

第2章 应用层

Web应用和HTTP协议

任课老师：周军海

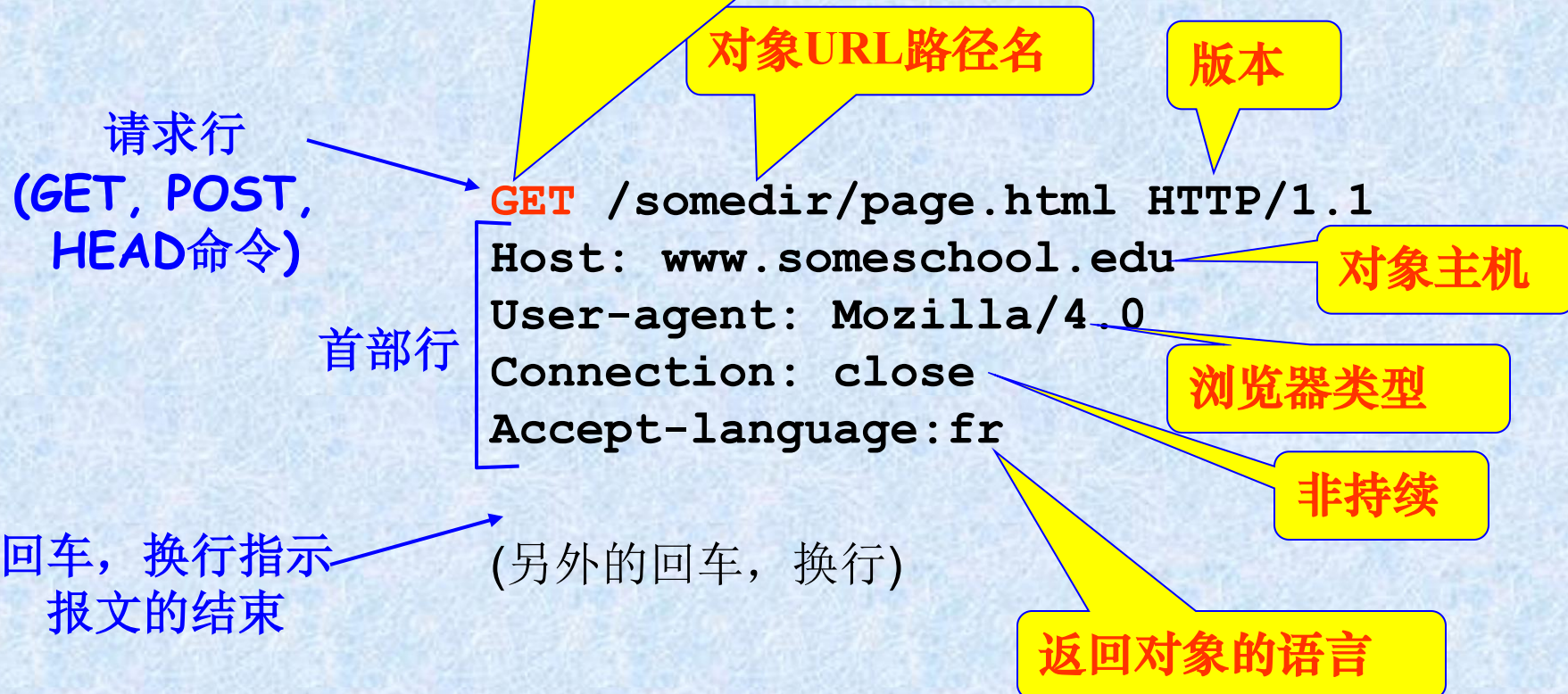
Email:rj_zjh@hnu.edu.cn

报文

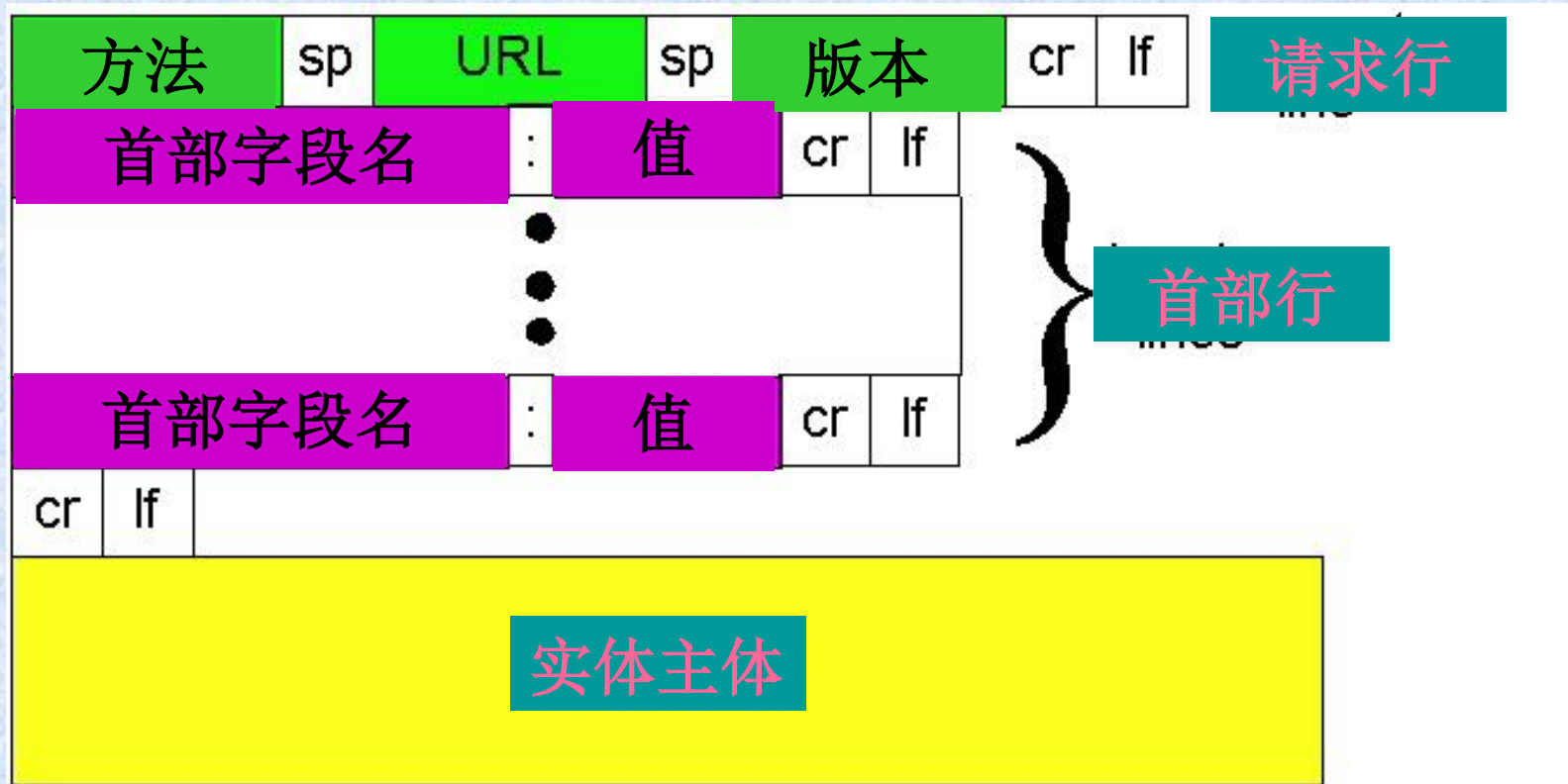
方法（命令）——

- ✓ GET: 请求一个对象。
- ✓ POST: 提交表单（添加信息）。
- ✓ HEAD: 请求返回对象响应报文首部

本形式，易读。



HTTP请求报文通用格式



“用表单生成的请求报文**不是必须**使用**POST**方法”
GET+扩展URL

2、HTTP 响应报文

服务器到客户机的回答

状态行
(版本、状态码、状态短语)

请求成功

发送日期

关闭连接

首部行

对象创建或修
改日期

对象长度

```
HTTP/1.1 200 OK
Connection: close
Date: Thu, 06 Aug 1998 12:00:15 GMT
服务器: Apache/1.3.0 (Unix)
Last-Modified: Mon, 22 Jun 1998 .....
Content-Length: 6821
Content-Type: text/html
```

```
data data data data data ...
```

实体：数据，如请
求的HTML文件

HTTP响应状态码

□ 在服务器到客户机响应报文中的首行。

200 OK

- 请求成功，请求的对象在这个报文后面

301 Moved Permanently

- 请求的对象已转移，新的URL在响应报文的Location:首部行中指定

400 Bad Request

- 请求报文不为服务器理解

404 Not Found

- 请求的文档没有在该服务器上发现

505 HTTP Version Not Supported

- 服务器不支持请求报文使用的HTTP版本

思考： P1, P4, P5

自行试验HTTP (客户机侧)

1. Telnet 到某个Web服务器上:

```
telnet csee.hnu.edu.cn 80
```

打开到位于**csee.hnu.edu.cn**的端口**80**(默认的**HTTP**服务器端口).
键入的任何东西将发送到位于**csee.hnu.edu.cn**的**80**端口

