Lunch Break SERIES

What Makes Application Delivery Controllers Tick

How ADCs work and how they can help you maximize your datacenter's computing power.

SPONSORED BY

VIRTUALIZATION REVIEW

BY BRIEN M. POSEY

Why You Should Already Have an ADC

Cloud and virtualization projects move around a billion bits at a whim, and Application Delivery Controllers ensure those projects move around intact and quickly.

The rapid adoption of technologies such as server virtualization and cloud services over the last few years have driven demand for Application Delivery Controllers (ADCs). I'll look at some of the ways in which ADCs have evolved, as well as some of the benefits to using them.

What Is an Application Delivery Controller?

Even if the term Application Delivery Controller seems foreign to you, you have likely seen or possibly even worked with one in the past. ADCs have existed in one form or another for many years. However, their features and capabilities have evolved to address the ever-changing needs of enterprise IT.

Previous ADCs weren't actually called Application Delivery Controllers. Instead, they were generally referred to as application proxies or load balancers. Such products were generally used to provide scalability and a degree of high availability to

Web sites or Web applications. Today, ADCs use a variety of techniques to ensure application availability, security and performance. Today's ADCs are actively involved in delivering applications, rather than simply acting as a load balancer or a proxy.

So, what capabilities should you look for? ADCs offer a number of advanced features that are designed to deliver applications in an efficient and secure manner. The actual feature set can vary widely from one vendor's product to the next, but there are some features that have become more or less standard.

Load Balancing

As I mentioned, first-generation ADC products served largely as load balancers for Web sites or for Web applications. Such products generally worked by sitting in an organization's DMZ and directing inbound traffic to a Web server. Typically, such products used techniques similar to DNS round robin to ensure the inbound traffic was evenly distributed among the available Web servers.

Today's ADCs still offer load-balancing capabilities, but these capabilities are far more sophisticated than they were in the past. For example, some ADCs offer global load-balancing services, which can automatically route inbound traffic to the datacenter that's in the closest geographical proximity to the person who's attempting to access the Web application.

This technology can also sometimes be used as a disaster recovery solution. If, for instance, an organization's primary datacenter were to fail, an ADC's global load-balancing capabilities might detect the failure and redirect inbound traffic to an alternate datacenter.

Application Layer Proxy

Although many of an ADC's features are related to performance, security is also a concern. One of the most common security features found in ADCs is an application layer proxy.

An application layer proxy is a proxy (or gateway) that has knowledge of specific applications and protocols. For example, an application layer proxy that's designed for use with Web applications would support protocols such as HTTP, HTTPS and FTP.

Most application layer proxies are designed to work as reverse proxies. This means that clients establish connectivity to the proxy (in the form of an application gateway) rather than directly to the application server. The proxy sends requests to the application server on the client's behalf and relays responses to the client. By doing so, the proxy is able to hide the application server from the client, thereby improving security.

Selective Compression

One the main jobs of an ADC is to ensure users receive a good level of performance when accessing applications. There are a number of different techniques used to improve overall performance, but nearly all of the ADCs on the market use a deduplication engine to compress traffic, thereby improving overall bandwidth usage.

One the main jobs of an ADC is to ensure users receive a good level of performance when accessing applications. There are a number of different techniques used to improve overall performance, but nearly all of the ADCs on the market use a deduplication engine to compress traffic, thereby improving overall bandwidth usage.

While virtually every ADC uses some form of data compression, some of the products on the market use what's known as selective compression. Selective compression generally refers to the ability to choose the method that's used for data compression. Some of the hardware appliances on the market include a dedicated board that's used for data compression, and the selective compression feature lets administrators choose whether to use dedicated hardware compression or the appliance's CPU to compress data.

Hardware-level compression is usually the default setting and works well for quickly compressing large volumes of data. However, the appliance's CPU may be able to achieve a higher compression ratio than the dedicated compression hardware. The problem with doing so is that CPU-bound compression might not be able to handle as high a volume of traffic as the dedicated compression hardware, and might also result in diminished overall performance for the appliance. Some selective compression models are adaptive in that responsibility for compression is dynamically switched between the appliance's CPU and the compression hardware based on the appliance's current workload.

Selective Caching of Dynamic Content

Another way in which some ADCs seek to improve performance with regard to the delivery of Web applications is through selective caching of dynamic content. Caching has been used for well over a decade as a way of speeding up the delivery of Web content. However, legacy caching solutions were designed to cache static Web content.

