Microsoft | Virtualization

F5 and Microsoft Delivering IT as a Service

F5's best-in-class, application-specific load-balancing and acceleration combined with Microsoft Windows Server 2008 R2 Hyper-V and System Center management provide the building blocks for a cloud-capable dynamic network

by

James Hendergart
F5, Business Development Manager

Jim Schwartz
Microsoft, Director, Windows Server Marketing

Published: December 2010

Contents

Executive Summary

To meet the constant demands to deploy, maintain, and grow a broad array of services and applications, IT organizations must continually evaluate the need to add new infrastructure. As organizations add more capacity to their infrastructure, capital and operational expenditures grow in proportion. In order to control these costs, improve business agility, and remain secure and in compliance, many of these IT organizations seek to transition to a truly dynamic virtualized, cloud-ready infrastructure.

Having realized the immediate benefits of virtualizing applications and consolidating servers, IT organizations are now taking a more holistic look at the entire datacenter. In order to fully optimize a dynamic virtual datacenter, strategic points of control must be orchestrated to uniformly manage compute, network, storage, and software resources. IT organizations must enable the system to automatically expand or contract system resources to optimize application load or resource demand through resource pooling and load balancing.

The more IT resources that are pooled within and across data centers, the more efficiency can be gained. As these pooled resources are shared across tenants, the management and security profiles for each tenant must be visible so management can be segmented and resources can be allocated properly. Because the ultimate goal is to be able to deliver IT as a service, the addition of self-service capabilities is an important step to creating a cloud-ready infrastructure.

Within a Microsoft® virtualized environment, Windows Server® with Hyper-V™ and Microsoft System Center are the foundation for the software and management infrastructure. They integrate with partner-provided storage and network infrastructure components to host and deliver business software applications that users need to accomplish their work. F5® extends the capability and value of a Microsoft virtualized environment by providing an intelligent, application-aware, and optimized network.

This paper provides a blueprint for integrating F5's BIG-IP® Application Delivery Controller (ADC) with Microsoft System Center to establish a dynamic network infrastructure. By making the network visible to System Center, System Center can automatically manage application load optimization policies.

This design enables System Center to automatically add or remove application virtual machine instances as application load thresholds are reached. In addition, this design enables F5's BIG-IP ADC to automatically adjust its load balancing to accommodate these changes. The net effect is that the application is always running at optimal load levels for users, maximizing performance and eliminating waste. The benefits of this design include:

- Application-aware load-balancing.
- Custom, application-specific traffic management.
- Increased availability of servers through offloading network-centric operations.
- Increased network visibility for better management decisions.
- Microsoft System Center Virtual Machine Manager Self-Service Portal integration for network-ready virtual machine deployment.
- Holistic data center management through flexible, power extensibility options including PowerShell™, .NET, and F5-authored System Center management packs.

Datacenter Evolution

The modern dynamic datacenter is a logical evolution from the traditional physical datacenter. Traditionally, to isolate workloads, corporate IT departments had deployed a single server or a single operating system (OS) image per workload. This physical datacenter is a well-understood, stable, and secure environment; however, it severely underutilizes the hardware averaging about 15 percent server utilization.

In an effort to increase server utilization, decrease management cost, and add dynamic capabilities, the traditional datacenter evolved into a virtualized datacenter. In a virtualized datacenter, multiple workloads are hosted side-by-side on the same piece of hardware. This increases server utilization levels from 15 percent to 50 percent in a typical virtualized datacenter. It should be noted that not all workloads are appropriate for virtualization, so not all workloads in a virtualized datacenter will be virtualized. However, managing multiple virtual machines on one piece of hardware when appropriate saves management costs by reducing the overall amount of hardware that needs managing. Although this reduces the infrastructure management profile, each service still needs to be managed individually.

The virtualized dynamic datacenter provides further opportunities to increase server utilization by providing the platform to launch a private cloud environment. A private cloud is basically an on-premise environment that also enables organizations to increase the utilization and efficiency of their overall environment through higher levels of infrastructure and management consolidation, and through automation. With a private cloud, organizations can provide the IT infrastructure as a service rather than using it only to host customer workloads.

From the infrastructure perspective, organizations are moving away from servicing boxes and towards a service-oriented infrastructure. With a service-oriented infrastructure, an organization is in a prime position to recognize specific opportunities when moving to the public cloud would be beneficial.

