

Ten Criteria for a Sound Investment in Application Switches

Application switches, also referred to as application delivery controllers (ADCs) or application front ends (AFEs), bring the following important benefits to IT organizations:

- Increased application availability and scalability
- Accelerated web services and application response times
- Support of data center consolidation and reduced power, cooling, and space requirements
- Faster application deployment times, as well as reduced ongoing management of application infrastructure
- Increased application, Extensible Markup Language (XML) web services, and data center security
- Better utilization of server resources through offloading of Secure Sockets Layer (SSL) Protocol and Transport Control Protocol (TCP) processing

Choosing the best application switching solution can be challenging because of the wide range of solutions now available from many vendors. How do you make a well-informed decision, achieving the greatest benefits of application switching and the lowest total cost of ownership (TCO)?

To help you, this document describes 10 primary criteria you should consider before investing in an application switching solution and why these criteria can be important to the success of your business. This document also provides a useful tool to assist you in the search for the right solution for your business.

Selection Criterion

IT organizations can use the criteria presented here to select an application switching solution. Table 1 at the end of this document uses these same criteria to compare the Cisco® Application Control Engine (ACE) solution to solutions from other vendors.

Performance and Scalability

Performance refers to a solution's throughput capacity and also the way the solution performs in the particular environment in which it is deployed. Scalability refers to the capability of the solution to adapt to the growing needs of the enterprise.

Customers need to evaluate the performance of the application switch being deployed from several viewpoints. First, consider raw bandwidth: How much throughput is actually available for application delivery? The device may claim a performance rate of 10 Gbps, but does performance degrade significantly if additional features within the solution are enabled or if the solution is deployed more broadly across multiple applications within the organization? Does the solution use purpose-built hardware that can process transactions, connections setup, and connections teardown quickly?

Also note that conserving servers, an advantage of high-performance solutions, has additional benefits: for instance, reducing the number of software licenses needed for servers, the amount of maintenance required, and power and cooling requirements, and freeing space in data centers.

Customers also need to understand how performance can be scaled over time. If a solution ceases to meet an enterprise's needs a year after it is deployed, more expense will be incurred if the solution needs to be replaced than if the solution can be upgraded by simply adding bandwidth through software licensing.

Enterprises and service providers should have the capability to scale the device's throughput capacity from lower to very high performance with simple software license upgrades. This kind of solution avoids the need to purchase, install, and test new hardware or forklift upgrade the entire system, which in addition to being costly can take weeks to accomplish. With a software-license-based scalable solution, organizations can also avoid application downtime and degradation associated with hardware-centric capacity upgrades.

In addition to scaling for performance, the device should scale to support multiple applications in a consolidated infrastructure. If more devices must be purchased to support more applications, then scalability is reduced. IT departments should also be able to use virtualization capabilities, discussed in detail the next section, to create additional virtual device instances on the existing application switch platform.

Virtualization

Environmental factors such as cooling, rack space, cabling, and power pose challenges for data centers. Server farms generate a large amount of heat, consume large amounts of power, and occupy considerable rack space in data centers. Application switching solutions that reduce the number of devices necessary to support a service or that offer more energy or space efficiency can mitigate these problems.

An important criterion in the new consolidated data center is the way the device handles virtualization.

Virtualization is the capability to logically partition a single physical application switch device into many virtual devices. Each virtual device should have all the capabilities of the actual physical device, resulting in virtual devices that are each isolated from other virtual devices in both their data and management. By using virtual devices, the application switch can deliver any service (for instance, server load balancing, acceleration, or security) across any application or department on a virtual device basis; for example, a service provider can allocate a virtual device for each customer. Virtualization can substantially reduce capital, cooling, rack space, and power requirements in the data center and enhance an organization's ability to scale its data center resources.

The virtualization implementation should be at the device-level, so that all aspects of the physical device can be virtualized. Some vendors limit their support of virtualization to access-control capabilities and do not offer true device- and service-level virtualization.

Look for the following characteristics when choosing an application switching solution that can address environmental concerns:

- The solution should support device and full service-level virtualization, allowing the deployment of virtual devices instead of physical devices and thus reducing incremental power and space requirements for the application delivery infrastructure and facilitating data center consolidation.
- The solution should need minimal increase in power, cooling, and space requirements to address the growing application needs of the business.
- Application switches should offer industry-leading performance with the lowest possible power and cooling requirements.

Ease and Speed of Application Deployment

Two compelling challenges faced by today's data centers and application teams are how to speed up application deployment cycles and how to reduce interdependency between IT organizations. Application switches can achieve these improvements through device virtualization, described earlier, and also through software configuration rollback and role-based administration capabilities supported by application switches.

