



Figure 2.12 ♦ Bottleneck between an institutional network and the Internet

an institution's access link to the Internet. By reducing traffic, the institution (for example, a company or a university) does not have to upgrade bandwidth as quickly, thereby reducing costs. Furthermore, Web caches can substantially reduce Web traffic in the Internet as a whole, thereby improving performance for all applications.

To gain a deeper understanding of the benefits of caches, let's consider an example in the context of Figure 2.12. This figure shows two networks—the institutional network and the rest of the public Internet. The institutional network is a high-speed LAN. A router in the institutional network and a router in the Internet are connected by a 15 Mbps link. The origin servers are attached to the Internet but are located all over the globe. Suppose that the average object size is 1 Mbits and that the average request rate from the institution's browsers to the origin servers is 15 requests per second. Suppose that the HTTP request messages are negligibly small and thus create no traffic in the networks or in the access link (from institutional router to Internet router). Also suppose that the amount of time it takes from when the router on the Internet side of the access link in Figure 2.12 forwards an HTTP request (within an IP datagram) until it receives the response (typically within many IP datagrams) is two seconds on average. Informally, we refer to this last delay as the “Internet delay.”

The to
until its rec
the delay b
calculation

whereas th
router) is

A traffic in
onds of de
Section 1.4
in Figure 2
Thus, the a
minutes, if
thing must

One p
Mbps. This
lates to neg
time will r
means that
costly prop

Now c
instead ins
trated in F
cache—typ
suppose th
clients and
the request
cache. Nev
fied by the
through the
0.6. Typica
of milliseco
second Inte

which is ju
even lower
to upgrade