Moving toward the Private Cloud

Microsoft recognizes that the more an enterprise can automate the pooling, allocation, and management of IT resources, the more IT can be delivered in a hosting model. In the hosting model, software capabilities are abstracted from physical server and storage devices, even from geographic sites. This enables software applications to leverage an IT footprint across a varied range of available IT resources – a footprint that is continually responding – even anticipating changes in demand. This delivery model for the IT infrastructure is:

- **Scalable:** The IT infrastructure has the agility to be grown gracefully by simply adding more capacity to the resource pools without requiring a redesign.
- **Elastic:** The IT resource allocation expands and contracts dynamically to maintain optimal performance and economy.
- **Multi-tenant:** The IT service is easily shared across users, groups, and business units.
- **Self-service:** The IT agility is increased by enabling faster responses to business requirements through the self-service provisioning of the IT infrastructure.

These benefits cannot be delivered by software alone. The combination of compute, storage, software, and networking solutions are required for the effective deployment of a service-oriented IT infrastructure. This is also referred to as the enterprise private cloud.

Key Concepts

The basis of this blueprint is the dynamic networking capabilities provided by F5's integration with Microsoft's management platform. In order to successfully apply this blueprint, a basic understanding of the dynamic networking concepts and self-service portal technologies is necessary.

Dynamic Resource and Performance Optimization

F5 and Microsoft virtualization technology work together to create a dynamically provisioned datacenter that uses resources more efficiently, improves application performance, and delivers a better user experience. Client traffic flows through the F5 BIG-IP Local Traffic Manager™ (LTM). The LTM makes intelligent routing decisions and application-specific performance improvements to maintain the best experience for all users.

F5 management packs add network visibility to the monitoring and reporting capabilities of System Center. This visibility is used by IT administrators to manage, monitor, and detect issues that require intervention. Intervention takes the form of reassigning or provisioning resources as well as taking resources offline for maintenance and troubleshooting. These management actions can be fully automated or performed manually depending on the goals of the administrator. This provides customers with the complete flexibility to integrate the F5 application delivery platform with Microsoft's management platform in a way most suitable to their operation processes.

With the proper thresholds set and the required preconfigured actions scripted, System Center Operations Manager (SCOM) can automatically detect application workloads outside of their optimal load levels and take appropriate actions such as instructing System Center Virtual Machine Manager (SCVMM) to deploy or remove additional instances of the application. These technologies all communicate together using Windows PowerShell and Windows Management Instrumentation (WMI).

In our dynamic application workload optimization example, when SCOM and/or SCVMM monitoring determines that an increase in network load is affecting application performance adversely, an automated script is triggered that deploys an appropriate number of new virtualized instances of the affected application using SCVMM.

When the new application instances are detected, SCOM uses the F5 management pack to share this information with BIG-IP LTM which begins to route traffic intelligently to the new nodes. This reduces the load on the other instances and increases the application response for all users.

When the user load decreases, the process is reversed. When the additional instances are taken offline, BIG-IP LTM redistributes the user traffic between the remaining virtual machines to eliminate wasteful underutilization.

Dynamic Network Infrastructure

A robust and integrated network infrastructure is critical for achieving the end-to-end benefits of virtualization within and between datacenters. As applications are virtualized and businesses take advantage of new capabilities for high availability and disaster recovery, the network infrastructure needs to be more robust and intelligent to accommodate the continuous re-allocation of IT resources. Application and virtual machine awareness is also critical for extending the dynamic datacenter capabilities. The following sections describe the key F5 BIG-IP Application Delivery Controller (ADC) attributes that add dynamic network infrastructure capabilities.

Intelligent Load Balancing

Load balancing is more critical in virtual environments due to the increase in the number of application servers and the added flexibility that the virtual environment requires. This makes traffic management through F5 devices even more important. Even though the fundamental operation that F5 performs at the network layers does not change, the frequency and types of network configurations required increases. Moreover, traffic might be managed within a pool of servers, across pools within a data center, or across distant data centers. Therefore, network traffic must be managed from the perspective of the entire pooled environment with the changes managed simultaneously between each environment that makes up the pool.

To support intelligent load balancing, SCOM and the F5 management pack work together to provide visibility into application server capacity and virtual machine resources. This information is used to drive intelligent load balancing decisions. This is especially useful in heterogeneous environments mixing physical and virtual server deployments, or in environments with assorted virtual machine resource configurations. These heterogeneous environments are common in pooled private cloud environments.

