Service Orchestration: The Key to the Evolution of the Virtual Data Center

By Jim Metzler, Cofounder, Webtorials Editorial/Analyst Division

Introduction

Data center managers are faced with an array of challenges. This includes the requirement to support the rapid growth in applications and storage with minimum budget increases, as well as the ever-increasing pressure to become more agile in order to respond to constantly changing business requirements. One approach that IT organizations have taken in order to respond to these challenges is to implement server virtualization. The primary factor that drove the initial implementations of server virtualization was the significant cost savings that result from hosting multiple virtual machines (VMs) on a single physical server. These savings include reducing the cost of servers, as well as the cost of the supporting real estate, power and cooling. While cost savings is still a significant driver, today there are two additional drivers of server virtualization. These drivers are the ability to provision VMs and the ability to move VMs among physical servers, both within a given data center and between disparate data centers, in a matter of seconds or minutes.

Unfortunately, while computing resources have become virtualized and dynamic, most of the rest of the IT infrastructure is still physical and static. As a result, deploying new services and/or making changes in the alignment of all the infrastructure resources (e.g., servers, networks, security and storage) that support a service typically requires numerous time-consuming manual tasks that span a number of technology and organizational boundaries. This combination of tedious manual processes and the inefficiencies created by organizational silos leads to escalating workloads and operating expenses. This combination also leads to having processes that don't complete in seconds or minutes, but in days or weeks.

This goal of this white paper is to demonstrate that service orchestration can automate the configuration of physical and virtual entities and can control the allocation of resources shared among a number of virtual services. Because it provides this functionality, service orchestration enables IT organizations to substantially lower cost and improve operational efficiencies by maximizing the leverage achievable from a shared pool of resources within the virtual data center infrastructure.

The Traditional Provisioning Process

The traditional process that IT organizations use to provision new servers typically takes weeks or months. To illustrate that process, consider a hypothetical company called BizCrit and assume that one of BizCrit's application development teams has determined that it needs four new servers. There are two primary phases to the server provisioning process that BizCrit uses. The first phase is to acquire the servers and place them in the appropriate data center. To begin that phase of the provisioning process, the development team has to document why it needs four additional servers as well as a number of other considerations, such as the security functionality required on the servers and whether or not load balancing is required. Their request goes to a review committee where it is analyzed from a variety of perspectives, including whether or not the added capacity is justified, whether or not there is budget for the new servers and whether or not there is space in the data center for the new servers. Assuming that the review committee approves the request, an order is placed for four new servers and the supporting peripherals; e.g., disks and cabinets. As the equipment that was ordered arrives, it is accepted by the BizCrit purchasing organization and stored. When all of the necessary equipment has been received, another BizCrit team integrates the equipment and ships the servers to the designated data center.

Server virtualization enables BizCrit IT to reduce the amount of time associated with the first phase of its server provisioning process from weeks to minutes. Unfortunately, server virtualization does not reduce the time associated with the second phase of BizCrit's server provisioning process: getting the new servers into production. That phase begins shortly after the servers have arrived at the designated data center. As part of the second phase, the network operations group requests a block of IP addresses and permission to assign the servers to one or more VLANs. Assuming that the network operations group receives the requested IP addresses, they assign addresses to each server and make the appropriate entries into DNS. Assuming that the network operations group receives permission to assign the servers to one or more VLANs, the group configures the VLAN(s). As part of this phase of the server provisioning process, the database operations group is responsible for anything related to the database access that has been requested by the application development team. This includes tasks such as configuring fibre channel access. Both the network and database operations groups provide input to the security operations team that is responsible for a variety of tasks, including inputting the appropriate access control lists (ACLs).

Phase two of BizCrit's server provisioning process can take weeks in part because there are so many activities involved and in part because the tools that are currently available are vendor specific. For example, a tool that can configure a Cisco router cannot configure a Juniper router and vice versa. In addition, the use of manual scripts is an acceptable approach in BizCrit's traditional environment in which the addition of new servers is somewhat rare. However, in order to realize the benefits of cloud computing, BizCrit has established the goal of implementing a dynamic data center. The use of manual scripts is not an acceptable approach in a dynamic data center environment in which a large number of VMs are constantly being provisioned and de-provisioned.

