The ABCs of ADCs

The Basics of Server Load Balancing and the Evolution to Application Delivery Controllers

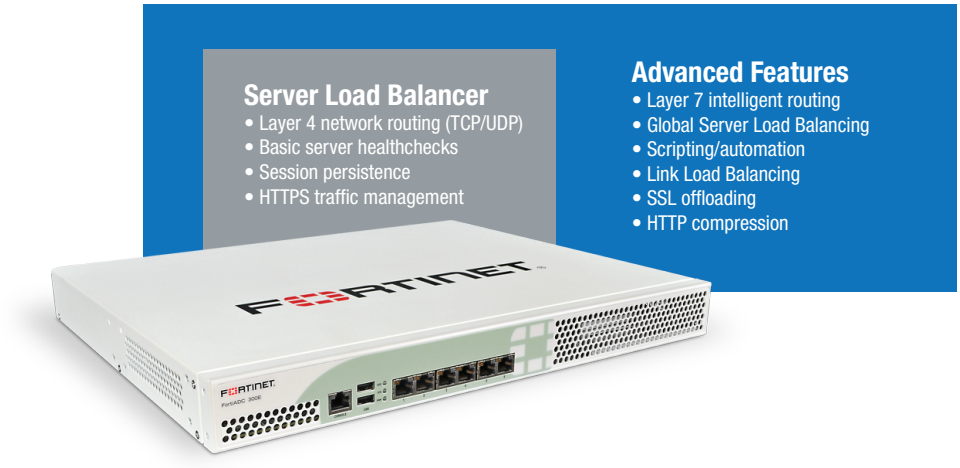
Introduction

Whether you need to expand an application from one server to two or need to deliver an application to millions of users across the globe, you're going to need an application delivery controller (ADC). ADCs provide basic application scalability, availability and reliability of earlier server load balancers and include advanced features for today's dynamic, content-rich applications like hardware-based secure traffic acceleration, HTTP compression and virtual environment integration.

Every ADC is a server load balancer first with advanced features layered on top of that core. So what is a server load balancer?

Application Delivery Controller

At its core, every ADC is first and foremost a server load balancer. ADCs build on this with advanced features that support today's complex application environments.



Business Challenges

- Application availability
- Application scalability
- Application performance
- Business continuity
- Data center cost reduction

Segments

- Small Business
- Medium Business
- Enterprise
- Data center
- MSP

The Basics of Server Load Balancing

As websites began to see increased traffic in the mid-1990s, single servers were reaching their limits to handle the capacity. Additional servers were required to expand applications along with technologies to make it appear to end users that they were accessing a single server.

The first method to address this scalability was DNS resolution, also referred to as "Round-robin DNS". This method assigns a group of unique internal IP addresses to servers behind a firewall to a single DNS name. When a user requested a resolution to a website name the DNS would respond back with multiple addresses in order, for example 10.1.0.10, 10.1.0.11 and 10.1.0.12. The next request made to the DNS would be supplied the same addresses, however they would be rotated so the second server would be first (10.1.0.11, 10.1.0.12 and 10.1.0.10). The DNS would continue to rotate through the servers for each sequential response.

Round Robin DNS was a simple solution that solved the issue of scalability by offering an almost limitless number of servers to be added to a DNS name. However without the capability to know the status of the server on the receiving end of the request, users could be sent to a server that was down or overloaded.