Virtual Machine Awareness

BIG-IP supports Live Migration for seamless user connection management as virtual machines are moved between physical hosts. Due to virtual machine awareness, BIG-IP is able to provide:

- Connection draining.
- Active session preservation of stateful applications.
- Recognition of new virtual machines that come online.

Application-Level Awareness

ADCs understand the application, security, and network context of the user. Because the ADC is application-aware, IT administrators are able to leverage it as a point of control in the datacenter.

Using IP geolocation technologies and application awareness to ensure the most appropriate application routing decisions, BIG-IP Global Traffic Manager™ (GTM) intelligently determines which cloud data center or application will best serve each individual request. It then provides global application availability regardless of the implementation model.

Availability

ADCs monitor application availability at various ports and protocols. This activity provides users with a good understanding of where the network and application health is at, reflecting a true image of the end users' actual experience. This information is used to pre-determine which server will receive each new client session. ADCs also support a number of client session persistence methods that enable them to maintain stateful user connections when necessary. This ensures that users enjoy uninterrupted, complete interactions, with backend applications during a user session.

Integration

BIG-IP integration with System Center and extensibility through PowerShell ensures that the monitoring and management of the virtual environment is seamless. All 250 plus network- related statistics captured by BIG-IP are available to SCOM to enrich the IT administrator's view of the environment and enable them to make better

management decisions. When using management packs to integrate systems, any System Center user can integrate BIG-IP in the same way that all other elements of their infrastructure are managed. This provides significant flexibility and a deep level of control over ADC and user connections.

Self-Service Portal

The Microsoft Server platform, comprised of Windows Server Hyper-V and System Center, combined with the System Center Virtual Machine Manager Self-Service Portal 2.0 (SCVMM SSP), helps IT organizations extend their existing investments to deliver on-premise private cloud computing. This increases IT agility enabling faster responses to business requirements through the self-service provisioning of the IT infrastructure.

The self-service portal, built on top of Windows Server Hyper-V and System Center, is a free, extensible, turnkey solution that empowers datacenters to dynamically pool, allocate, and manage resources to enable on-premise private cloud computing. Using this solution, IT organizations create agile, virtualized infrastructures, facilitate business agility and operational efficiencies, and reduce management complexity.

The key capabilities of the self-service portal include:

- **Automation and guidance:** With step-by-step instructions and technical best practices to help assess, plan, and design an on-premise cloud infrastructure.
- **Organization and business-unit onboarding:** With automated workflows to onboard line-of-business applications to a virtualized, shared resource pool.
- **A dynamic provisioning engine:** In conjunction with System Center and Hyper-V, SSP can rapidly provision a virtualized infrastructure.
- **End-user self-service capabilities:** For requesting infrastructure in a self-service model for their applications and services.

F5 Networks extends the SCVMM SSP to the network by providing integration with their leading Application Delivery Controller (ADC), BIG-IP. Customers investing in the F5 and Microsoft platforms for a more agile, dynamic datacenter will reap the benefits of this extendable solution by being able to provision applications for traffic management through BIG-IP at the time of virtual machine provisioning. This reduces the number of manual processes and provisioning steps required to provision and de-provision virtual machines in an enterprise private cloud.

The F5 solution for SCVMM SSP provides:

- **Simplified bulk virtual machine provisioning:** Enables customers to easily and quickly prepare their network for the deployment of virtual machines in bulk.
- **Reduced virtual machine deployment time:** Increases the agility of an enterprise private cloud infrastructure by reducing the time to deploy virtual machines for applications relying on the ADC.
- **Multi-tenant support:** Extends the reliability, availability, and scalability benefits of the ADC to multiple tenants.
- **Network self-service:** Adds application delivery networking to the IT resources available in an enterprise private cloud.

Solution Components

An optimized networking solution for enterprise private cloud Hyper-V deployments can be created by combining Microsoft and F5 technologies. This section discusses the relevant technologies required.

Microsoft Technologies

Microsoft virtualization provides an architecture that can profoundly affect nearly every aspect of the IT infrastructure management lifecycle. It can drive greater efficiencies, flexibility, and cost-effectiveness throughout the organization.