The Promise and Challenge of Cloud Computing

There is considerable confusion about the exact definition of what is meant by the phrase *cloud computing*. For example, the January 2009 edition of The Cloud Computing Journal contained an article[1] that had twenty one definitions of cloud computing. In contrast to the confusion around the definition of cloud computing, the goal of cloud computing is exceptionally clear. As stated in a recent report[2], the goal of cloud computing is to achieve an order of magnitude improvement in the cost effective, elastic provisioning of IT services. That report also identified the characteristics that are typically associated with a cloud computing solution. Three of those characteristics are virtualization, the dynamic movement of IT resources and automation.

An article in Network World[3] validated that the adoption of cloud computing can lead to dramatic cost savings. In that article, Geir Ramleth, the CIO of Bechtel, stated that he had benchmarked his organization against some providers of public cloud computing solutions such as Amazon. Ramleth declared that the price that Amazon charges for storage is roughly one fortieth of his internal cost for storage. The decision that Ramleth reached was not that he was going to rely on public cloud computing providers to supply all of his IT requirements. Rather, he decided that Bechtel would adopt characteristics of cloud computing within its internal IT environment that are similar to what have been deployed by cloud computing service providers.

As previously mentioned, in order to realize the benefits of cloud computing, BizCrit has established the goal of implementing a dynamic data center. In order to achieve that goal, BizCrit has adopted a phased implementation strategy that involves the following steps:

1. Server virtualization for resource consolidation and energy savings.

2. Pervasive virtualization of the data center - including servers, storage and the network.

3. Integrated management of physical and virtual resources with automated provisioning and control of virtual entities.

4. Automated, on-demand data center resource provisioning and allocation to meet user demand.

---

[1] Twenty-One Experts Define Cloud Computing, http://cloudcomputing.sys-con.com/node/612375
[2] http://www.webtorials.com/content/2009/11/a-guide-for-understanding-cloud-computing.html
[3] The Google-ization of Bechtel, Carolyn Duffy Marsan, Network World, October 28, 2008

BizCrit is comfortable with its ability to implement server virtualization. However, BizCrit foresees problems as it attempts to adopt pervasive virtualization of the data center. That problem is that its traditional approach to network design and the associated labor-intensive management tools and manual scripts that it uses to control and manage the IT infrastructure will not be able to keep pace with the frequent changes that are required. One limitation of this approach is that the workload for BizCrit's operational support staff will spiral out of control due to the constant stream of configuration changes that are needed to maintain access and security models in conjunction with dynamic provisioning and movement of VMs. The second limitation is that even if BizCrit's IT organization had enough support staff to implement the necessary configuration changes, the time to implement these changes is typically measured in days and weeks. In order to implement a truly dynamic data center, these configuration changes must be made in the same amount of time that it takes to provision or move a VM (i.e., seconds or minutes) and these changes must be made without disrupting user access or creating security vulnerabilities.

The key capability that BizCrit needs to implement in order to be able to support a dynamic data center is the automation of as much of the provisioning and configuration processes as possible across the end-to-end infrastructure. Coordination of automated configuration changes across servers, storage, security and networking elements implies that there must be some form of integration that spans the management systems used for each of these domains and that provides a single point of control for a range of application services that draw on a shared pool of resources.

Orchestration of Network Services

Orchestration refers to the automated arrangement, coordination and management of computer systems, storage, security and networks in order to efficiently deliver application services to end users. Referring back to the section of this white paper entitled *The Traditional Provisioning Process,* orchestration has the same impact on the time it takes for phase two of BizCrit's server provisioning process as server virtualization had on phase one of the process: it reduces it from weeks to minutes. The way that orchestration works is that it creates an executable process that involves centrally controlled message exchanges among network entities. As part of the orchestration of network services, the orchestration system provides a layer of abstraction between the application services and the infrastructure. This layer of abstraction is sometimes referred to as Network Services Virtualization (NSV).

