



F5 White Paper

Automating the Data Center

This paper discusses an alternative architecture that supports data center automation and dynamic provisioning without operating system virtualization.

by Lori MacVittie
Technical Marketing Manager, Application Services



Contents

Introduction	3
<hr/>	
Data Center Automation	4
The Vision	4
Virtualization Options	5
– Operating System Virtualization	5
– Server Virtualization	6
<hr/>	
The Dynamic Data Center	7
Dynamic Provisioning with Server Virtualization	7
– Application Architecture	7
– Application Delivery	8
– Monitoring	9
– Automation Engine	10
<hr/>	
Conclusion	11



Introduction

The concept of virtualization has existed for a long time. Indeed, the very basics of load balancing use the concept of virtualization in order to make multiple web or application servers *appear* to be a single server with massive capacity. Hence the reason most Application Delivery Controllers (ADCs) use the terms *virtual IP address* and *virtual server* to describe the components comprising its configuration.

In recent years a new concept of virtualization has developed—one in which one server appears as many servers. This concept, usually referred to as operating system (OS) virtualization, has become an increasingly important tool for IT to accomplish its goals, especially those involving consolidation and improving the efficiency of computing resources (Table 1).

But as virtualization of all kinds becomes standard fare in IT, management of resources—both physical and virtual—has become more and more difficult. IT budgets are strained while available staff dwindles. IT must do more with less, and thus it is imperative to automate the data center.

Reasons to Virtualize Servers

Cut costs via server consolidation	81%
Improve disaster recovery and backup plans	63%
Provision computing resources to end users more quickly	55%
Offer more flexibility to the business	53%
Provide competitive advantage	13%

Table 1: (Respondents chose up to three) SOURCE: CIO Research

Data center automation provides a codified system for executing pre-defined policies that automatically provision and deprovision resources as necessary. Automation of these processes reduces the time required of IT staff to monitor and manage critical application infrastructure, and thus reduces the overall costs of implementing a given solution.

Data center automation requires intelligent components with the ability to codify policies and execute a process or workflow. While OS virtualization can certainly be one of those components, it is not necessarily a requirement.



Data Center Automation

The Vision

The vision of the automated data center is a lofty one. The automated data center is expected to dynamically provision computing resources in the most efficient manner while also meeting business needs by enforcing service level agreements (SLAs). This goal is a difficult one, as it requires management of resources as well as performance.

Current data center architectures have been built to support performance and capacity requirements, which leave available resources sitting idle. With the growing emphasis on reducing IT budgets, it becomes imperative that IT implement a solution that can not only continue to meet SLAs, but can become more efficient in its use and distribution of computing resources.

Computing resources must be coordinated among a number of systems at the application and network layers of the architecture in order to be properly distributed. In addition, they must be dynamic. If Server A has resources available, then it should be possible to provision those resources for Application B when its primary server(s) are at or near capacity. Similarly, when capacity needs to reduce, the computing resources of Server A should be deprovisioned or reassigned to another application as needed. This process should be automated in case the benefits of the dynamic provisioning system are offset by the cost of manual intervention.

A second form of data center automation is one that does not necessarily duplicate applications across multiple servers, virtual or physical, but one that distributes applications intelligently across multiple resources such that the capacity and performance needs of the application are met by the appropriate hardware. This may result in applications being moved from one server to another throughout the day. This enables IT to use hardware with a range of processing capabilities without sacrificing application performance. This type of data center automation is often referred to as “utility computing.”



Virtualization Options

Current trends indicate that many believe automation should be accomplished through a combination of workflow systems and OS virtualization. Other views propose that existing solutions already support dynamic provisioning without the need to invest heavily in virtualization technologies.

Operating System Virtualization

Operating system virtualization essentially turns one server into many. Its advantages lie in the ability to quickly provision new systems simply by loading up the correct image and its underlying system agnosticity; that is, a virtual machine can run a different operating system than the server on which it is deployed. This type of virtualization makes it easier to increase capacity, as it is easy to provision additional computing resources.

But OS virtualization—whether implemented via VMware, Microsoft, or Citrix solutions—is disruptive and requires modifications to servers as well as the network. If you have implemented a data center using commodity hardware to reduce expenses may find that new, more powerful hardware is required in order to implement a virtualized environment. Each virtual server requires its own IP address, which may require a new network scheme.

This type of virtualization also does nothing to address IT's top challenge when virtualizing the data center: balancing server load and enforcing SLAs (Table 2). Therefore it becomes necessary to introduce additional solutions to the architecture—one that is capable of addressing those challenges by providing load balancing and SLA enforcement. If the data center does not have a solution currently in place that is capable of providing these functions, it becomes necessary to invest in such a solution, thereby increasing the total cost of acquisition and implementation of a virtualized data center.

Operating system virtualization is the primary technology upon which utility computing is built. While SOA (service oriented architecture) is often touted as a form of utility computing, it does not match the versatility and flexibility offered by OS virtualization technologies. The ability to deploy alternative operating systems and application servers on virtually any hardware platform in a matter of minutes makes OS virtualization invaluable for implementing a true utility computing architecture.

