Migrating Tier 1 Application Workloads to AWS with F5

Enterprises deploying applications on AWS may have concerns about performance and security. The F5 BIG-IP platform alleviates these with high availability, performance, and security.

Introduction

Enterprises seeking to migrate or deploy production workloads in the public cloud are most likely to adopt Amazon Web Services (AWS) as their preferred platform. According to Gartner, the deployment of AWS compute capacity is five times that of the 14 other members in their magic quadrant combined.[1] As the public cloud platform service of choice, it is also the most likely cloud platform to be selected by enterprise customers for their mission-critical applications.[2] This paper will review many of the identified inhibitors to broader adoption of the public cloud for Tier 1 applications, identify potential solutions, and suggest how advanced Application Delivery Controller (ADC) technologies are a critical element of broader cloud adoption.

Rapid, Flexible Deployment of Infrastructure and Services

Tier 1 applications are not static resources with clearly defined performance parameters. Enterprises may unexpectedly need to adapt to spikes in demand for a particular application. Traditional IT organizations respond to this demand by provisioning infrastructure and application resources to meet peak rather than average demand. This type of resource overkill was necessary because of the inflexibility of IT services and the need to invest in fixed assets. Additionally, procurement and deployment of new infrastructure, networking, and associated application resources is a time-consuming process, so unexpected changes in usage patterns would need to be planned for well in advance. Attempting to provision for average demand rather than peak left companies vulnerable to the financial hits of outages, reduced demand, and poor user experience. The promise of the cloud, whether private or public, is the flexible, on-demand allocation of resources to address spikes in demand.

Many enterprises are already exploring or have implemented private cloud solutions to provide pooled resources for dynamic allocation—although at a substantial cost. Public clouds provide even greater benefit to an enterprise, enabling greater flexibility in responding to specific application demands without requiring a large reserve pool or resources.

The challenge in adopting a hybrid private-public cloud model is the federation of networking, access, and application resources between locations. The traditional perception is that a public cloud model exposes the application directly to the Internet and that certain application workloads, for example public-facing websites, are safer candidates than others for deployment in public clouds. The migration of internal applications is perceived as riskier.

Flexible IT Infrastructure Model

In many enterprise environments, pre-deployment testing and validation require large, parallel test and development labs—considerable and ongoing expenses in infrastructure, software licenses, and manpower. Moving many of these workloads to the AWS cloud makes compelling business sense for customers. AWS rapidly spins up application resources for testing and development and then recovers these assets when no longer needed, providing an enterprise customer a great deal of flexibility while reducing infrastructure overhead and potentially leading to substantial cost savings.

Realize Economies of Scale

An enterprise has substantial purchasing power in the procurement and allocation of software and hardware resources, but the financial benefits of a proven infrastructure model like AWS's are obvious. AWS has already made the necessary investments in infrastructure and software assets to enable an enterprise to benefit immediately from an ever-declining infrastructure cost curve. Private cloud deployments may be a necessary interim step in adopting public cloud infrastructures in that they provide much of the same flexibility, rapid provisioning, and pooled resources to allow enterprises greater flexibility in how to support varying workloads. The efficiencies and rapid deployment of application services that a private cloud provides can be easily expanded to the AWS cloud.

Why Enterprises Are Hesitant to Move Mission-Critical Applications to the Public Cloud

Performance and Scale Concerns

There is a widespread misperception that application workloads will simply not perform as well in the public cloud as they do in the enterprise data center. As more and more large businesses successfully deploy their mission-critical applications in the AWS cloud, that idea is fading. It stems from the idea that the AWS cloud is subject to the performance limits of fixed external assets, as well as limits explicitly placed on individual resources in the cloud service. Addressing these concerns is essential to migration—and, more important, to initial deployment of application workloads into AWS. Features such as auto-scaling can address these concerns by enabling the application and its associated resources to scale horizontally to address unexpected traffic spikes.

Cloud application deployment of mission-critical applications is further hampered by the perception that cloud service level agreements (SLAs) include availability guarantees that are lower than enterprise IT organizations require. The stated SLA for Amazon EC2 is 99.95 percent which, although lower than the ideal five nines availability desired by many enterprises, is still best in class among cloud service providers. The key point in evaluating current IT SLAs is to take a look at current SLAs defined for the application; determine the business impact of an outage; and find ways to mitigate the potential effect of an application outage. The 99.95 percent SLA provided by AWS means just slightly more than four hours of downtime per year. The key here is to weigh the costs and benefits of moving to a public cloud infrastructure against the effect of the 99.95 percent SLA.

Security

The perception that cloud application services are insecure is another significant impediment. Concern about application resources being accessible over the Internet is a major concern with regard first to adoption of cloud resources, and then to moving primary workloads to the cloud. Addressing security threats such as malware, cross-site scripting, and DDoS threats are critical to deploying a Tier 1 application in the AWS cloud.