2. 键入一个GET HTTP请求:

```
GET /xygk.htm
```

向**HTTP**服务器发送最小的**GET**请求

3. 得到由HTTP服务器发送的响应报文!

```
C:\ 选择 C:\Windows\system32\cmd.exe
HTTP/1.0 403 Forbidden
Server: squid
Mime-Version: 1.0
Date: Wed, 21 Sep 2022 03:14:57 GMT
Content-Type: text/html
Content-Length: 1359
X-Squid-Error: ERR_ACCESS_DENIED 0
X-Cache: MISS from Hnu-Cnc-Proxy
Via: 1.0 Hnu-Cnc-Proxy (squid)
Connection: close

<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01//EN" "http://www.w3.org/TR/html4/strict.dtd">
<html><head>
  <meta http-equiv="Content-Type" content="text/html;
  charset=iso-8859-1">
  <title>ERROR: The requested URL could not be retrieved</title>
  <style type="text/css"><!--BODY{backg
round-color:#ffffff;font-family:verdana,sans-serif}PRE{font-family:sans-serif}--></style>
</head><body>
  <h1>ERROR</h1>
  <h2>The requested URL could not be retrieved</h2>
  <hr>
  <p>The following error was encountered while trying to retrieve the URL
: <a href="http://58.20.46.196/xygk.htm">http://58.20.46.196/xygk.htm</a></p>
  <blockquote>
    <p><b>Access Denied.</b></p>
  </blockquote>
  <p>Access control configuration prevents your request from being allowed at this time. Please contact your se
rvice provider if you feel this is incorrect.</p>
  <p>Your cache administrator is <a href="mailto:security@hnu.edu.cn?subj
ect=CacheErrorInfo%20-%20ERR_ACCESS_DENIED&amp;body=CacheHost%3A%20Hnu-Cnc-Proxy%0D%0AErrPage%3A%20ERR_ACCESS_DENIED%0D%0AErr%3A%20%5Bnone%5D%0D%0ATimeSt
amp%3A%20Wed,%2021%20Sep%202022%2003%3A14%3A57%20GMT%0D%0A%0D%0AClientIP%3A%20120.227.23.20%0D%0A%0D%0AHTTP%20Request%3A%0D%0AGET%20%2Fxygk.htm%20HTTP%2F
0.9%0A%0D%0A%0D%0A">security@hnu.edu.cn</a>.</p>
  <br>
  <hr>
  <div id="footer">Generated Wed, 21 Sep 2022 03:14:57 GMT by Hnu-Cnc-Proxy (squid)</div>
</body></html>
```

遗失对主机的连接。

2.2.4 用户与服务器交互：Cookie

HTTP服务器是**无状态**的，不保存客户信息。

- **Cookie**: 允许Web站点**跟踪**、识别用户；服务器可以限制用户访问，或把内容与用户身份关联。

许多重要的Web站点使用cookies。

- **包括四个部分**

- 1)在HTTP响应报文中有一个cookie 首部行
- 2)在HTTP请求报文中有一个cookie 首部行
- 3)用户主机中保留有一个 cookie 文件并由浏览器管理
- 4) Web站点的后端数据库保存cookie

例

- Susan总是从相同的PC访问因特网
- 她首次访问一个特定的电子商务站点
- 当起始HTTP请求到达站点时，站点产生一个独特的ID，并为此ID在后端数据库中生成一个表项

工作过程

客户机

服务器



❑ Cookie用途

- ✓ 身份认证
- ✓ 虚拟购物车(跟踪用户购买的物品)
- ✓ 推荐广告
- ✓ 用户会话状态 (Web e-mail)

❑ Cookie缺陷

- ✓ 站点可以知道用户许多信息
- ✓ 不利于用户隐私保护

2.2.5 HTTP内容

- ✓ **传输Web页面所含对象：**如HTML文件、JPEG文件、Java小应用程序等等。
- ✓ **可以传输其他类型的文件：**如传输XML（可扩展标记语言）文件。
- ✓ **P2P文件共享中：**作为文件传输协议使用。
- ✓ **用于流式存储的音频和视频。**

2.2.6 Web缓存

□ Web缓存器(Web cache): 也叫代理服务器。

能够代表初始web服务器来满足HTTP请求的网络实体。

✓ 保存最近请求过的对象的副本。

✓ 可在客户机或服务器工作，也可在中间系统工作。

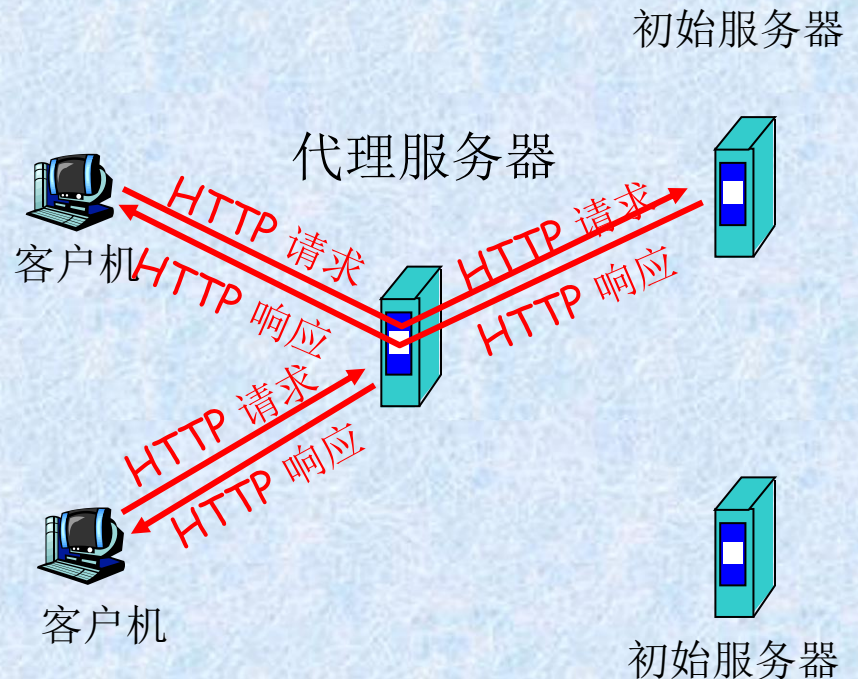
□ 初始（原始）服务器(origin server): 对象最初存放并始终保持其拷贝的服务器。

目标: 代替初始服务器满足HTTP请求。

使用Web缓存器

客户机通过Web缓存器请求对象。

- 用户配置浏览器: 所有Web访问经由Web缓存器
- 浏览器向Web缓存器发送所有HTTP请求
 - 对象在缓存中: Web缓存器返回对象
 - 不在: Web缓存器向初始服务器发出请求, 接收对象后转发给客户机