The problem with this approach is that today's Web applications produce dynamic content that's generated on the fly. Legacy Web caching solutions are unable to effectively cache dynamic content.

Some ADCs attempt to solve this problem through a dynamic content caching engine. The idea behind this is that even if a Web page consists primarily of dynamic content, there's likely to be at least some static content. For example, a page's header might be static and some of the graphics might be static. These static elements can be cached and then added to the dynamic content and served to clients. The overall performance gain varies based on the volume of content that's able to be cached.

Traffic Shaping

One of the things that ADCs typically do in an effort to improve overall performance is to perform traffic shaping. Traffic shaping (which is sometimes referred to as packet shaping) is a technique for manipulating network traffic flows in order to achieve a certain level of application performance.

In any organization there is a finite amount of Internet bandwidth available. Technologies such as compression can be used to make more efficient use of the available bandwidth, but ultimately there's a limited amount of Internet bandwidth that must be shared by multiple users, applications and services.

If there isn't enough bandwidth to service all of the competing traffic sources then network contention occurs and latency goes way up. This latency directly impacts the responsiveness of Web applications. Traffic shaping seeks to solve this problem by manipulating traffic streams in a way that ensures mission-critical Web applications receive a sufficient amount of Internet bandwidth.

The most common way for an ADC to perform traffic shaping is by using the Quality of Service (QoS) protocol. QoS is a networking standard that's supported by Windows OSes and by most modern network hardware. It allows bandwidth to be throttled or reserved, which makes it possible to ensure that Web applications receive a sufficient amount of bandwidth.

Customizations

The various features found in modern ADCs go a long way toward making Web applications scalable and highly available. It's important to remember, however, that large organizations rarely have just one Web application.

Some ADCs are equipped with a rules engine that allows Web applications to be prioritized. In some cases it may also be possible to create a series of rules that are designed to enhance security (similarly to the way firewall rules work).

There are even ADCs that contain programmatic interfaces. These interfaces make it possible to develop custom code that can be used to extend the device's functionality. For example, an organization might use custom code to extend an ADC's functionality to support a proprietary protocol.

The various features found in modern ADCs go a long way toward making Web applications scalable and highly available. It's important to remember, however, that large organizations rarely have just one Web application.

Efficient End-User Experience

As you can see, an ADC can help to ensure application availability and to scale applications in a way that ensures an efficient end-user experience. Of course, the actual feature set tends to vary from one vendor's product to the next, so it's a good idea to compare the competing solutions before making a purchasing decision. VR

Brien Posey is a freelance writer and seven-time Microsoft MVP who has written and contributed to many thousands of articles and dozens of books on a wide variety of IT topics. Posey was CIO for a national chain of hospitals and health-care facilities and also served as a network administrator for some of the country's largest insurance companies, as well as for the Department of Defense at Fort Knox. When he isn't busy writing, Posey enjoys exotic travel, scuba diving and racing his cigarette boat. You can visit his personal Web site at brienposey.com.