Most networks experience periodic changes such as the addition of new sites or new applications, and the data center application switching solution should be able to adapt to these changes easily. The IT administrator should be able to roll back any virtual device to a previous configuration. The IT administrator should also be able to easily save an instance of an application in service from one virtual device and gracefully reuse it as new instances of existing applications are deployed in other virtual devices, without affecting any other applications serviced by the device.

Using role-based administration, IT personnel and organizations can provision and manage multiple virtual devices within a single application platform in parallel. With this capability, IT departments can deploy applications much more quickly than if each group (for example, the server team and the security team) has to provision the application switches serially.

The application switch architecture should accommodate expanding business requirements without the need for new hardware, which can take weeks to purchase, qualify, install, and test. With such an architecture, organizations circumvent the application downtime and degradation associated with hardware-centric capacity upgrades.

Application Acceleration

Numerous studies indicate that users will wait up to 10 seconds for an application to load after logging in. A longer delay poses the risk that users will attempt to circumvent the system. Users' rejection of business applications translates into poor returns on application deployment investments as well as degraded user productivity, customer service, and, ultimately, sales and revenue.

Application switching products should use a range of acceleration capabilities to boost remote end-user application response times. Some of the more advanced technologies include compression, flash-forward, and delta encoding features. These functions minimize distance-imposed latency when application requests are served to remote users by reducing the number of round-trip data transfers and messages required for any HTTP-based application. These functions also optimize bandwidth by delivering to the client just the differences between cached original pages and updated new pages. Customers using these acceleration technologies can achieve up to 300 percent improvement in response times.

According to the 451 Group, XML accounted for 15 percent of data center traffic in 2005, and by 2008 XML expected to account for 50 percent of data center traffic. An XML message is 3 to 10 times larger than an equivalent binary message, which makes servers and infrastructure vulnerable to overload as XML traffic increases. General-purpose servers are expensive resources that should not be used for computationally intensive XML functions. Hence, another key differentiator in choosing a solution is whether the solution can accelerate XML applications. Most solutions can accelerate web-based applications. Some solutions, however, can also accelerate XML applications.

Security

Application switches should provide an additional layer of security and should act as a last line of defense for the servers in the data center, performing deep packet inspection. The application switching solution must not disrupt the current security environment, nor should it create any new security vulnerabilities.

The application switching solution should provide an integrated data center firewall that protects against protocol and denial-of-service (DoS) attacks and encrypts mission-critical content. The data center firewall should protect the data center and critical applications from malicious traffic with features such as the following:

- TCP/IP normalization and RFC compliance
- HTTP deep packet inspection of HTTP header, URL, and payload
- Bidirectional Network Address Translation (NAT) and Port Address Translation (PAT)
- Support for static, dynamic, and policy-based NAT and PAT
- Access control lists (ACLs) to selectively allow traffic between ports

The application switching solution should also provide an application-layer firewall to prevent attacks embedded in application payloads, including zero-day attacks. Most traditional firewalls do not protect against application-layer attacks. An application-layer firewall secures mission-critical applications and protects against identity theft, data theft, application disruption, and fraud. It uses features such as efficient inspection, filtering, and fix up of popular data center protocols such as HTTP, Real-Time Streaming Protocol (RTSP), Domain Name System (DNS), FTP, and Internet Control Message Protocol (ICMP) to defend web-based applications and transactions against known and unknown attacks by professional hackers.

With XML traffic is increasing in data centers, it is imperative that the application switching solution also provide XML security.

Application Availability and Uptime

Maintaining application uptime and availability is a major concern of IT administrators. To increase application availability, the application switching solution should use best-in-class Layer 4 load balancing and Layer 7 application switching dynamic and adaptive algorithms coupled with highly available system software and hardware. These features together should offer many configuration options for intelligent failover and redundancy across the application switches, across the virtual devices, and across the data centers. The application switching solution should also offer an extensive set of application health probes to help ensure that traffic is forwarded to the most available server. It should also allow servers to be added or maintained without service disruption.

Solution Architecture

The architecture of the application switch determines the switch's capability to support the following:

- Increased traffic load
- Additional servers
- New applications

Application switches typically have one of two main types of architecture: server based or hardware based.

Server-based application switches usually have general-purpose hardware running a standard or modified freeware operating system such as Linux or FreeBSD. This server-based approach reduces the need for code development and enables faster turnaround of new features; however, it can diminish performance, with customers having to trade performance for features. In addition, this approach requires the operating system to be redesigned to increase scalability or performance, which can result in an unstable product.

In evaluating products, customers should note whether the solution has been designed for a specific function. The right solution architecture uses purpose-built hardware designed for the application switching solution rather than general-purpose processor hardware. Purpose-built hardware performs tasks much more quickly and efficiently than a general-purpose processor. Purpose-built hardware provides the highest performance for transactions and for connections setup and tear down, as well as state-of-the-art virtualized application switching solutions with minimal power, cooling, and space requirements. General-purpose hardware with software devices requires more rack space and, in most cases, additional devices to deliver higher performance, thus also requiring more power and cooling resources.

