



# 第 6 章 应用层

---



# 应用层

---

6.1 域名系统 DNS

6.4 万维网 WWW

6.5 电子邮件



# 应用层协议的特点

---

- 每个应用层协议都是为了解决某一类应用问题，而问题的解决又往往是通过位于不同主机中的多个应用进程之间的通信和协同工作来完成的。应用层的具体内容就是规定应用进程在通信时所遵循的协议。
- 应用层的许多协议都是基于客户服务器方式。客户(client)和服务器(server)都是指通信中所涉及的两个应用进程。客户服务器方式所描述的是进程之间服务和被服务的关系。客户是服务请求方，服务器是服务提供方。



# 6.1 域名系统 DNS

## 6.1.1 域名系统概述

---

- 许多应用层软件经常直接使用**域名系统** DNS (Domain Name System)，但计算机的用户只是间接而不是直接使用域名系统。
- 互联网采用层次结构的命名树作为主机的名字，并使用**分布式的**域名系统 DNS。



## 6.1.1 域名系统概述

---

- 名字到 IP 地址的解析是由若干个域名服务器程序完成的。域名服务器程序在专设的结点上运行，运行该程序的机器称为**域名服务器**。



## 6.1.2 互联网的域名结构

- 域名的结构由标号序列组成，各标号之间用点隔开：

... . 三级域名 . 二级域名 . 顶级域名

- 各标号分别代表不同级别的域名。



# 顶级域名 TLD

## (Top Level Domain)

---

### (1) 国家顶级域名：

.cn 表示中国，.us 表示美国，.uk 表示英国，等等。

### (2) 通用顶级域名：最常用顶级域名是：

.com	（公司和企业）	.net	（网络服务机构）
.org	（非营利性组织）	.edu	（美国专用的教育机构）
.gov	（美国的政府部门）	.mil	（美国军事部门）
.int	（国际组织）		



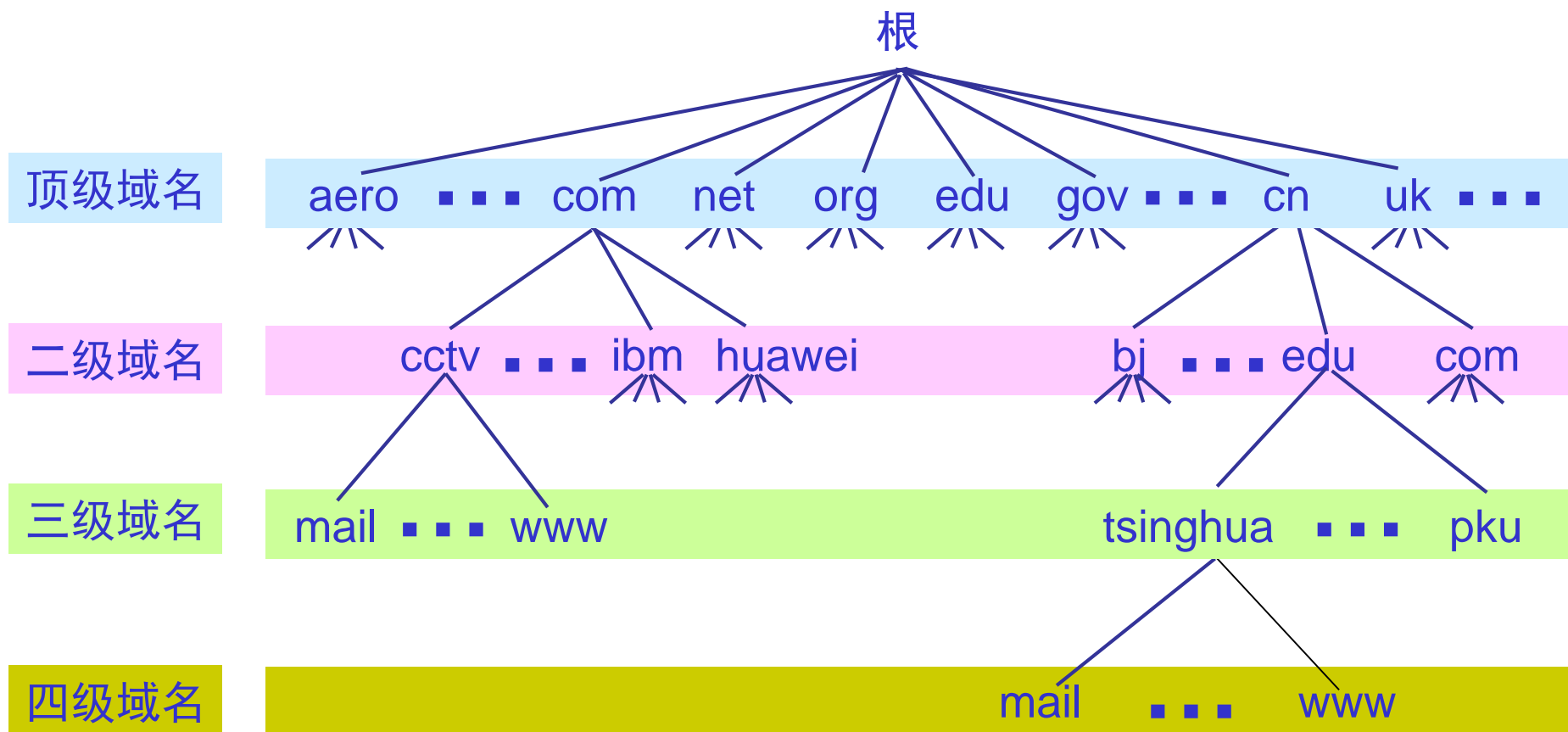
## 顶级域名 TLD （续）

---

(3) 基础结构域名(infrastructure domain): 这种顶级域名只有一个, 即 arpa, 用于反向域名解析, 因此又称为反向域名。



# 互联网的域名空间

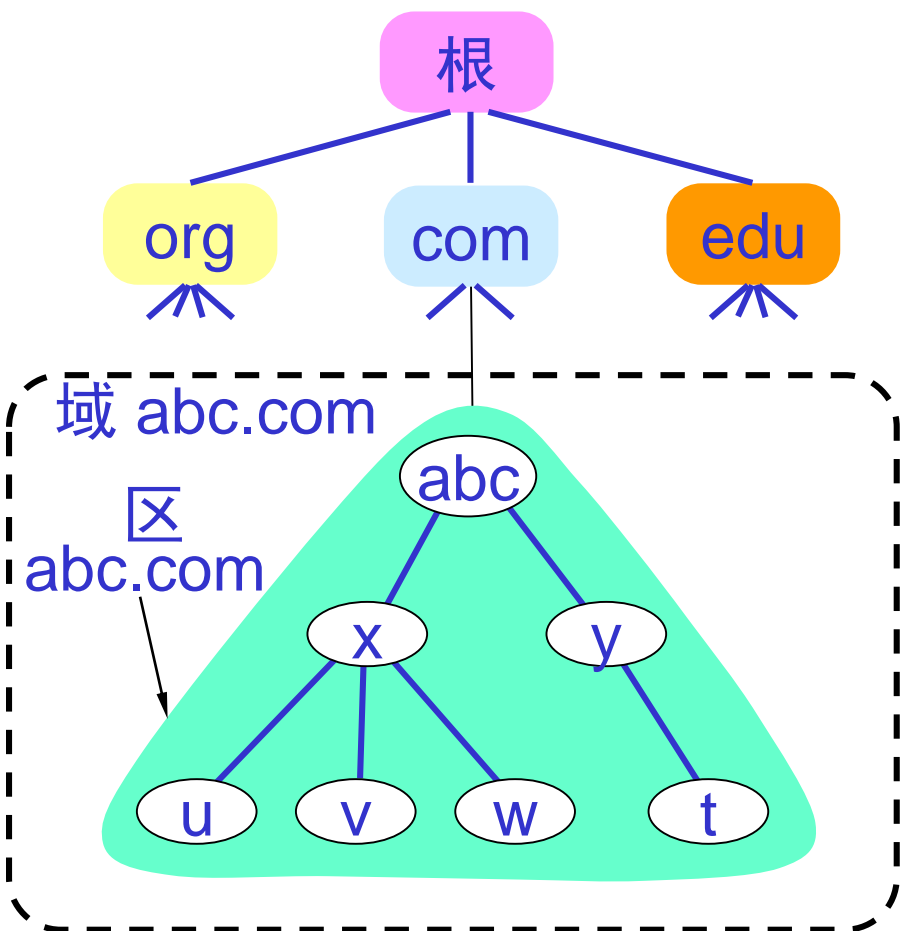




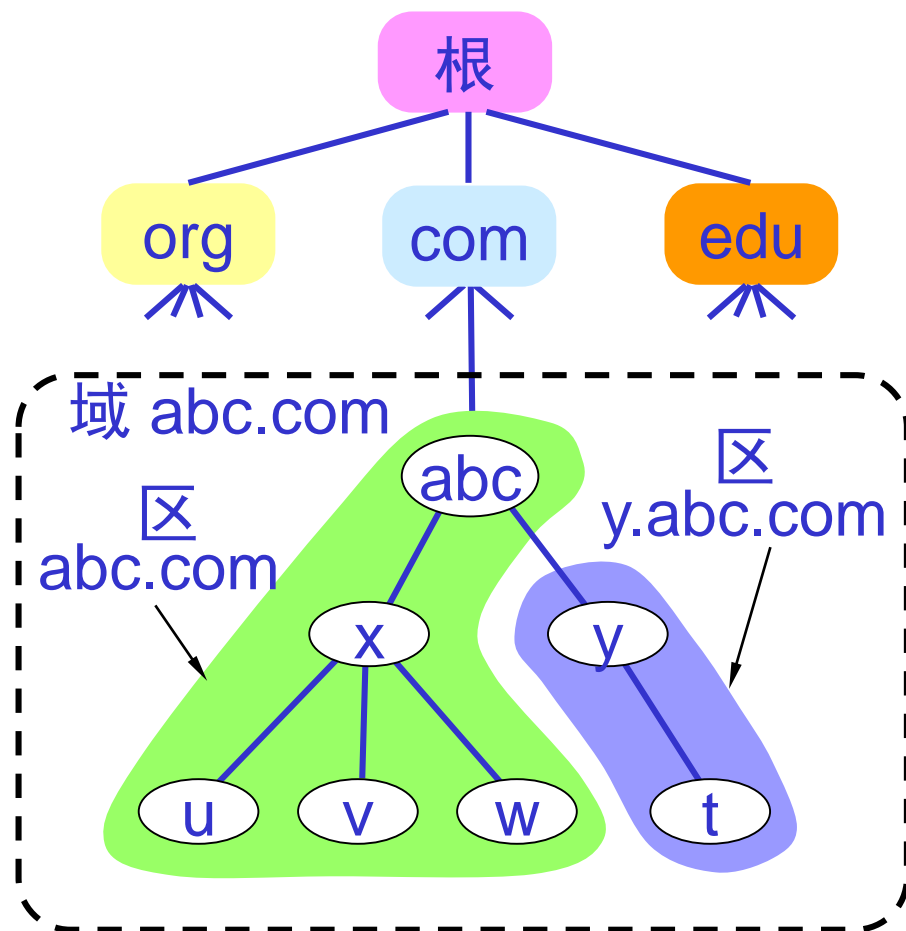
## 6.1.3 域名服务器

- 一个服务器所负责管辖的（或有权限的）范围叫做**区**(zone)。
- 各单位根据具体情况来划分自己管辖范围的区。但在一个区中的所有节点必须是能够连通的。
- 每一个区设置相应的**权限域名服务器**，用来保存该区中的所有主机的域名到IP地址的映射。

