1. **HTTP概述**

HTTP是因特网的多媒体信使，将图片、文本文件、HTML页面、MPEG电影、WAV音频文件、Java小程序和其他资源从世界各地的Web服务器上搬运到用户的Web浏览器中。

客户端向服务器发送HTTP请求，服务器会在HTTP响应中回送所请求的数据。HTTP客户端和服务器构成了万维网的基本组件。

**资源**

Web服务器中存放了各种Web资源；所有能够提供Web内容的东西都是Web资源；

**媒体类型**

因特网上有数千种不同的数据类型，HTTP给每种要通过Web传输的对象都打上了名为MIME类型的数据格式标签。

MIME类型是一种文本标记，表示一种主要的对象类型和一个特定的子类型，中间由一条斜线来分隔，如：

HTML格式的文本文档由text/html类型来标记

普通的ASCⅡ文本文档由text/plain类型来标记

JPEG格式的图片为image/jpeg类型

Gif格式的图片为image/gif类型

URI

服务器资源名被称为统一资源标识（URI）；给定了URI，HTTP就可以解析出对象；

URI有两种形式，分布成为URL和URN

URL

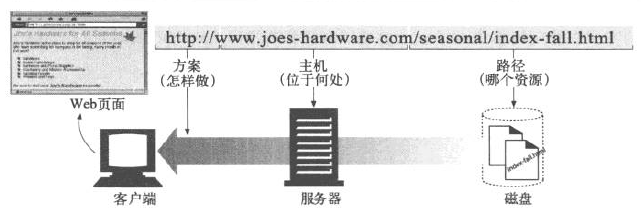
统一资源定位符（URL）是资源标识符最常见的形式。URL描述了一台特定服务器上某资源的特定位置。它们可以明确说明如何从一个精确、固定的位置获取资源。

大部分URL都遵循一种标准格式，这种格式包含最主要的三个部分：

第一部分为方案，规定如何访问指定资源，告诉复杂解析URL的应用程序应该使用什么协议；这部分通常就是http协议(http://)