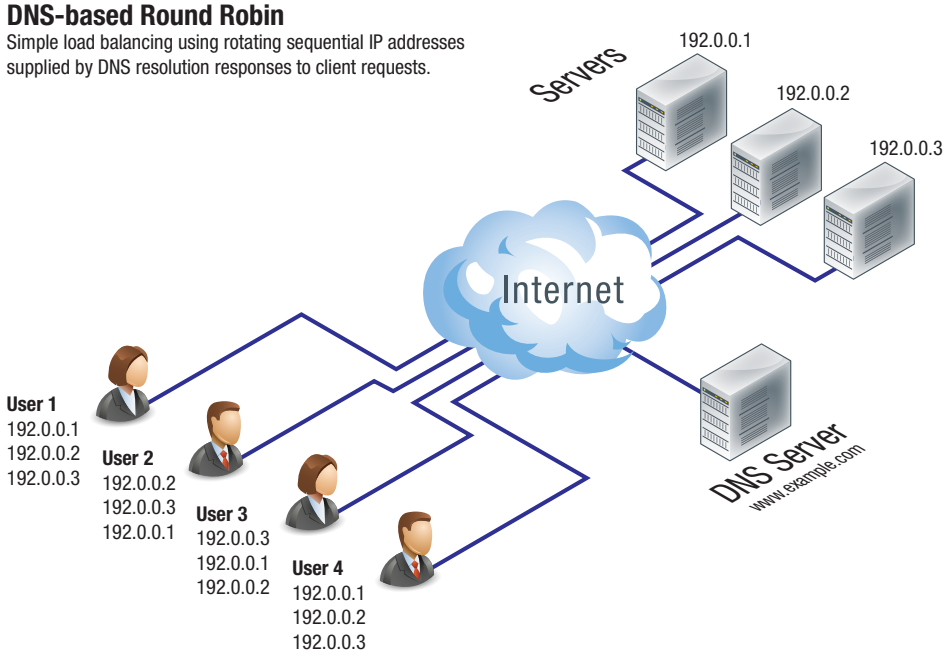
Soon many software-based approaches for load balancing became available to address the issue of server availability, usually as part of an operating system or application software. These systems created clusters of servers that were constantly in contact with one another to share information on server status, connections and other means to provide forms of server health-checking. Connection requests would be handed to the first available server to then be routed to the best available one (either itself or another server in the cluster). This worked well for smaller applications with less than 10 servers. Larger applications saw dramatic performance decreases with each new server due to the continuous need for servers to stay in contact with each other. This limited capacity combined with proprietary software led to the need for a new solution that could reliably scale and support multiple applications.

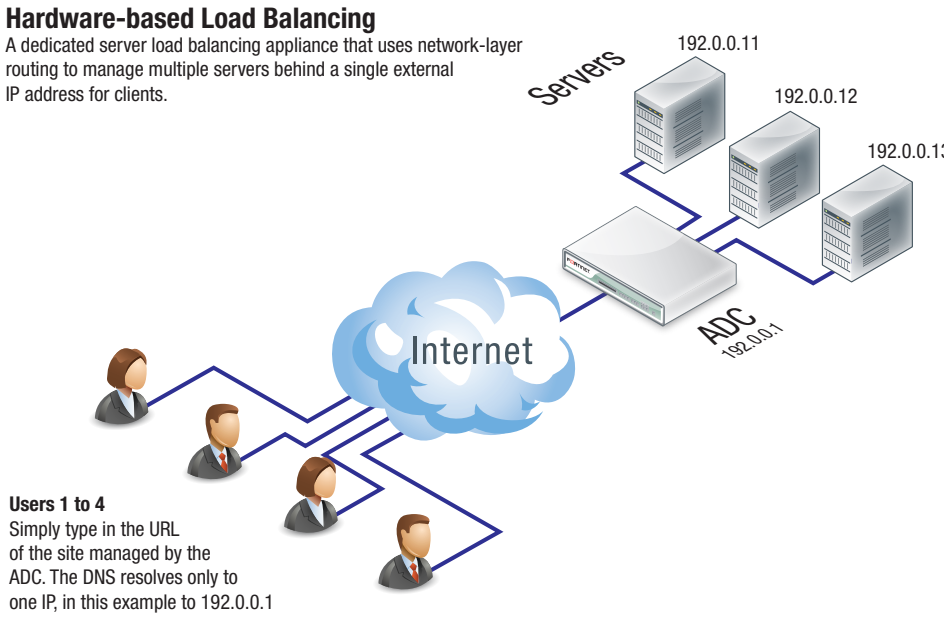
The Hardware-based Load Balancer

Beginning in the late 1990s, manufacturers introduced the first hardware-based load balancing appliances. By separating load balancing from the applications themselves, the appliances could rely on using network layer techniques like network address translation (NAT) to route inbound and outbound traffic to servers. Another key component that was introduced was server health-checking. At predefined intervals, the load balancer would check on the status of the server to determine if it was available and what its traffic load was. If a server was down, traffic would be directed to operational servers. If a server was overloaded, traffic would be redirected until it was back below set thresholds to receive new requests.