How F5 Enables Adoption of AWS for Tier 1 Applications

Availability and Performance

Enterprise customers view their applications as critical business assets responsible for generating revenue, improving employee productivity, or ensuring customer satisfaction. Unexpected outages or a poor user experience can have a direct effect on the revenue potential of organizations and their ability to respond to the challenges inherent in the deployment of an application. The work necessary to address these concerns is an essential part of the overall deployment of an application, and it affects the movement of resources to the public cloud. Redirection and modification of traffic between the data center and the AWS cloud is an important part a successful cloud implementation; however the federation of application resources between the data center and AWS cloud can be perceived as a complex undertaking with items like identity management, automated resource provisioning, and seamless redirection of users to cloud resources.
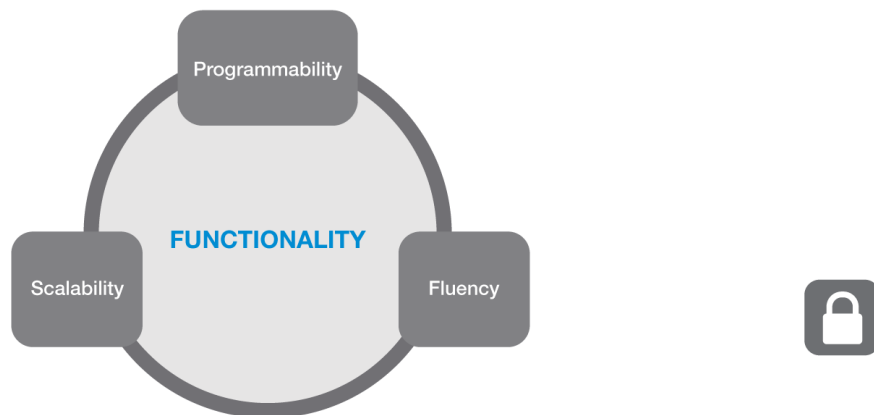

Figure 1: F5 provides a fast, available, and secure experience in the AWS cloud.

Scalability to Meet Application Performance Requirements

A key benefit in moving application workloads to public cloud platforms is the ability to scale the application beyond provisioned capacity in the data center. Historically, in order to meet occasional spikes in application demand, enterprises would need to design their application infrastructure to meet this peak demand. Operations like cloud bursting can enable an enterprise to scale their application resources in an on-demand basis and pay only temporarily for the resources utilized during these occasional spikes. To effectively burst to the AWS cloud, a number of very tightly choreographed operations need to occur so that this operation is transparent to end users:

- When a pre-set threshold is reached, resources will be provisioned and deployed in the cloud.
- The new application server resources are automatically added to the pool of resources.
- Users are redirected to the application resources in the AWS cloud.
- When application loads fall below the pre-set threshold, the cloud infrastructure resources are deprovisioned and all connections are redirected back to the data center.
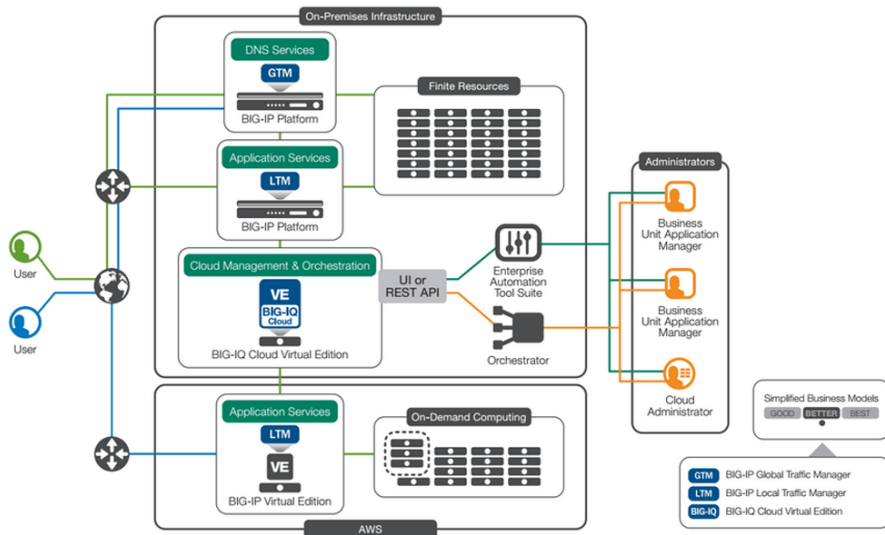

Figure 2: F5 provides a simplified transition between the data center and AWS.