# 区的不同划分方法举例



(a) 区 = 域



(b) 区 < 域

# 树状结构的 DNS 域名服务器

根域名服务器

根域名服务器

顶级域名服务器

org 域名服务器

com 域名服务器

edu 域名服务器

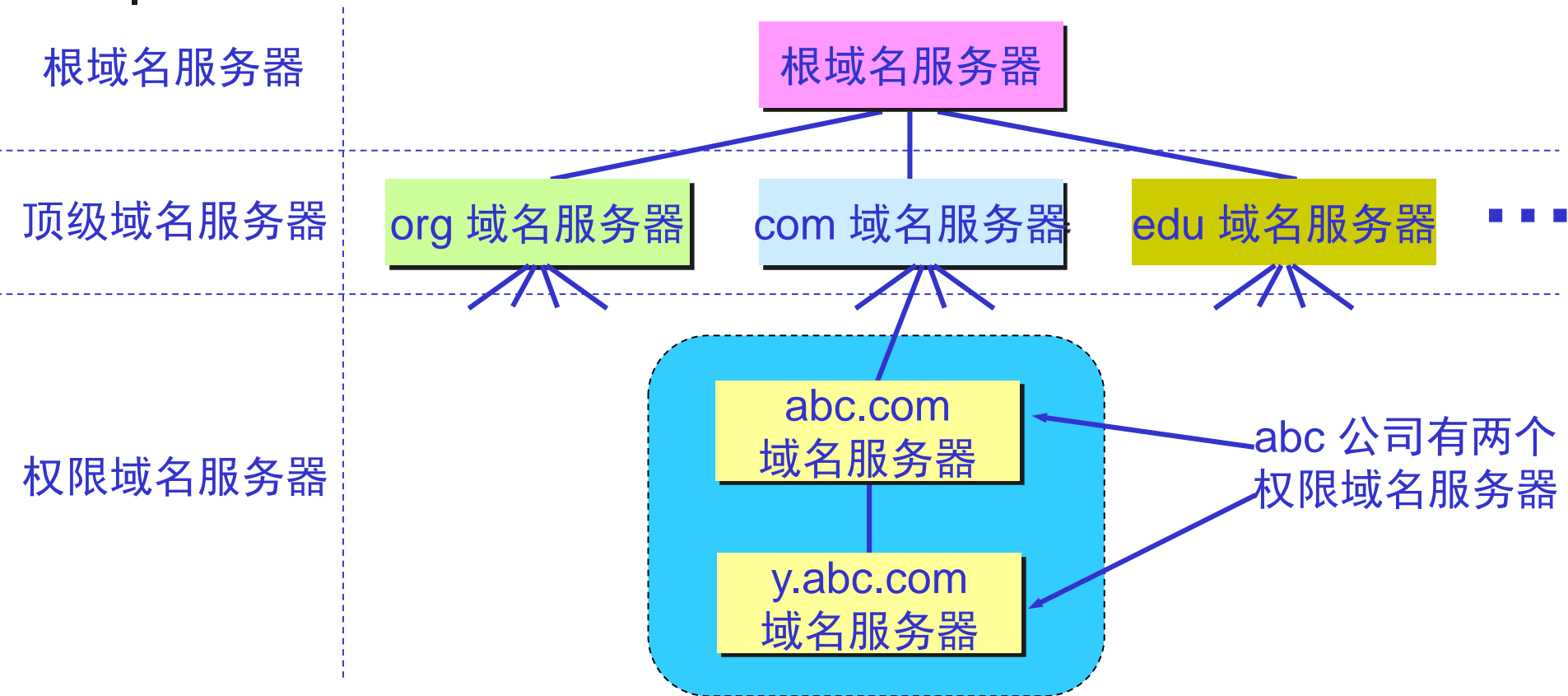
...

权限域名服务器

abc.com  
域名服务器

y.abc.com  
域名服务器

abc 公司有两个  
权限域名服务器





# 域名服务器有以下四种类型

---

- 根域名服务器
- 顶级域名服务器
- 权限域名服务器
- 本地域名服务器



# 根域名服务器

## —最高层次的域名服务器—

- 根域名服务器是最重要的域名服务器。所有的根域名服务器都知道所有的顶级域名服务器的域名和 IP 地址。
- 不管是哪一个本地域名服务器，若要对互联网上任何一个域名进行解析，只要自己无法解析，就首先求助于根域名服务器。
- 在因特网上共有13 个不同 IP 地址的根域名服务器，它们的名字是用一个英文字母命名，从 a 一直到 m（前13 个字母）。

# 根域名服务器共有 13 套装置 (不是 13 个机器)

- 这些根域名服务器相应的域名分别是  
a.rootservers.net  
b.rootservers.net  
...  
m.rootservers.net
- 到 2021 年全世界已经安装了 1375 个根域名服务器机器，分布在世界各地。
- 这样做的目的是为了更方便用户，使世界上大部分 DNS 域名服务器都能就近找到一个根域名服务器。

# 根域名服务器 L 的地点分布图 (2012 年5月)



**根域名服务器 L 分布在世界 150 个地点**

- 根域名服务器并不直接把域名直接转换成 IP 地址。
- 在使用迭代查询时，根域名服务器把下一步应当找的顶级域名服务器的 IP 地址告诉本地域名服务器。





# 顶级域名服务器和权限域名服务器

---

- 顶级域名服务器:这些域名服务器负责管理在该顶级域名服务器注册的所有二级域名。
- 权限域名服务器:负责一个区的域名服务器。



# 本地域名服务器

---

- 本地域名服务器对域名系统非常重要。
- 当一个主机发出 DNS 查询请求时，这个查询请求报文就发送给本地域名服务器。
- 每一个互联网服务提供者 ISP，或一个大学，甚至一个大学里的系，都可以拥有一个本地域名服务器。

