White Paper

Application Delivery in the Cloud: Minimizing Disruption and Maximizing Control

Whether serving applications in a traditional or cloud environment, IT organizations want to do so with minimum user disruption and maximum control and manageability. F5 solutions support and enhance application delivery in the cloud, such as in VMware-based solutions, just as they do in traditional environments: by ensuring security, availability, and scalability.

by Simon Hamilton-Wilkes
Solution Engineer - VMware Alliance

Contents

Introduction

Cloud computing offers the promise of IT resources on demand with minimal management and capital overhead by enabling real-time scaling, rapid application deployment, and IT agility. It delivers on this promise by building on server virtualization, now a widespread technology. Cloud computing enables organizations to isolate multiple entities or tenants, and provides an interface for virtual machines and applications to be self-provisioned, with an organization being billed only for the resources it uses. As businesses flock to the cloud, the IT administrators that facilitate its adoption want uncompromised security, successful migrations, and total flexibility—all without increasing complexity or having to do more with fewer resources.

Recent breaches and unplanned outages have attracted a great deal of negative publicity and incurred significant costs for some organizations; this has brought the need for IT security into sharp relief. Application downtime, whether caused by such a breach or a more traditional outage, still results in a slew of negatives: lost productivity, damaged reputation, financial costs, and more. To keep these hindrances at bay, both security and high availability are critical to deploying enterprise-class applications in a cloud model.

However, the need for security is often in conflict with organizations' desire to migrate workloads to cloud services. In cloud services, internal data and applications that are usually contained within the corporate network perimeter reside instead on third-party systems and frequently need to be accessible from the public Internet. Diverse attacks intended to manipulate vulnerabilities in web pages, applications, and operating systems, such as distributed denial-of-service (DDoS) attacks, can affect legitimate users' access or result in stolen or modified data.

To successfully adopt a cloud model, IT organizations must be agile and flexible enough to achieve application migration. This process—and cloud computing itself—is quite new to most enterprises. Most of today's cloud environments are the result of new enterprise initiatives to deploy a private cloud, and to experiment with migration in the form of hybrid cloud models and cloudbursting. As the pace of cloud adoption accelerates, a high degree of flexibility will be key to architecting environments that provide for highly secure and available application delivery and migrations.

F5 products are the application glue that seals the gaps in cloud computing for the enterprise. They offer strong integration with many cloud platforms, such as VMware and its industry-leading VMware vSphere product suite that provides the hypervisor layer for data center server consolidation. It also offers cloud services abstracted from one or more vSphere environments through VMware vCloud Director.

F5 provides the automation and flexibility IT organizations need in their infrastructure to reap and maintain the benefits of a cloud initiative. Integrating F5 products with a cloud deployment such as VMware ensures security, scalability, and manageability across environments and systems.

Cloud Migration

F5 can address anything from a big bang migration to just the migration of a single application to the cloud with BIG-IP® Global Traffic Manager™ (GTM), which provides multiple-site, context-aware load balancing between data centers, whether conventional or cloud-based. This enables the hybrid cloud model, in which applications can be active/active in both environments, active/passive to provide disaster recovery and business continuity, and cloudbursting.

Cloudbursting is a form of migration that involves provisioning private cloud capacity for median load rather than peak. When traffic spikes—for instance, during a new product launch—application workloads can be dynamically provisioned into an external cloud, and organizations can use intelligent traffic management to distribute and share the traffic load between each cloud accordingly. BIG-IP GTM works in conjunction with BIG-IP® Local Traffic Manager™ (LTM) instances deployed in each location; together they provide clear visibility of overall application traffic load. This coordination not only enables the application to burst capacity intelligently out to a second cloud infrastructure, but it provides administrators with the intelligence they need to recognize subsiding traffic and facilitate deprovisioning when the surge has ended.

In addition to traffic characteristics, bursting events can take into account other business logic, such as different providers' billing models, or be primed to pre-stage capacity prior to a big announcement. F5 products can granularly control and dynamically manage traffic during migrations with context such as application user identity, geolocation, and specific groups of users. This flexibility complements VMware's vCloud Connector, which addresses the steps involved with moving VMware vApps, VMs, and templates, but doesn't address end user application traffic.

In addition, the BIG-IP® WAN Optimization Manager™ (WOM) module can provide faster TCP transport between clouds or to the corporate network with TCP optimization, caching, and deduplication, ultimately yielding significant savings in bandwidth costs and transport times.

