upguard.com

Reverse Proxy Vs. Load Balancer | UpGuard

Kaushik SenChief Marketing Officer

10-12 minutes

A reverse proxy server (or reverse proxy) facilitates a user's requests to a web server/application server and the server's response.

A load balancer receives user requests, distributes them accordingly among a group of servers, then forwards each server response to its respective user.

From the brief definitions above, it's clear that reverse proxies and load balancers have some overlapping functionalities.

For example, they both act as intermediary nodes that forward requests and responses in a client-server model.

While they appear similar at face value, they are two distinct pieces of architecture with varied roles in network security.

This article unpacks the differences between reverse proxies and load balancers in detail to clear up any confusion between the pair.

What is a Reverse Proxy Server?

A reverse proxy server is an intermediary device or application between a user and a web server. Reverse proxies are type of proxy server designed to enhance web servers' security by ensuring that users never communicate directly with the origin

server.

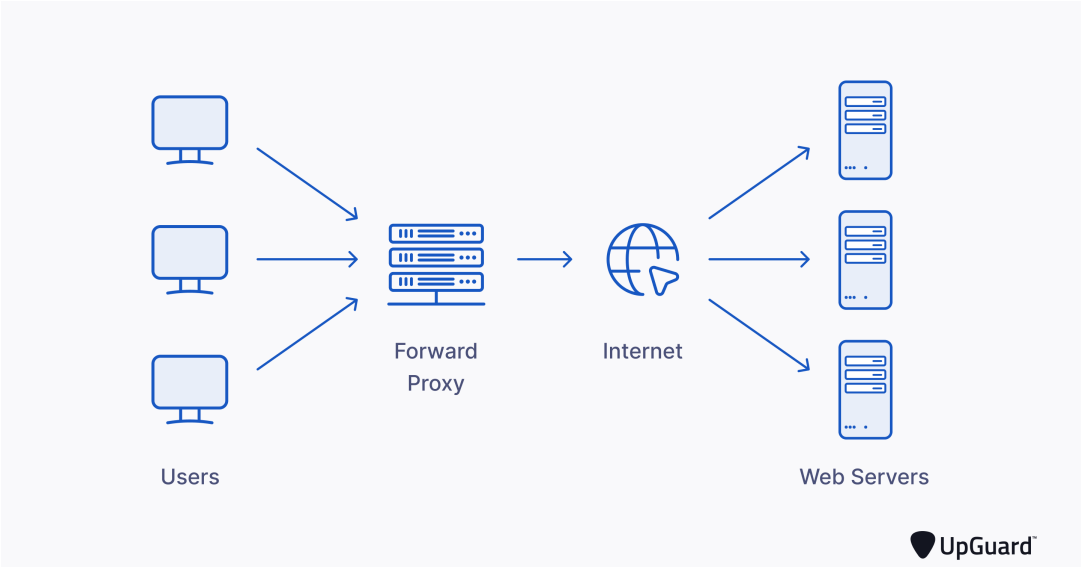
How Does a Reverse Proxy Work?

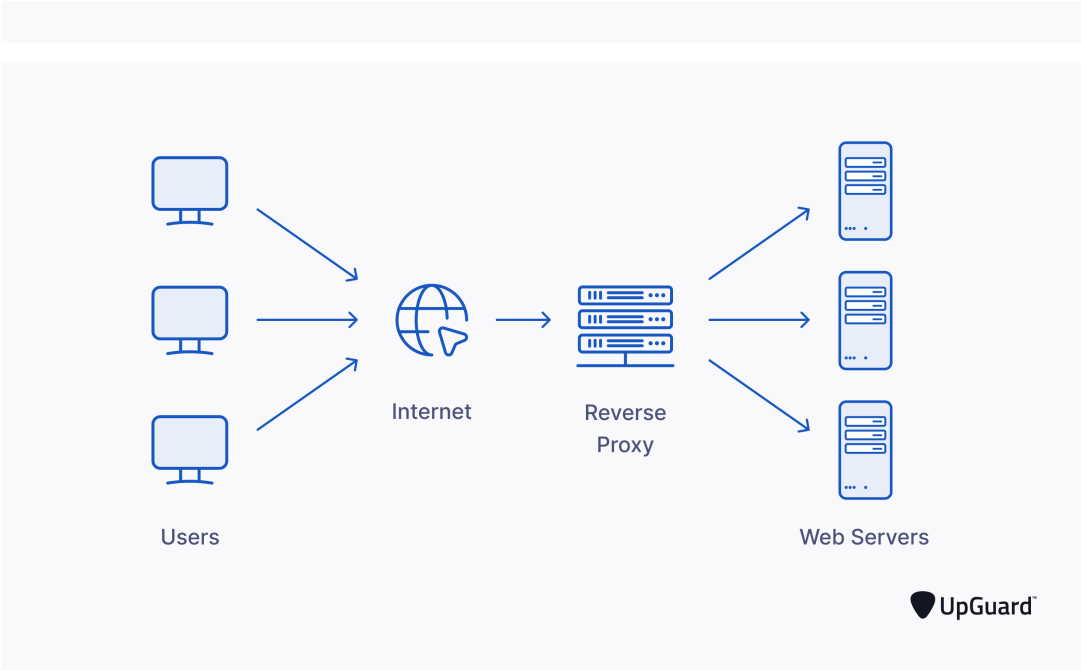
A typical reverse proxy operates as follows:

1. A user makes an HTTP request (via a firewall), e.g. enters a website's URL into their web browser.
2. The reverse proxy receives the user's request.
3. The reverse proxy either ‘allows’ or ‘denies’ the user’s request.
4. If allowed, the reverse proxy forwards the request to the web server. If denied, it sends the user an error or redirect message.
5. The web server sends a response (website data) back to the reverse proxy.
6. The reverse proxy forwards the server’s response to the user.

Reverse proxies receive their name from the way they are configured – in ‘reverse’ to a standard proxy server (or forward proxy). Forward proxies sit between a user and the Internet, whereas reverse proxies sit between the network’s edge and the Internet.

The opposing configurations of forward and reverse proxies are depicted below.





Types of Reverse Proxies

Reverse proxies are available in both hardware and software forms.

There are many popular open-source reverse proxy software solutions, such as:

- Apache HTTP Server
- Apache Traffic Server
- HAProxy
- NGINX
- Pound
- Træfɪk
- Varnish HTTP Cache

Web Application Firewalls (WAFs) are a type of reverse proxy, commonly deployed in commercial use cases. A WAF monitors web traffic and protects an organization's web applications against cyber attacks, such as SQL injections, sensitive data theft, cross-

site scripting, and other vulnerabilities.

Another type of reverse proxy is a Layer 7 load balancer. Layer 7 load balancers distribute web requests across multiple servers to improve client-side network performance and user experience.

Reverse Proxy Benefits

While a reverse proxy's main objective is to protect server-side operations, its functionality provides benefits to both the client and server.

Enhanced Security

- **Threat Protection:** Due to their placement at the network's edge, reverse proxies prevent malicious clients from directly accessing and exploiting any vulnerabilities within an organization's internal network.
- **Privacy:** Reverse proxies conceal the origin server's IP address for added data security, preventing Internet Service Providers (ISPs), web services, and data centers from monitoring their traffic and potentially causing data breaches.
- **Filtering:** Their ability to blacklist certain client IP addresses and limit the number of connections clients can initiate helps prevent DDoS attacks.

Load Balancing

Organizations can use a reverse proxy to distribute traffic evenly and efficiently across multiple backend servers.

Load balancing helps prevent site shutdowns as traffic can be rerouted to an alternative server (instead of relying on one server) in the event of a network outage or DDoS attack.

Web Acceleration

Reverse proxies enable faster web server response times, improving site load times and user experience.

They use varied web acceleration techniques to achieve greater speeds.

- **Caching:** Each time a web server sends a response, reverse proxy stores a local copy of it. Next time the user makes the same request, the proxy can respond directly to the user instead of forwarding the request to the web server, decreasing the response time and also alleviating the load on the server.
- **Compression:** Reverse proxies use compression algorithms to reduce the bandwidth required to send the server's response, increasing traffic speed.
- **SSL/TLS Offloading** – Reverse proxies can perform SSL encryption and authentication on all incoming requests and responses.

  As proxy acts as the endpoint to the SSL connection instead of the web server, the web server can serve content much faster.

What is a Load Balancer?

A load balancer acts as an intermediary between a user and a group of servers. Load balancers are used to alleviate strain on high-traffic servers. They route client requests to the most appropriate server, maximizing network speed and efficiency. By ensuring operational efficiency, load balancers help organizations establish a scalable foundation for their IT infrastructure.

How Does a Load Balancer Work?