A standard Microsoft virtualization implementation is typically structured using the Windows Server 2008 R2 Hyper-V role to enable virtualization and Windows Server Failover Cluster to handle high availability and disaster recovery requirements. System Center Virtual Machine Manager (SCVMM) is typically used to simplify virtualization management.

Windows Server 2008 R2 Hyper-V

Hyper-V is the hypervisor-based virtualization technology from Microsoft that is integrated into all Windows Server 2008 R2 operating systems. As a virtualization solution, Hyper-V enables users to take maximum advantage of the server hardware by providing the capability to run multiple operating systems within virtual machines on a single physical server. The availability of Hyper-V as a role in a mainstream Windows operating system provides several key advantages:

Features/Capabilites	Benefits
Built-in Windows Server 2008 R2 technology	Enables enterprises to easily leverage the benefits of virtualization without adopting a new technology.
Broad device driver support	The new 64-bit micro-kernelized hypervisor architecture leverages the broad device driver support in the Windows Server 2008 R2 parent partition to extend support to a broad array of servers, storage types, and devices.
Symmetric Multiprocessor (SMP) support	Hyper-V supports SMP on virtual machines.
Host high availability	Windows Server 2008 R2 clustering provides high availability to virtual machines to minimize unplanned downtime.
Shared storage high availability	Microsoft MPIO dynamically routes I/O to the best path and protects against connection failures at any point between a Hyper-V host and shared storage including NICs/adapters, switches, or array ports.
Easy virtual machine migration	Live migration capability is available to support business continuity during planned and unplanned downtime and over distances.
Volume Shadow Copy support (VSS)	Provides a robust host-based backup of virtual machines by leveraging the existing Windows VSS-based infrastructure.
Easy extensibility	Easy extensibility is available by using the standards-based Windows Management Instrumentation (WMI) interfaces and APIs.
Simplified integrated management	With its tight integration into the Microsoft System Center family of products, customers have end-to-end physical and virtual infrastructure management capability for Hyper-V environments.

Table 1: Hyper-V features

System Center Operations Manager

System Center Operations Manager (SCOM) 2007 R2 uniquely enables customers to reduce the cost of data center management across server operating systems and hypervisors through a single, familiar, and easy-to-use interface. Through numerous views that show the state, health and performance information as well as alerts generated according to some availability, performance, configuration or security situation being identified, operators can gain rapid insight into the state of the IT environment, and the IT services running across different systems and workloads. The key capabilities of SCOM 2007 R2 include:

Features/Capabilites	Benefits
End-to-end service management	Provides end-to-end service management that is easy to customize and extend for improved service levels across the IT environment. This enables Operations and IT Management teams to identify and resolve issues affecting the health of distributed IT services.
Best-of-breed for Windows	Includes expertise from the Microsoft server, client, and application teams, providing customers with knowledge and capabilities to drive greater efficiency.
Deep environmental insight	Gain deep insight from the perspective of the private cloud resources as a pool, the individual hardware within, or even the workload inside each virtual machine.
Increased efficiency and control	Automates routine, redundant tasks, and provides intelligent reporting and monitoring to help increase efficiency and enable greater control of the IT environment.

Table 2: System Center Operations Manager features

System Center Virtual Machine Manager

Microsoft System Center Virtual Machine Manager (SCVMM) 2008 R2 is enterprise-class management software that enables administrators to easily and effectively manage both the physical and virtual environments from a single management console and avoid the complexity of using multiple consoles typically associated with managing an IT infrastructure. The key capabilities of SCVMM 2008 R2 include:

Features/Capabilites	Benefits
Enterprise-class management suite	Manages both Hyper-V and VMware® ESX® virtualization environments.
Intelligent virtual machine placement	Provides support for the intelligent placement of virtual machines.
System Center Operations Manager 2007 integration	Integrates with System Center Operations Manager 2007 to provide proactive management of both virtual and physical environments through a single console by leveraging PRO.
Native P2V/V2V migration	Provides native capability for physical-to-virtual migrations and virtual-to-virtual migrations.
Failover integration	Provides integration with failover clustering to support high availability and the live migration of virtual machines.
Automation	Provides easy automation capabilities leveraging Windows PowerShell.
Private enterprise cloud self-service portal	The self-service portal, built on top of Windows Server Hyper-V and System Center, is a free, extensible, turnkey solution that empowers datacenters to dynamically pool, allocate, and manage resources to enable on-premise private cloud computing.

