



White Paper

The Optimized and Accelerated Cloud

As more organizations begin moving applications into the cloud, congestion will become an increasingly critical issue. F5 offers solutions for optimizing and accelerating applications in the cloud, making them fast and available wherever they reside.

by Alan Murphy

Technical Marketing Manager



Contents

Introduction	3
<hr/>	
Optimizing the Cloud	3
Virtualization—One Small Step for Optimization	3
Cloud Optimization: More Than Resources	5
<hr/>	
Accelerating the Cloud	6
Speeding Things Up	6
Moving Applications Into the Cloud: Wait Your Turn	7
<hr/>	
Conclusion	8



Introduction

It would not be a stretch to say that a move toward cloud computing is on the list of many IT departments business goals for 2010. Yet analyst reports, press articles, and blog posts about the proliferation of cloud computing in the tech industry reflect different opinions on the actual impact that cloud computing will have on enterprise IT. Building a cloud solution will have effects ranging from saving a business money all the way to changing the way IT functions today. And, in spite of the positive or negative impact on IT, modern cloud computing is definitely here to stay.

One of the key challenges facing early cloud adopters is how to architect a cloud solution that is optimized and designed for their particular environment and their specific application needs. Today it's become easy to choose a cloud provider for an external cloud, or buy off-the-shelf solutions to build an internal cloud, and start using new services right away. However, using those services in an efficient manner—one that satisfies the business needs with this new technology—is not nearly as simple or straightforward. Designing and deploying an optimized cloud solution is a cross-platform and cross-technology endeavor.

F5® offers solutions to manage application availability in the cloud as if those applications were installed in the local data center. Using products and technology that focus on the applications themselves rather than re-architecting the network, F5 can extend application optimization and acceleration into any cloud solution, be it internal or external, public or private, hosted or managed.

Optimizing the Cloud

Virtualization—One Small Step for Optimization

Although not required for cloud deployments, virtualization is often the first step many IT departments take toward building their cloud solution. Virtualization provides a much more agile and dynamic platform in the data center from which cloud components, such as PaaS and SaaS, can be built. Even if IT departments are using virtualization solely for physical server replacements, moving virtual machines around the data center—a feature offered by most enterprise-class virtualization platforms—provides critical progress toward agility and creating an internal cloud.



F5 offers application delivery solutions designed to fit in the enterprise virtualization topology. F5 BIG-IP® Local Traffic Manager™ (LTM) Virtual Edition (VE)—F5’s first virtual Application Delivery Controller (ADC) solution—is a full-feature BIG-IP LTM virtual machine that runs on VMware’s enterprise-class ESX 4.0 virtual platforms. BIG-IP LTM VE can be packaged with individual, application-specific virtual machine bundles on ESX servers and can move around the data center (or outside to an external provider) with that application bundle. BIG-IP LTM VE provides the same level of optimization and availability to those application as traditional BIG-IP LTM, regardless of where the apps are running. Like BIG-IP LTM VE, BIG-IP hardware appliances also plug into virtual management suites from VMware and Microsoft to extend their virtual platform management into the Application Delivery Network (ADN) and vice versa.

Cloud providers, on the other hand, almost exclusively build their elastic computing cloud services on top of virtualized server platforms. The same flexibility virtualization brings to the enterprise data center is felt exponentially by cloud providers as they are provisioning and de-provisioning thousands of computing resources every minute for their customers. With BIG-IP LTM hardware appliances, customer application traffic can be managed at a granular level by the cloud provider. The provider can provision network resources as computing and system resources are dynamically deployed to match each customer’s requirements. When used with BIG-IP LTM hardware, BIG-IP LTM VE extends control of the ADN into the virtual network and virtual platforms, both in the enterprise data center and in the cloud.

The true benefit of virtualizing each end of the data center spectrum is seen when enterprise customers opt to move entire application services outside the primary data center. When building out secondary or more data centers, virtualization enables systems and applications to be easily—and often transparently—moved to those new data centers. Once the decision to use external cloud providers for application services has been made, IT departments can easily move virtualized systems and applications to their cloud providers as needed, based on either resource and business needs (as with an architecture change) or on immediate need (as in a disaster recovery scenario).

