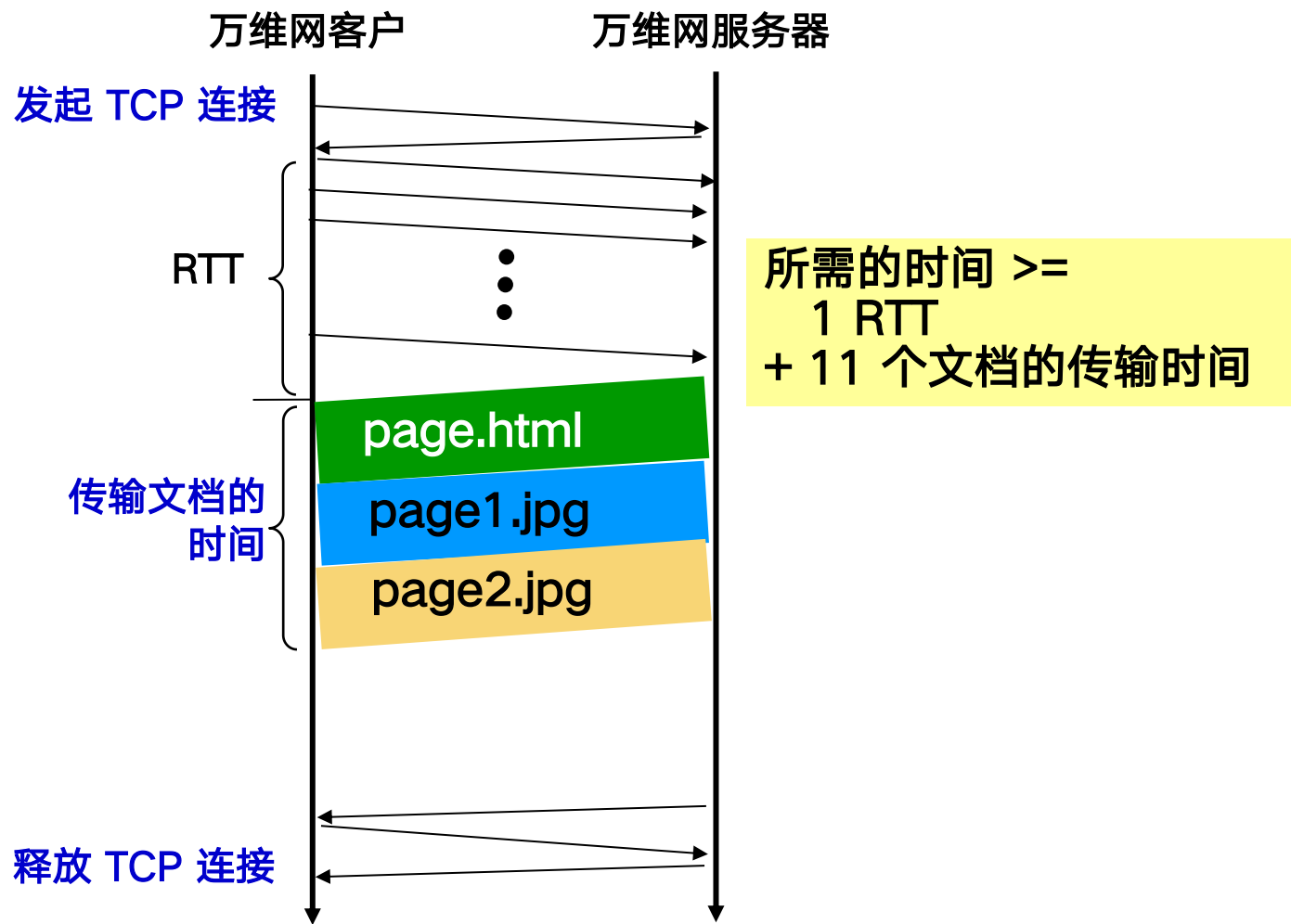
持续连接：非流水线方式

- 客户在收到前一个响应之后才能发出下一个请求。
- 缺点：TCP 连接空闲状态。



持续连接：流水线方式

- 客户在收到响应报文**之前**就能够接着发送新的请求报文。
- 连续的多个请求报文到达服务器后，服务器就可连续发回响应报文。
- 下载效率提高。



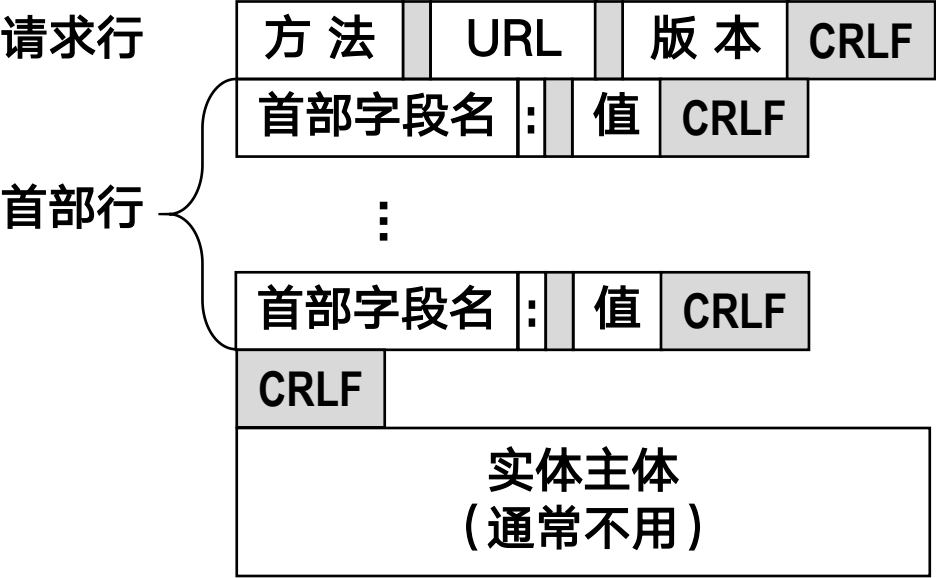
04 超文本传输协议HTTP(HyperText Transfer Protocol)

