

Web 代理 (proxy) 服务器是网络的中间实体。代理位于客户端和服务端之间, 扮演“中间人”的角色, 在各端点之间来回传送 HTTP 报文。本章介绍了所有与 HTTP 代理服务器有关的内容, 为代理特性提供的特殊支持, 以及使用代理服务器时会遇到的一些棘手的问题。

本章主要内容如下:

- 对 HTTP 代理进行解释, 将其与 Web 网关进行对比, 并说明如何部署代理;
- 给出一些代理所能提供的帮助;
- 说明在现实网络中是怎样部署代理以及如何将网络流量导向代理服务器;
- 说明如何配置浏览器来使用代理;
- 展示 HTTP 的代理请求, 说明它们与服务器请求的区别, 以及代理是如何微妙地改变浏览器行为的;
- 解释如何通过 via 首部和 TRACE 方法来记录报文传输路径上的代理服务器链;
- 描述基于代理的 HTTP 访问控制方法;
- 解释代理如何与客户端和服务端进行交互, 每个客户端和服务端支持的特性和使用的版本都可能有所不同。

6.1 Web 的中间实体

Web 上的代理服务器是代表客户端完成事务处理的中间人。如果没有 Web 代理, HTTP 客户端就要直接与 HTTP 服务器进行对话。有了 Web 代理, 客户端就可以与代理进行对话, 然后由代理代表客户端与服务器进行交流。客户端仍然会完成对事务的处理, 但它是通过代理服务器提供的优质服务来实现的。

HTTP 的代理服务器既是 Web 服务器又是 Web 客户端。HTTP 客户端会向代理发送请求报文, 代理服务器必须像 Web 服务器一样, 正确地处理请求和连接, 然后返回响应。同时, 代理自身要向服务器发送请求, 这样, 其行为就必须像正确的 HTTP 客户端一样, 要发送请求并接收响应 (参见图 6-1)。如果要创建自己的 HTTP 代理, 就要认真地遵循为 HTTP 客户端和 HTTP 服务器制定的规则。

6.1.1 私有和共享代理

代理服务器可以是某个客户端专用的, 也可以是很多客户端共享的。单个客户端专用的代理被称为私有代理。众多客户端共享的代理被称为公共代理。

- 公共代理

大多数代理都是公共的共享代理。集中式代理的费效比更高, 更容易管理。某些