With products such as BIG-IP® Global Traffic Manager™ (GTM), BIG-IP® Edge Gateway™, and BIG-IP® Access Policy Manager™ (APM), F5 helps customers move their application assets outside the enterprise data center by enabling them to maintain control over their public cloud resources. IT departments have the

Leaving the Data Center

61 percent of IT organizations consider it important that cloud providers have similar BIG-IP infrastructure to their current infrastructure.

Source: TechValidate
TVID: FD7-FE4-9F4



flexibility to route traffic and control user access to applications in the cloud as if those applications were residing in local data centers. F5 provides the flexibility to dynamically route user and application traffic between internal and external locations without requiring expensive hardware purchases or leases in the cloud.

Cloud Optimization: More Than Resources

Despite focusing on computing resources such as CPU, RAM, and other hardware, building an optimized cloud isn't just about virtualizing hardware resources. Much like any data center build-out, every part of the application service stack—servers, applications, the network, user access to services, and more—needs to be optimized for a cloud to be adaptable, extensible, and manageable. More so than with a traditional data center, managing application optimization is a critical and required component of any cloud deployment. By design, the elastic nature of a cloud has many more dynamically moving parts than a physical data center—parts that need to be tracked and managed in real time.

Consolidating Service Resources

Consolidation is essential in optimizing any data center deployment. Over time, it is easy for data centers to become unmanageable, especially as more and more hardware is used to try and solve application issues. Consolidation is one of the key reasons virtualization became so prevalent within the enterprise data center. For the cloud, consolidation comes from both trimming down IT assets in the local data center and from using on-demand resources in the cloud.

F5's deep application delivery product base helps IT departments consolidate their application resources inside and outside the primary data center. Simple load balancing is important when consolidating application resources because it helps reduce the number of physical or virtual machines required to deliver an application to a user. Advanced load balancing is also critical for on-demand resources, as it helps guarantee that only the resources needed at any given time are being used.

Together, BIG-IP GTM and BIG-IP LTM take advanced load balancing with dynamic resources further by managing automated provisioning with virtual platforms. As CPU and RAM resources are needed by a particular application, BIG-IP LTM notifies the virtual management platform that a specific application is requesting those resources, and only those necessary application-specific virtual machines are spun



up to provide those resources. Likewise, when resources are no longer needed by BIG-IP LTM, those same virtual machines are spun down and the hardware resources are put back into the cloud pool. Those services can exist within the local data center or externally, in a remote or cloud-based data center.

Accelerating the Cloud

Speeding Things Up

One of the initial drawbacks to moving mission-critical applications into the cloud has been maintaining speed—guaranteed application response time and available bandwidth—typically required with user-based applications. When an application is in-house, the enterprise IT department has control over the application and, therefore, has the flexibility to make necessary changes. Optimizing the LAN network, optimizing WAN connections, and even offloading application services to BIG-IP LTM and the BIG-IP® WebAccelerator™ product are all options available to the IT department to help speed up applications when they are in the data center.

Many of those acceleration options are not available once an application is moved into the cloud. The enterprise IT department has little control over (and often little visibility into) the LAN conditions at the cloud provider's hosted data center. This lack of control has always been an important issue with traditional hosting provider networks, but it becomes more critical with cloud providers because of the nature of dynamic provisioning. There's no guarantee that another cloud customer on the same LAN won't spin up 1,000 virtual machines, pushing massive amounts of application traffic over the LAN all at once. In fact, this specific scenario has led to criticism of cloud providers, particularly when malicious customers have used the flexibility and limitless computing power of the cloud to launch denial-of-service attacks.

To address these challenges, focusing on application acceleration, rather than network acceleration, is important with cloud-based application deployments. Since control over the provider's LAN will never be an option, implementing BIG-IP LTM VE can be critical in taking back control over a particular part of the shared LAN. In the data center, IT can configure BIG-IP LTM VE for applications that are about to move to the cloud and then provision BIG-IP LTM VE with those applications in the cloud data center, managing the ADN for applications even when they are in the cloud. Deploying BIG-IP LTM VE into the cloud along with



the applications will enable a known set of performance standards and configurations, such as application availability SLAs, to be provisioned in the cloud and will bring some level of application control over the remote network back to the enterprise. BIG-IP LTM VE is just one example of maintaining as much control over applications as possible, even once they've moved into the cloud. Managing application delivery over core components of the cloud-based network with BIG-IP LTM VE enables the enterprise IT administrator to guarantee application access times and helps control speeds and feeds into and out of the remote cloud network.