HTTP的报文格式

HTTP是面向文本的，其报文中的每一个字段都是一些ASCII码串，并且每个字段的长度都是不确定的。

- 空格
- CRLF 回车换行

HTTP请求报文格式



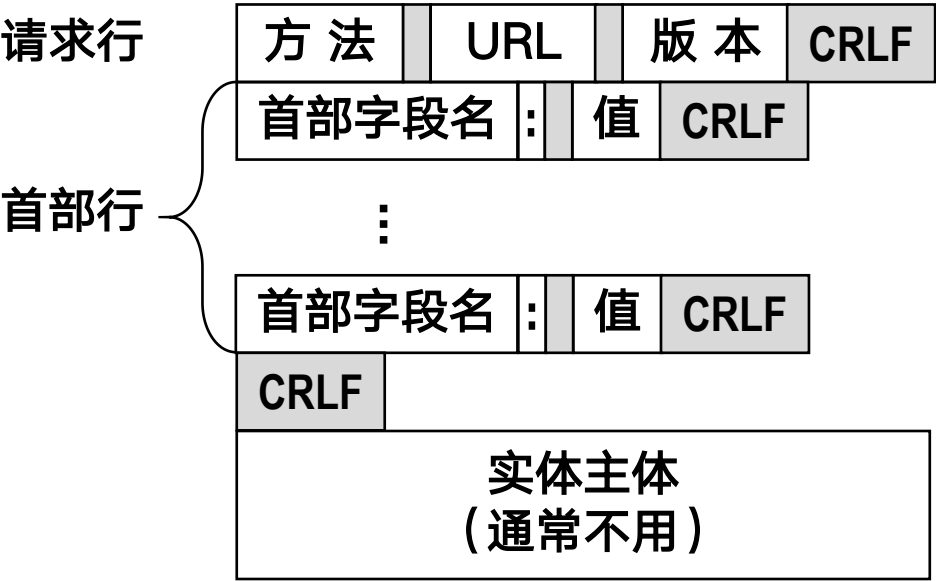
04 超文本传输协议HTTP(HyperText Transfer Protocol)

HTTP的报文格式

HTTP是面向文本的，其报文中的每一个字段都是一些ASCII码串，并且每个字段的长度都是不确定的。

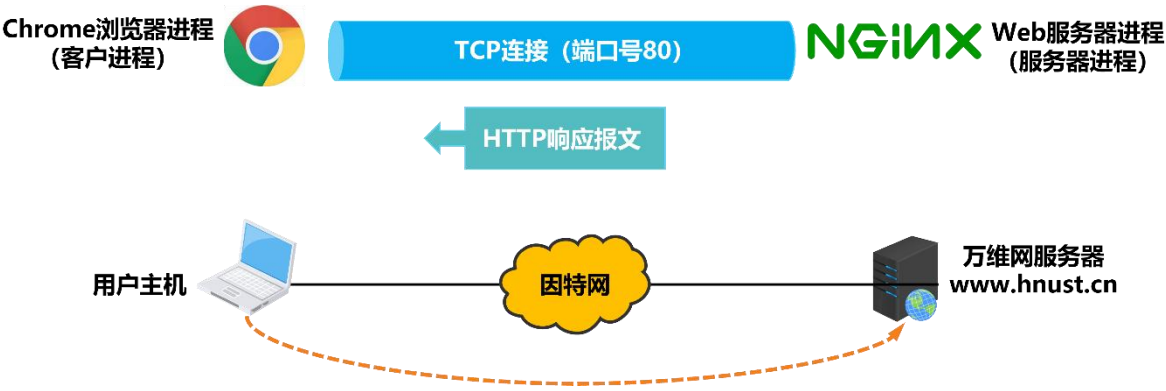
- 空格
- CRLF 回车换行

HTTP请求报文格式



```
GET /index.htm HTTP/1.1
Host: www.hzau.edu.cn
Connection: close
User-Agent: Mozilla/5.0
Accept-Language: cn
```

指明方法“GET”，URL，HTTP版本
指明服务器的域名
告诉服务器发送完请求的文档后就可以释放连接
告诉服务器浏览器的类型及版本
告诉服务器用户希望优先得到中文版本的文档
报文的最后需要由一个空行



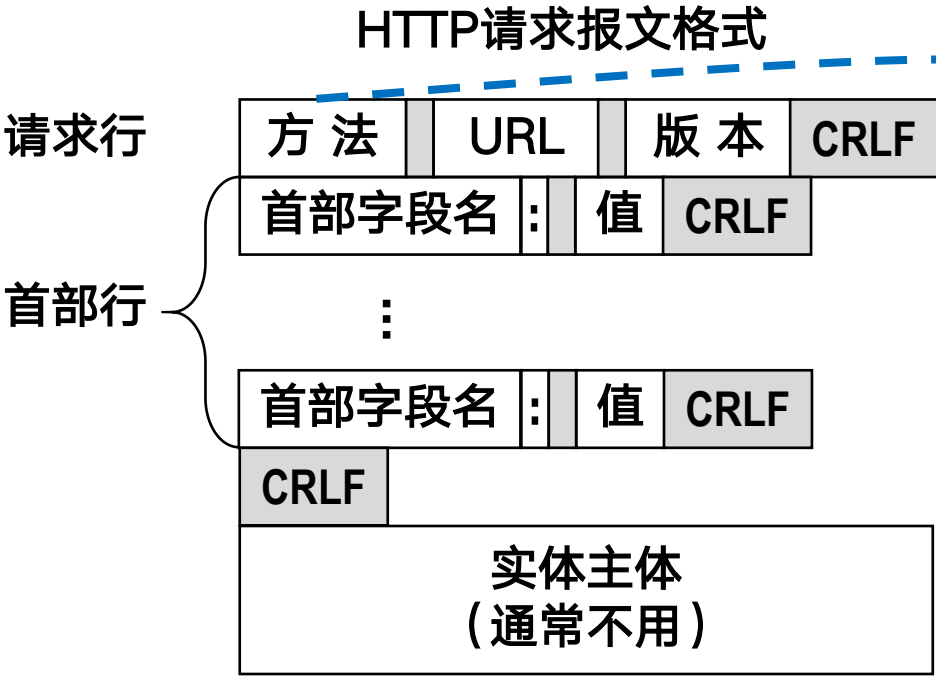
04 超文本传输协议HTTP(HyperText Transfer Protocol)