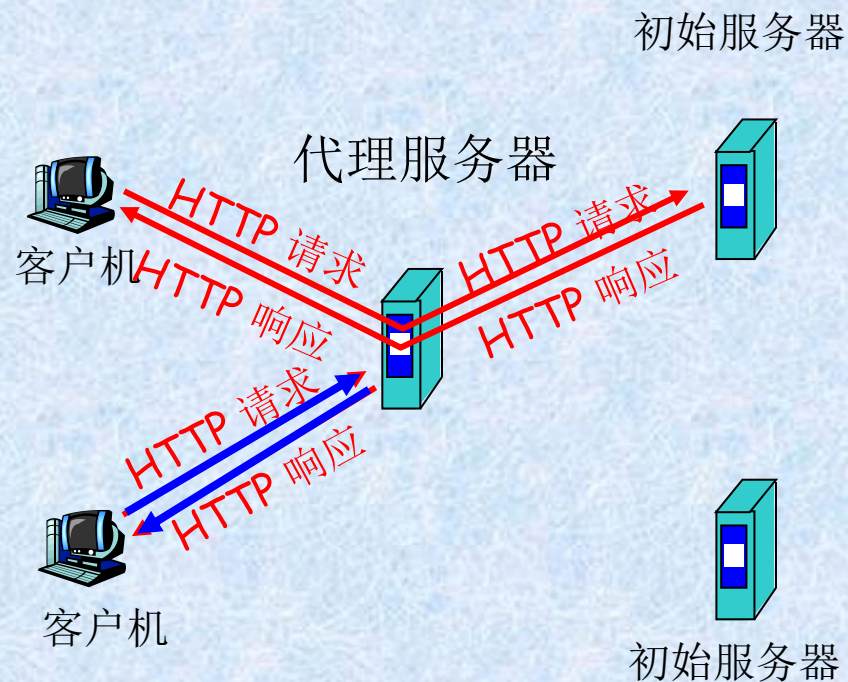


具体操作过程

例：假设浏览器请求对象

`http://www.someschool.edu/campus.gif`

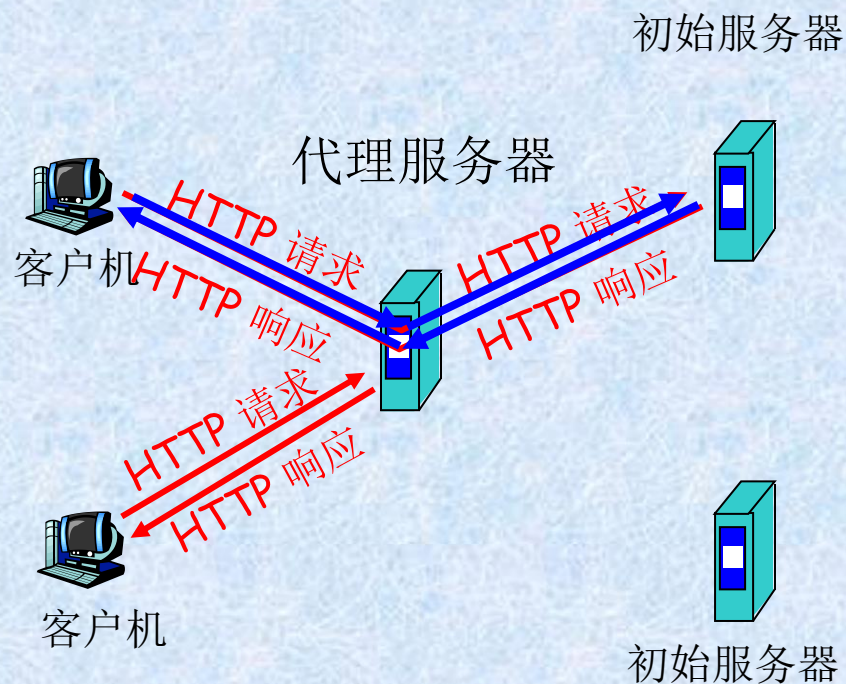
- **浏览器**：建立一个到Web缓存器的TCP连接，并向Web缓存器发送一个对该对象HTTP请求
- **Web缓存器**：检查本地是否有该对象的拷贝。
- ✓ **有**：就用HTTP响应报文向浏览器转发该对象



✓ **没有：** 与该对象的初始服务器打开一个TCP连接。

客户机 \longleftrightarrow Web缓存器 \longleftrightarrow 初始服务器

- ◆ Web缓存器在TCP连接上发送获取该对象的请求。
- ◆ 初始服务器收到请求，向Web缓存器发送该对象的HTTP响应
- ◆ Web缓存器接收该对象，存储一份在本地中，并通过HTTP响应报文向浏览器转发该对象（通过已经建立的TCP连接）。



说明

- Web缓存器既可以是服务器也可以是客户机：
 - ✓ 当它接收浏览器请求并发回响应时，是服务器；
 - ✓ 当它向初始服务器发出请求并接收响应时，是客户机

Web缓存优点

- 减少客户机请求的响应时间:

客户机 \longleftrightarrow Web缓存器 \longleftrightarrow 初始服务器

高速链路

- 减少机构内部网络与因特网连接链路上的通信量:

降低开销, 改善各种应用的性能。

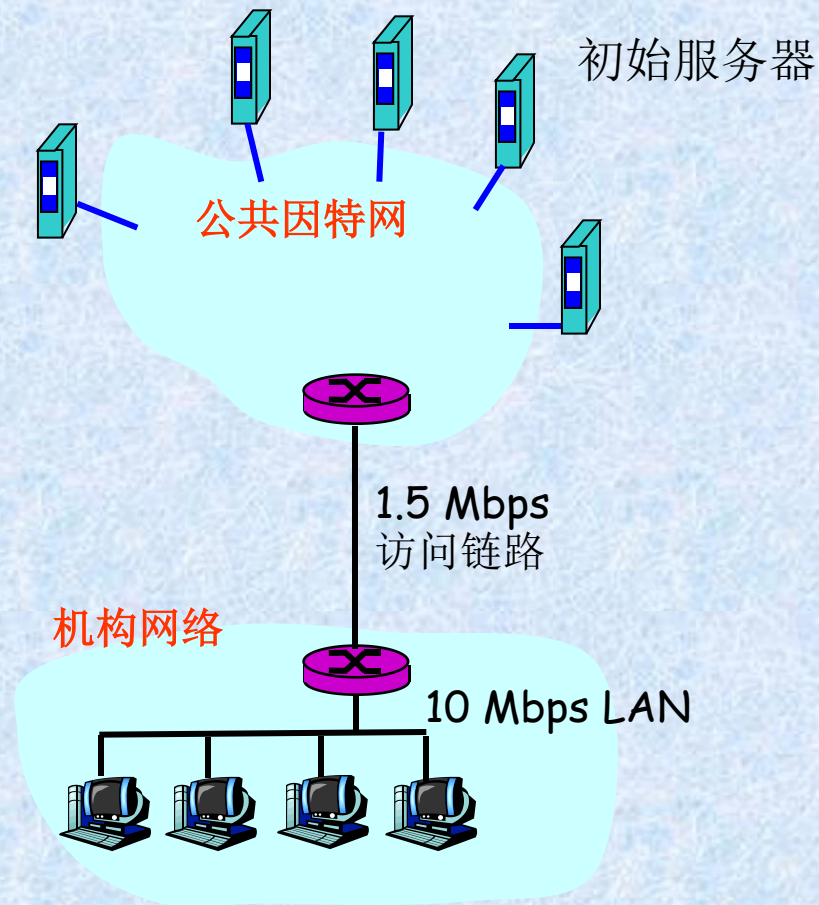
例1，无Web缓存

□ 包括两个网络：**机构的内部网络和因特网。**

✓ **机构内部网络：**是一个高速的局域网。

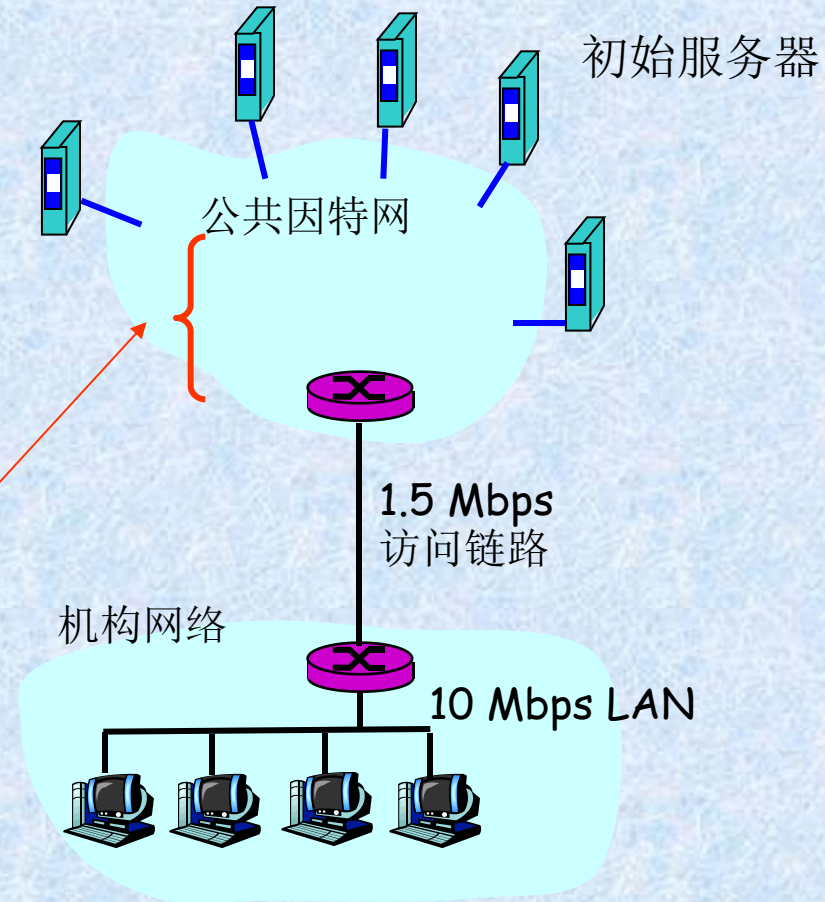
