What is an application delivery controller?

ADCs have gained traction within the last decade, largely due to increased demand for legacy load balancing appliances to handle more advanced application delivery requirements and improve application performance.

Application delivery controllers (ADCs) are purpose-built networking appliances whose function is to improve the performance, security, and resiliency of applications delivered over the web.

This white paper introduces many of the core services ADCs provide and explains how they benefit both users and application administrators.

Application delivery

Applications have evolved significantly over the years. The term "delivery" is now generally accepted as the means of bringing an application to the user in this new era of mobility and cloud. Business applications have moved away from desktop-bound software installed on a local server accessed by users across the LAN. Modern applications need to work across all types of networks, and at locations beyond the confines of the physical workplace.

ADCs, which are widely deployed as a key fixture in the organization, help applications adapt to the networks and protocols that are in place today. They also ensure that applications perform optimally, are always available, and don't present any security risks either to the user or business.

Application availability

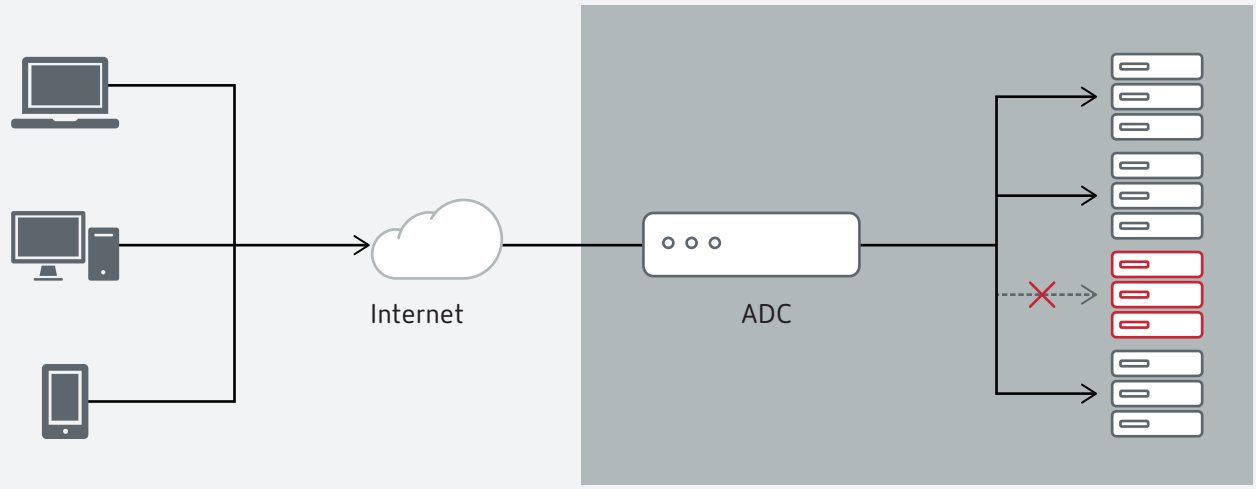
The average consumer expects the devices and applications they interact with on a daily basis to always work, and for information to be instantly available on demand. These expectations have carried over to the types of devices and applications that they use. To satisfy today's workers, business applications need to be as intuitive and easy to use as the ones they rely on for personal tasks and entertainment.

Many employees are no longer restricted to using locked-down, company-owned equipment, and can use personal devices to work whenever they choose. With people working at any time of the day or night, IT must make certain that workplace servers and applications are available around the clock. Businesses invest heavily in IT infrastructure to ensure that employees always have access to applications and information when they are needed.

Of course, servers can fail for a number of reasons ranging from mechanical problems to over-utilization and security breaches. If a server goes down, applications running on it become unusable or inaccessible.

IT organizations can plan for these occurrences by building fault tolerance into their environments. Deploying additional servers in the datacenter or at a co-located site are typical failover strategies. ADCs can help ensure high availability of applications by providing seamless failover. This is done by balancing application workloads across a cluster of active servers in one or multiple sites.

Figure 1. Active monitoring to ensure requests are sent to available servers



How server load balancing helps

Server load balancing helps distribute traffic across a cluster of servers to optimize utilization, improve responsiveness, and increase availability.