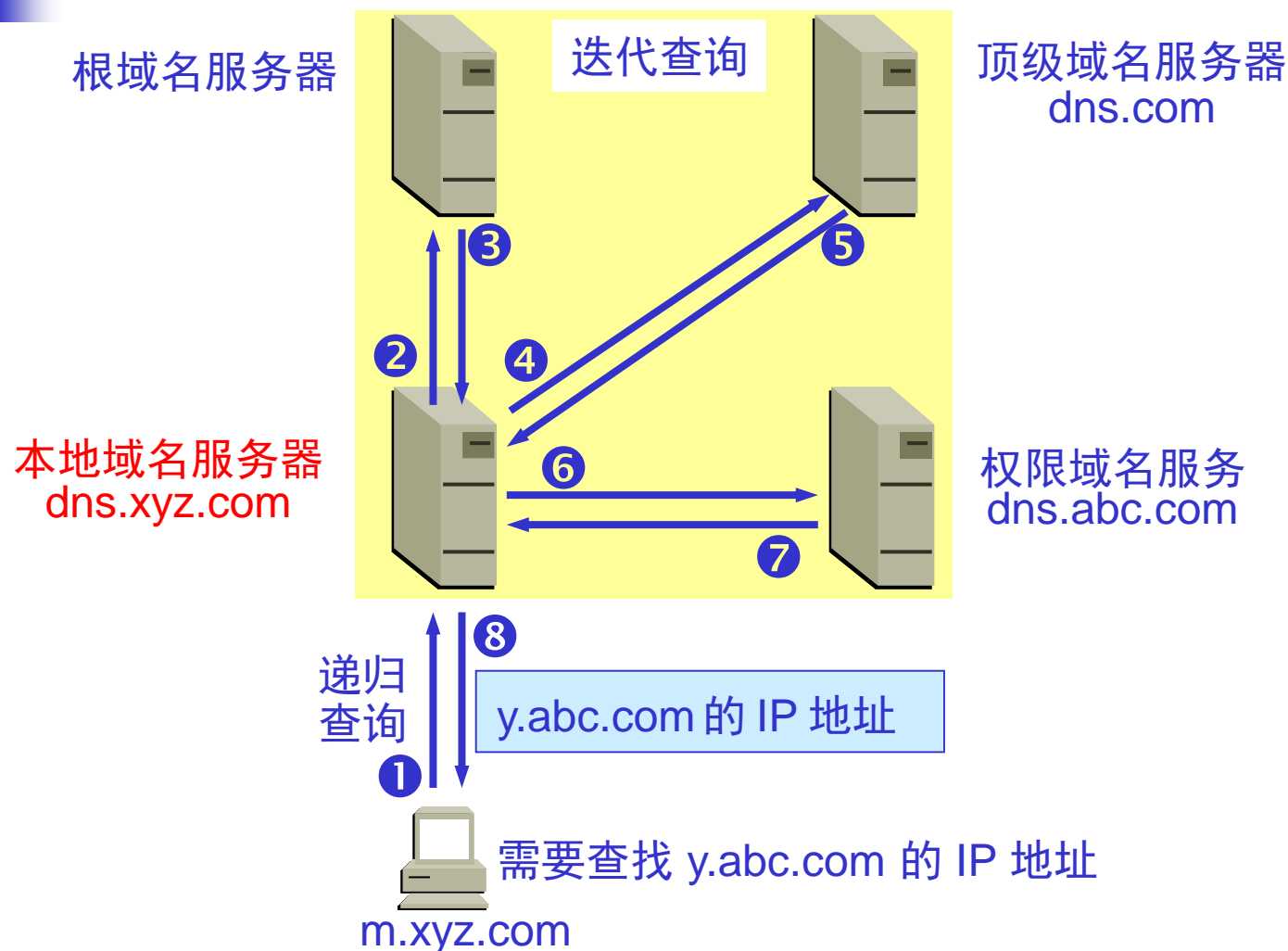


# 域名的解析过程

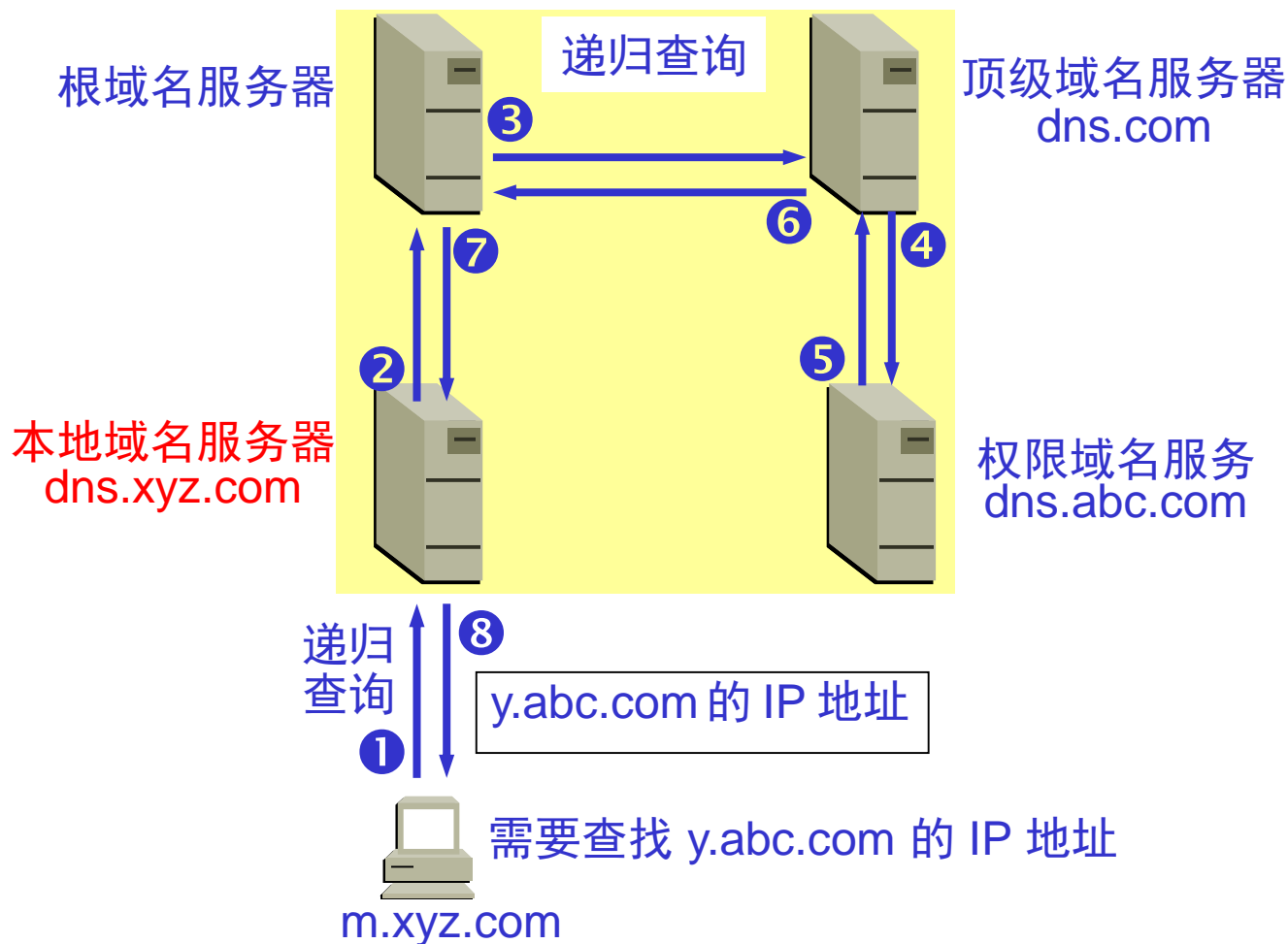
---

- 主机向本地域名服务器的查询一般都是采用**递归查询**。
- 本地域名服务器向根域名服务器的查询通常是采用**迭代查询**。

# 本地域名服务器采用迭代查询



# 本地域名服务器采用递归查询 (比较少用)



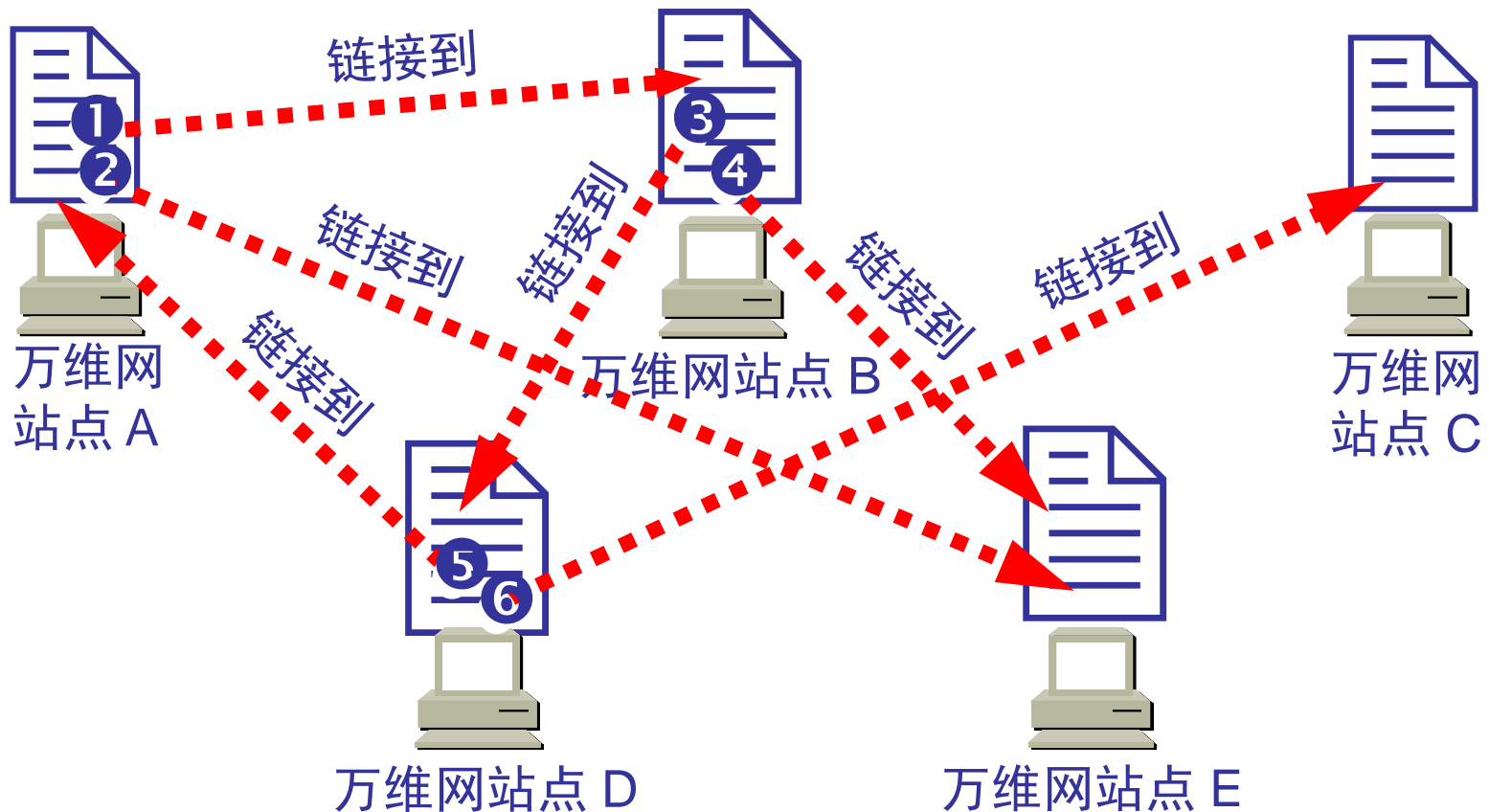
## 6.4 万维网 WWW

### 6.4.1 万维网概述

---

- **万维网** WWW (World Wide Web)是一个大规模的、联机式的信息储藏所。英文简称为Web。

# 万维网提供分布式服务





# 万维网工作方式-客户服务器方式

---

- 浏览器就是在用户计算机上的万维网客户程序。
- 万维网文档所驻留的计算机则运行服务器程序，因此这个计算机也称为万维网服务器。
- 客户程序向服务器程序发出请求，服务器程序向客户程序送回客户所要的万维网文档。





# 万维网必须解决的问题(1)

---

怎样标志分布在整个因特网上的万维网文档？

- 使用**统一资源定位符** URL (Uniform Resource Locator)来标志万维网上的各种文档。



## 万维网必须解决的问题 (2)

---

用何协议实现万维网上各种链接？

- 在万维网客户程序与万维网服务器程序之间进行交互所使用的协议，是超文本传送协议 HTTP (HyperText Transfer Protocol)。



## 万维网必须解决的问题 (3)

---

怎样使各种万维网文档都能在网上各种计算机上显示出来, 同时使用户清楚地知道在什么地方存在着超链?