HTTP的报文格式

HTTP是面向文本的，其报文中的每一个字段都是一些ASCII码串，并且每个字段的长度都是不确定的。

空格

CRLFCRLF回车换行

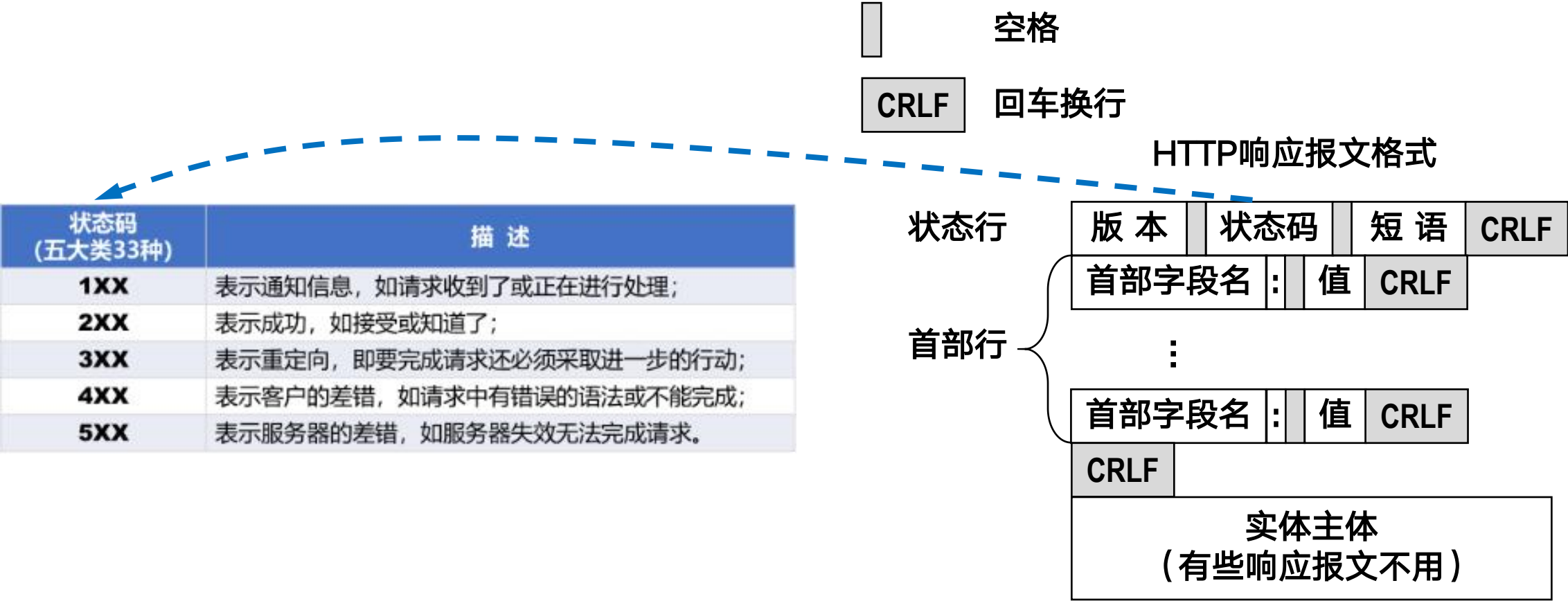


方 法	描 述
GET	请求URL标志的文档
HEAD	请求URL标志的文档的首部
POST	向服务器发送数据
PUT	在指明的URL下存储一个文档
DELETE	删除URL标志的文档
CONNECT	用于代理服务器
OPTIONS	请求一些选项信息
TRACE	用来进行环回测试
PATCH	对PUT方法的补充，用来对已知资源进行局部更新

04 超文本传输协议HTTP(HyperText Transfer Protocol)

HTTP的报文格式

HTTP是面向文本的，其报文中的每一个字段都是一些ASCII码串，并且每个字段的长度都是不确定的。



04 超文本传输协议HTTP(HyperText Transfer Protocol)

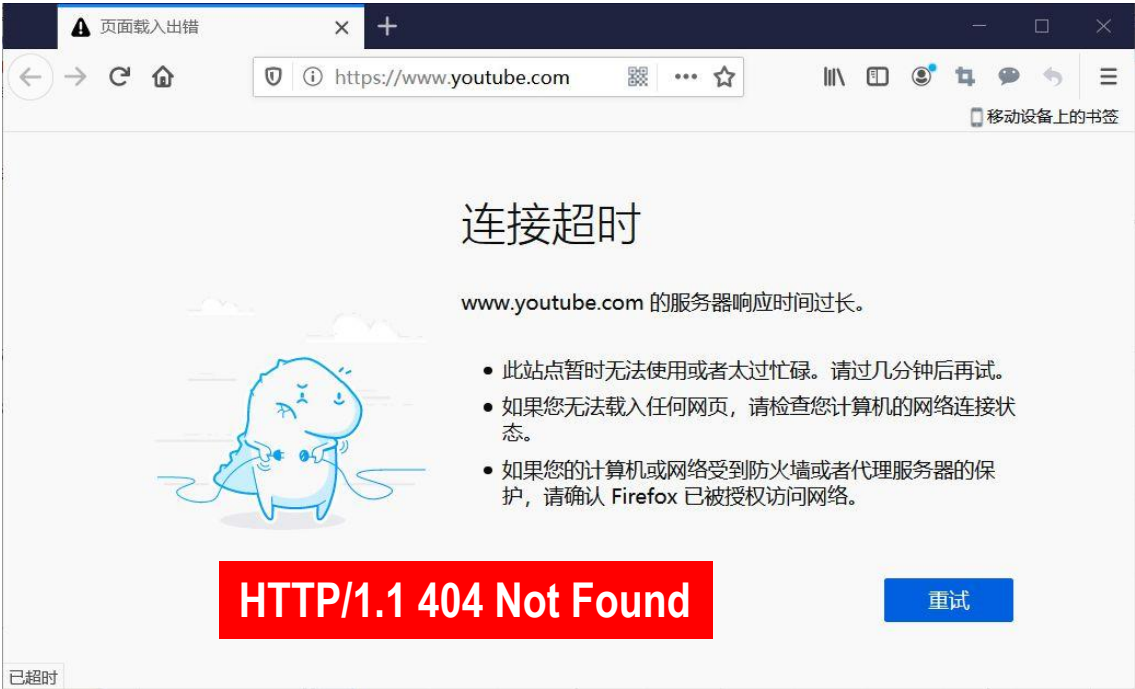
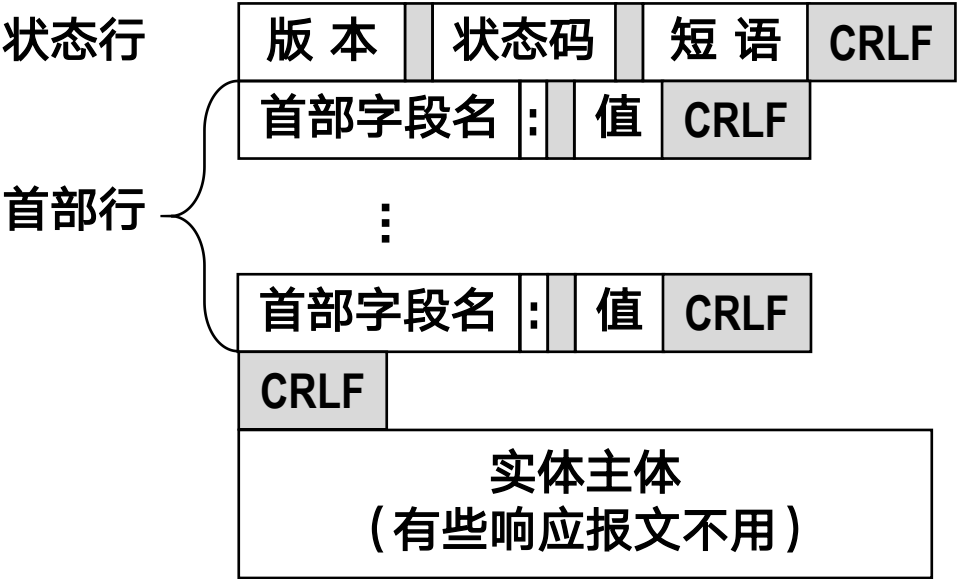
HTTP的报文格式