DNS-based Round Robin

Simple load balancing using rotating sequential IP addresses supplied by DNS resolution responses to client requests.





Applications could now scale and users would have reliable connections. The only limiting factor was the capacity of the hardware itself. In most cases, organizations that migrated from DNS-based or software load balancing saw an average 25% increase in server performance, reducing the need to add new servers to add more capacity.

The Application Delivery Controller

Simple load balancing is no longer sufficient to meet the basic needs of most organizations. Today web servers aren't just delivering static content, they're delivering dynamic, content-rich applications. Businesses are using web based applications to deliver mission critical functionality to employees and customers.

Over the past 10 years load balancers have evolved into Application Delivery Controllers (ADCs). These new devices understand application specific traffic and can optimize application server performance by offloading many of the compute-intensive tasks that would otherwise bog down CPUs that could be better occupied elsewhere. A common comparative analogy used to describe the role of SLBs is to compare them to a "network traffic cop". We'll use this analogy to describe the incremental advantages of an ADC over a server load balancer.

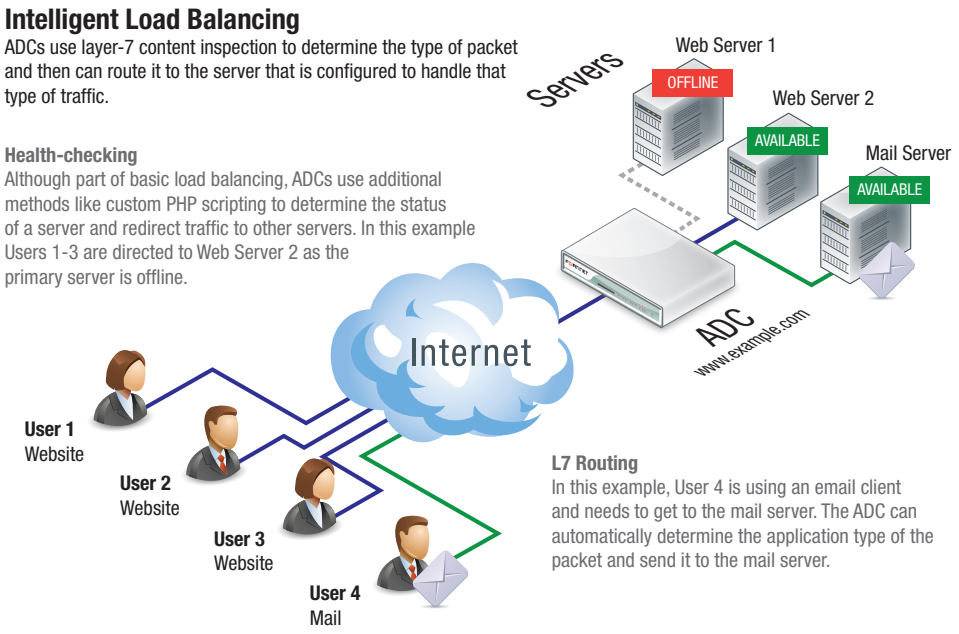
> Hardware-based load balancers with network-level traffic management were the forerunners of modern application delivery controllers

Intelligent Load Balancing

When a car is disabled on an interstate highway, a traffic cop will direct cars around the disabled lane. Similarly, an SLB can direct network traffic away from a slow or disabled server. But, the highway, much like the data center, is only a means to the end. What's really important to you is the destination (or, the "application", in data center terms). And every destination is unique, each with its own priority and value to the data center operators and the users accessing applications.

For example, you may take a different route to get to your office than you do to your grocery store. And getting to the office in a timely manner probably has a higher priority. When you get into your car, you want to get to your destination as expediently as possible. What we need today is a traffic cop who cannot only clean up the congestion after it happens, but can actually prevent the traffic jam from occurring in the first place. That's the role of the application delivery controller. In addition to load balancing traffic, what distinguishes ADCs from server load balancers is their ability to intelligently route users to their application and content destinations efficiently and intelligently, based on business priorities and goals.

Referring to the analogy above, imagine the ADC is the ultimate traffic cop; one who would not only redirect you



around the disabled lane, but would know where you were going, take into consideration the time of day, and know where the location is within the surrounding city. With that information, he would give you directions that would take you directly to your destination, bypassing stoplights, construction and any delays along the way.