- 使用 超文本标记语言 HTML (HyperText Markup Language) 。



## 万维网必须解决的问题(4)

---

怎样使用户能够很方便地找到所需的信息？

- 为了在万维网上方便地查找信息，用户可使用各种的搜索工具（即搜索引擎）。

## 6.4.2 统一资源定位符 URL

### 1. URL的格式

- 统一资源定位符 URL 是对可以从因特网上得到的资源的位置和访问方法的一种简洁的表示。
- URL 的一般形式是：

<协议>://<主机>:<端口>/<路径>

http —— 超文本传送协议 HTTP



# 使用 HTTP 的 URL

---

- <http://finance.sina.com.cn/g/20151208/09397071495.html>

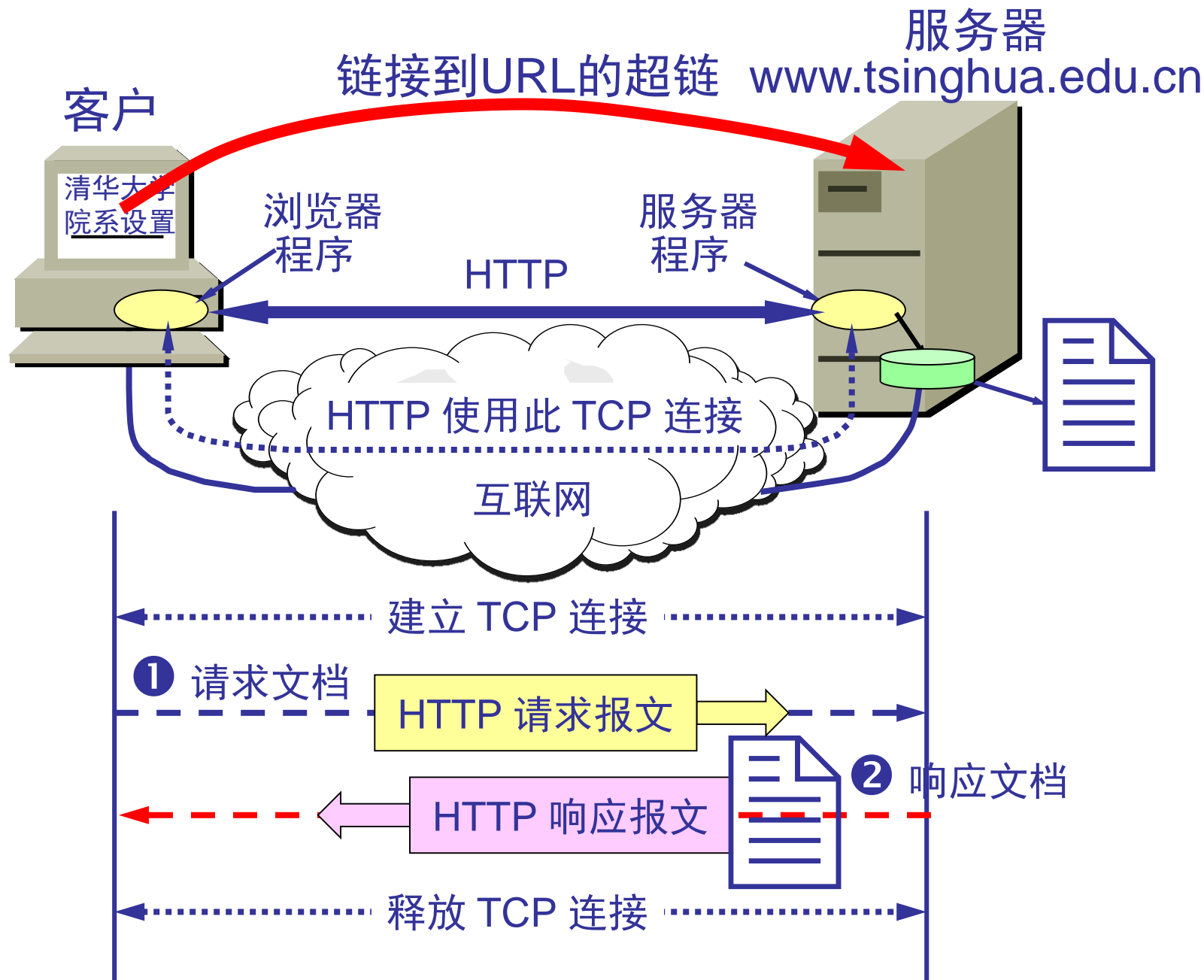
## 6.4.3 超文本传送协议 HTTP

### 1. HTTP 的操作过程

---

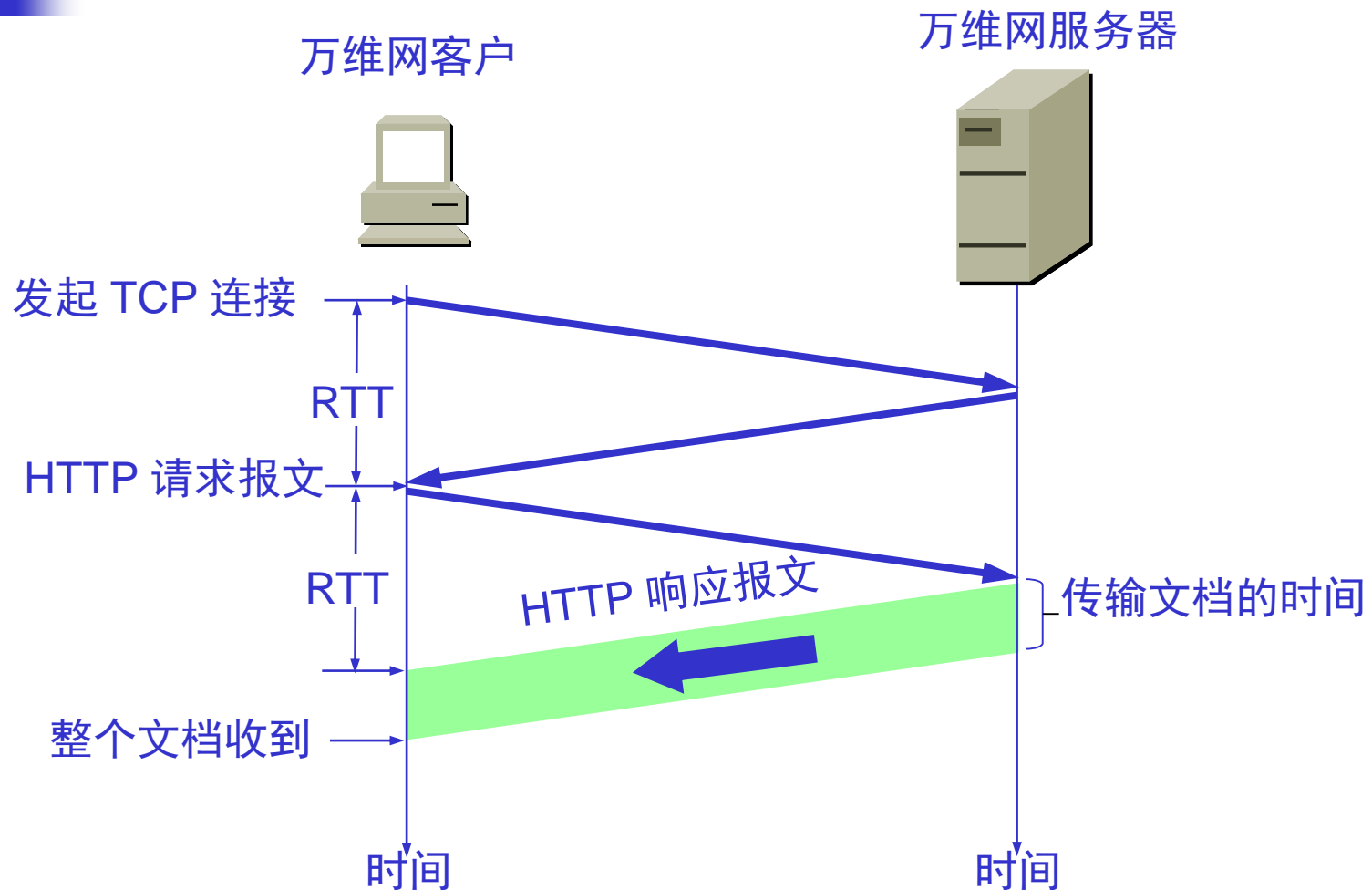
- HTTP 协议定义了浏览器怎样向万维网服务器请求万维网文档，以及服务器怎样把文档传送给浏览器。

# 万维网的工作过程





# HTTP/1.0请求一个万维网文档所需的时间





# 持续连接(persistent connection)

- HTTP/1.1 协议使用持续连接。
- 万维网服务器在发送响应后仍然在一段时间内保持这条连接，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。
- 目前一些流行的浏览器（例如，IE 11.0）的默认设置就是使用 HTTP/1.1。



# 持续连接的两种工作方式

---

- 非流水线方式：客户在收到前一个响应后才能发出下一个请求。这比非持续连接的两倍 RTT 的开销节省了一个 RTT 时间。
- 流水线方式：客户在收到 HTTP 的响应报文之前就能够接着发送新的请求报文。一个接一个的请求报文到达服务器后，服务器就可连续发回响应报文。



# HTTP/2

---

- 是协议 HTTP/1.1 的升级版本。
- 服务器可以并行发回响应（使用同一个 TCP 连接）。
- 允许客户复用 TCP 连接进行多个请求。
- 把所有的报文都划分为许多较小的二进制编码的帧，并采用了新的压缩算法，不发送重复的首部字段，大大减小了首部的开销，提高了传输效率。
- 向后兼容。