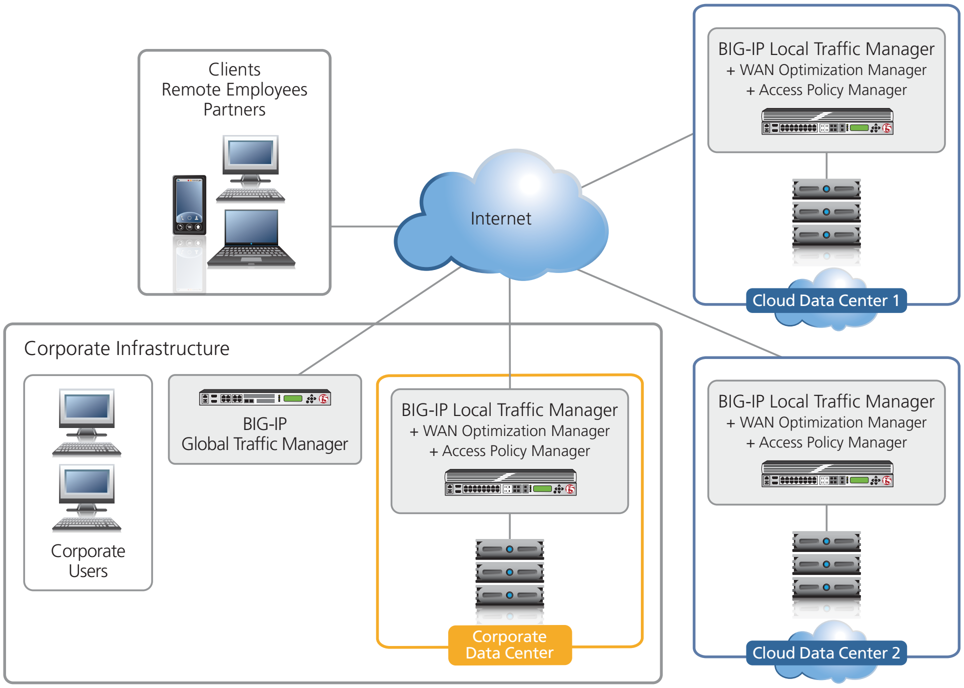


Figure 1: F5 products improve the efficiency and security of a cloud environment.

Solving Cloud Security Challenges

While the cloud often seems to be an ephemeral thing, it is actually a contained network environment with traditional perimeters just like the traditional corporate LAN or DMZ. The main networking difference between a cloud built on vCloud Director and a traditional private data center environment is that rather than using virtual LANs (VLANs) to segment IP subnets or security zones, multiple tenants are separated by an encapsulation protocol used at the physical host layer. Most cloud platforms offer such isolation through encapsulation, and in the case of vCloud Director, VMware uses vCloud Director Network Isolation to ensure guest virtual machines have no access to other organizations' traffic.

Within this secure logical organization, administrators need to pass and control traffic between application tiers and security zones, and between the cloud environment, the Internet, and physical corporate sites. Cloud software can supply tools for this, though they may have limited feature sets and require different management skills than those used at a corporate level.

The BIG-IP system alleviates this difficultly. It runs on F5's TMOS® operating system, providing strategic points of control in the network, and is a high-performance application layer proxy. The BIG-IP system performs the application load balancing, caching, acceleration, and security that are instrumental to successful application delivery.

Key BIG-IP system features include an optimized application (layer 7) proxy, and extensible and open control-plane and data-plane APIs. The application layer proxy terminates the TCP connection from the client and sets up a separate session to the server, which offers full isolation of the network session. This allows for far more than just the parameter verification used in stateful firewalls, as the session is examined at layer 7 rather than just layer 4. In the 1990s, a T1 was considered high speed, and hardware could easily keep up. Connection speeds then exceeded hardware performance, so the market transitioned to stateful inspection at layer 4. But now a virtual appliance can achieve wire speeds at the gigabit level, and hardware appliances can achieve the fastest Internet connection speeds.

BIG-IP LTM is a Certified Network Firewall by ICSA Labs, an internationally recognized independent evaluator of security solutions. In addition to performing as a stateful firewall, BIG-IP LTM provides application networking intelligence. Its ability to handle a high connection count helps combat today's more intelligent attacks, which are often based on exhausting application or firewall connection counts rather than raw bandwidth.