Moving Applications Into the Cloud: Wait Your Turn

Another concern with moving applications into the cloud is how to actually get those applications to the cloud provider and have them production-ready, online, and accepting connections as soon as possible. Many cloud providers permit creating new systems—from the OS all the way through the application—from scratch, but not all of them support bringing in pre-built applications and/or virtual machines (VMs). For those providers that do offer a “bring-your-own” option, the only way to move those virtual machines into the cloud are via methods such as FTP or even sneakernet.

Moving large virtual machines over the public internet is far from efficient. Applications typically require more than just one virtual machine to operate; they may have multiple services spread across multiple VMs (the standard tiered application approach still holds true for VMs). VMs will often also include some type of data store, be it local storage, which can drastically increase the size of the VM, or a more traditional database—either of which add complexity to moving applications into the cloud. This dilemma holds true for internal data centers as well. As new virtualized data centers are created throughout the enterprise, moving virtual machines among these data centers still carries the challenge of efficiency.

For both internal and external virtual data centers, BIG-IP LTM can help drastically increase network performance and transfer speeds of high-quantity data such as VM images. Internally, BIG-IP LTM can bridge multiple data centers, creating an optimized and secure tunnel for all IP-based traffic. In fact, BIG-IP LTM can provide the same degree of bit-level optimization across the private iSessions tunnel as it does when optimizing application traffic, including TCP optimization, caching, compression, and even data de-duplication. If multiple similar VMs (or even identical



VMs in the case of “gold” base images that never change) are moved between data centers, BIG-IP LTM de-duplication can drastically diminish the amount of data flowing through the tunnel.

This same model holds true for external cloud providers. If an enterprise chooses a cloud provider that extends BIG-IP LTM as part of its service offering, a customer can configure the same private dedicated tunnel between its cloud provider and its own local data center. An enterprise can then use this tunnel as a transfer device for optimizing large data transfers as needed, or it can leave this tunnel in place and use it for day-to-day application traffic, such as what would be required for cloudbursting.

Another benefit of linking multiple data centers together with BIG-IP devices is traffic management. This model gives IT full control on where application connections are created and terminated. Using BIG-IP GTM, application users may be directed to the internal data center for authentication and authorization and then either re-directed to the cloud provider via BIG-IP GTM or re-routed through the primary data center to the cloud provider using BIG-IP LTM.

Conclusion

Cloud computing is an interesting change in the data center because it simultaneously offers huge benefits and brings huge risks to enterprise IT. The benefits come from providing a much more dynamic and agile computing environment for any application, as well as helping IT move expensive assets outside its own data center. Asset movement is risky however, because it also means moving control of those assets to a third party, and, to a certain extent, shifting responsibility for traditional application performance issues such as response time, bandwidth usage, and availability.

F5 helps enterprise IT maintain as much control as possible over its applications once they are moved into the cloud. By providing core application availability and optimization solutions to the cloud, F5 products enable IT departments to build and deploy the same application delivery tools they use in their own data centers. Application availability and optimization is maintained with both physical and virtual instances of BIG-IP LTM (the former on hardware, the latter with BIG-IP LTM VE running on VMware) and modules such as BIG-IP WebAccelerator.

White Paper

The Optimized and Accelerated Cloud

Advanced solutions such as iSession tunnels and data de-duplication handle WAN optimization; BIG-IP LTM and BIG-IP GTM maintain availability as applications move in and out of the cloud and between data centers.

It is critically important to factor in application performance and availability when moving to a cloud-based environment. Most of the benefits of cloud computing are lost if and when the applications begin to suffer. F5's advanced application awareness across all products can help maintain application standards—even if the applications are floating around in the cloud.

F5 Networks, Inc. 401 Elliott Avenue West, Seattle, WA 98119 888-882-4447 www.f5.com

F5 Networks, Inc.
Corporate Headquarters
info@f5.com

F5 Networks
Asia-Pacific
apacinfo@f5.com

F5 Networks Ltd.
Europe/Middle-East/Africa
emeainfo@f5.com

F5 Networks
Japan K.K.
f5j-info@f5.com