Figure 1 provides a conceptual model of an orchestration system including its relationships with the infrastructure of the virtual data center. Polices define the relationship between users, computing resources, security and network services. These policies are automatically translated in real-time into device configurations that dynamically provision the necessary resources and/or modify the resource pool to bring it into alignment with the service definition. User access information is updated in user directories and network access control system databases. The virtual server requirements of the service are communicated to the hypervisor management system. The hypervisor management system provides the required VMs and also can provision the networked or virtualized storage required by the VMs from a storage pool that has been placed at the disposal of the virtualized infrastructure. Storage provisioning from the pool can performed via storage plugins in the hypervisor manager. The hypervisor manager will also be able to make the necessary configuration changes for virtual switches and virtual appliances under its control. Network elements not under the control of the hypervisor management system (e.g., routers, switches, firewalls, and other external devices that provide secure access to networked services) are controlled on an individual basis by the orchestration system. The orchestration system also communicates with traditional network management solutions, such as fault and performance management systems, which can help to monitor services and assure that the service level specified in the policies are being delivered for each network service.

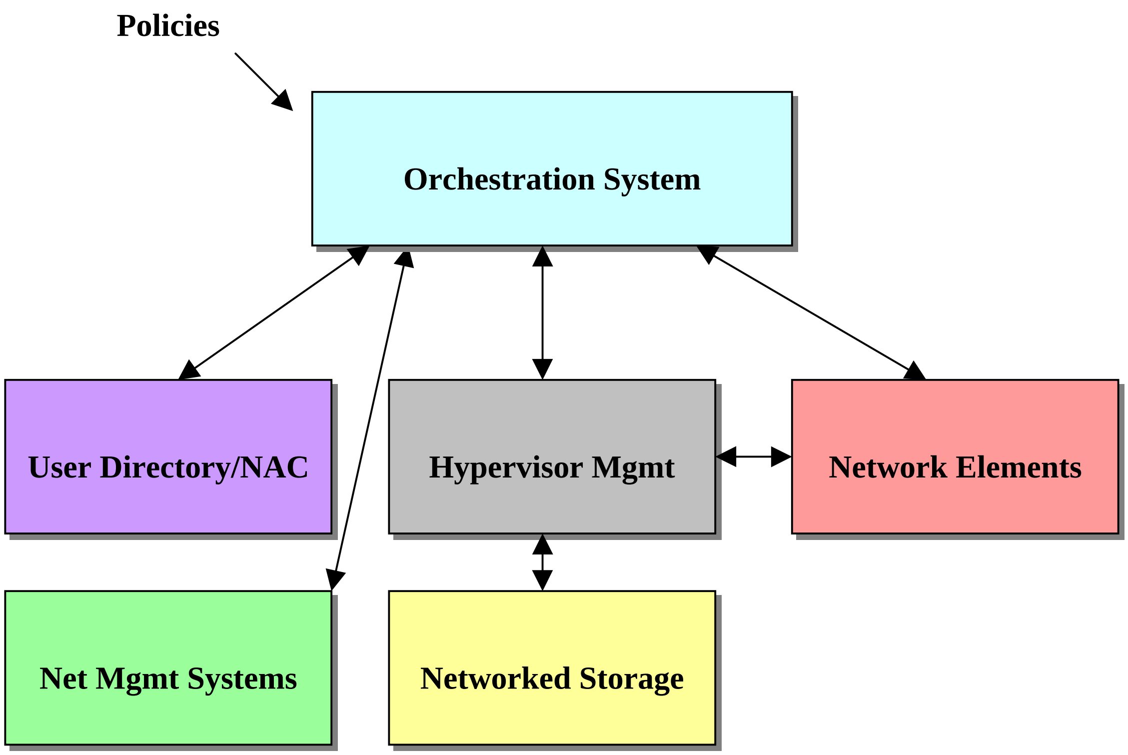

Figure 1: Orchestration System Exercising Centralized Control of the Virtual Data Center

Requirements for an Orchestration System

In order to enable IT organizations to implement a fully dynamic data center, an orchestration solution needs to be capable of automating all of the processes that support the entire life cycle of an IT service. This includes service creation and deployment, monitoring performance and utilization as well as the adjustment of resource allocations. It also includes the eventual return of those resources to a resource pool from which they can be drawn to support other services. In order to be effective throughout the service life cycle, the orchestration platform must possess a number of key attributes. That includes:

- **Infrastructure Awareness:** The system must be able to discover all of the physical and virtual entities within the shared infrastructure pools that are used to deliver virtualized services.