HTTP是面向文本的，其报文中的每一个字段都是一些ASCII码串，并且每个字段的长度都是不确定的。

空格

CRLF 回车换行

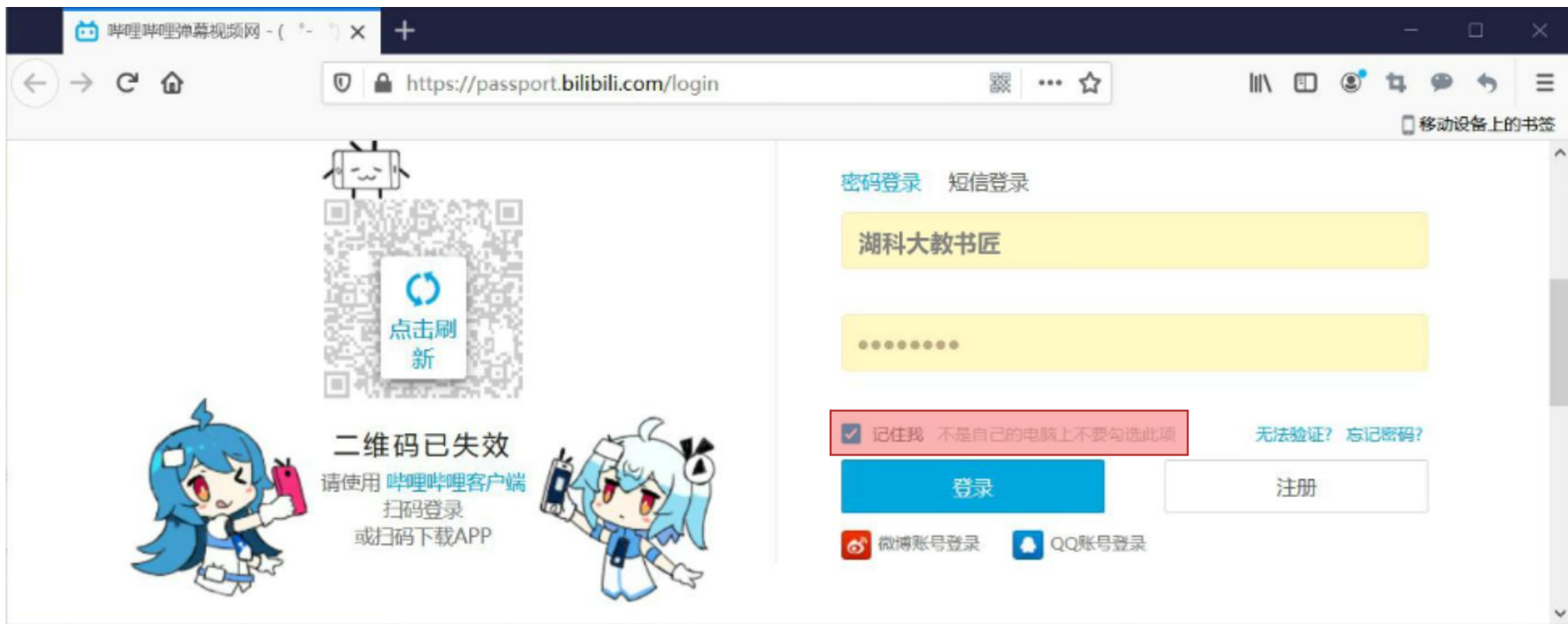
HTTP响应报文格式



05 使用Cookie在服务器上记录用户信息

早期的万维网应用非常简单，仅仅是用户查看存放在不同服务器上的各种静态的文档。因此HTTP被设计为一种**无状态**的协议。这样可以简化服务器的设计。

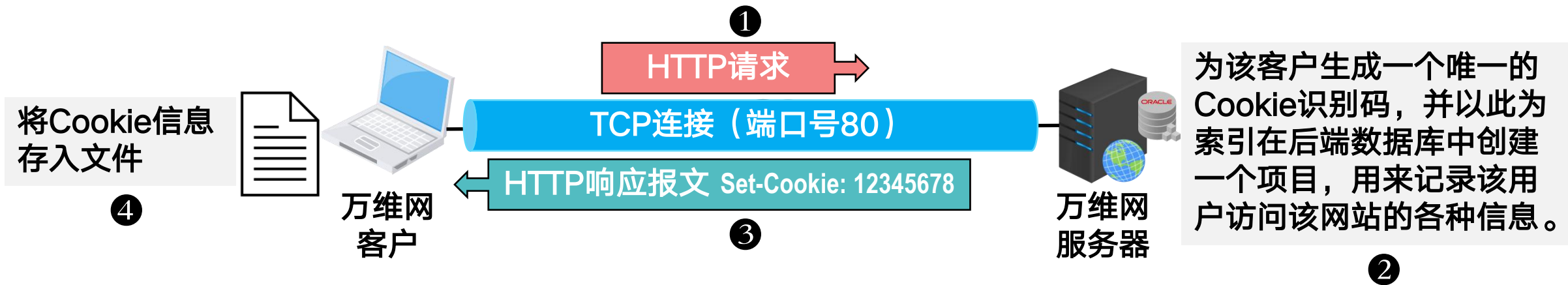
现在，用户可以通过万维网进行各种复杂的应用，如网上购物、电子商务等。这些应用往往需要万维网服务器能够识别用户。