To address web security, organizations can deploy BIG-IP® Application Security Manager™ (ASM)—a web application firewall (WAF)—to mitigate application layer attacks, such as SQL injection, cross-site scripting, and even application misconfiguration. A WAF is required to comply with PCI and HIPAA standards, and recommended by VMware in the vCloud Director reference architecture for its web portal. While BIG-IP ASM downloads attack signatures on a daily basis, it does not rely on them like a conventional intrusion prevention system (IPS), as its position within the traffic flow enables it to validate protocol compliance and filter out abnormal activity. In a learning mode, it can discover what "normal" should be for a given environment, and whenever application changes are made.

When vulnerabilities in the server operating system or applications are discovered, organizations can use this layer 7 proxy functionality to virtually patch exploits on the fly—providing a rapid solution to problems in production until (or instead of) lengthy development and testing of a permanent fix can be completed. This type of response can be manual, or automated as part of a feedback loop with an external security scanning service or software. A similar feature can be used to examine outgoing traffic, which addresses data loss prevention security concerns by examining traffic for patterns that correspond to certain data types that should not legitimately be traversing the firewall (credit card and Social Security numbers are prime examples). When it observes this type of data, BIG-IP ASM can modify that content on the fly to either be masked or blocked, as well as trigger an alert.

If cloud-hosted applications are not public-facing, but rather for an organization's internal use only (or a combination of the two), secure remote access will be required. BIG-IP® Access Policy Manager™ (APM) offers this remote access through a rich framework of authentication options and endpoint inspection, in addition to SSL VPN for remote users, regardless of their device type. With the Visual Policy Editor, administrators can create complex policies in an easily understandable and configurable visual policy flow chart. With BIG-IP APM, organizations also have the option to provide a custom webtop portal, as well as support for VMware View for VDI.

Another common security concern is maintaining control and ownership of the enterprise directory, and most are not willing to deploy a replica in a third-party environment for authentication. BIG-IP APM uses a federated authentication model, in which BIG-IP LTM redirects non-authenticated sessions back to the corporate site for authentication, then permits the returning request once the session has been successfully authenticated by BIG-IP APM at the home site. This protection could be extended to many types of sensitive data that a company prefers not to hold a copy of in the cloud, by maintaining and controlling access only through the primary site, with a WAN-accelerated and encrypted tunnel between the two sites.

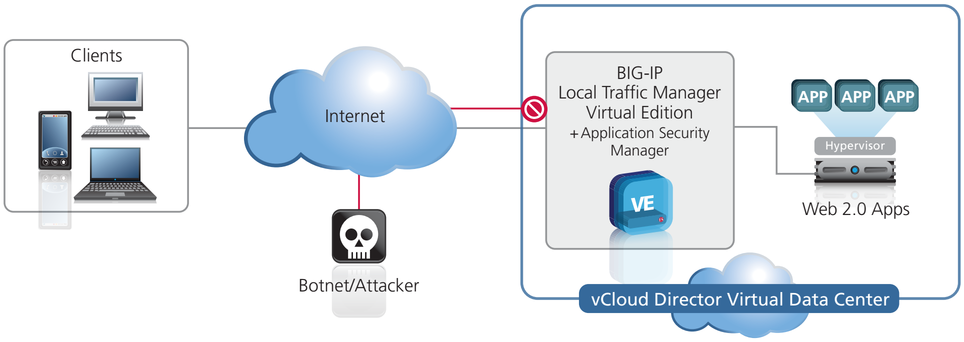


Figure 2: BIG-IP virtual editions provide firewall and application security to tenants, in addition to application delivery.