An ADC employs algorithms and policies to determine how inbound application traffic is distributed. Round robin, which forwards client requests to each server in turn, is a fairly rudimentary form of load balancing. This method assumes all servers are the same; it does not take into account health or responsiveness. An administrator can implement additional policies that direct an ADC to check for a number of criteria before determining where an inbound request should be sent. The ADC can inspect packet headers for keywords or requested file types, and direct the request to the appropriate server based on this information.

ADCs are also heavily relied upon for their monitoring capabilities. They can check a server's health and operability beyond the standard ping. If monitoring indicates a server is experiencing an issue, or that specific health criteria needed to ensure a server's reliability are not being met, the ADC will route traffic to an alternate server, avoiding a potential disruption (see Figure 1).

ADCs can also provide real-time and historical analysis of all user and network traffic, including metrics for round-trip times, bandwidth usage, and datacenter and WAN latency. This information can assist help desk staff by minimizing the time they spend identifying the cause of an issue, and help users by providing faster resolution.

Load balancing servers across multiple sites

Load balancing is a critical service in any high-traffic datacenter, but an ADC can also redirect traffic to a cluster of servers located in an entirely different datacenter. This is called global server load balancing. The servers in the other datacenter can be front-ended by another ADC, which works in tandem with the first appliance. These sites can be configured in either active-passive or active-active mode. In the latter, both sites are actively supporting inbound traffic. Each ADC detects which datacenter is closest to a given user, and routes the client request to a server in that datacenter. This process minimizes latency and round trip times for the user's request and ensures a better experience.

This configuration also supports business continuity if a datacenter suffers a shutdown. When traffic is routed to that datacenter, the ADC will divert it to an available ADC in a co-located site that can direct traffic to a viable server resource.

Application performance

If applications do not perform to users' expectations, their productivity can be severely compromised. An ADC can employ an array of mechanisms to improve application performance, especially over mobile and high-latency networks.

SQL database load balancing is one mechanism that can deliver performance gains. SQL load balancing uses many of the same techniques employed for load balancing TCP traffic, but applies this intelligence at the database level. It uses policy-driven logic for each SQL transaction, improving the number of requests and connections that can be handled within the database cluster.

Other common app performance optimization services offered by an ADC are offloading of server-intensive tasks, connection multiplexing, compression, and caching.

SSL and TLS are mainstays for doing business on the web. Managing traffic encrypted with new ciphers is very CPU intensive. ADCs can handle exceedingly high volumes of encrypted and unencrypted traffic. The ADC manages certificates and decrypts traffic before it reaches the server.

TCP multiplexing is an effective method for handling high volumes of inbound server requests. TCP multiplexing maintains active connections between the ADC and the servers. As traffic hits the ADC, it routes requests using these open channels, which eliminates the inefficient "open-close" overhead for each transaction that can negatively impact server performance.

Performance optimization on mobile networks

ADCs can also provide performance benefits across mobile networks. Web pages designed for high-speed Internet links often fail to deliver the same user experience on a mobile device connecting over a bandwidth-constrained network.

Several creative mechanisms enable an ADC to optimize web content delivery over mobile networks. Domain sharding is one example. Connection-layer optimization is applied to a single domain. Content on each page is broken down into a sequence of subdomains that allow a larger number of channels to be opened simultaneously, which decreases page load time and improves performance.

ADCs have visibility into the content that is being delivered, and can further optimize delivery of web pages containing large images by converting GIF files into more-efficient PNG formats.

The other large components of a web page include extensive scripts and cascading style sheet (CSS) files, which ADCs can compress by removing unnecessary characters and white space. When compressed, files traverse the network at a much faster rate, so download times are significantly reduced.

Application and user security

Delivery over the web has introduced new threats and vulnerabilities that traditional LAN-bound applications never had to contend with. As workers become more mobile and require remote access to applications and data, IT must devise more stringent safeguards against external attacks and data leakage.

ADCs serve as the natural entry point or gateway to the network. They authenticate each user attempting to access an application. If the application is SaaS based, the ADC can validate a user's identity using an on-premises Active Directory data store that eliminates the need to store credentials in the cloud. Not only is this process more secure, it also enhances the user experience by providing single sign-on capabilities across multiple applications.