Applying this analogy to users requesting applications and content from a data center, an advanced ADC will route users to destination servers based on a variety of criteria that the data center manager implements using policies and advanced application-layer knowledge to support business requirements. And, much like our example traffic officer, an advanced ADC will ensure that the users get to the applications based on their specific needs while protecting the network and applications from security threats.

> Intelligent load balancing provides administrators the capability to create rules that route traffic based on business rules and network traffic conditions

Advanced Features of an ADC

Among the advanced acceleration functions present in modern ADCs are SSL offloading technology, data compression, TCP and HTTP protocol optimization and virtualization awareness.

Much in the same way that a highway commuter lane has fewer cars with higher occupancy to reduce congestion, advanced ADCs offload servers by reducing the bandwidth utilization required to deliver application data from the data center to the desktop. ADCs offer compression to remove non-essential data from traversing network links. This helps to deliver maximum bandwidth utilization to support more traffic and avoids the need for network upgrades.

By offloading and accelerating SSL encryption, decryption and certificate management from servers, ADCs enable web and application servers to use their CPU and memory resources exclusively to deliver application content and thus respond more quickly to user requests. Our smarter traffic cop comes to the rescue again, this time eliminating distractions that prevent you from concentrating on the driving tasks at hand. Web-based applications consist of a variety of different data objects which can be delivered by different types of servers. ADCs provide application-based routing using file types to direct users to the server (or group of servers) that is set up to handle their specific information

requests, such as ASP or PHP applications. User requests can be routed to different servers by sending requests for static file types (jpg, html, etc.) to one server group, and sending user requests for dynamic data to other servers optimized for that purpose. Like the ultimate traffic cop, the ADC knows the optimal path for each destination.

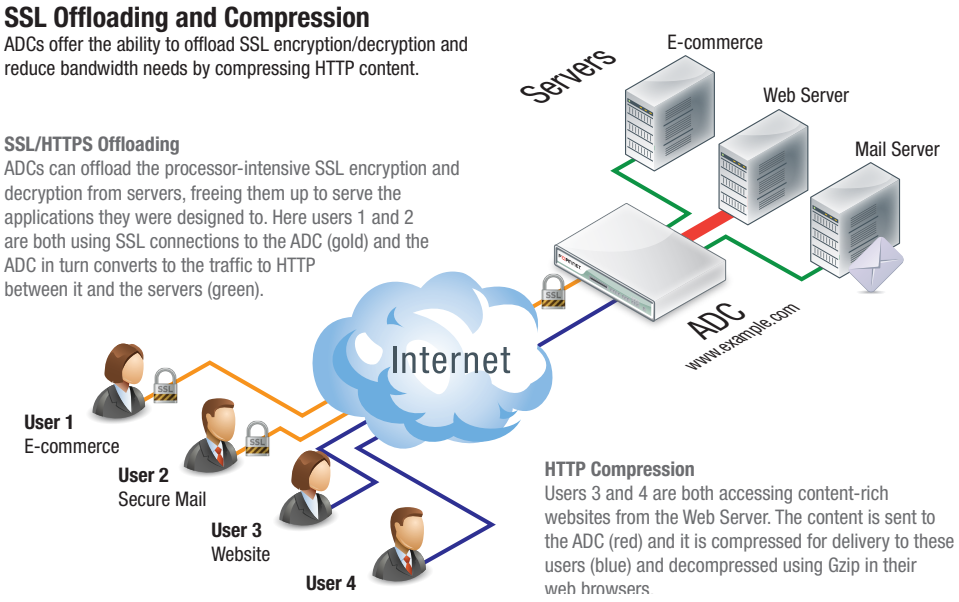
Transaction-based applications require connections to the same server in order to operate correctly. The best-known example of this is the "shopping cart" problem when you establish a session with one server to add an item to your cart and then are load balanced to a different server to checkout. If you don't have a persistent connection to the original server, you'll find your cart is empty.