- **Topology Awareness:** Once the elements of the shared resource pools are identified, the system must build a topological model that includes both the physical and logical relationships between the physical and virtual entities.

- **Service Definition and Mapping:** Application services abstraction must allow services to be defined based on business policies and a set of policy guidelines for different types of services. For example, policy guidelines for a service might include restrictions on user access, required security measures and auditing procedures, average and peak workloads, and service level priorities. Once the service has been defined, the system must to able to map the service definition to the required set of resources within the available pool.

- **Service Provisioning and Infrastructure Control:** The orchestration system implements or provisions the service by configuring the entities from the shared infrastructure pools. In order to do this, the system must understand the associations between the resources in different technology domains and must be able to implement automated and coordinated configurations of the physical and virtual entities across the multiple domains. Integration across technology domains has to be based on communications with one or more hypervisor management systems, networked storage management systems, and a number of other element management systems and platforms. Communications can be based on industry standard APIs (e.g., SOAP/REST), vendor specific APIs, orchestration software agents on network devices, or traditional device management interfaces (e.g., CLIs).

- **Resource Allocation and Reclamation:** Over the course of its life cycle, a service will require a number of adjustments to the resources allocated to it. The resource levels may change due to growth or fluctuations in user demand for the service, or due to changes in the policies that govern access to resources. As these changes occur, the system must be able to dynamically re-allocate resources, trigger requests for additional physical resources to be added to the pool, and return unneeded resources to the pool so that they become available to other services.

- **Service Isolation:** In order to provide the required level of security in a dynamic environment, the system must be able to utilize and control the complete gamut of security measures that can provide service isolation in a shared resource environment. In order to be applicable in a Cloud Computing Environment, the orchestration system needs to be able to isolate the virtual data centers of each of the tenants drawing on the shared resource pool, whether those tenants are different companies or different organizations within the same company.

- **Promote Collaboration among Server, Storage, and Network Management Domains:** Traditionally, IT departments are organized along functional lines, with separate teams responsible for server, storage, and network operations. Server virtualization has already had an impact on these domains by giving the server organization increased levels of responsibility. Orchestration has the potential to improve collaboration among technology domains by providing a common shared interface into the resource pool of the virtual data center. As such, orchestration provides a platform for inter-domain collaboration in terms of defining policies, service templates, and provisioning workflows.

As shown in Figure 2, an orchestration system that meets these requirements can support multiple phases in the evolution of the data center towards a Cloud Computing model. That follows both for enterprise IT organizations looking to deploy a private cloud computing model as well as for providers of public cloud computing solutions.