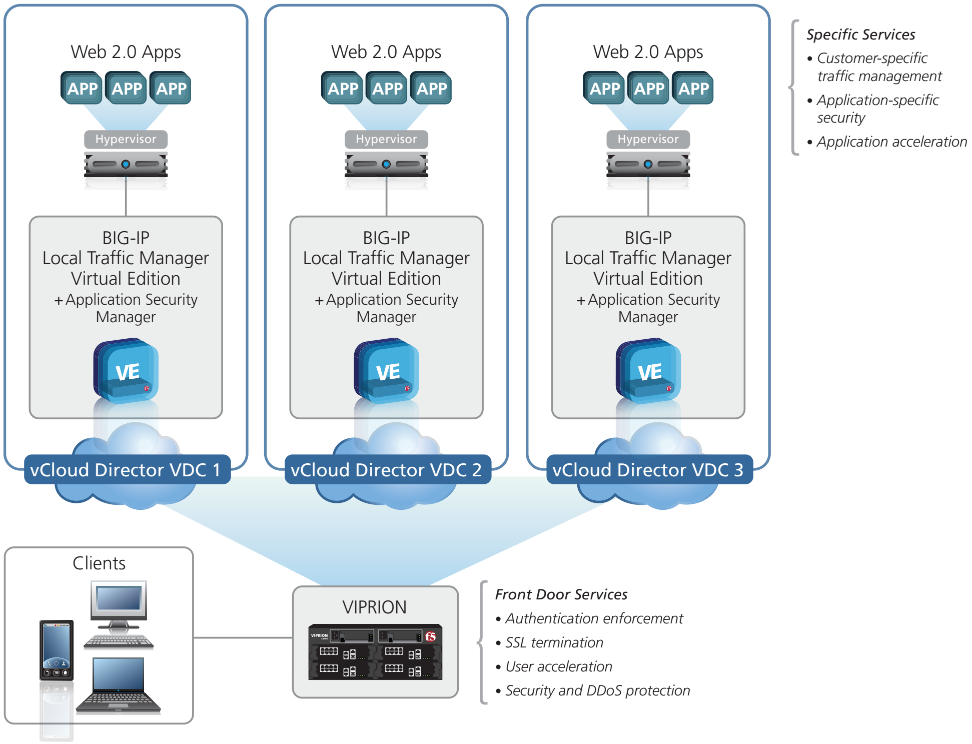
Flexibility in the Dynamic Data Center

One of the most unique aspects of cloud computing is the constant allocation (and reallocation) of a data center's underlying infrastructure resources in response to business demands. These changes must be accommodated without affecting security and compliance, and day-to-day operations must be straightforward enough that an administrator can perform them accurately.

The complexity and dynamism of a cloud environment demands more automation and better orchestration than a conventional data center. In anticipation of this shift, F5 and VMware both provide a range of APIs for orchestration and automation engines. Using F5's iControl® API, administrators can automate the control plane, allowing for system configuration and modification. The F5 iRules® scripting language offers full event-driven access to the data plane. Administrators can use it to perform deep traffic analysis, manipulation, and customization of application traffic.

Similarly, VMware products such as VMware vSphere and VMware vCloud Director provide API interfaces. Additionally, VMware vCenter Orchestrator enables administrators to develop complex operational workflows that leverage the APIs of their organizations' existing solutions, as well as third-party solutions, including F5. vCenter Orchestrator is bundled with vCenter, and is an extensible and key part of a VMware-based private cloud. Many third-party modules can be added to control storage and network components, and to permit input and output from workflows to ticketing systems and email. Role-based access control goes hand in hand with logging to provide an audit trail, and instrumentation and analysis are available for capacity planning as well as troubleshooting when problems do arise.

This ability to log, monitor, and analyze infrastructure also applies to security, where such analytics can be used to spot anomalies in traffic, and identify application performance issues, problems with quality of service, and service level agreement violations. Programmatic APIs are the key element to all of this functionality. Not only do they enable administrators to report on statistics locally, but also to export them into other systems for correlation, root cause analysis, and more accurate automated response and alerting.



Specific Services
- *Customer-specific traffic management*
- *Application-specific security*
- *Application acceleration*

Front Door Services
- *Authentication enforcement*
- *SSL termination*
- *User acceleration*
- *Security and DDoS protection*

Figure 3: Example of a hybrid deployment model with VIPRION® chassis and BIG-IP virtual editions.

Service Providers

Thus far, our cloud perspective has been that of a single organization or tenant of a public cloud service, deploying services specific to their own organization. A typical deployment model in this case would include BIG-IP LTM Virtual Edition (VE), running on a virtual machine within the cloud. Both the physical and virtual editions of BIG-IP LTM function as a network firewall and an Application Delivery Controller (ADC); in a bare-bones tenant architecture, BIG-IP LTM enables tenants' administrators to control and log traffic from a single appliance (or a high availability pair).