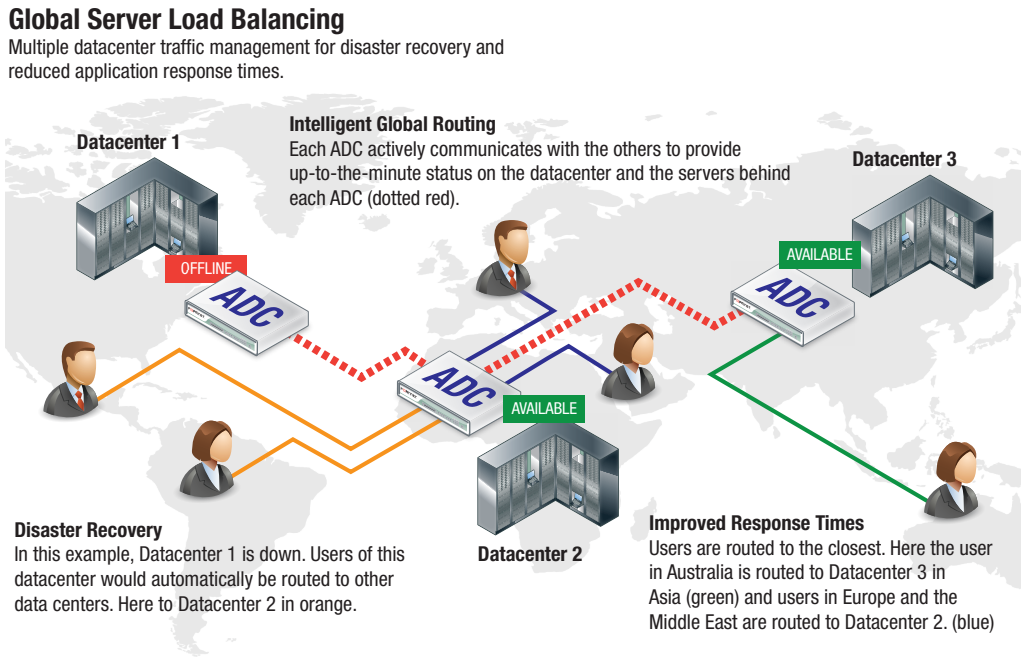
Advanced features like SSL offloading, HTTP compression and content-aware routing separate ADCs from basic load balancers

ADCs use session state with HTTP headers and cookies to ensure that users and servers remain "persistent". The ADC uses the cookie within the HTTP header to ensure that users continue to be directed to the specific server where the session state information resides. Without this capability, if the user went to a different server, the previous transaction history would be lost, and the user would need to start the transaction over. Once again, the ultimate traffic cop saves the day by understanding the application, network conditions and your priorities.

Global Server Load Balancing for ADCs solves the complex problem of scaling applications across multiple data centers for disaster recovery or to improve application response times for geographically dispersed users. Using a DNS-based approach combined with configurable business rules, user requests are resolved to the closest, best performing or lowest-cost data centers. If a data center is down due to a natural disaster or planned maintenance, automatically users are routed to a different data center until the primary data center is back online.

Link Load Balancing intelligently manages multiple wide-area links (WAN) to the internet from the ADC to improve application response times, reduce bandwidth needs and to provide redundancy should a link fail. If an internet connection becomes congested or is offline, traffic is automatically routed to the remaining links.





Finally, today's ADCs need to operate in and manage virtual environments. Advanced ADCs offer deep resource management of virtual environments and not just basic health-checking for server availability. With this tight virtual integration, the ADC can make load balancing decisions based on the status of the virtual machines and the servers they run on.

The Future of ADCs

Just as ADCs have replaced server load balancers, new technologies and new application delivery needs will shape the future of the ADC. Trends in network security, SDN, device consolidation, cloud/virtualization and other future developments will impact the evolution of these devices.

Fortinet sees network security as the major factor shaping the ADC market in the coming years. As network threats continue to get more sophisticated, most of these new attacks are targeted at the applications themselves like SQL Injection and Cross-Site Scripting. Inclusion and/or close coupling with additional security platforms like firewalls and Web Application Firewalls (WAFs) will help to minimize these risks. Most advanced ADCs have some form of security and some include basic WAF services. We expect that this trend will continue with the ADC playing a key role in helping prevent application-layer threats.

> Global Server Load Balancing and Link Load Balancing are important features for routing traffic between multiple data centers

We also see SDN as a game-changing technology that has the potential to reshape the IT industry, as well as ADCs. The adaptive, flexible environment that SDN enables will require an ADC that supports features like customized scripting and comprehensive APIs.

