

INDEM DOCUMENTATION SERIES

👉 The Proxy Server 👈

K. Ashwin Kumar

June 28, 2006

❖ *Doorbell rings, I'm not list'nin',
From my mouth, drool is glist'nin',
I'm happy—although
My boss let me go—
Happily addicted to the Web.*

❖ *All night long, I sit clicking,
Unaware time is ticking,
There's beard on my cheek,
Same clothes for a week,
Happily addicted to the Web.*

1 Introduction

Old ideas aren't the trendiest, but they're often the best. 📖 *Proxy servers, this is a concept that sounds trendy and cutting-edge, but its roots are also in dusty old library science.* Remember in college when we first needed to check out a book housed in your university's locked stacks? Since we weren't allowed to go into this secured part of the library, a staff member acted as our proxy and retrieved the book for us.

All too often, of course, this process took longer than if we'd been able to go to the shelf and get the book yourself. But suppose that each time librarians retrieved a book for one student, they also made several copies, keeping them at the front desk for other users who requested the same title. The result would've been an ideal blend of fast service and airtight security.

This analogy explains the two main functions of a proxy server. First, the proxy server acts as an intermediary, helping users on a private network get information from the Internet when they need it, while ensuring that network security is maintained. Second, a proxy server may store frequently requested information in a local disk cache, rapidly delivering it to multiple users without having to go back to the Internet to get it.

2 The Layered Approach

A proxy server is a piece of software which serves different purposes. It can act as gateway to connect the local network to the internet and or a firewall to provide protection from outside.

Sometimes I start explaining about proxy server to my friends by saying that “proxy server provides protection from outside network”. This explanation will raise a question “If proxy server provides protection from outside, then what does firewall do?”. ☞ *It should be clear that normal packet-filtering firewall operates in Network layer of OSI model, while proxy works at Application layer.*

Packet-filtering firewalls have the advantage of speed, and they require no special configuration on the part of end-user applications. On the other hand, creating complex access rules can be difficult. Further, all packet filters can do is grant or deny access based on a packet’s apparent source or destination address. Hackers can fool such firewalls by forging source addresses via IP packet spoofing. Since client-server connections are direct, hackers can also use packet sniffers to discern a network’s address structure with relative ease.

In our library analogy, the equivalent of a packet-filtering firewall would be the librarian keeping a list of trusted students, then allowing only those individuals into the locked stacks to retrieve books. This might make book retrieval faster, but it would require that a list be created and maintained. It would also be vulnerable to impostors who turn up at the front desk bearing fake IDs.

Proxy servers are different. They break the direct link between client and server (or, if you will, between the student and the valuable book). They start by performing network address translation, mapping all of a network’s internal IP addresses to a single “safe” IP address. Since the latter is the only address the untrusted network is aware of, spoofing attacks are no longer possible.

Because they operate at the Application layer of the OSI model, proxy servers can do a lot more. Any given proxy server includes a collection of application-specific proxies: an HTTP proxy for Web pages; an FTP proxy; an SMTP/POP proxy for e-mail; a Network News Transfer Protocol (NNTP) proxy for news servers; a RealAudio/RealVideo proxy; and more. Each of these proxies accepts only packets generated by services it is designed to copy, forward, and filter.

Application-specific proxies are almost infinitely configurable. For example, they can be set to block access to certain Web servers at all times, let only certain users play RealAudio files, permit FTP downloads but not uploads, or keep users of any organization from logging on to their personal accounts until after 5 p.m. Proxy servers can also bar specific MIME types and, in conjunction with a third-party plug-in such as SurfWatch, even filter content.

Proxy servers also do a superior job of logging network traffic, and can ensure that connectivity is always available for certain traffic types. For example, a small office might be connected to the Internet at all times for Web browsing via a single dial-up connection; a proxy server could automatically bring up a second dial-up connection when a user starts a long download via FTP.