Organizations can use this same virtual appliance to achieve load balancing and acceleration for their applications, between application tiers, and for VLAN separation of multiple application traffic. More often, however, a tenant architecture would include a pair of virtual appliances performing internal traffic optimization and direction running just BIG-IP LTM, while another pair on the edge of the cloud environment would include the BIG-IP ASM module atop BIG-IP LTM for web application firewalling, virtual patching, and data loss prevention. The exact architecture and sizing for any given environment is dependent on the bandwidth and connection handling required. As capacity requirements grow, administrators can deploy hardware devices, particularly at the edge, and gradually scale as needed. It may not be possible for cloud tenants to deploy their own hardware within the cloud provider's data center, in which case renting from the provider may be the solution. Hardware at the edge is particularly valuable where DDoS handling capacity may be required.

The F5 VIPRION® chassis-based platform provides the ultimate performance for the most demanding environments, and enables cloud providers (or larger-scale enterprises) to run multiple instances of BIG-IP LTM on a single hardware device. This functionality, called Virtual Clustered Multiprocessing (vCMP), allows for multi-tenancy. Each tenant or BIG-IP LTM instance can be assigned hardware resources, fully isolated from each other, to meet the highest throughput, security, load balancing, web firewalling, and traffic management demands. Providers can leverage their own instance to protect the administration portal, and to support applications such as billing.

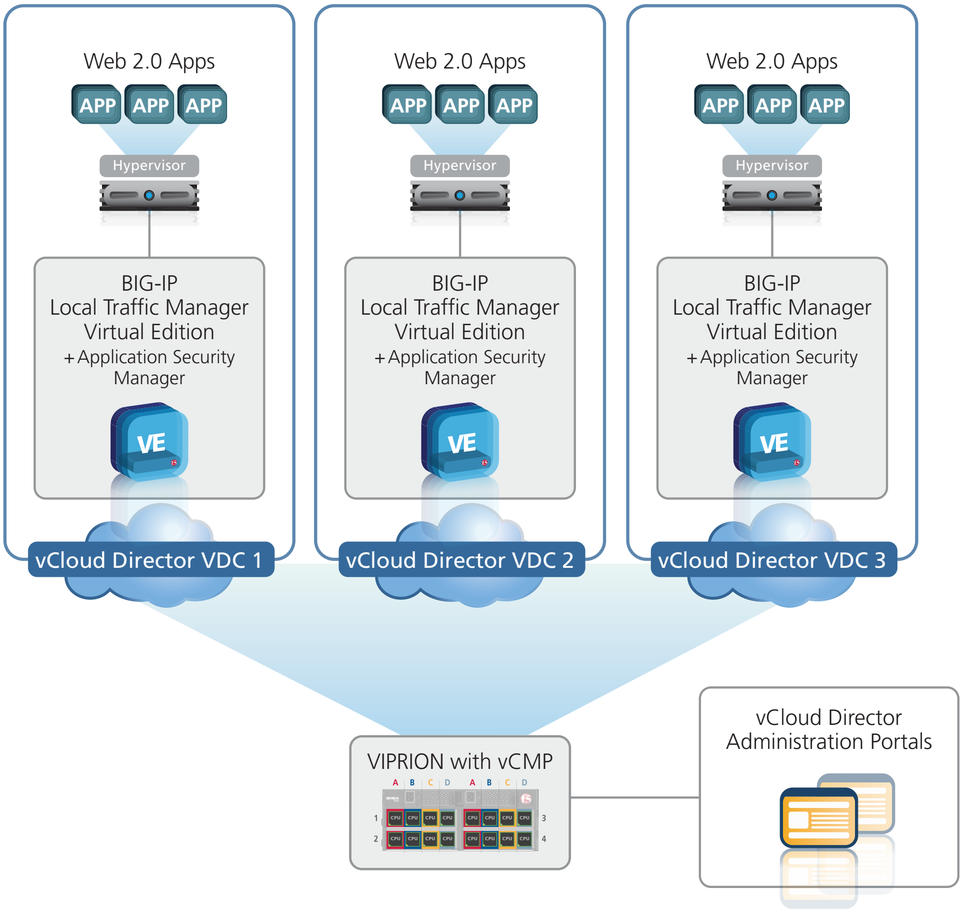


Figure 4: A VIPRION chassis with vCMP protects the administration portal and high-traffic customers.

Conclusion

F5 enhances cloud environments and applications with a range of functionality focused on superior application layer performance. Whether in physical or virtual form, F5 products seamlessly scale to accommodate all of an organization’s capacity needs. F5 products provide complete open APIs that enable automation, migration, security, and scalability, as well as increased visibility, into cloud-based applications. Together, F5 and VMware solutions for the cloud provide optimum user experience and superior cost benefit.

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