其路由器与因特网上的路由器通过一条1.5Mbps的链路连接。

✓ **初始服务器：**与因特网相连，遍布全球。



假设

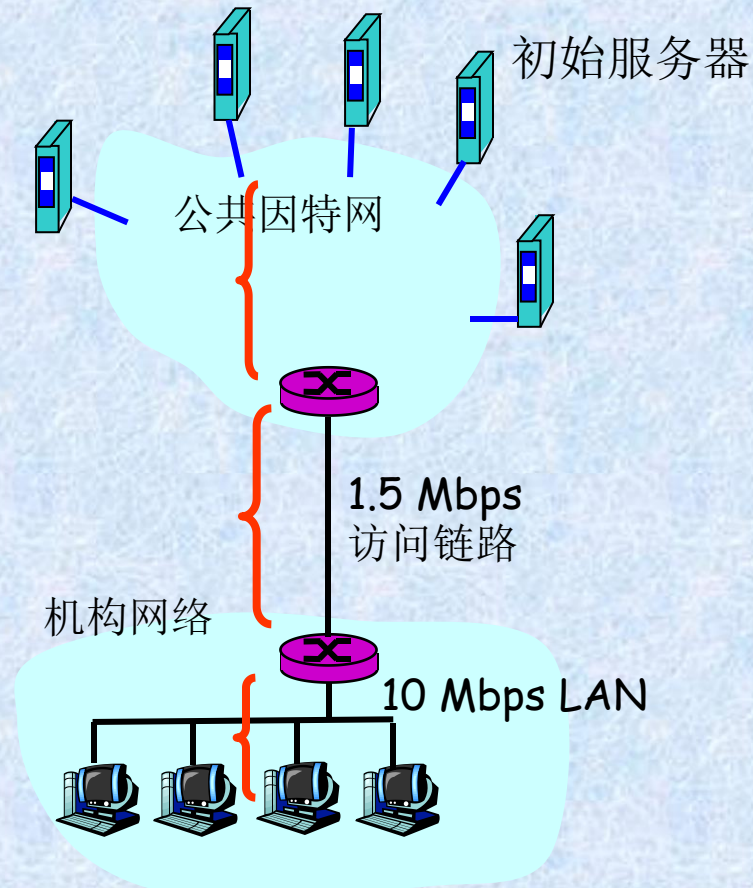
- ✓ 平均对象长度为**100kb**
- ✓ 机构内浏览器对初始服务器上对象的平均请求率=**15/sec**
- ✓ **机构内**的HTTP报文小，**传输时延忽略**
- ✓ 从因特网路由器转发HTTP请求报文，到收到其响应报文的时间平均=**2s（因特网时延）**