Table 3: System Center Virtual Machine Manager features

F5 Technologies

The F5 technology required to create this cloud-capable network is the BIG-IP Local Traffic Manager (LTM).

BIG-IP Local Traffic Manager

The BIG-IP product family is a system of integrated application delivery services that work together on the same best-in-class hardware. From load balancing, SSL offload, and Web acceleration to application security, access control, and much more, a single BIG-IP device can do the work of a dozen single-purpose appliances.

Features/Capabilites	Benefits
Plan for growth and avoid downtime	Advanced load balancing and comprehensive health monitoring help customers seamlessly add more virtual servers and direct traffic.
Gain ultimate flexibility	Deploy both virtual and physical BIG-IP LTM components to create a fully flexible and adaptable Application Delivery Network.
Secure applications and data	From powerful network and protocol-level security to application attack filtering, BIG-IP LTM deploys a suite of security services to protect the customer's most precious resources — the applications that run their business.
Reduce servers, bandwidth, and management costs	Advanced TCP connection management, TCP optimization, and server offloading enable customers to optimize the utilization of their existing infrastructure — tripling server capacity and reducing bandwidth costs by up to 80 percent. By using fewer servers, less bandwidth, less power, and less cooling, while reducing the time spent managing the infrastructure; customers can significantly reduce operational costs.

Features/Capabilites	Benefits
Take control over application delivery	The F5 TMOS® platform gives customers complete control of the connection, packets, and payload for applications. Using F5's event-driven iRules®, customers can customize how they intercept, inspect, transform, and direct inbound and outbound application traffic. The F5 iControl® API makes it easy to integrate with third-party management systems using PowerShell, .NET and System Center management packs.

Table 4: BIG-IP LTM features

F5 Online Community

DevCentral is global user community that helps users get more from F5 technologies. Comprised of over 60,000 IT professionals from around the world and managed by F5 staff, DevCentral helps users learn how to use F5 products, get tips from peers, share best practices, and connect with others.

Dynamic Network Blueprint

This blueprint is for integrating F5's BIG-IP Application Delivery Controller (ADC) with Microsoft System Center to establish a dynamic network infrastructure. By making the network visible to System Center, application load optimization policies can be automatically managed by System Center. This design enables System Center to automatically add or remove application virtual machine instances as application load thresholds are reached, and F5's BIG-IP application delivery controller to automatically adjust its load balancing to accommodate. The net effect is that the application is always running at optimal load levels for users maximizing performance and eliminating waste.