Migrating Tier 1 Workloads

Tier 1 enterprise applications require features such as persistence, acceleration, an advanced traffic decision engine (such as the F5 iRules scripting language), and unified access. Successful application migration ensures that this functionality is preserved during the migration of any critical application. The best way to ensure that it remains intact is to rely on an industry-proven application delivery architecture. When an enterprise IT environment migrates a workload to the AWS cloud, F5 BIG-IP Application Delivery Controllers provide the aforementioned elements that ensure a successful public cloud deployment.

Flexible Pricing Models

BIG-IP technology has been a part of the marketplace under the bring-your-own-license (BYOL) model for nearly a year. An option that many enterprises have chosen to deploy, BYOL provides a platform for those that need to deploy steady state, permanent application workloads in the AWS cloud without worrying about potential limitations presented by licensing. And for customers looking to deploy BIG-IP ADCs to evaluate and test pre-production workloads, F5 now offers an hourly pricing option. In addition to benefiting from the flexibility that hourly pricing provides for test and development environments, organizations can now use cloud bursting operations in a dramatically more efficient manner because there are no concerns over licensing BIG-IP virtual editions. BIG-IP virtual editions can now be activated to support intermittent spikes in demand, with customers paying for the use of the BIG-IP platform only when it is active.

Simplified Deployment

AWS Cloud Formation Templates (CFTs) are a scripted method of automating the deployment of all server, storage, networking, and compute resources in the public cloud. They provide a repeatable way to rapidly deploy the same configuration multiple times within AWS. Ensuring that the networking, server instance, and application components are provisioned properly and installed in a simplified and repeatable manner with scripted output that is customizable can aid in the deployment of a CFT. Similar to the functionality of CFTs, F5 provides iApps Templates so enterprises can deploy the application delivery services necessary to support their applications.

Federated Public/Private Cloud Deployments

Integrating public cloud resources seamlessly with the existing private cloud gives enterprises a much-needed on-demand platform to scale rapidly. The combination of BIG-IP ADCs and AWS CFTs creates an integrated cloud configuration that can be used to rapidly and transparently deploy additional application resources. Key advantages to this federated cloud configuration include the seamless redirection of application users, geolocation and acceleration technologies, and secure SSL VPN connections. The user experience remains unaffected whether application resources are delivered from a private cloud or public cloud. Transparent and continuous use of private and public cloud resources can be based on demand, on whether a project is new or already in place, or on the specific location of the requester. The F5 application delivery portfolio dynamically scales beyond fixed data center boundaries and enables a user to connect to an application regardless of location. The F5 BIG-IQ platform provides a management layer that enables organizations to provision BIG-IP services in the AWS cloud.
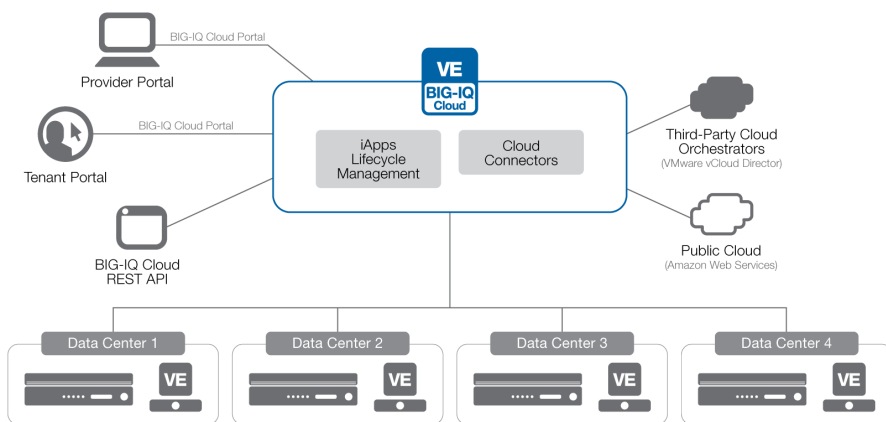

Figure 3: The BIG-IQ platform manages federated private and public cloud environments.

The BIG-IQ platform and its associated application services provide an enterprise with a comprehensive management platform that can enable rapid and repeatable application deployment in both private and public cloud environments.

Conclusion

Adoption of public cloud services has grown exponentially over the past few years, and AWS has been the consistent market leader for these services. Many IT startups and some mature media companies have even deployed entirely in the AWS cloud with significant success. As enterprises undertake plans to migrate more critical applications to the cloud, the proven benefits of application delivery using the BIG-IP platform can be easily ported to cloud application workloads. Doing so will address many of the fundamental concerns that enterprise customers have regarding greater adoption of public cloud. Now that flexible licensing models will be available for the BIG-IP platform in the AWS marketplace, enterprises can plan, stage, and deploy applications in AWS with minimal financial exposure and begin to execute strategies to benefit from the efficiencies of the public cloud.

---

[1] Source Gartner: Magic Quadrant for Cloud Infrastructure as a Service August 19 2013

[2] Source Gartner: Magic Quadrant for Cloud Infrastructure as a Service August 19 2013