第二部分给出了服务器的因特网地址即服务器的位置（如：[www.joes-hardware.com](http://www.joes-hardware.com)）

第三部分是资源路径，说明了请求的是服务器上哪个特定的资源（如：/specials/saw-blade.gif）

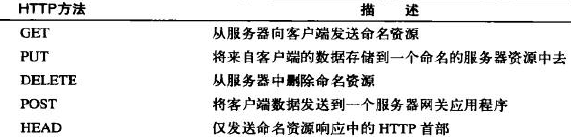


事务

一个HTTP事务由一条（从客户端发往服务器的）请求命令和一个（从服务器发回客户端）的响应结果组成。

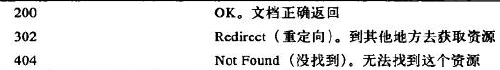
方法

每条HTTP请求都包含一个方法，常见的HTTP方法：



状态码

每条HTTP响应报文返回时都会携带一个状态码。状态码是一个三位数字的代码，告知客户端请求是否成功，或者是否需要采取其他行动。



一个web页面，通常不是单个资源，而是一组资源的组合

报文

从web客户端发往web服务器的HTTP报文称为请求报文；从服务器发往客户端的报文称为响应报文；

报文包含以下三个部分：

起始行

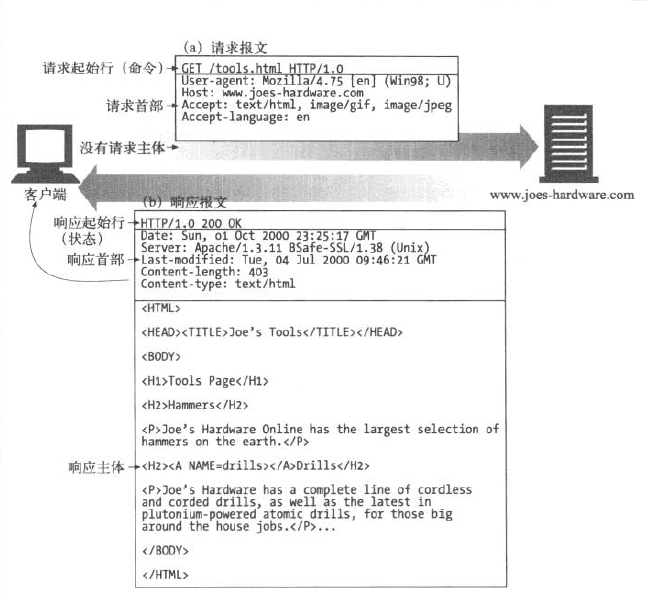
报文的第一行就是起始行，在请求报文中用来说明要做些什么，在响应报文中说明出现了什么情况

首部字段

起始行后面有零个或多个首部字段。每个首部字段都包含一个名字和一个值，为了便于解析，两者之间用冒号来分隔。首部以一个空行结束。

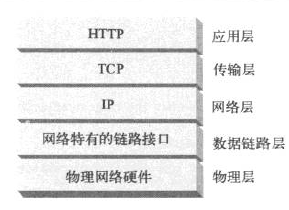
主体

空行之后，就是可选的报文主体，包含了所有类型的数据。请求主体中包含了要发送给web端服务器的数据；响应主体中装载了要返回给客户端的数据。起始行和首部都是文本形式其都是结构化的，而主体中可以包含任意的二进制数据（如图片、视频、音轨、软件程序），也可以包含文本。



连接

HTTP是应用层协议，报文通过传输控制协议TCP/IP来传输



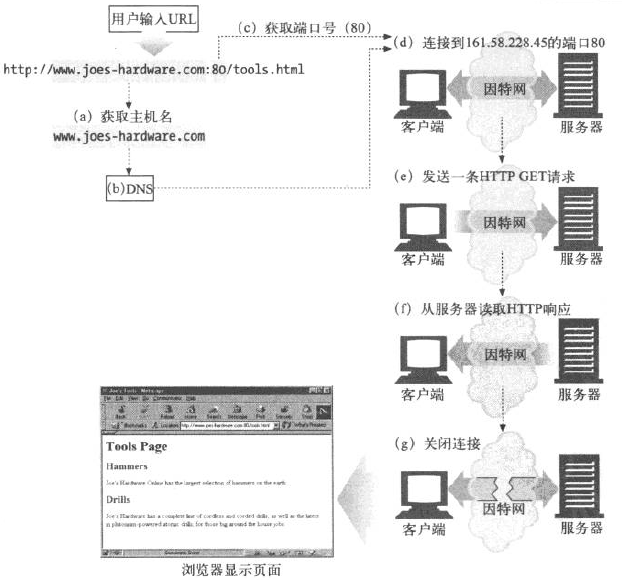
在HTTP客户端向服务器发送报文之前，需要用网际协议（IP）地址和端口号在客户端和服务器之间建立一条TCP/IP连接

对于使用IP地址和端口号的URL（如：http://192.168.10.42:8081/index.html），可以直接得到IP和端口号

对于使用域名或主机名的URL（如：http://www.netscape.com:80/index.html）,可以通过域名服务（DNS）机制将主机名转换为IP地址。

对于没有端口号的URL，默认端口号为80；

有了IP地址和端口号，客户端就可以通过TCP/IP进行通信



基本的浏览器连接通信过程：

1. 浏览器从URL中解析出服务器的主机名
2. 浏览器将服务器的主机名转换成服务器的IP地址
3. 浏览器将端口号（如果有的话）从URL中解析出来
4. 浏览器建立一条与web服务器的TCP连接
5. 浏览器向服务器发送一条HTTP请求报文
6. 服务器向浏览器回送一条HTTP响应报文
7. 关闭连接，浏览器显示文档