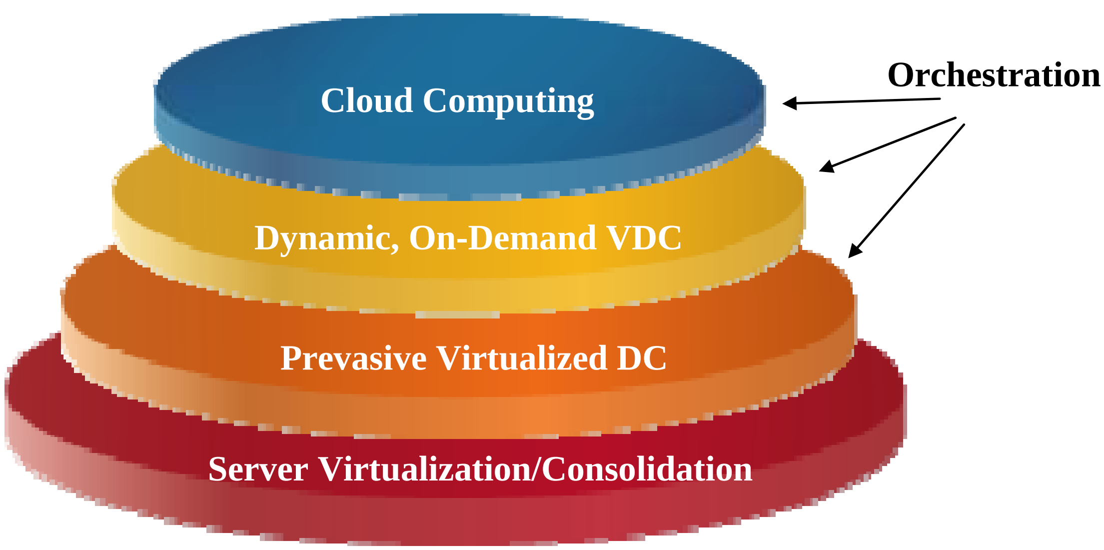

Figure 2: The Role of Orchestration Systems in the Evolution of Virtual Data Centers Summary

Server virtualizaion has dramatically reduced the amount of time it takes an IT organization such as BizCrit's to implement new computer resources. An application, however, is comprised of more IT functionality than just compute resources and the traditional approach to provisioning or modifying an application is a very lengthy, complicated, multi-step process involving coordination across the diverse technology domains that support the application. IT organizations will not experience the agility that they need to support rapidly changing business requirements and that is promised by server virtualization, if they don't implement the ability to automate the provisioning and operation of network services across the end-to-end virtual and physical infrastructure of the data center.

The goals of an orchestration system are to maximize the business leverage that can be gained from virtualization and to greatly simplify the management of the pervasive virtualization of data center resources. Therefore, virtualization and orchestration are highly complementary technologies that together enable the implementation of a dynamic data center model. An orchestration system achieves its goals by automating the provisioning and operation of network services. It does this by initiating the creation of VMs through the hypervisor management system. As the VMs and supporting storage resources and virtual switches and appliances are coming online, the orchestration engine defines and deploys the network access and security models across all required infrastructure devices (e.g., routers, switches, firewalls, load balancers) required to deliver the service to the specified end-users. The entire process can be completed in a matter of seconds or minutes, including the setup and deployment of IP addresses, network routes, VPNs, VLANs, ACLs, the deployment of security certificates as well as the configuring of firewall rules and DNS entries.

About the Webtorials® Editorial/Analyst Division

The Webtorials® Editorial/Analyst Division, a joint venture of industry veterans Steven Taylor and Jim Metzler, is devoted to performing in-depth analysis and research in focused areas such as Metro Ethernet and MPLS, as well as in areas that cross the traditional functional boundaries of IT, such as Unified Communications and Application Delivery. The Editorial/Analyst Division's focus is on providing actionable insight through custom research with a forward looking viewpoint. Through reports that examine industry dynamics from both a demand and a supply perspective, the firm educates the marketplace both on emerging trends and the role that IT products, services and processes play in responding to those trends.

For more information and for additional Webtorials® Editorial/Analyst Division products, please contact Jim Metzler at jim@webtorials.com or Steven Taylor at taylor@webtorials.com.

Published by Webtorials Editorial/Analyst Division
www.Webtorias.com

Division Cofounders:
Jim Metzler
jim@webtorials.com
Steven Taylor
taylor@webtorials.com

Professional Opinions Disclaimer
All information presented and opinions expressed in this publication represent the current opinions of the author(s) based on professional judgment and best available information at the time of the presentation. Consequently, the information is subject to change, and no liability for advice presented is assumed. Ultimate responsibility for choice of appropriate solutions remains with the reader.

Copyright © Webtorials
For editorial and sponsorship information, contact Jim Metzler or Steven Taylor. The Webtorials Editorial/Analyst Division is an analyst and consulting joint venture of Steven Taylor and Jim Metzler.