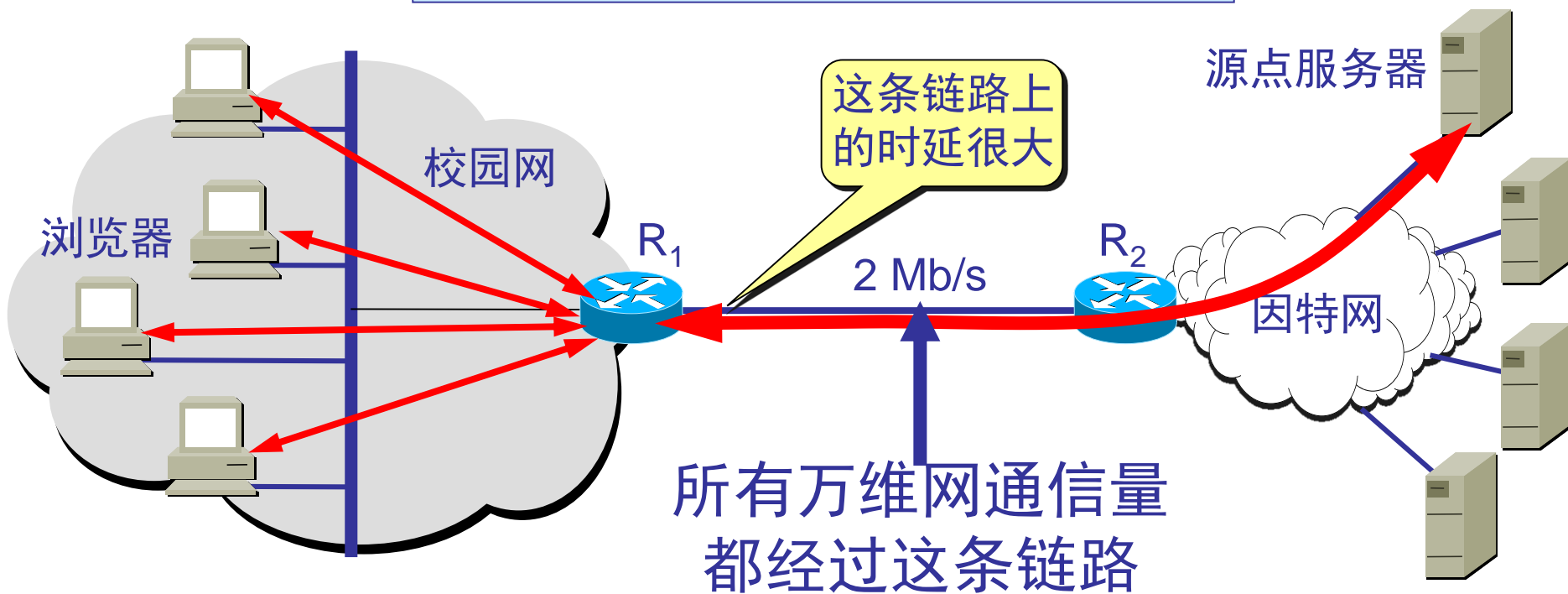


# 代理服务器(proxy server)

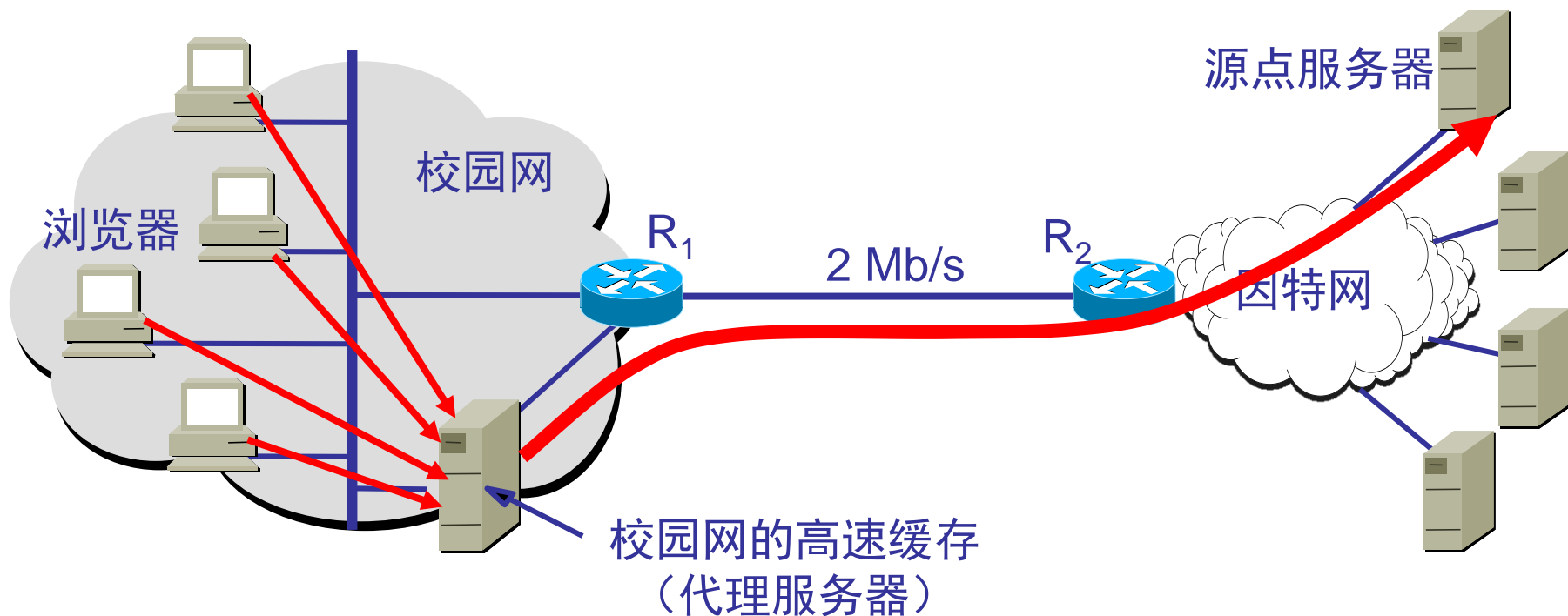
- **代理服务器**(proxy server)又称为万维网高速缓存(Web cache), 它代表浏览器发出 HTTP 请求。
- 万维网高速缓存把最近的一些请求和响应暂存在本地磁盘中。
- 当与暂时存放的请求相同的新请求到达时, 万维网高速缓存就把暂存的响应发送出去, 而不需要按 URL 的地址再去因特网访问该资源。

# 使用高速缓存可减少访问因特网服务器的时延

没有使用高速缓存的情况



# 使用代理服务器





### 3. HTTP 的报文结构

---

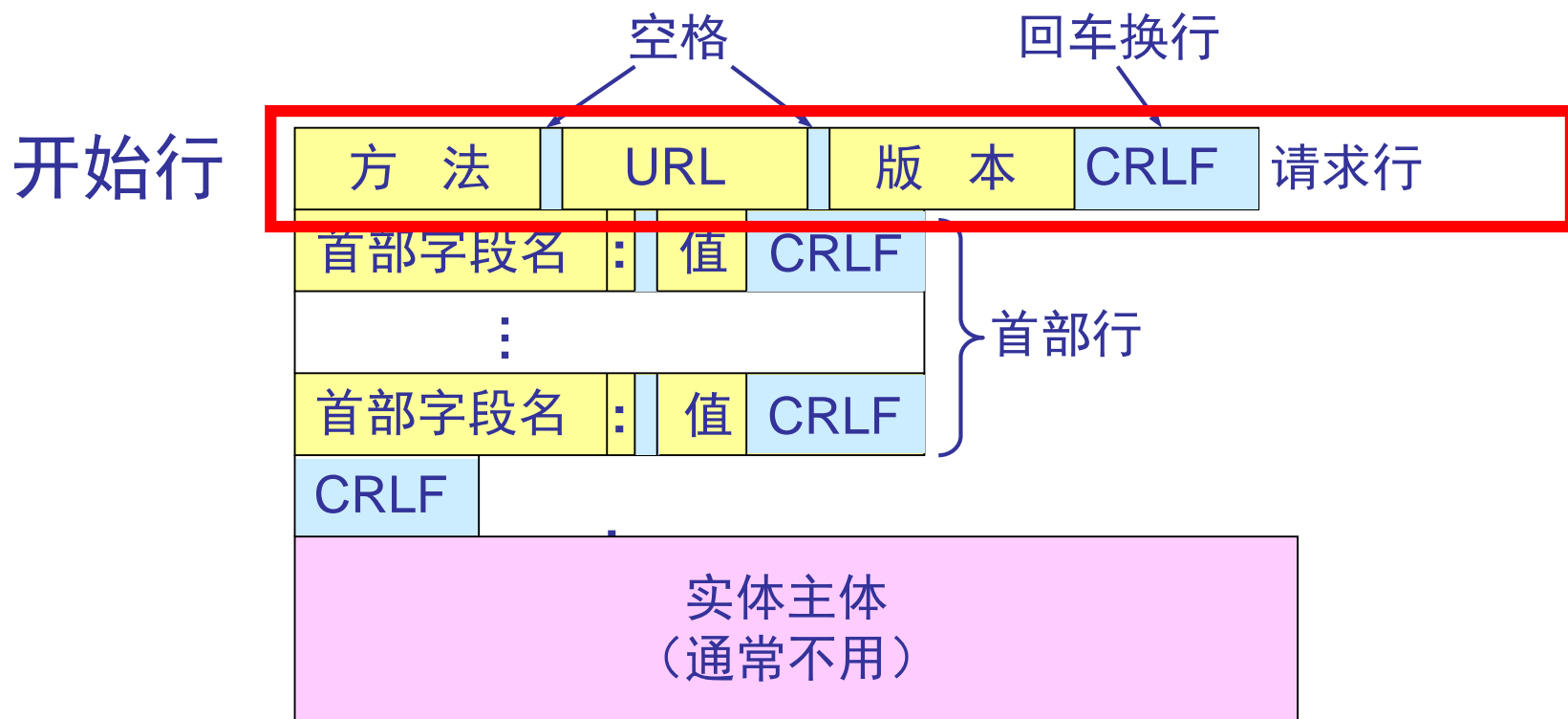
- HTTP是面向文本的, 每一个字段都是一些 **ASCII 码** 串, 各个字段长度都不确定。

HTTP 有两类报文:

- 请求报文——从客户向服务器发送请求报文。
- 响应报文——从服务器到客户的回答。

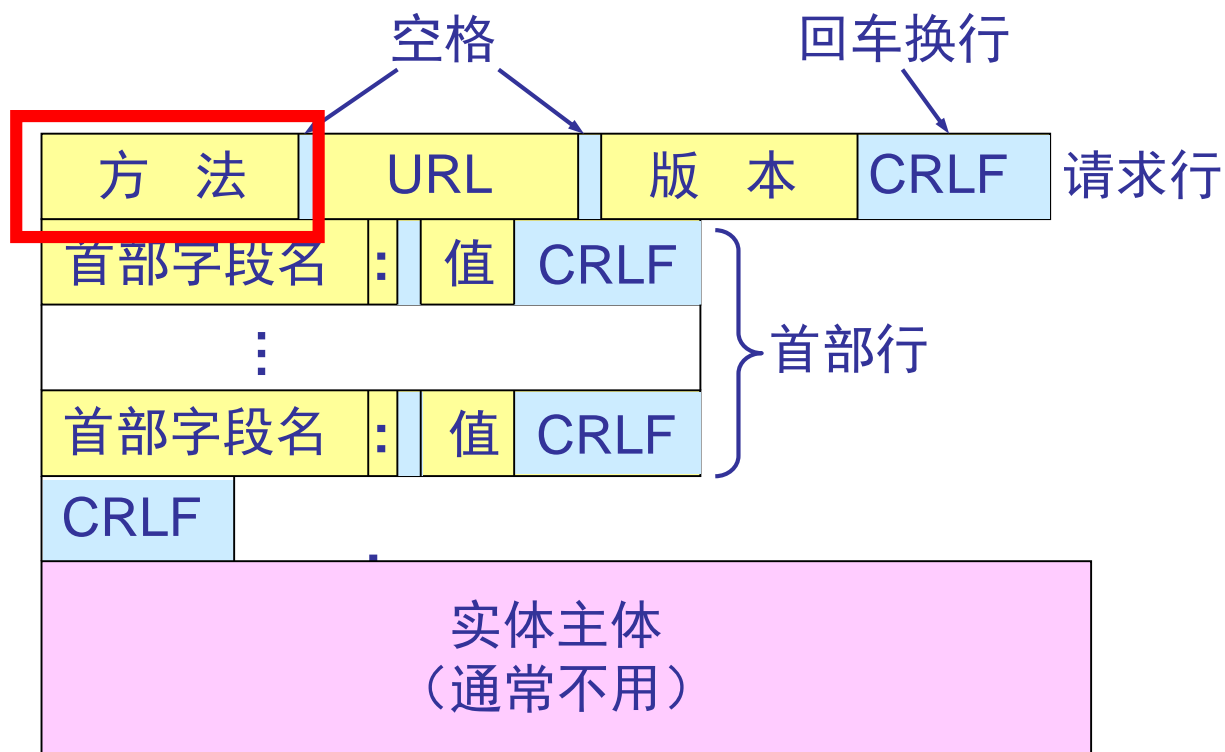


# HTTP 的报文结构（请求报文）



报文由三个部分组成，即开始行、首部行和实体主体。  
在请求报文中，开始行就是请求行。

# HTTP 的报文结构（请求报文）



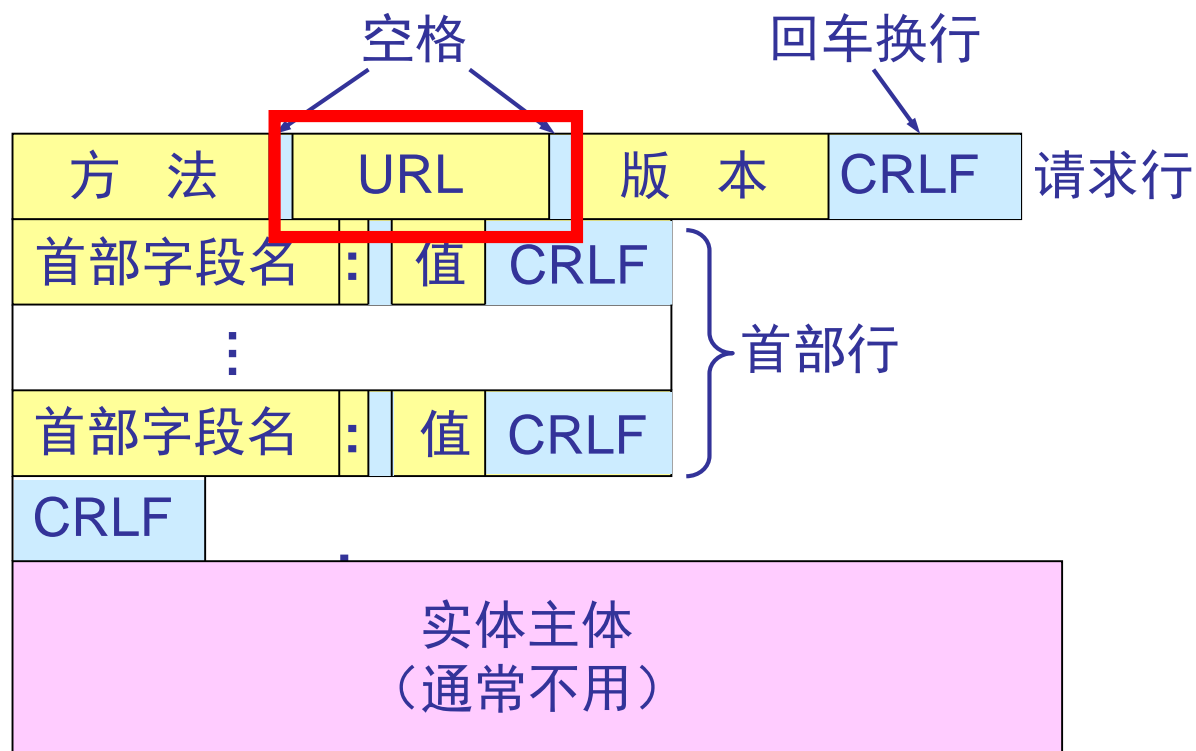
“**方法**”是面向对象技术中使用的专门名词。所谓“方法”就是**对所请求的对象进行的操作**，因此这些方法实际上也就是一些**命令**。因此，请求报文的类型是由它所采用的方法决定的。



# HTTP 请求报文的一些方法

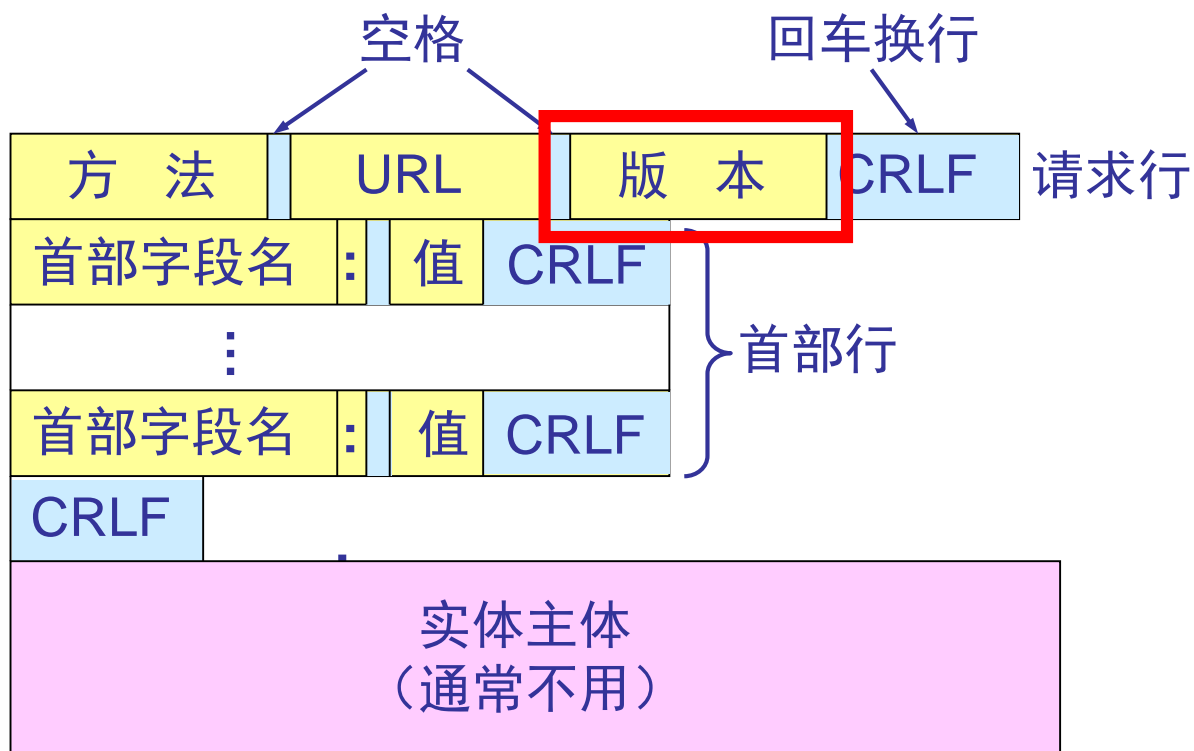
方法（操作）	意义
OPTION	请求一些选项的信息
GET	请求读取由 URL 所标志的信息
HEAD	请求读取由 URL 所标志的信息的首部
POST	给服务器添加信息（例如，注释）
PUT	在指明的 URL 下存储一个文档
DELETE	删除指明的 URL 所标志的资源
TRACE	用来进行环回测试的请求报文
CONNECT	用于代理服务器