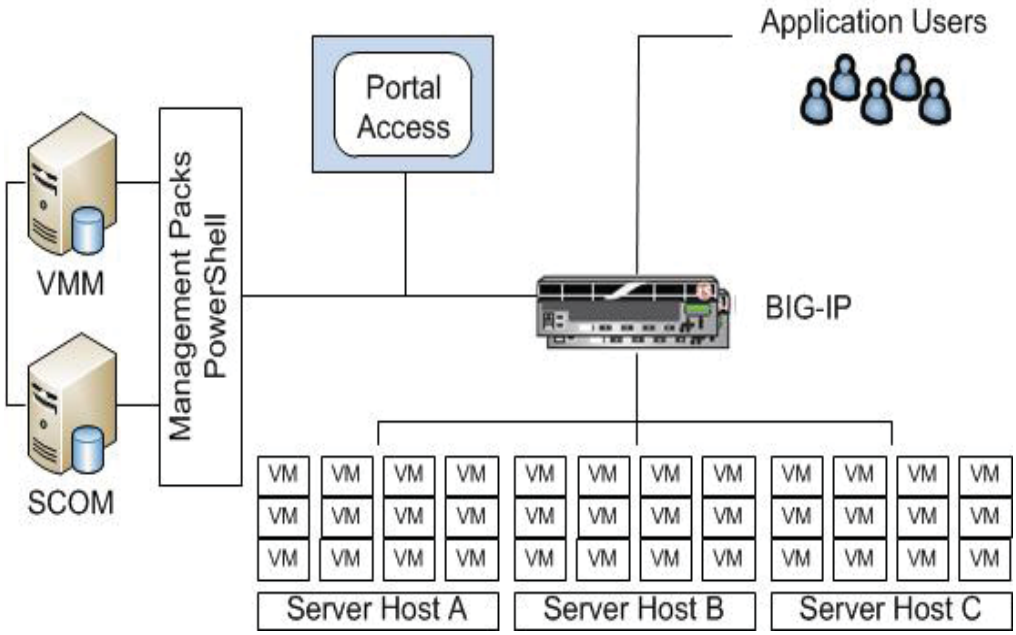

Figure 1: Dynamic Network Blueprint

In this design as shown in Figure 1 above, F5's BIG-IP ADC is managing network traffic for servers A, B, and C which could each be hosting a variety of applications in the form of virtual machines. To provide network visibility, the F5 management pack is applied to enable two-way communications between SCOM and the F5 BIG-IP ADC. This two-way communication as well as the entire Microsoft management platform communication is handled through PowerShell. It is this two-way communication that enables the dynamic network to work.

The System Center Virtual Machine Manager Self-Service Portal (SCVMM SSP) is accessed through an intranet Web site. Its ability to provide self-service provisioning of network resources is enabled by a script that is applied to the Self-Service Portal management interface when setting up the portal.

SCOM Integration Management Pack

The F5 management pack integrates the BIG-IP ADC with SCOM through the PowerShell API providing two-way communications. The management pack can be applied by simply running the installer on the SCOM host server. The F5 management pack installer can be found at: http://devcentral.f5.com/mpack.

SSP Integration Script

F5's BIG-IP ADC and VMM SSP integration is provided through scripting that is applied using the SSP management interface as seen in Figures 2 and 3 below. From the Settings tab in SSP, the DeployVM action is appended with specific script(s). The scripts run after the virtual machine is deployed and configure the BIG-IP to manage traffic to that node. Contact MicrosoftPartnership@F5.com at F5 for assistance in developing a custom script.

The following figure demonstrates what the integration scripting looks like when it is added as a task to the DeployVM action.

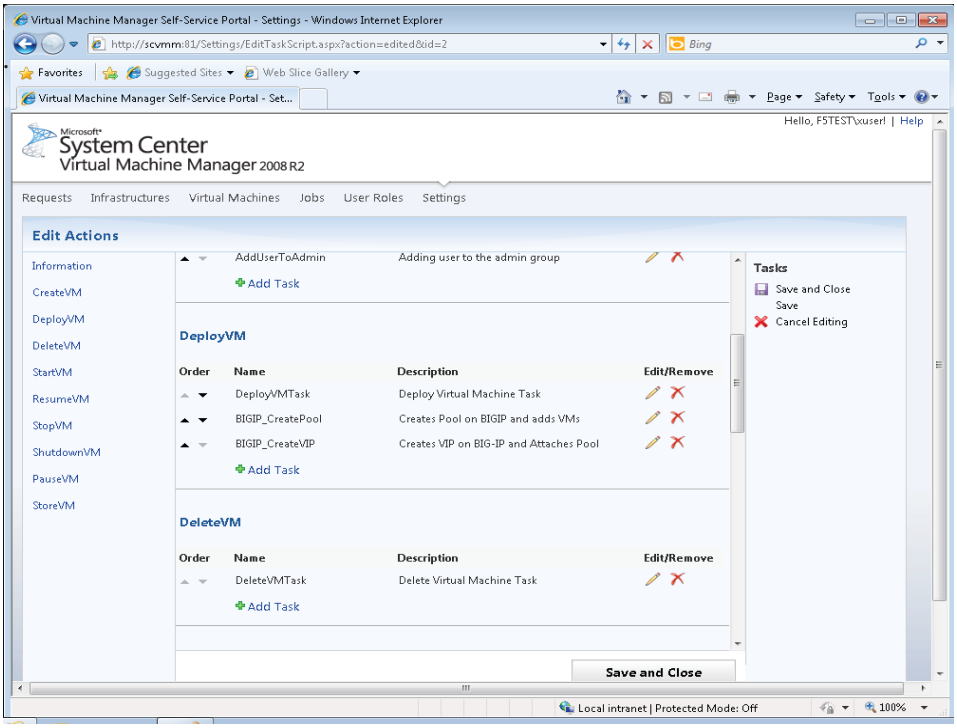

Figure 2: Integration scripting added as a task to the Deploy VM action

The following figure shows the user request view. Here, load balancing is shown as a service option for the requested virtual machine.