Optimized Server Operations

Application switches should offload server functions that can be handled better and more effectively by the network, allowing the server to do what it does best: processing and serving information to its users. TCP communications management functions and SSL encryption can be moved to application switches so that the servers can devote their computing cycles entirely to their primary mission: quickly fulfilling user requests for application content.

The offloading of communications management and SSL encryption, by optimizing each server's performance, also reduces an organization's need to make additional capital investment in server capacity. It also reduces delays in application availability resulting from server inefficiencies.

Module and Appliance Form Factors

The application switch vendor should provide a flexible solution by offering switches with various form factors to address the different deployment requirements of customers. Application switches should be available as appliances for discrete data center deployments and also as modules either for high-performance data center modular Ethernet switches or modules for carrier-grade modular routers. Modules provide the high availability, quality of service (QoS), and other features of high-performance data center routing switches along with application switching; the use of modules also results in fewer devices to manage. The appliance form factor is useful for small-scale deployments.

End-to-End Professional Services and Support

Increasingly, business application delivery requires not only the best features and performance but also a strategic partner that can advise about, install, and support its solution throughout the entire

product lifecycle. Enterprises and service providers should invest in a vendor that delivers award-winning global support 24 hours a day, every day, and also offers award-winning solution lifecycle services, including planning, design, implementation, operation, and optimization services.

Table 1. Cisco ACE Solution Compared to Solutions from Other Vendors

Selection Criterion	Cisco ACE	Vendor _____		Vendor _____	
		Yes	No	Yes	No
Performance and Scalability	Cisco ACE provides up to 64 Gbps				
Highest raw throughput rate?	X				
No significant degradation of throughput after features are enabled?	X				
No significant degradation of performance after solution is deployed more broadly?	X				
Licensable performance scalable over time as business needs increase?	X				
Scalable to support multiple applications with isolation?	X				
Virtualization, Power, Cooling, and Space Requirements	Cisco ACE provides up to 90% power and cooling savings compared to leading competitor*				
True device-level virtualization?	X				
Capability to add new applications with minimal increase in power and cooling requirements?	X				
Minimal space requirements?	X				
Ease and Speed of Application Deployment	Cisco ACE provides up to 70% faster application deployment**				
Role-based administration support?	X				
Configuration rollback support?	X				
Application Acceleration	Cisco ACE provides up to 300% improvement in response times				
Latency reduction and bandwidth optimization mechanisms such as flash-forwarding and delta-encoding support?	X				
Hardware-based compression support?	X				
XML acceleration support?	X				
Security					
Does not create security vulnerabilities?	X				
Deep packet inspection?					
Data center firewall?	X				
Application-layer firewall?	X				
XML security support?	X				
Application Availability and Uptime					
Failover between virtual devices?	X				
Failover between application switches?	X				
Failover between data centers?	X				
Solution Architecture					
Specialized hardware instead of general-purpose hardware?	X				

Optimized Server Operations	Cisco ACE provides 80% additional server processing capacity**				
SSL encryption offloading?	X				
TCP server offloading?	X				
Module and Appliance Form Factors					
Modular form factor available?	X				
Standalone 1RU form factor available?	X				
End-to-End Professional Services and Support					
24-hours-a-day global support available?	X				
Award-winning professional services available?	X				

* According to Miercom report.

** Based on Cisco internal IT department use.



Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV
Amsterdam, The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at www.cisco.com/go/offices.

CCDE, CCENT, Cisco Eos, Cisco Lumin, Cisco Nexus, Cisco StadiumVision, the Cisco logo, DCE, and Welcome to the Human Network are trademarks; Changing the Way We Work, Live, Play, and Learn is a service mark; and Access Registrar, Aironet, AsyncOS, Bringing the Meeting To You, Catalyst, CCDA, CCDP, CCIE, CCIP, CCNA, CCNP, CCSP, CCVP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Cisco Unity, Collaboration Without Limitation, EtherFast, EtherSwitch, Event Center, Fast Step, Follow Me Browsing, FormShare, GigaDrive, HomeLink, Internet Quotient, IOS, iPhone, iQ Expertise, the iQ logo, iQ Net Readiness Scorecard, iQuick Study, IronPort, the IronPort logo, LightStream, Linksys, MediaTone, MeetingPlace, MGX, Networkers, Networking Academy, Network Registrar, PCNow, PIX, PowerPanels, ProConnect, ScriptShare, SenderBase, SMARTnet, Spectrum Expert, StackWise, The Fastest Way to Increase Your Internet Quotient, TransPath, WebEx, and the WebEx logo are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the United States and certain other countries.

All other trademarks mentioned in this document or Website are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (0805R)