# HTTP 的报文结构（请求报文）



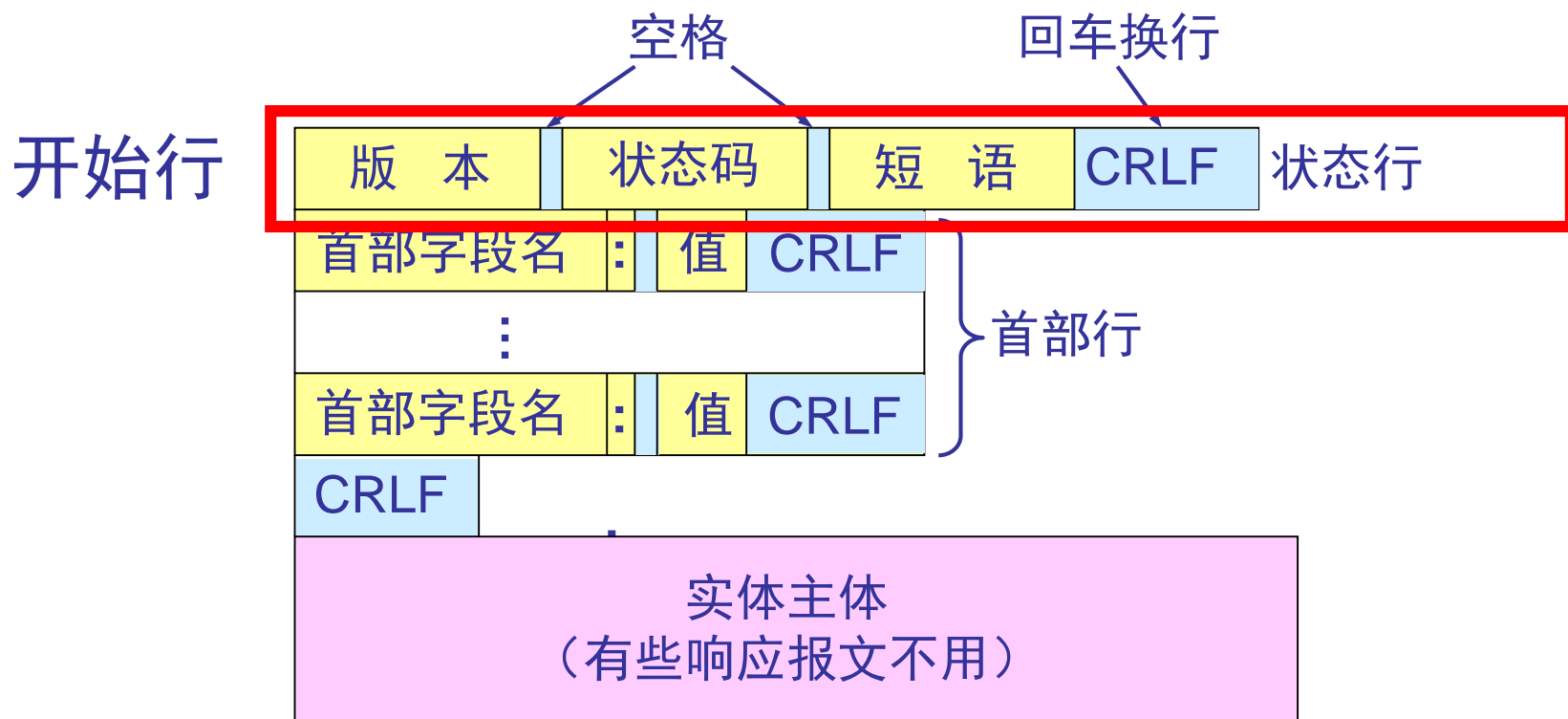
“URL”是所请求的资源的 URL。

# HTTP 的报文结构（请求报文）



“版本” 是 HTTP 的版本。

# HTTP 的报文结构（响应报文）



响应报文的开始行是**状态行**。

状态行包括三项内容，即 **HTTP** 的版本，**状态码**，以及解释状态码的**简单短语**。



## 4. 在服务器上存放用户的信息

---

- 万维网站点使用 Cookie 来跟踪用户。
- Cookie 表示在 HTTP 服务器和客户之间传递的状态信息。
- 使用 Cookie 的网站服务器为用户产生一个唯一的识别码。利用此识别码，网站就能够跟踪该用户在该网站的活动。

## 6.4.4 万维网的文档

### 1. 超文本标记语言 HTML

---

- 超文本标记语言 HTML (HyperText Markup Language) 是一种制作万维网页面的标准语言。
- HTML 文档是以 .html 或 .htm 为后缀。



# HTML 文档中标签的用法

<HTML>

<HEAD>

<TITLE>一个 HTML 的例子</TITLE>

</HEAD>

<BODY>

<H1>HTML 很容易掌握</H1>

<P>这是第一个段落。</P>

<P>这是第二个段落。</P>

</BODY>

</HTML>

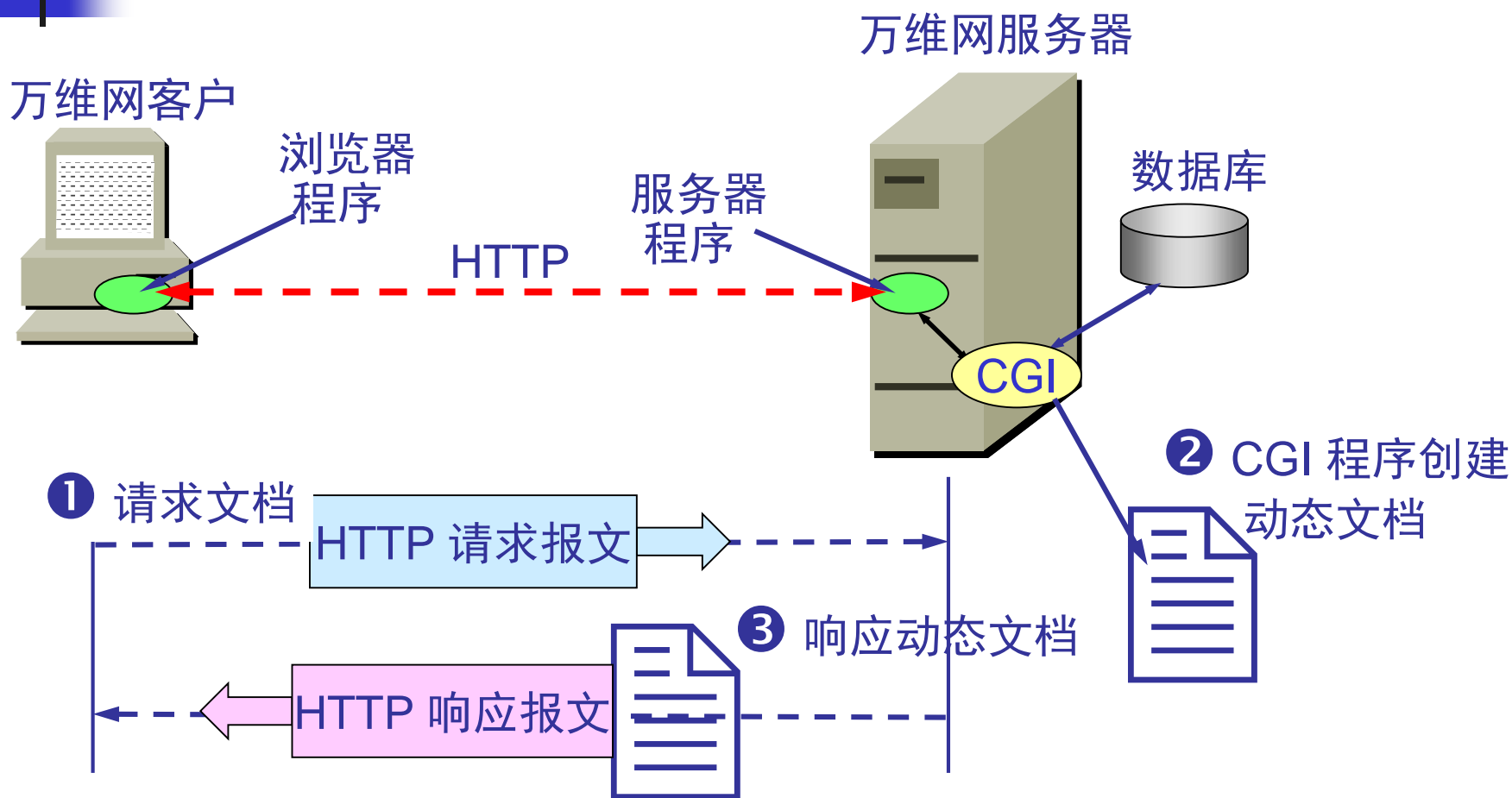


## 2. 动态万维网文档

---

- **静态文档**是指该文档创作完毕后就存放在万维网服务器中，在被用户浏览的过程中，内容不会改变。
- **动态文档**是指文档的内容是在浏览器访问万维网服务器时才由应用程序动态创建。

# 扩充了功能的万维网服务器



# 通用网关接口 CGI

(Common Gateway Interface)

- CGI 是一种标准，它定义了动态文档应如何创建，输入数据应如何提供给应用程序，以及输出结果应如何使用。
- CGI 程序的正式名字是 CGI 脚本 (script)。

