

GET 方法指导服务器生成和返回 URI(Uniform Resource Identifier, 统一资源标识符)标识的内容。URI 是相应的 URL 的后缀, 包括文件名和可选的参数。^①

请求行中的 version 字段表明了该请求遵循的 HTTP 版本。最新的 HTTP 版本是 HTTP/1.1 [37]。HTTP/1.0 是从 1996 年沿用至今的老版本 [6]。HTTP/1.1 定义了一些附加的报头, 为诸如缓冲和安全等高级特性提供支持, 它还支持一种机制, 允许客户端和服务器的同一条持久连接(persistent connection)上执行多个事务。在实际中, 两个版本是互相兼容的, 因为 HTTP/1.0 的客户端和服务器的简单地忽略 HTTP/1.1 的报头。

总的来说, 第 5 行的请求行要求服务器取出并返回 HTML 文件/index.html。它也告知服务器请求剩下的部分是 HTTP/1.1 格式的。

请求报头为服务器提供了额外的信息, 例如浏览器的商标名, 或者浏览器理解的 MIME 类型。请求报头的格式为

header-name: header-data

针对我们的目的, 唯一需要关注的报头是 Host 报头(第 6 行), 这个报头在 HTTP/1.1 请求中是需要的, 而在 HTTP/1.0 请求中是不需要的。代理缓存(proxy cache)会使用 Host 报头, 这个代理缓存有时作为浏览器和管理被请求文件的原始服务器(origin server)的中介。客户端和原始服务器之间, 可以有多个代理, 即所谓的代理链(proxy chain)。Host 报头中的数据指示了原始服务器的域名, 使得代理链中的代理能够判断它是否可以在本地缓存中拥有一个被请求内容的副本。

继续图 11-24 中的示例, 第 7 行的空文本行(通过在键盘上键入回车键生成的)终止了报头, 并指示服务器发送被请求的 HTML 文件。

2. HTTP 响应

HTTP 响应和 HTTP 请求是相似的。一个 HTTP 响应的组成是这样的: 一个响应行(response line)(第 8 行), 后面跟随着零个或更多的响应报头(response header)(第 9~13 行), 再跟随一个终止报头的空行(第 14 行), 再跟随一个响应主体(response body)(第 15~17 行)。一个响应行的格式是

version status-code status-message

version 字段描述的是响应所遵循的 HTTP 版本。状态码(status-code)是一个 3 位的正整数, 指明对请求的处理。状态消息(status message)给出与错误代码等价的英文描述。图 11-25 列出了一些常见的状态码, 以及它们相应的消息。

状态代码	状态消息	描述
200	成功	处理请求无误
301	永久移动	内容已移动到location头中指定的主机上
400	错误请求	服务器不能理解请求
403	禁止	服务器无权访问所请求的文件
404	未发现	服务器不能找到所请求的文件
501	未实现	服务器不支持请求的方法
505	HTTP 版本不支持	服务器不支持请求的版本

图 11-25 一些 HTTP 状态码

① 实际上, 只有当浏览器请求内容时, 这才是真的。如果代理服务器请求内容, 那么这个 URI 必须是完整的 URL。