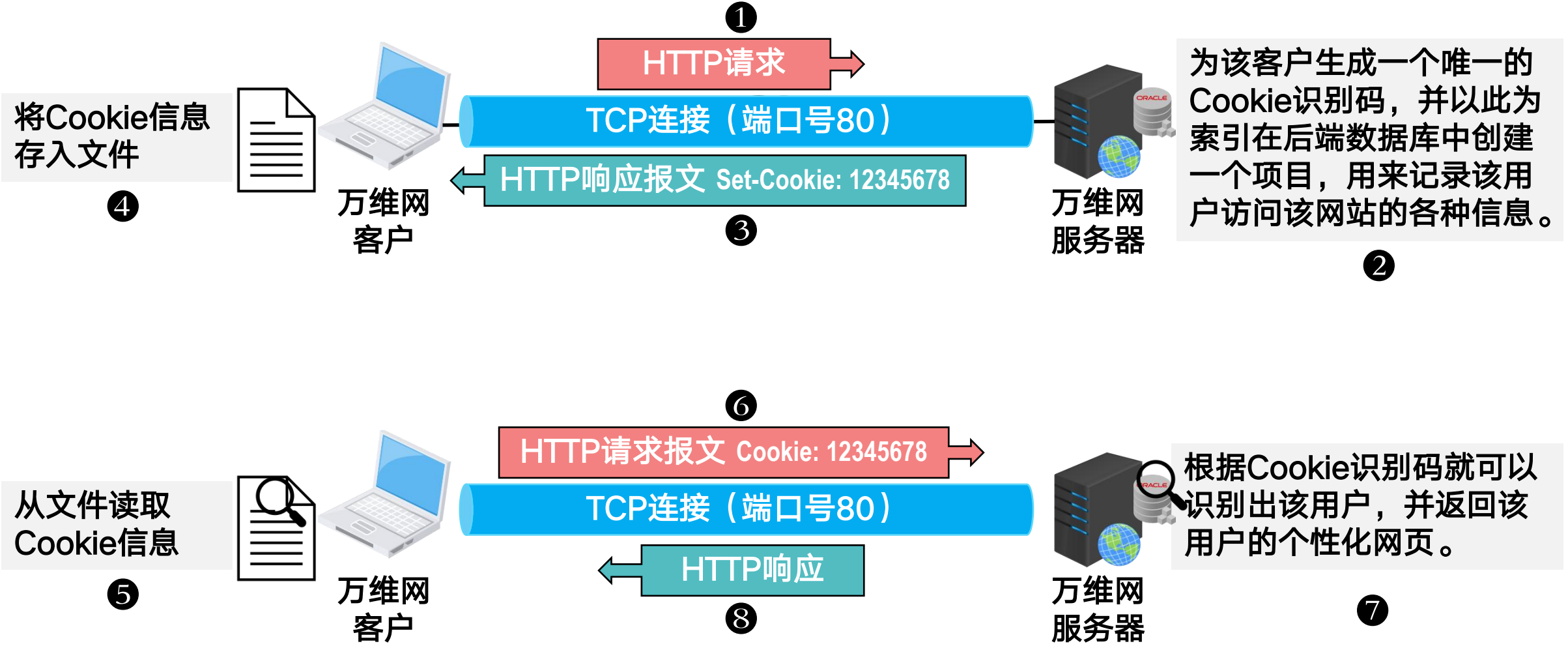
05 使用Cookie在服务器上记录用户信息

- 早期的万维网应用非常简单，仅仅是用户查看存放在不同服务器上的各种静态的文档。因此HTTP被设计为一种**无状态**的协议。这样可以简化服务器的设计。
- 现在，用户可以通过万维网进行各种复杂的应用，如网上购物、电子商务等。这些应用往往需要万维网服务器能够识别用户。
- Cookie提供了一种机制使得万维网服务器能够“记住”用户，而无需用户主动提供用户标识信息。也就是说，**Cookie是一种对无状态的HTTP进行状态化的技术**。