We predict that ADCs will be a point of service and feature aggregation as opposed a device that is subsumed by another. The ADC is a critical routing hub that is difficult to replace with another device and will continue to stand as a primary network component in the modern data center.

FortiADC Application Delivery Controllers

The FortiADC line of hardware and virtual Application Delivery Controllers provide unmatched Server Load Balancing performance whether you need to scale an application across a few servers in a single data center or serve multiple applications to millions of users around the globe.

With included SSL Offloading, HTTP Compression, Global Server Load Balancing, Firewall and Link Load Balancing, they offer the performance, features and security you need at a single-all inclusive price. Advanced models include 10-GE SFP+ ports, hardware-based SSL ASICs, dedicated management channels and dual power supplies to meet the demands of datacenter environments with L4 throughput up to 50 Gbps. FortiADCs include:

- Advanced server load balancing for scalability and resilience of your infrastructure by distributing application load over multiple servers.
- Caching of static content to reduce the load on the server and network infrastructure, increasing application responsiveness and reducing delivery delays.
- Dynamic HTTP Compression to accelerate network performance without using vital server resources.
- Hardware and software-based SSL Offloading to reduce the performance impact on your server infrastructure.
- Link Load Balancing to distribute traffic over multiple ISPs to increase resilience and reduce the need for costly bandwidth upgrades.
- Global Server Load Balancing to manage traffic across multiple geographical locations for disaster recovery and improved application response times.

FortiADC Benefits

When you choose a FortiADC for your application delivery needs you'll be guaranteed the security, performance and interoperability you need today and in the future.

Security: Fortinet is a leader in network security and unified threat management. Our FortiADC products build on that expertise to ensure your applications and users are protected from the latest network and application threats.

Performance: All of Fortinet's appliances and virtual products are built to perform. Our latest FortiADC appliances offer up to 50 Gbps of L4 throughput for data center and MSP environments.

Interoperability: When you buy a FortiADC, you get an integrated application delivery solution. All our products are designed to leverage and seamlessly interoperate with other Fortinet products and services like FortiGate, FortiManager and FortiAnalyzer. We optimize and test our products to minimize bottlenecks to increase overall performance between platforms when used together in a secure application delivery network.

FortiADC Application Delivery Appliances

With Layer-4 throughput starting at 2.7 Gbps through to 50 Gbps, Fortinet has an ADC to meet needs of almost any application environment.

Summary

Server load balancing grew out of the need to scale websites in the 1990s and is the foundation of today's modern application delivery controller. Building on this core of server load balancing, the advanced features of ADCs not only scale applications, they intelligently provide application availability.

Features such as SSL offloading, HTTP compression and intelligent policy-based layer 7 routing, distinguish a basic server load balancer from a modern ADC.

As applications needed to expand across multiple data centers, features like Global Server Load Balancing and Link Load Balancing were introduced to manage inter-site traffic and WAN links between multiple locations.

The ADC will continue to evolve with new features like virtual environment management, integrated security services and SDN support.

About Fortinet

Fortinet (NASDAQ: FTNT) helps protect networks, users and data from continually evolving threats. As a global leader in high-performance network security, we enable businesses and governments to consolidate and integrate stand-alone technologies without suffering performance penalties. Unlike costly, inflexible and low-performance alternatives, Fortinet solutions empower customers to embrace new technologies and business opportunities while protecting essential systems and content. Learn more at www.fortinet.com.

GLOBAL HEADQUARTERS
Fortinet Inc.
899 Kifer Road
Sunnyvale, CA 94086
United States
Tel: +1.408.235.7700
Fax: +1.408.235.7737
www.fortinet.com/sales

EMEA SALES OFFICE
120 rue Albert Caquot
06560, Sophia Antipolis,
France
Tel: +33.4.8987.0510
Fax: +33.4.8987.0501

APAC SALES OFFICE
300 Beach Road 20-01
The Concourse
Singapore 199555
Tel: +65.6513.3730
Fax: +65.6223.6784

LATIN AMERICA SALES OFFICE
Prol. Paseo de la Reforma 115 Int. 702
Col. Lomas de Santa Fe,
C.P. 01219
Del. Alvaro Obregón
México D.F.
Tel: 011-52-(55) 5524-8480