As usual, though, the flip side of extensive configurability is complexity. Client applications such as Web browsers and RealAudio players must often be reconfigured to be made aware of proxy servers. In addition, as new Internet services become available and use new protocols and ports, new proxies must be written to support them. The process of adding users and defining permissions can also be complicated, though some proxy servers ease this task by working with Lightweight Directory Access Protocol (LDAP) information.

3 Application Level and Circuit Level

Proxy servers are available for common Internet services; for example, an HTTP proxy is used for Web access; an FTP proxy is used for file transfers. Such proxies are called "application-level" proxies or "application-level gateways," because they are dedicated to a particular application and protocol and are aware of the content of the packets being sent. A generic proxy, called a "circuit-level" proxy, supports multiple applications. For example, SOCKS is a generic IP-based proxy server that supports TCP and UDP applications.

4 Other Proxies

Without being called a proxy specifically, the Internet's e-mail system (SMTP) is an example of a proxy server because it stores and forwards messages. E-mail is not sent directly from client to client without going through the mail server. Likewise, the Internet's Usenet news system (NNTP) forwards messages to neighboring servers.

5 Proxy has "CACHE" W

A proxy server's basic caching function works much like what's built into a Web browser, with the exception that the contents of the proxy server cache are available to multiple users. Whenever one user on the local network retrieves pages from the Internet, the pages are stored locally, which dramatically speeds access (see Figure 1). For example, Novell claims that when its BorderManager FastCache is configured to run from RAM, it is capable of processing more than 5,000 hits per second.

Some proxy servers offer read-ahead caching, which is capable of loading images and other objects embedded on a Web page into a cache before a Web browser has requested them. Caches may also be preloaded via a mechanism known as the last-modified multiplier. With the last-modified multiplier, a proxy server examines the creation dates of frequently requested pages, learning when updates are likely to occur and retrieving the pages when appropriate. And of course proxy servers also let administrators schedule batch retrieval of Web pages during any time of day when network traffic is known to be light.

Reverse caching is an additional feature of some proxy servers. In reverse caching, the cache server not only stores pages from the Internet for the benefit of local users, but it also stores local pages for the benefit of Internet users.

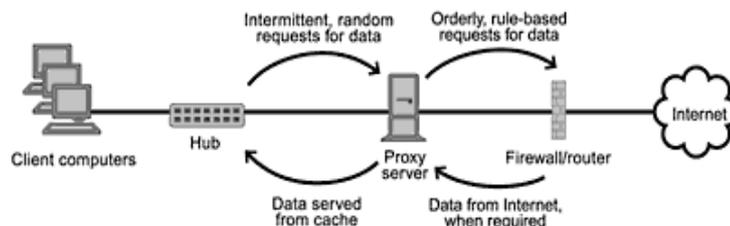


Figure 1: Proxy servers offer many features, but they are most commonly associated with caching. Caching gets the most out of any Internet connection by converting random, intermittent HTTP requests into an efficient, rule-based stream.

6 Cache as a Part of Proxy

Cache servers used to be viewed as nice-to-have items you got for free when you purchased a proxy server. Now that the Internet is growing steadily more congested and more and more clients have broadband connections, the terms "cache server" and "proxy server" may not be used quite so interchangeably. Proxy servers will continue to offer caching as one of their features. However, the increasing demand for specialized caching means that cache servers will gain more visibility as separate products. For example, the CacheQube from Cobalt Networks (Mountain View, CA) is an appliance that can simply be connected between a LAN and a router to provide transparent caching. The Streaming Media Cache from Inktomi (San Mateo, CA) and MediaMall from InfoLibria (Waltham, MA) are caches designed specifically for handling streaming audio and video.

What Is Caching?



Caching stores content objects from a Web server closer to the user, where they can be retrieved more quickly



- **Forward Proxy Cache**
 - Cache server acts as a proxy for client requests



- **Reverse Proxy Cache**
 - Cache server acts as a proxy for the origin server