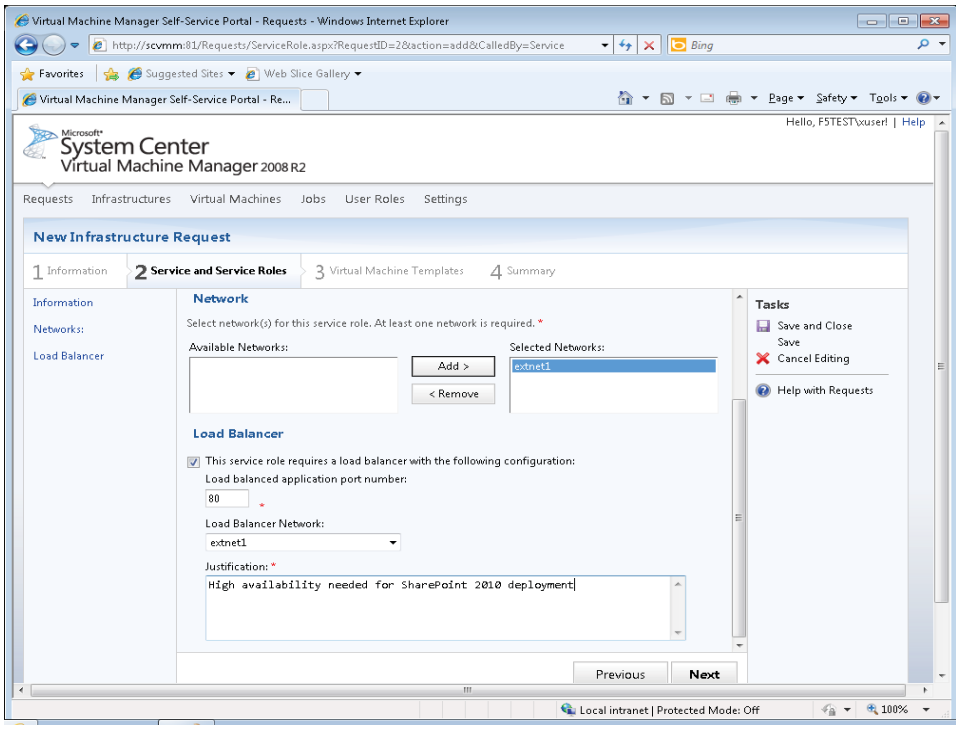

Figure 3. User request view showing load balancing as a service option

Conclusion

F5's best-in-class application-specific, load-balancing and acceleration capabilities combined with the Microsoft management platform provide the basic building blocks for a cloud- capable dynamic network. To be cloud-ready, the network infrastructure needs to be:

- Scalable
- Elastic
- Multi-tenant
- Self-service

F5 BIG-IP LTM supports each of these. It is scalable. Scaling is supported through load-balancing and off-loading. Off-loading is critical to scaling because it makes server applications more available for their core features versus being tied up doing network-centric operations.

F5 BIG-IP LTM is elastic. The ADC provides visibility and control over client access to servers so that as an application's footprint expands and contracts, the ADC, virtual machines, and management software are all in sync.

F5 BIG-IP LTM has multi-tenant capabilities. When IT resources in an organization are pooled for sharing across tenants, the applications are handled like services that can be enabled or deployed and managed with established patterns per user and per group, and across multiple groups in an organization.

F5 BIG-IP LTM supports self-service. This is supported through integration with SCVMM SSP enabling network provisioning capabilities without requiring heavy involvement by core IT. Through the SCVMM SSP workflow, core IT just supervises and approves provisioning as requests come through.

These characteristics provide economies of scale, operational excellence through patterned operations that are repeatable, error resistant, and automatic. IT can keep up with changing business demands because IT is delivered holistically as a service, not just as independent applications. The whole IT infrastructure is service-oriented in nature.

Additional Information

To implement the blueprint provided by this paper, please contact F5 or a Microsoft certified partner to discuss specific needs. To learn more about the technologies involved, please see the links below.

Microsoft virtualization solutions:

http://www.microsoft.com/virtualization/solutions

BIG-IP product family:

http://www.f5.com/products/big-ip/

Join the online community for F5 – Microsoft solutions:

http://tinyurl.com/f5msdcgrp

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
info.asia@f5.com

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

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