总的响应时间

浏览器从请求一个对象到接收到的时间：三部分和

- ✓ 局域网时延
- ✓ 接入链路时延(两个路由器间)
- ✓ 因特网时延



- ✓ **局域网时延**：与流量强度有关（**比特到达率/推出率**）

$$(15\text{请求/s}) \times (100\text{kb/请求}) / (10\text{Mbit/s}) = 0.15$$

强度为0.15的通信量最多数十毫秒的时延，可**忽略**

- ✓ **接入链路时延**：与接入**流量强度**（路由器之间）有关

$$(15\text{请求/s}) \times (100\text{kb/请求}) / (1.5\text{Mbit/s}) = 1$$

强度接近1，链路**时延非常大**或无限增长。

- ✓ **总响应时间** = 接入链路时延 + 因特网时延
= （分钟 + 2 sec）

请求时间长，用户难接受。

改进方法一

- ✓ 增加接入链路的速率：

如从1.5Mbps增加到10Mbps，使链路上的流量强度减少到0.15，链路时延也可以忽略了。

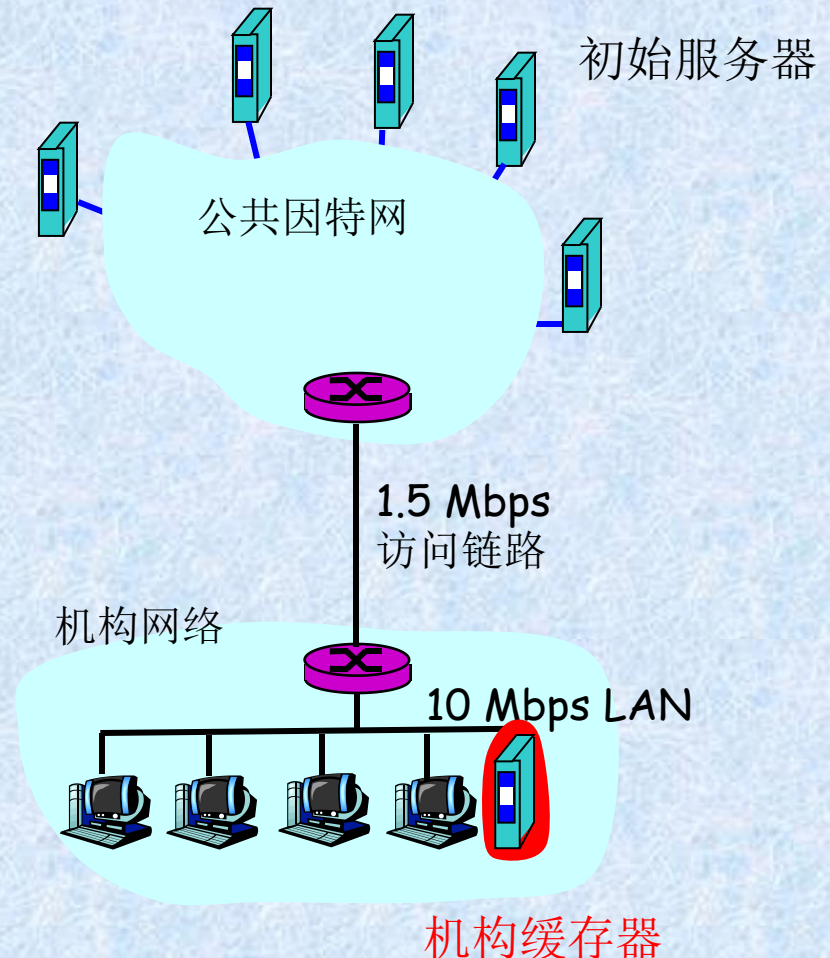
总响应时间=因特网时延=2秒钟

- ✓ 投资较大，成本昂贵。

改进方法二

在机构网络中安装一个Web缓存器。

- ✓ Web缓存器的命中率：缓存器满足请求的比率（0.2~0.7）。
- ✓ 设命中率为0.4。

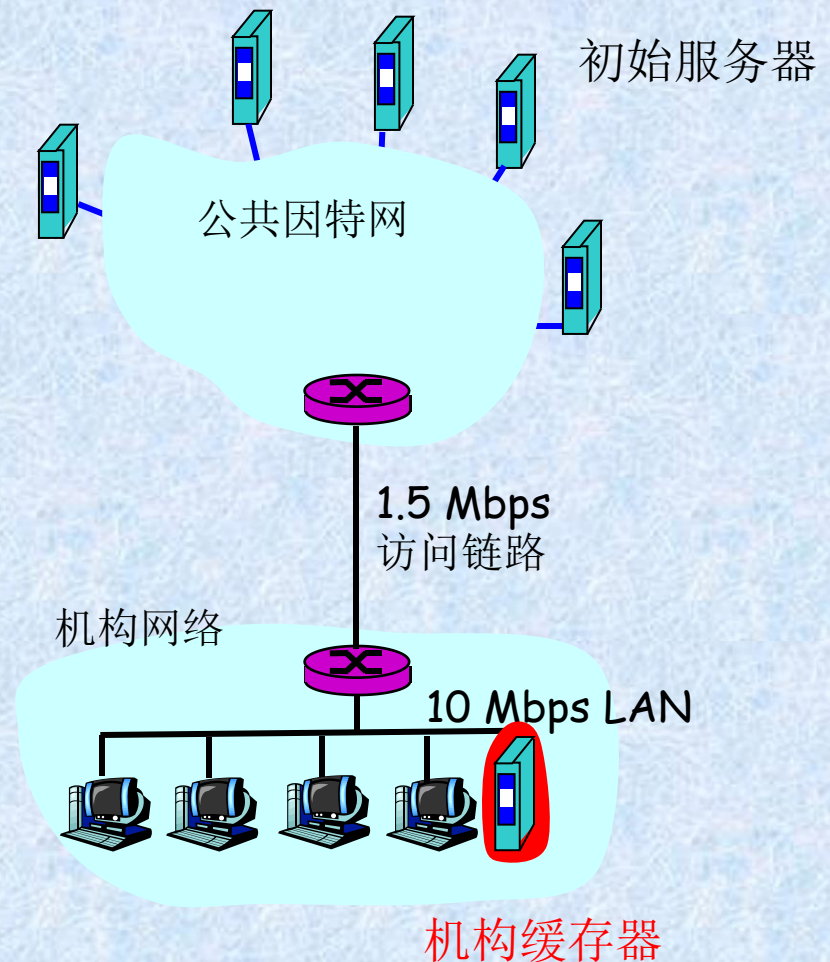


改进方法二

100kb/10Mbps

- ✓ 局域网时延：客户机缓存器位于同一局域网，40%的请求几乎会立即得到响应，时延约10ms。

剩下的60%请求需要通过访问初始服务器才能满足。



✓ 接入链路时延:

只有60%的请求对象通过接入链路传送, 流量强度从1.0减小到0.6。

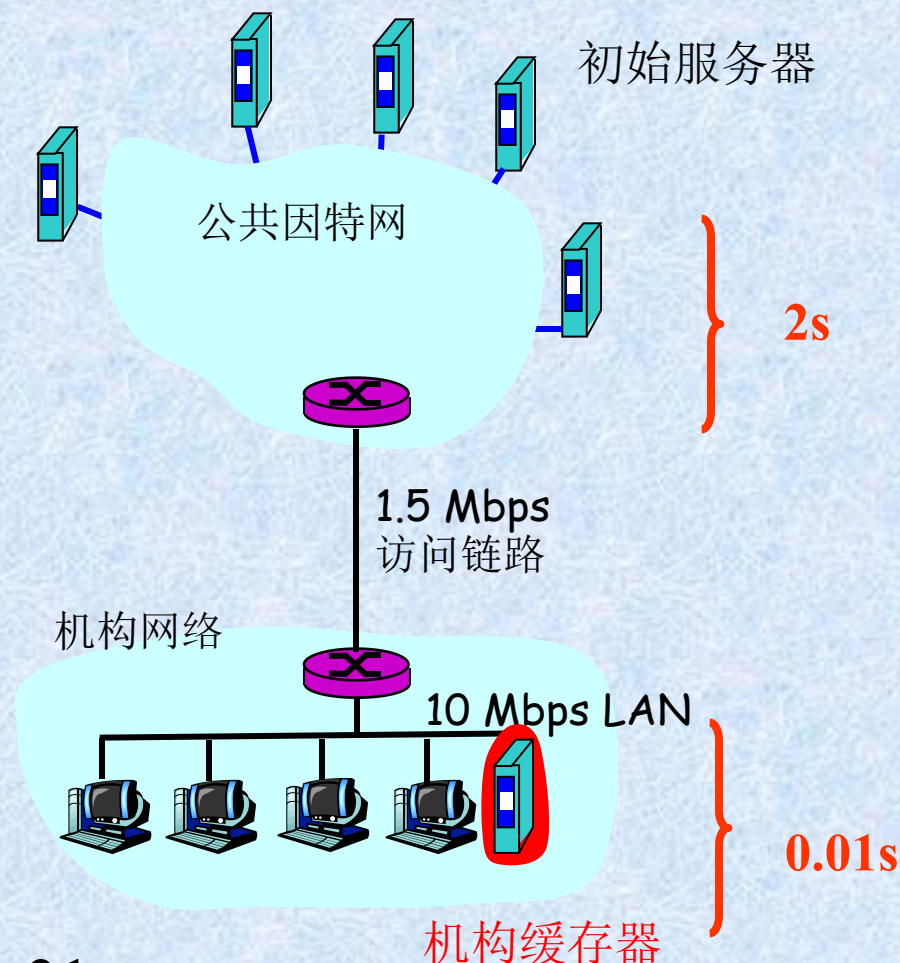
通常, 在1.5Mbps链路上, 当流量强度小于0.8时, 链路时延很小, 可忽略。

✓ 平均时延为:

$$0.4 \times (0.01s) + 0.6 \times (0.01s + 2s) = 1.21s$$

Web缓存器减少响应时延, 成本低

思考: P9



2.2.7 条件GET方法

□ 高速缓存：

- ✓ 减少响应时间；
- ✓ 存放在缓存中的对象拷贝可能是旧的。即保存在初始Web服务器中的对象可能已经被修改。

□ 条件GET方法：

- ✓ 使缓存器能够证实其保存的对象是否为最新。
- ✓ 如果缓存中是最新对象版本，可继续使用，初始Web服务器就不需重新发送该对象。

条件GET方法使用

- Web服务器回发**响应报文**：包括对象的最后修改时间
Last-modified: date1
- 缓存检查Web服务器中的该对象是否已被修改，发送一个**条件GET请求报文**：
If-modified-since: date1
- ✓ 告诉服务器，仅当**自指定日期之后该对象被修改过**，才发送该对象。
- ✓ 若Web服务器中的该对象未被修改，则响应报文含有**304 Not Modified**，并且实体为空。

例

缓存器

服务器

缓存器将对象转发到浏览器，并保存对象到本地（包括对象的最后修改时间）。

GET /fruit/kiwi.gif HTTP/1.1
Host:www.exotiquecuisine.com

HTTP/1.0 200 OK
Last-modified: date1
<data>

一周后，用户再次请求该对象（仍保留在缓存器中），缓存器发送一个条件GET，检查该对象是否已被修改

GET /fruit/kiwi.gif HTTP/1.1
Host:www.exotiquecuisine.com
If-modified-since: date1

对象未修改

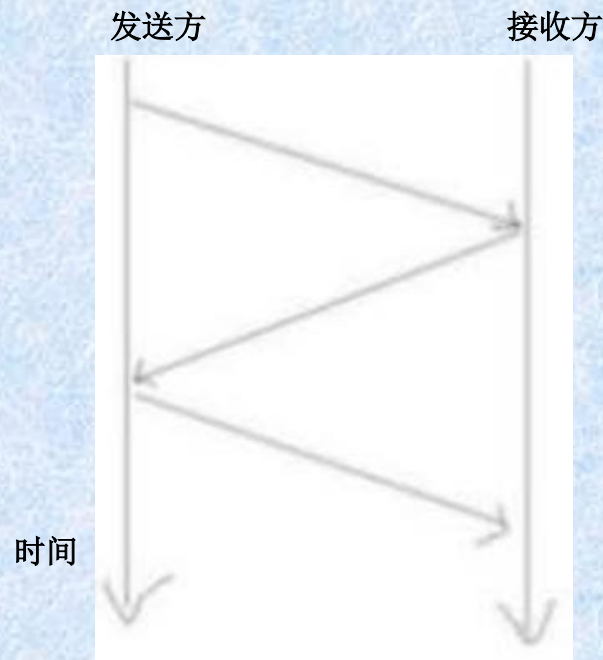
304 Not Modified
实体为空

对象未修改，缓存器可以继续使用该对象的拷贝，并转发给用户浏览器。

复习题：R14

2、下图描述的协议要素是

I、语法 II、语义 III、时序



☐ A 仅 I

☒ C 仅 III

☐ B 仅 II

☐ D I、II、III

提交