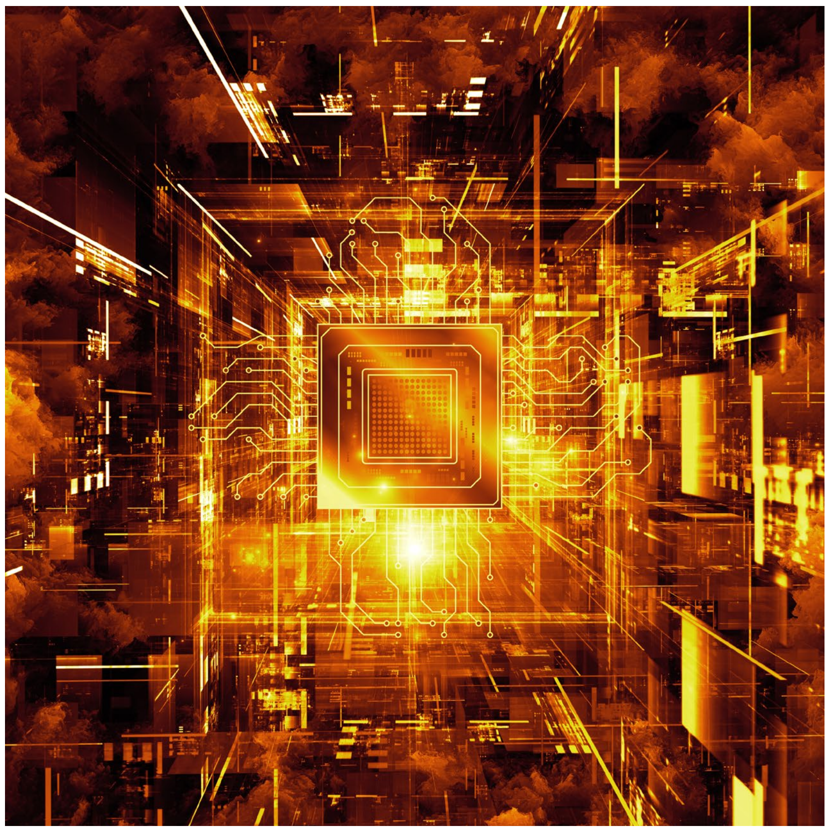
BY KARL TRIEBES

Don't Abandon Your Purpose-Built Network Solutions Just Yet

A vendor's view: Purpose-built hardware often fulfills a need that general-purpose appliances can't adequately address.

Imagine taking your car to the shop for major repairs and finding that the mechanic uses just one tool for everything, from changing a flat tire to replacing your transmission. He explains that he doesn't need all those expensive, special-purpose tools anymore now that he's found this one, All-Powerful Tool.

It sound ridiculous, but similar images come to mind when I hear the suggestion that network appliance vendors need to abandon their purpose-built platforms and deliver their solutions on general-purpose hardware. The argument is that raw compute power has increased to the point that specialized hardware solutions are no longer necessary. But before we write off purpose-built network solutions, let's revisit the reasons why they exist in the first place and consider how significantly they improve an enterprise's ability to deliver mission-critical applications.



Today's advanced Application Delivery Controllers (ADCs) evolved from software-based load-balancing solutions designed a decade or more ago to distribute comparatively unsophisticated applications and light user loads across low-speed physical networks. As hardware, OSes, applications, and networks grew infinitely more sophisticated and mobile devices became pervasive, software-based load balancers couldn't keep pace and were soon moved to hardware platforms. To handle today's massive amounts of traffic and to deliver mission-critical applications and data to users from virtually any location and on any device, enterprises need advanced ADC solutions they can trust to be fast, secure, and available—not to mention scalable and fault tolerant—to accommodate rapidly changing business needs.

It's tough for a network vendor to deliver on these requirements and achieve the highest levels of performance with a solution that's based on a general-purpose platform. By definition, a general-purpose machine is designed to support many types of applications and workloads and specialize in nothing. And while it's true that technological advances have made today's general-purpose machines magnitudes more powerful than those of a decade or two ago, raw processing power alone doesn't equate to performance. Even if it did, it would be tough to cite it as a reason to do away with purpose-built solutions because they, too, benefit from those same technological advances.

Architecture, not raw processing power, is where performance strides can be made, and that's what gives purpose-built network solutions a distinct advantage. Because vendors of such solutions can choose not only the hardware components (such as CPUs, RAM and networking devices) but also leverage customized hardware (such as ASIC and FPGAs) to add value, offload processing, and relieve architectural bottlenecks, they're able to provide fully integrated, high-performance, predictable, and highly reliable solutions.

Hardware accelerators integrated into such solutions are specifically designed to greatly speed certain computational processes, such as cryptographic operations. When performed on a general-purpose server, these operations can consume 30 percent or more of the server's CPU and memory cycles—and that's just one example. Acting as a proxy between client and server, a purpose-built ADC can aggregate millions of client requests into hundreds of server connections and cache them for reuse; it can intelligently manage and prioritize SSL sessions; it can apply intelligent compression to data—a task so compute-intensive when performed on a server that it can actually degrade rather than boost performance. These specialized hardware components, carefully integrated into the architecture, enable purpose-built network solutions to scale to carrier-class levels of performance.

In the end, if it were feasible for network vendors to optimize these compute-intensive tasks on general-purpose platforms, don't you think they would? After all, it would certainly make these solutions easier to architect and more affordable to manufacture. When performance counts, however--and increasingly that seems to be the critical need—enterprises gain nothing by abandoning purpose-built network solutions. So when you hear such suggestions, remember why these solutions were built in the first place, not on a whim to benefit vendors, but rather to help enterprises improve efficiency, control their costs, and meet their most difficult and challenging needs: managing massive amounts of traffic in the delivery of mission-critical applications and data. **VR**

Karl Triebes is executive vice president of product development and CTO for F5 Networks Inc., responsible for overseeing the company's technology roadmap and leading the F5 engineering team. He was previously CTO at Foundry Networks.

