

由于重定向和负载均衡是共存的，所以本章也涵盖了负载均衡的话题。大多数重定向部署都包含了某些形式的负载均衡。也就是说，它们可以将输入报文的负载分摊到一组服务器中去。反之，因为输入报文一定会在分担负荷的服务器之间进行某种分布，所以任意形式的负载均衡中都包含了重定向。

20.2 重定向到何地

从客户端向目标发送 HTTP 请求，目标对其进行处理的角度来看，服务器、代理、缓存和网关对客户端来说都是服务器。很多重定向技术都可用于服务器、代理、缓存和网关，因为它们具有共同的，与服务器类似的特征。其他一些重定向技术是专门为特定类型的端点设计的，没有通用性。本章稍后的小节会介绍一些通用及专用的重定向技术。

Web 服务器会根据每个 IP 来处理请求。将请求分摊到复制的服务器中去，就意味着应该把对某特定 URL 的每条请求都发送到最佳的 Web 服务器上去（最靠近客户端的、或负载最轻的或采用其他优化策略选择的服务器）。重定向到某台服务器就像将所有需要给汽车加油的司机都送到最近的加油站去一样。

代理希望根据每个协议来处理请求。在理想情况下，某个代理附近的所有 HTTP 流量都应该通过这个代理传输。比如，如果某代理缓存靠近各种不同的客户端，那么理想情况下，所有请求都应流经这个代理缓存，因为代理缓存上会存储常用的文档，可以直接提供，从而避免通过更长、更昂贵的路径连接到原始服务器。重定向到代理就像从一条主要通路（无论它通往何处）上将流量分流到一条本地快捷路径上去一样。

449

20.3 重定向协议概览

重定向的目标是尽快地将 HTTP 报文发送到可用的 Web 服务器上去。在穿过因特网的路径上，HTTP 报文传输的方向会受到 HTTP 应用程序和报文经由的路由设备的影响，参见以下示例。

- 配置创建客户端报文的浏览器应用程序，使其将报文发送给代理服务器。
- DNS 解析程序会选择用于报文寻址的 IP 地址。对不同物理地域的不同客户端来说，这个 IP 地址可能不同。
- 报文经过网络传输时，会被划分为一些带有地址的分组；交换机和路由器会检查分组中的 TCP/IP 地址，并据此来确定分组的发送路线。
- Web 服务器可以通过 HTTP 重定向将请求反弹给不同的 Web 服务器。