Generally, a load balancer operates as follows:

1. A user makes a request, e.g. an HTTP request – entering a website's URL into their web browser.

2. The load balancer receives the user's request.
3. The load balancer sends the request to a single server in a group of different servers.
4. The selected server sends a response (website data) back to the load balancer.
5. The load balancer forwards the server’s response to the user.

   While the above example explains how a load balancer handles web requests, load balancers can support many other protocols, depending on their type.

   The way the load balancer selects which server to forward a user’s request to depends on which algorithm it uses.

   Load Balancing Algorithms

   A load balancing algorithm is a set of rules that determines which server is selected within a group of different servers.

   There are several types of load balancing algorithms.

- **Hash:** Selects a server based on a predefined key, e.g. the client’s IP address.
- **Least Connections:** Sends requests to the server dealing with the fewest existing client connections.
- **Least Response Time:** Uses a combined formula of fastest response times and fewest active connections.
- **Power of Two Choices:** Randomly chooses two servers then selects the server with the fewest active connections.
- **Round Robin:** Distributes requests across servers in a sequential manner.s.

Types of Load Balancers

Load balancers are categorized by which digital system they use and the specific layer of the Open Systems Interconnection (OSI) Model through which they operate.

Load Balancers by Digital System

Hardware Load Balancer Device

Hardware load balancer devices are often found in data centers. They are physical devices that usually operate on the Transport Layer (L4) or Application Layer (Layer 7).

Software Load Balancer (SLB)

Software load balancers are purchasable as load balancer as a service (LBaaS), e.g., as a feature of an application delivery controller (ADC), or can be installed directly onto a server.

Load Balancers by OSI Layer

Application Load Balancer (Level 7 Load Balancer or Reverse Proxy)

An application load balancer operates on Layer 7 of the OSI model – the highest layer.

It distributes web requests based on application-level variables, such as URLs, HTTP headers, and SSL.

A Layer 7 load balancer is a reverse proxy as it handles requests on the application level – the layer through which HTTP operates.

Gateway Load Balancer

A gateway load balancer operates on Layer 3 (L3). As all traffic flows through a single entry and exit point, enabling easy scalability.

Global Server Load Balancer

A global server load balancer can connect to servers all over the world. It responds to user requests from the server geographically closest to the requesting user.

Network Load Balancer (Level 4 Load Balancer)

A network load balancer operates on layer 4 (L4).

It distributes traffic based on network variables, including IP protocol, source IP, source port, destination IP, and destination port.

Load Balancer Benefits

A load balancer's ability to distribute user requests efficiently across multiple high-traffic servers provides many mutual benefits to the client-server model.

Enhanced User Experience

Load balancers perform health checks to identify server outages and then reroute user traffic to a functioning server.

They conduct health checks by intercepting error response messages to the user or by sending direct requests to the server which require a specific response to signal the server is healthy.

The load balancers' intervention in the event of a server error means that users experience far fewer error messages and avoid response lag.

Greater Reliability

Load balancers are implemented across multi-server deployments, ensuring requests are distributed evenly and efficiently.

By preventing server overload and traffic bottlenecks, load balancers provide greater reliability to users.

The availability of many servers instead of a single server ensures user requests are still fulfilled, even in the event of an outage.

Session Persistence

As HTTP/S is a stateless protocol, session persistence is not intuitive but is necessary for many applications to perform effectively.

For example, web applications, such as e-commerce sites, rely on session persistence to keep shopping baskets active.

Load balancers ensure a specific user's (i.e. from a particular IP address) requests are always sent to the same server during a session.

What's the Difference Between a Reverse Proxy and a Load Balancer?

Reverse proxies and load balancers both enhance the performance of application delivery networks, but the roles they play in this optimization aren't quite the same.

A reverse proxy is specifically a Level 7 load balancer, dealing exclusively with web requests. A load balancer can operate on Levels 3-7 of the OSI model, handling numerous types of requests on top of web requests, e.g., DNS, SSL, TCP.

A reverse proxy can perform additional roles to that of a load balancer. For example, a reverse proxy can also:

- Operate as a WAF
- Perform web acceleration, e.g. caching, TLS/SSL offloading, compression

- Provide cybersecurity mechanisms, e.g. threat protection, IP concealment, web filtering

  A load balancer’s main role is to distribute user requests across multiple servers. A reverse proxy can be used to facilitate requests between users and a single server.