BY ERIC BEEHLER

The ABCs of ADCs

ADCs serve a new generation of business applications, with specific benefits when implemented in the cloud.

With the push in the past few years to scalable, service-oriented applications that primarily rely on Web services, we now have the ability to scale the delivery of those applications as required. This is good, because the ability of users to add more servers as applications demand them helps drive acceptance of virtualization, Web-based services and UIs. In turn, that ability is also enabling vendors to cook up new ways to deliver their applications to multiple devices—from desktops to smartphones.

When many datacenters first began delivering Web-based applications, we relied on IP load balancers to make sure traffic was shared properly between servers. Currently, however, multiple tiers of services and increased security needs are making life more complicated. IP load-balancing systems are graduating to a new category called Application Delivery Controllers (ADCs). ADCs load balance incoming traffic, but also perform many other advanced functions centered on optimization, acceleration and security.

Led by the need for large Web sites to deliver extremely fast service to end users from increasingly complex back-ends, ADCs are leading a new expansion to Web-based application delivery. When a big site delivers e-commerce or is working with multimedia content, it needs to squeeze every bit of performance it can out of its platform to avoid being victimized by slow loading. In addition, real-time insight into the uptime and performance of the server farm allows the site to appear fully functional, even when there are problems on the back-end. ADCs help these deliveries by providing deep inspection of the incoming requests, enabling specialized services served by tailored farms of servers, providing tremendous performance with off loading of encryption, and keeping availability up, no matter what the status of a specific server is.

What Does an ADC Do?

When considering adding an ADC for a datacenter solution, the question a potential buyer often asks is: "Can I just do this with what I have?" The list of specific features will certainly sound familiar. For example, caching and TCP tweaking are all things that major Web servers can provide. In addition, you may already have some sort of Layer-4 load balancing in place for existing applications. So what are the compelling benefits of ADCs?

ADCs provide much more granular control over all those areas that affect performance. Additionally, they provide functionality that wouldn't be possible with just a Web farm and an IP load balancer.

For example, ADCs can maintain a pool of server links, providing ready-to-use connectivity. This is called TCP multiplexing, or TCP offloading, and isn't all that different from the concept of a database connection pool, which reduces the expense of building server connections and tearing them down multiple times during even a single page request. The performance gain is noticeable, and is certainly a timesaver when working with dynamic Web components. Some vendors estimate as much as a 66 percent performance improvement using this technology.

Application Layer-7 abilities are also baked into ADCs with HTTP-aware load balancing, which enables users to specialize their server sets by splitting their responsibilities. With this capability, the device isn't just looking forward based on IP, but is also looking deep into the HTTP request to understand what content is being requested and where to find it. With today's dynamic applications stitched together from many specialized server sources, Layer-7 inspection and direction is currently the most sought-after feature of ADCs.

Security has moved to the forefront of Web-based applications, and fully encrypted SSL connections are becoming the norm for many applications that hold sensitive data. Offloading SSL is nothing new, but the increased use of encryption is putting a strain on some servers, especially those that run in virtualized environments. ADCs are providing performance gains for lots of fully encrypted traffic. By acting as the proxy for SSL traffic, ADCs apply specialized hardware to encryption/decryption computations, which is much more efficient than dedicating the general-purpose CPU of a Web server to this function.

There's also the possibility that users can be greener with their ADCs. For example, SSL encryption offloading could save you up to 60 percent CPU utilization on your Web servers, as shown in research by F5 Networks Inc. With this kind of efficiency, you may be able to actually shut down some servers in your Web farm. If you're dealing with a big number—say hundreds of servers—this could represent a significant savings in datacenter power consumption and cooling.

As datacenter consolidation eliminates more and more servers, network performance is becoming a factor at many companies. Browser-based applications used to take a backseat to fat-client application performance, but now critical applications live in the browser. This means you have to think about HTTP performance in the same way as a commercial Web site like Amazon.com. HTTP compression and caching is critical to making remote Web applications perform like local client applications. ADCs are designed to cache the static content portions of Web sites so those back-end servers can serve up dynamic content. The expectation of users that their Web pages be usable within a couple seconds also holds true at sites such as Google and Facebook.

