

存器,并且将所有校园网上的用户浏览器配置为指向它。或者,一个主要的住宅 ISP(例如 AOL)可能在它的网络上安装一台或多台 Web 缓存器,并且预先配置其配套的浏览器指向这些缓存器。

在因特网上部署 Web 缓存器有两个原因。首先,Web 缓存器可以大大减少对客户请求的响应时间,特别是当客户与初始服务器之间的瓶颈带宽远低于客户与 Web 缓存器之间的瓶颈带宽时更是如此。如果在客户与 Web 缓存器之间有一个高速连接(情况常常如此),并且如果用户所请求的对象在 Web 缓存器上,则 Web 缓存器可以迅速将该对象交付给用户。其次,如我们马上用例子说明的那样,Web 缓存器能够大大减少一个机构的接入链路到因特网的通信量。通过减少通信量,该机构(如一家公司或者一所大学)就不必急于增加带宽,因此降低了费用。此外,Web 缓存器能从整体上大大减低因特网上的 Web 流量,从而改善了所有应用的性能。

为了深入理解缓存器带来的好处,我们考虑在图 2-12 场景下的一个例子。该图显示了两个网络,即机构(内部)网络和公共因特网的一部分。机构网络是一个高速的局域网,它的一台路由器与因特网上的一台路由器通过一条 15Mbps 的链路连接。这些初始服务器与因特网相连但位于全世界各地。假设对象的平均长度为 1Mb,从机构内的浏览器对这些初始服务器的平均访问速率为每秒 15 个请求。假设 HTTP 请求报文小到可以忽略,因而不会在网络中以及接入链路(从机构内部路由器到因特网路由器)上产生什么通信量。我们还假设在图 2-12 中从因特网接入链路一侧的路由器转发 HTTP 请求报文(在一个 IP 数据报中)开始,到它收到其响应报文(通常在多个 IP 数据报中)为止的时间平均为 2 秒。我们非正式地将该持续时延称为“因特网时延”。

总的响应时间,即从浏览器请求一个对象到接收到该对象为止的时间,是局域网时延、接入时延(即两台路由器之间的时延)和因特网时延之和。我们来粗略地估算一下这个时延。局域网上的流量强度(参见 1.4.2 节)为

$$(15 \text{ 个请求/秒}) \times (1\text{Mb/请求}) / (100\text{Mbps}) = 0.15$$

然而接入链路上的流量强度(从因特网路由器到机构路由器)为

$$(15 \text{ 个请求/秒}) \times (1\text{Mb/请求}) / (15\text{Mbps}) = 1$$

局域网强度为 0.15 的通信量最多导致数十毫秒的时延,因此我们可以忽略局域网时延。然而,如在 1.4.2 节讨论的那样,如果流量强度接近 1(就像在图 2-12 中接入链路的情况那样),链路上的时延会变得非常大并且无限增长。因此,满足请求的平均响应时间将在分钟的量级上。显然,必须想办法来改进时间响应特性。

一个可能的解决办法就是增加接入链路的速率,如从 15Mbps 增加到 100Mbps。这可以将接入链路上的流量强度减少到 0.15,这样一来,两台路由器之间的链路时延也可以忽

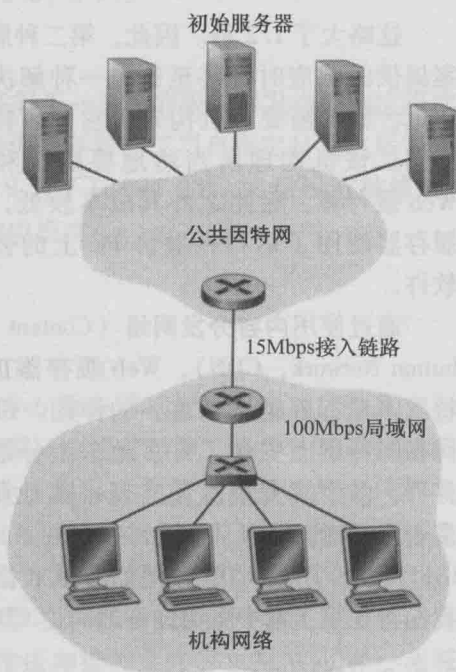


图 2-12 一个机构网络与因特网之间的瓶颈