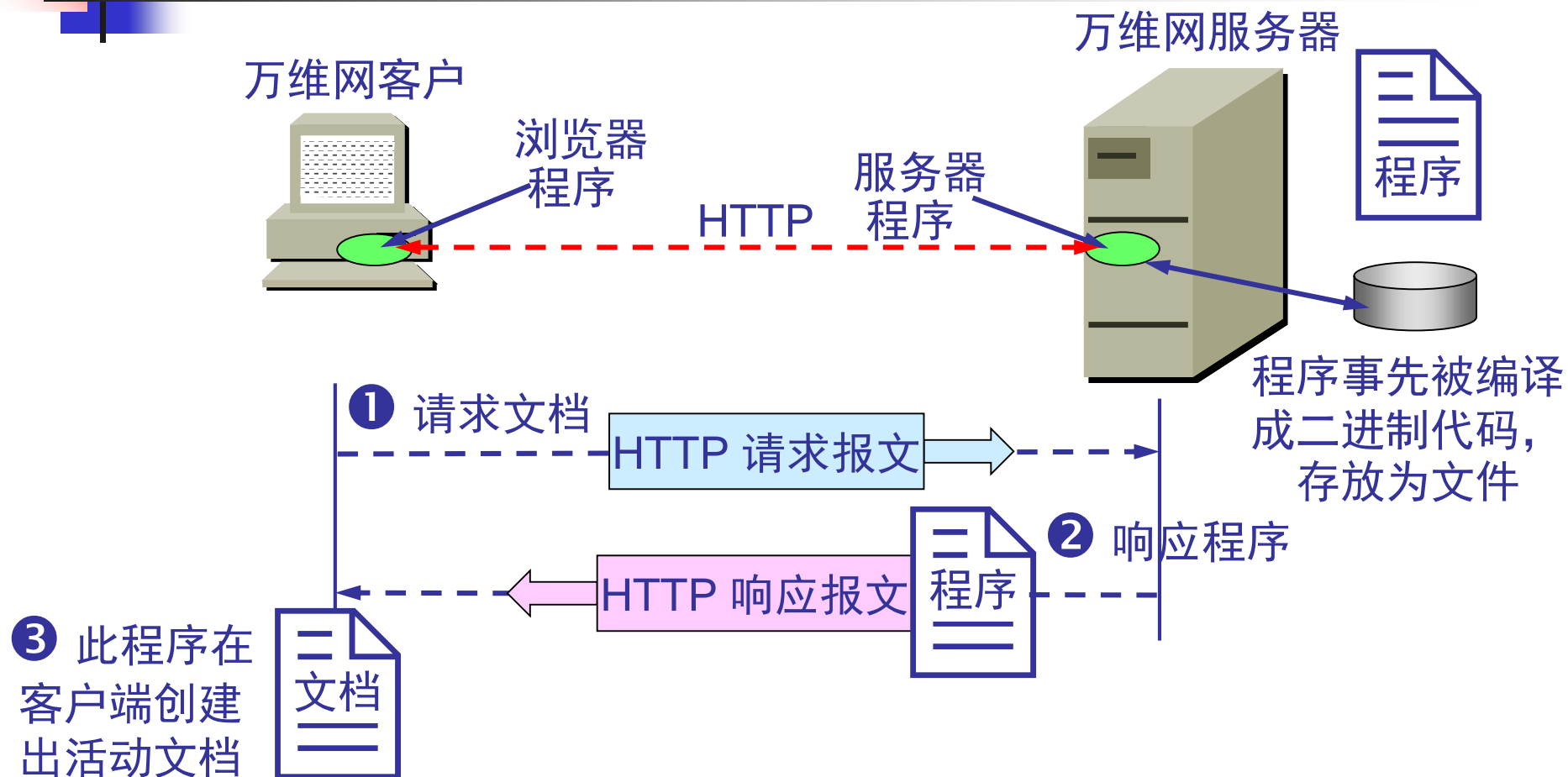


### 3. 活动万维网文档

---

- 动态文档的无法及时刷新屏幕以及像动画之类的显示效果。
- **活动文档**(active document)技术把所有的工作都转移给浏览器端。
- 每当浏览器请求一个活动文档时，服务器就返回一段程序副本在浏览器端运行。

# 活动文档在客户端创建





# 用 Java 技术创建活动文档

---

- 由美国 Sun 公司开发的 **Java** 语言是一项用于创建和运行活动文档的技术。
- 在 Java 技术中使用 “**小应用程序**” (applet) 来描述活动文档程序。

## 6.4.5 万维网的信息检索系统

### 1. 全文检索搜索和分类目录搜索

- 在万维网中用来进行搜索的程序叫做**搜索引擎**。
- **全文检索搜索引擎**是一种纯技术型的检索工具。





# 分类目录搜索

- **分类目录搜索引擎**并不采集网站的任何信息，而是利用各网站向搜索引擎提交的网站信息时填写的关键词和网站描述等信息，经过人工审核编辑后，如果认为符合网站登录的条件，则输入到分类目录的数据库中，供网上用户查询。



## 6.4.6 博客、微博

---

- 博客是万维网日志(web log)的简称。也有人把 blog 进行音译，译为“部落格”，或“部落阁”。还有人用“博文”来表示“博客文章”。
- 微博就是微型博客(microblog)，又称为微博客。

# 6.5 电子邮件

## 6.5.1 概述

---

- 电子邮件(e-mail)是因特网上使用得最多的和最受用户欢迎的一种应用。



# 电子邮件的一些标准

---

- 发送邮件的协议：SMTP 和 MIME
- 读取邮件的协议：POP3 和 IMAP
- MIME 在其邮件首部中说明了邮件的数据类型(如文本、声音、图像、视像等)，使用 MIME 可在邮件中同时传送多种类型的数据。

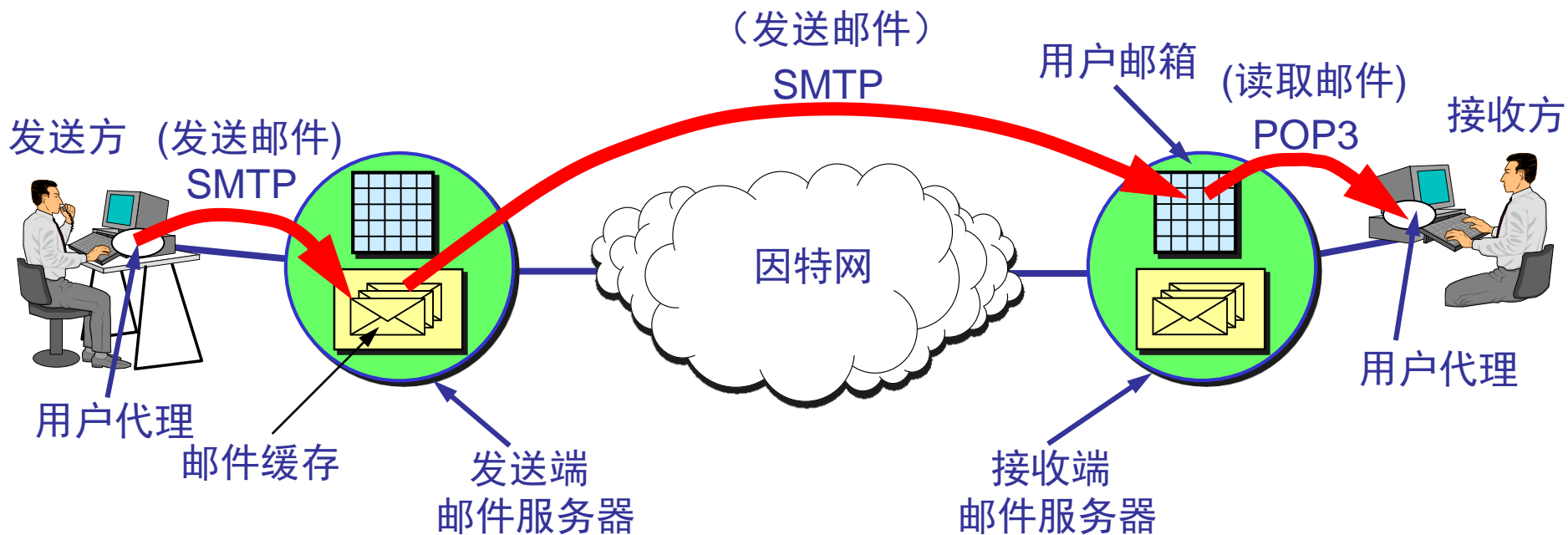


# 电子邮件的主要构件

---

- 用户代理 UA (User Agent)就是用户与电子邮件系统的接口，是电子邮件客户端软件。
- 邮件服务器的功能是发送和接收邮件，同时还要向发信人报告邮件传送的情况（已交付、被拒绝、丢失等）。

# 电子邮件的最主要的组成构件





## 6.5.2 简单邮件传送协议 SMTP

---

- SMTP 所规定的就是在两个相互通信的 SMTP 进程之间应如何交换信息。
- 由于 SMTP 使用客户服务器方式，因此负责发送邮件的 SMTP 进程就是 SMTP 客户，而负责接收邮件的 SMTP 进程就是 SMTP 服务器。



## 6.5.3 电子邮件的信息格式

---

- 一个电子邮件分为信封和内容两大部分。
- RFC 5322 只规定了邮件内容中的首部(header)格式，而对邮件的主体(body)部分则让用户自由撰写。





## 6.5.4 邮件读取协议

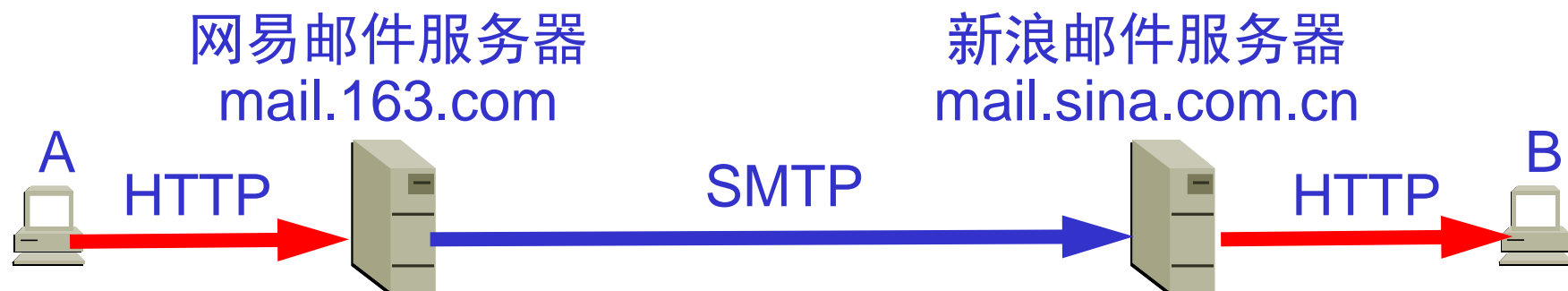
### POP3 和 IMAP

---

- 邮局协议 POP 是一个非常简单、但功能有限的邮件读取协议，现在使用的是它的第三个版本 POP3。

## 6.5.5 基于万维网的电子邮件

- 电子邮件从 A 发送到网易邮件服务器是使用 HTTP 协议。
- 两个邮件服务器之间的传送使用 SMTP。
- 邮件从新浪邮件服务器传送到 B 是使用 HTTP 协议。



## 6.5.6 通用因特网邮件扩充 MIME

### 1. MIME 概述

---

- SMTP 限于传送 7 位的 ASCII 码。
- MIME 并没有改动 SMTP 或取代它。
- MIME 的意图是继续使用目前的 [RFC 822] 格式，但增加了邮件主体的结构，并定义了传送非 ASCII 码的编码规则。

# MIME 和 SMTP 的关系





# 课后作业

---

- 习题6-02,6-03,6-08,6-12,6-16,6-20