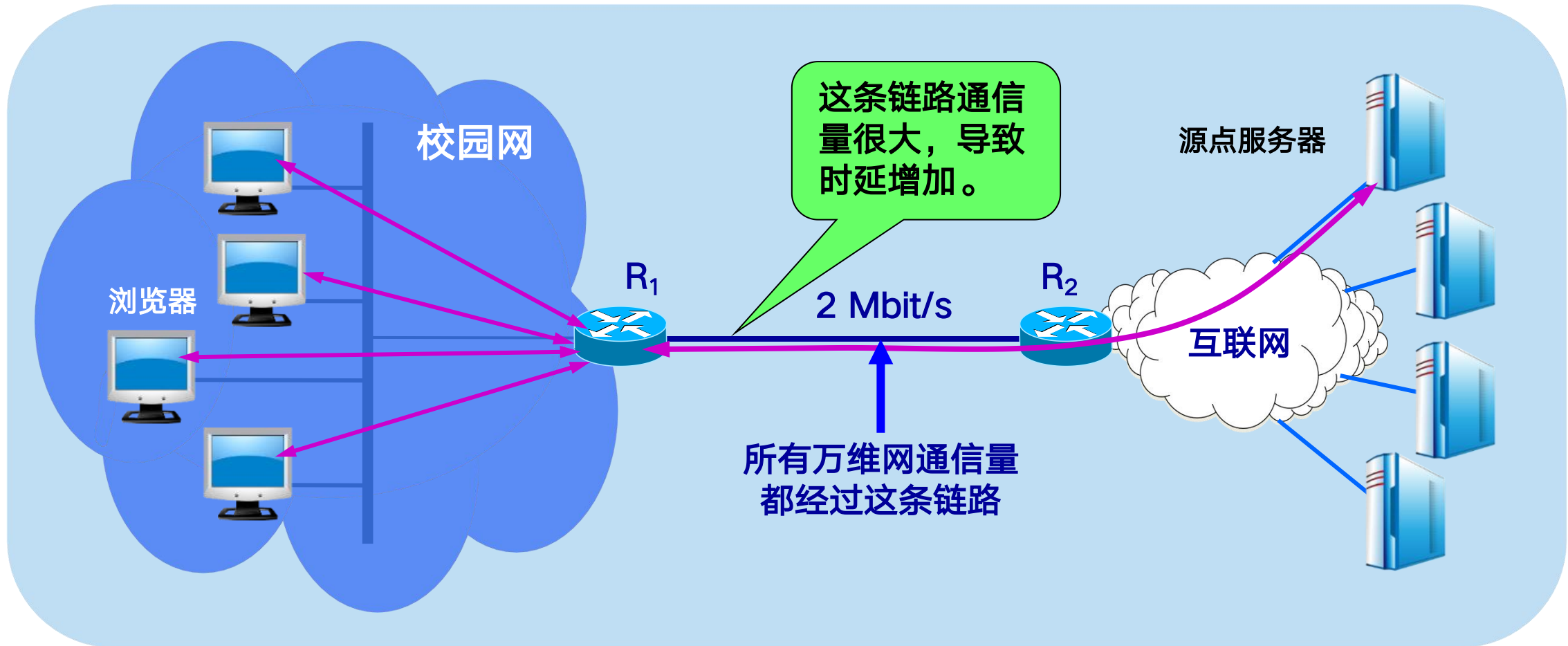
05 使用Cookie在服务器上记录用户信息



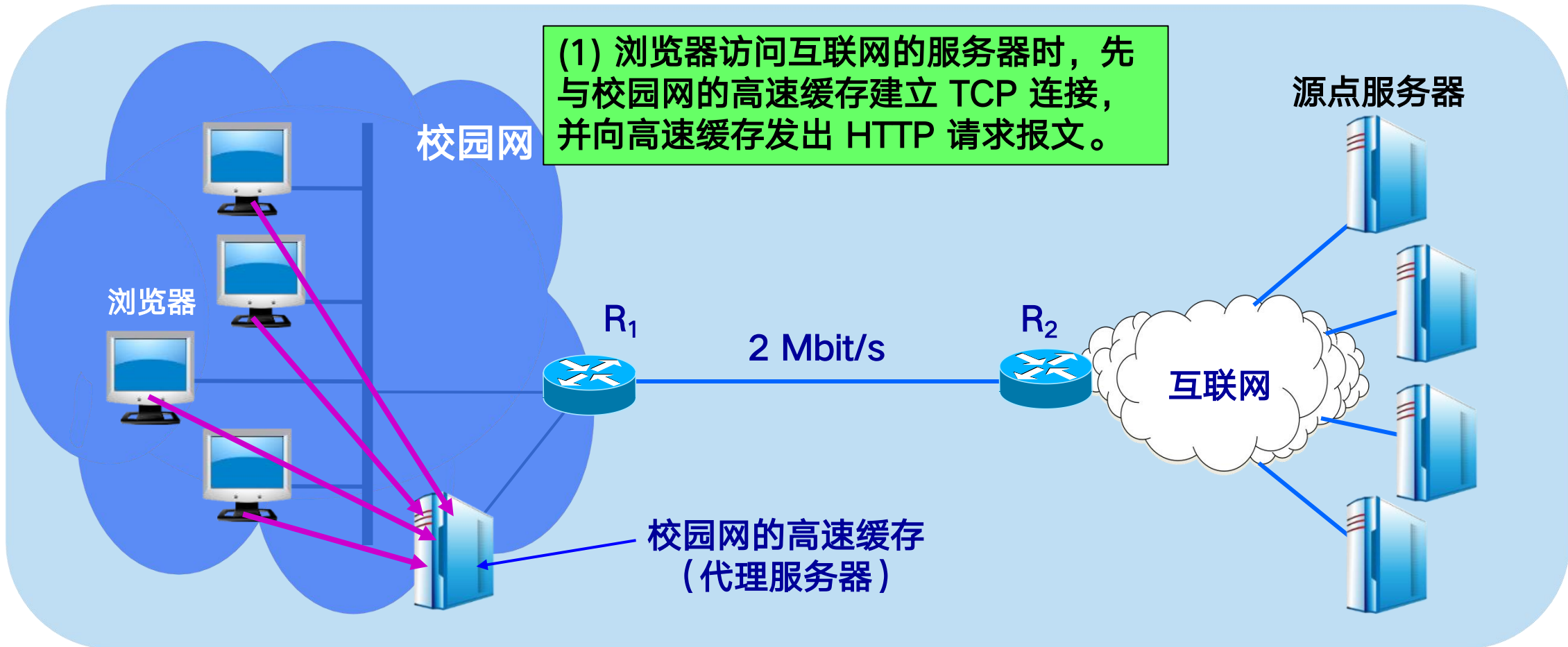
05 使用Cookie在服务器上记录用户信息



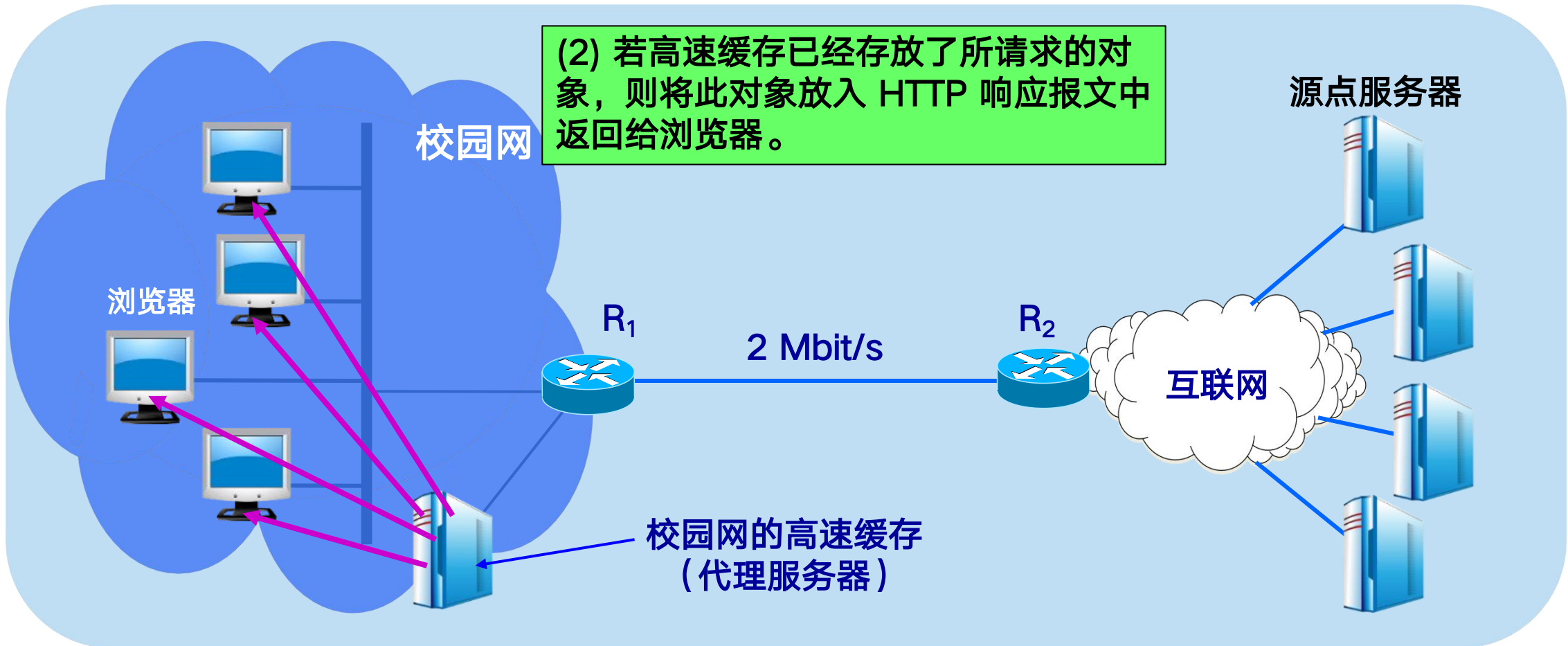
不使用高速缓存的情况



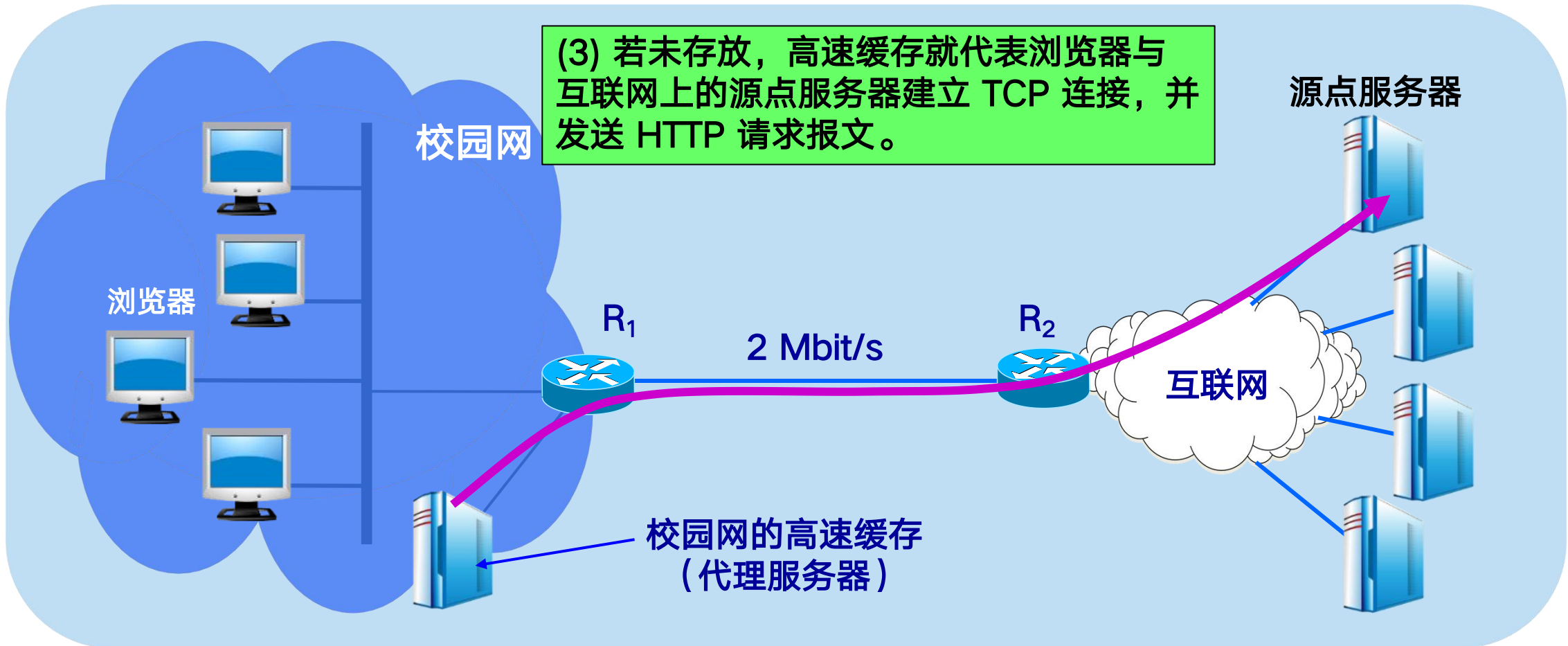
使用高速缓存的情况



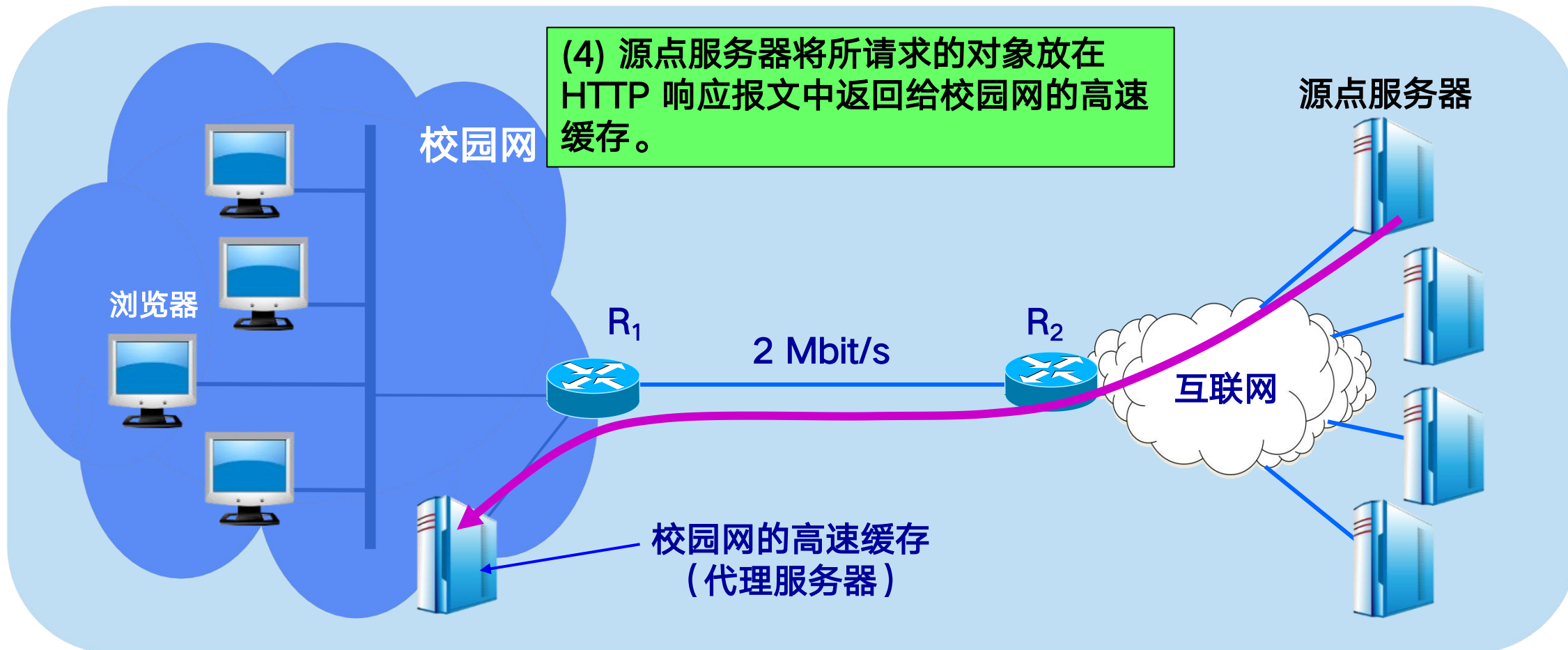
使用高速缓存的情况



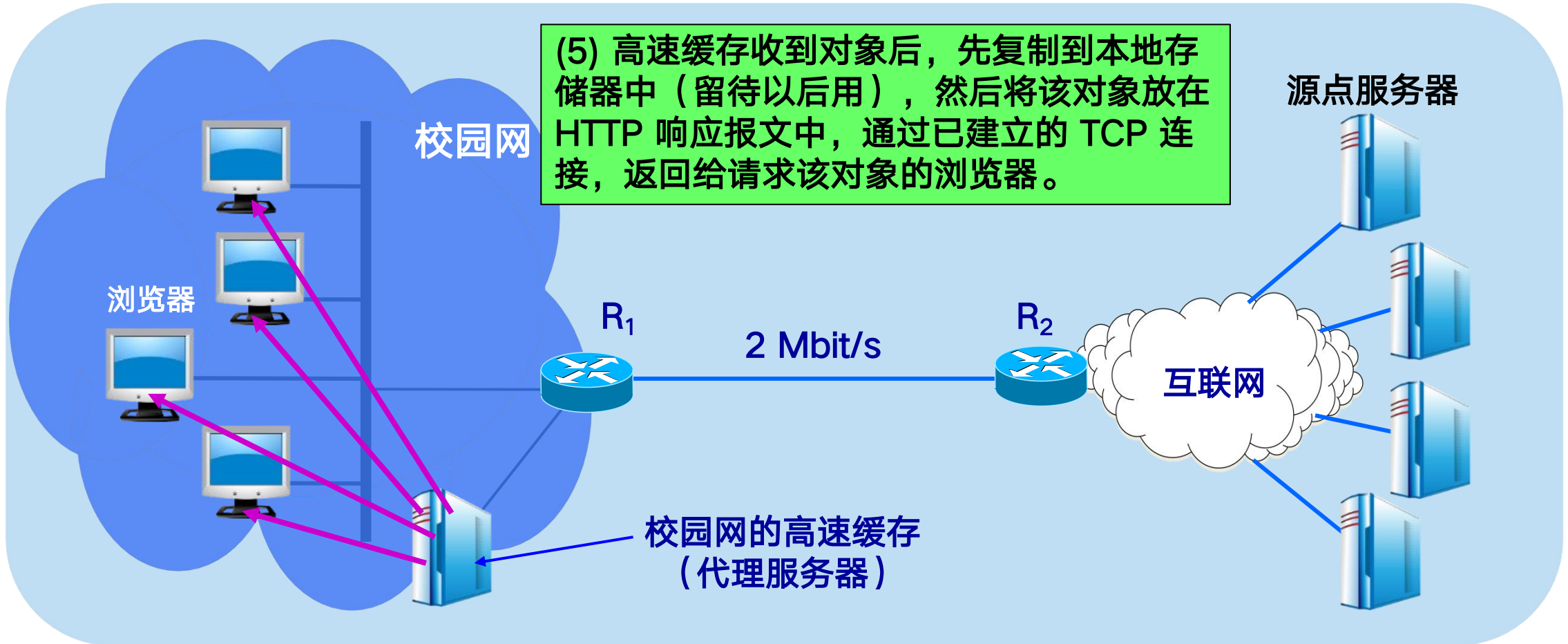
使用高速缓存的情况



使用高速缓存的情况



使用高速缓存的情况



01 万维网

【2015年 题40】某浏览器发出的HTTP请求报文如下

```
GET /index.html HTTP/1.1  
Host: www.test.edu.cn  
Connection: Close  
Cookie: 123456
```

下列叙述中，错误的是（**C**）。

- A. 该浏览器请求浏览index.html
- B. index.html存放在www.test.edu.cn上