Browser-based applications used to take a backseat to fat-client application performance, but now critical applications live in the browser. This means you have to think about HTTP performance in the same way as a commercial Web site like Amazon.com.

In addition, the compression of various Web parts can provide multiples of increased load-time improvement. This is of particular importance to internally built applications, where developers may not think about performance the way a commercial Web site developer would. Allowing an ADC to handle compression rather than dealing with large graphics on a main page can help self-regulate the performance of application delivery. It can also help tremendously in bandwidth-challenged locations, where compression can make an application usable without a large amount of manual tweaking to the code or images. There are also Web server CPU savings to consider, as offloaded compression can increase CPU utilization by as much as 30 percent in some cases.

Where Do You Deploy an ADC?

In proposing ADC solutions to management, it's essential to be able to cost-justify them. What are the scenarios that will show the greatest return on investment? According to Matthias Machowinski, directing analyst for Enterprise Networks and Video at research firm Infonetics Research: "If you're in IT for a telecom company or a financials business, you're already implementing ADCs."

Machowinski goes on to explain the growing ADC market: "The general uptick in the economic recovery is feeding much of the needed upgrades to corporate datacenter functions that have been ignored due to budget cuts."

Now is a good time to review your software needs. You may have put off upgrading your financials, ERP or HR software for a few years, but those behemoths are due for a refresh, and it should be done via Web services.

Any application that relies on a Web farm of servers, has a complex back-end and conforms to the service-oriented architecture (SOA) model of services is a candidate for ADC intervention. Add to this list servers that have been consolidated and virtualized. Further, add anything that must be bound to a server and might be eating CPU cycles, such as SSL, and you've pretty much covered the ADC application waterfront.

Not only are you providing scalable solutions, but you may be expanding into a more non-traditional delivery such as a co-location host or on top of cloud services. The ADC technology is also readily available to you in those arenas. Some cloud vendors are offering ADC as a Service (AaaS). This is really no different than using the ADC in your datacenter, except it's geared specifically for network OSes from cloud vendors such as Amazon Web Services LLC or Rackspace US Inc.

The benefits of ADC in the cloud are the same as they are in your datacenter. This is especially true if you're deploying internal applications via the cloud. In cases such as these, performance over the WAN becomes even more critical, and cloud ADC solutions should be a strong consideration for any scalable solutions.

The benefits of ADC in the cloud are the same as they are in your datacenter. This is especially true if you're deploying internal applications via the cloud. In cases such as these, performance over the WAN becomes even more critical, and cloud ADC solutions should be a strong consideration for any scalable solutions.

A varied combination of ADC solutions is also a possibility. Cisco, F5 and Citrix all tout combined hardware/software solutions that stack virtual ADCs with their hardware counterparts. Vendors such as Zeus also integrate with some cloud solutions. Just remember that many of these vendors are big networking companies that may have more than one solution for sale. Take the complete solution under advisement, but make sure it's not overkill for your needs.

Breaking Down Old Boundaries

The old boundaries between hardware, software and networks are disappearing quickly as we move from client/server solutions to a Web-services methodology. ADCs are perfect examples of hardware/software appliances that provide acceleration, uptime and security by reaching deep into apps hosted on multiple machines. This concept is common for giants such as Google, but may not be familiar to every sys admin or Web developer.

Moving these common functions out of code, off servers and into common devices that front-end entire Web applications provides advantages that are becoming commonplace in some industries. Gartner Inc. expects the proliferation of ADCs to continue, with spending already increasing from $874 million in 2005 to $2.1 billion in 2010. Infonetics sees ADC implementation as a general trend, along with 10GB switches and other core-network technologies.

You should definitely be considering ADCs as critically as you're considering your next core switch or set of servers, because ADCs may be the next piece of core equipment you'll add to your datacenter. VR

Eric Beehler has certifications from CompTIA (A+, N+, Server+) and Microsoft (MCITP: Enterprise Support Technician and Consumer Support Technician, MCTS: Windows Vista Configuration, MCDBA SQL Server 2000, MCSE+I Windows NT 4.0, MCSE Windows 2000 and MCSE Windows 2003). He has authored books and white papers, and co-hosts CS Techcast, a podcast aimed at IT professionals. He provides consulting, managed services and training through his co-ownership in Consortio Services LLC.