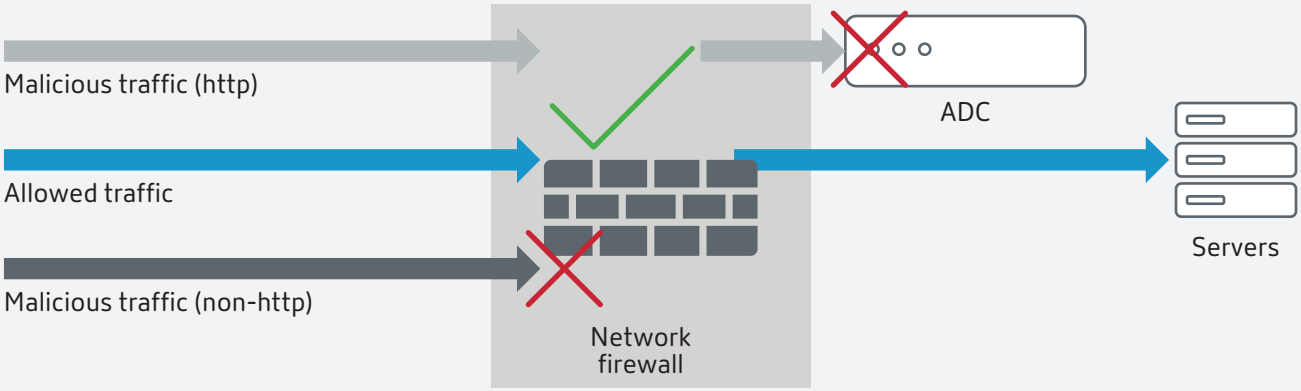
SAML, the XML-based protocol, is now widely used to simplify the application login process. The ADC can act as a SAML agent, authorizing users via any data stores where their identity can be confirmed. Some applications allow the use of credentials from sites such as Facebook or Google+ to validate identity before granting access. ADCs can act as a SAML identity or service provider in this respect.

Distributed Denial-of-Service (DDoS) attacks have become rampant.[1] Web properties, specifically, are being targeted with the intent of overwhelming their servers and disrupting their ability to conduct business. The ADC can implement rate-limiting measures to protect internal server resources from being targeted by these specially designed attacks. When an unusually massive surge of inbound requests occurs, the ADC can throttle these requests and minimize the amount of available bandwidth they consume, or reject the request entirely.

ADCs have converged load balancing and advanced Layer 7 protection, which traditionally were only available as standalone solutions. Application firewalls can inspect data packet headers for suspicious content or malicious scripts that may not be detected by network firewall (See Figure 2).

An ADC can support both positive and negative security models. When an ADC is placed in "learning" mode, it can analyze traffic to determine usage patterns that signify normal behavior. If a malicious inbound request is sent, for example, using SQL injection or cross-site scripting, it will automatically flag that request and

Figure 2. Layer 7 protection beyond a network firewall's capabilities



block it. It can also employ signature-based protection via integration with third party security providers such as Qualys. Combining these protection methods allows the ADC to use a comprehensive hybrid security model for applications and users.

What's next for ADCs?

ADCs already provide tremendous value to IT organizations ensuring the secure delivery of applications and data to the user. However, they are expected to continue advancing as applications evolve. Software-defined networking has placed increased demands on ADCs to function "as a service." As network protocols become more application centric, ADCs must also adapt and become more "self-automated" to provide seamless optimization and protection for every type of application.

For more information, you can reference the following videos:

How an ADC helps optimize Microsoft environments

What is Citrix ADC?

1 Jonathan Keane. DDoS Attacks Hit Record Numbers in Q2 2015. Digital Trends. August 19, 2015. http://www.digitaltrends.com/computing/ddos-attacks-hit-record-numbers-in-q2-2015/

CITRIX®

Enterprise Sales
North America | 800-424-8749
Worldwide | +1 408-790-8000

Locations
Corporate Headquarters | 851 Cypress Creek Road Fort Lauderdale, FL 33309, United States
Silicon Valley | 4988 Great America Parkway Santa Clara, CA 95054, United States