

HTTP 协议

一、超文本传输协议

Web 服务器和浏览器通过 HTTP 协议在 Internet 上发送和接收消息。HTTP 协议是一种请求-应答式的协议——客户端发送一个请求，服务器返回该请求的应答。HTTP 协议使用可靠的 TCP 连接，默认端口是 80。HTTP 的第一个版本是 HTTP/0.9，后来发展到了 HTTP/1.0，现在最新的版本是 HTTP/1.1。HTTP/1.1 由 [RFC 2616](#) 定义（pdf 格式）。

本文只简要介绍 HTTP 1.1 的相关知识，但应该足以让你理解 Web 服务器和浏览器发送的消息。如果你要了解更多的细节，请参考 RFC 2616。

在 HTTP 中，客户端/服务器之间的会话总是由客户端通过建立连接和发送 HTTP 请求的方式初始化，服务器不会主动联系客户端或要求与客户端建立连接。浏览器和服务器都可以随时中断连接，例如，在浏览网页时你可以随时点击“停止”按钮中断当前的文件下载过程，关闭与 Web 服务器的 HTTP 连接。

1.1 HTTP 请求

HTTP 请求由三个部分构成，分别是：方法-URI-协议/版本，请求头，请求正文。下面是一个 HTTP 请求的例子：

```
GET /servlet/default.jsp HTTP/1.1
Accept: text/plain; text/html
Accept-Language: en-gb
Connection: Keep-Alive
Host: localhost
Referer: http://localhost/ch8/SendDetails.htm
User-Agent: Mozilla/4.0 (compatible; MSIE 4.01; Windows 98)
Content-Length: 33
Content-Type: application/x-www-form-urlencoded
Accept-Encoding: gzip, deflate

userName=JavaJava&userID=javaID
```

请求的第一行是“方法-URI-协议/版本”，其中 GET 就是请求方法，/servlet/default.jsp 表示 URI，HTTP/1.1 是协议和协议的版本。根据 HTTP 标准，HTTP 请求可以使用多种请求方法。例如，HTTP 1.1 支持七种请求方法：GET, POST, HEAD, OPTIONS, PUT, DELETE, 和 TRACE。在 Internet 应用中，最常用的请求方法是 GET 和 POST。

URI 完整地指定了要访问的网络资源，通常认为它相对于服务器的根目录而言，因此总是以“/”开头。URL 实际上是 [URI](#) 一种类型。最后，协议版本声明了通信过程中使用的 HTTP 协议的版本。

请求头包含许多有关客户端环境和请求正文的有用信息。例如，请求头可以声明浏览器所用的语言，请求正文的长度，等等，它们之间用一个回车换行符号（CRLF）分隔。

请求头和请求正文之间是一个空行（只有 CRLF 符号的行），这个行非常重要，它表示请求头已经结束，接下来的是请求的正文。一些介绍 Internet 编程的书籍把这个 CRLF 视为 HTTP 请求的第四个组成部分。

在前面的 HTTP 请求中，请求的正文只有一行内容。当然，在实际应用中，HTTP 请求正文可以包含更多的内容。

1.2 HTTP 应答

和 HTTP 请求相似，HTTP 应答也由三个部分构成，分别是：协议-状态代码-描述，应答头，应答正文。下面是一个 HTTP 应答的例子：

```
HTTP/1.1 200 OK
Server: Microsoft-IIS/4.0
Date: Mon, 3 Jan 1998 13:13:33 GMT
Content-Type: text/html
Last-Modified: Mon, 11 Jan 1998 13:23:42 GMT
Content-Length: 112

<html>
<head>
<title>HTTP 应答示例</title></head><body>
Hello HTTP!
</body>
</html>
```

HTTP 应答的第一行类似于 HTTP 请求的第一行，它表示通信所用的协议是 HTTP 1.1，服务器已经成功地处理了客户端发出的请求（200 表示成功），一切顺利。

应答头也和请求头一样包含许多有用的信息，例如服务器类型、日期时间、内容类型和长度等。应答的正文就是服务器返回的 HTML 页面。应答头和正文之间也用 CRLF 分隔。