- C. 该浏览器请求使用持续连接
- D. 该浏览器曾经浏览过www.test.edu.cn

01 万维网

【2015年 题40】某浏览器发出的HTTP请求报文如下

```
GET /index.html HTTP/1.1  
Host: www.test.edu.cn  
Connection: Close  
Cookie: 123456
```

下列叙述中，错误的是（**C**）。

- A. 该浏览器请求浏览index.html
- B. index.html存放在www.test.edu.cn上
- C. 该浏览器请求使用持续连接
- D. 该浏览器曾经浏览过www.test.edu.cn

解析

请求行：指明方法GET，URL，HTTP版本
首部行的开始：指明服务器的域名

01 万维网

【2015年 题40】某浏览器发出的HTTP请求报文如下

```
GET /index.html HTTP/1.1
Host: www.test.edu.cn
Connection: Close
Cookie: 123456
```

解析

请求行：指明方法GET，URL，HTTP版本
首部行的开始：指明服务器的域名

告诉服务器发送完请求的文档后就可释放连接，即非持续连接；
若是持续连接方式，取值应为keep-alive，而不是Close

下列叙述中，错误的是（**C**）。

- A. 该浏览器请求浏览index.html
- B. index.html存放在www.test.edu.cn上
- C. 该浏览器请求使用持续连接
- D. 该浏览器曾经浏览过www.test.edu.cn

01 万维网

【2015年 题40】某浏览器发出的HTTP请求报文如下

```
GET /index.html HTTP/1.1
Host: www.test.edu.cn
Connection: Close
Cookie: 123456
```

下列叙述中，错误的是（**C**）。

- A. 该浏览器请求浏览index.html
- B. index.html存放在www.test.edu.cn上
- C. 该浏览器请求使用持续连接
- D. 该浏览器曾经浏览过www.test.edu.cn

解析

请求行：指明方法GET，URL，HTTP版本
首部行的开始：指明服务器的域名

告诉服务器发送完请求的文档后就可释放连接，即非持续连接；
若是持续连接方式，取值应为keep-alive，而不是Close

这是服务器为浏览器生成的Cookie识别码，表明该浏览器曾经访问过www.test.edu.cn