Top Challenges to Virtualization Success

Reasons to Virtualize Servers

Cut costs via server consolidation	81%
Improve disaster recovery and backup plans	63%
Provision computing resources to end users more quickly	55%
Offer more flexibility to the business	53%
Provide competitive advantage	13%

Table 2: (Respondents chose up to three) SOURCE: CIO Research

Server Virtualization

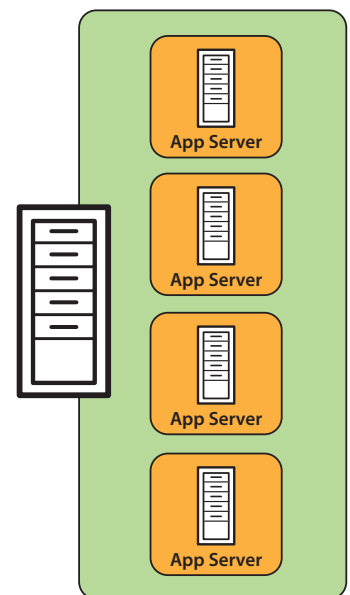
The alternative to operating system virtualization is server virtualization. This type of virtualization is the same virtualization technique that has been used successfully for many years in the data center. Server virtualization makes multiple servers appear to be a single server, and is primarily used in application delivery (load balancing) to provide scalability, reliability, and support application performance requirements.

Server virtualization cannot, of course, automate the data center on its own. Just as solutions based on OS virtualization require additional components, through which the process of provisioning and deprovisioning is accomplished, so too does a solution based on server virtualization. Server virtualization also requires some modification in the way applications are deployed, but like its OS-based counterpart, also increases the ease and speed with which resources can be provisioned once the system is in place.

An automated data center that is based on server virtualization requires a lower cost of implementation. Additional software or hardware upgrades are not required, and the absence of virtual machines frees up computing resources usually dedicated to management.

Server virtualization also addresses the needs inherent in building a utility computing-based architecture. By ensuring that applications are deployed on every physical server, real-time decisions can be made to determine the appropriate server for each application, based on computing power and resources.

Server virtualization makes many physical servers appear as one.



The Dynamic Data Center

Dynamic Provisioning with Server Virtualization

By using proven server virtualization and open, standards-based integration mechanisms, you can create a complete, automated data center solution.

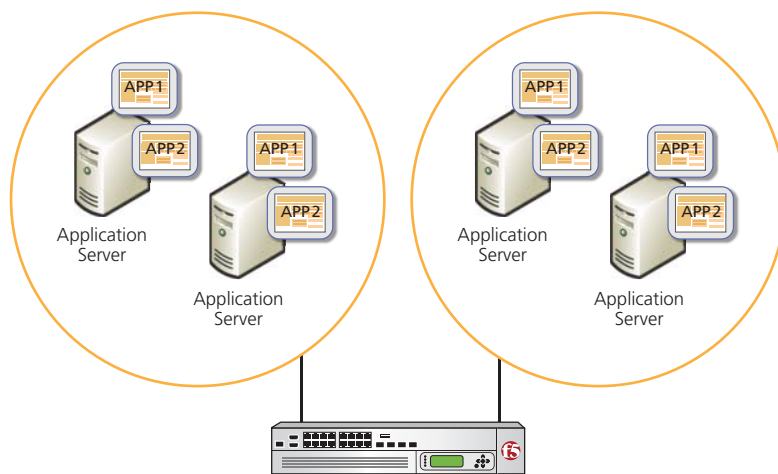


Figure 1: Typical high-availability application architecture

In this scenario, resources (applications) are virtualized behind an Application Delivery Controller. Each “pool” of servers is configured to serve specific applications. This architecture is commonly used to support high-availability and application performance requirements. This model does not optimally support emerging trends requiring higher efficiency in the data center, as the resources dedicated to delivering *App 1* and *App 2* cannot easily be reallocated to *App 3* and *App 4* if the need arises.

The goal of dynamic provisioning in the data center is to use available resources as efficiently as possible. This means that the typical architecture must be modified in order to achieve that goal, regardless of whether operating system or server virtualization forms the basis of the implementation.

Application Architecture

When implementing an architecture based on OS virtualization, images must be created and software must be installed on each physical server to support the automated provisioning of applications through virtual machine images. In a server virtualization scenario, applications are deployed across all physical servers that may be provisioned to serve them.

In both models, additional storage requires additional costs. Server virtualization requires the application to be deployed across all usable resources, which translates into more storage used for the deployments. OS virtualization requires more if images are stored locally—which they often are, due to performance—because not only is the application stored, but the application server is also stored, as well as the virtual image.

Storage virtualization can reduce the impact in both scenarios because it can be shared across all physical and virtual servers, but may degrade application performance due to latency introduced by remote storage access. Storage virtualization provides all instances of applications with access to the same storage and therefore files. Virtual images, EAR (Enterprise Application Resource) files, images, configuration. All can be stored in one place and accessed by all applications through a storage virtualization solution. This makes deployment as well as creation of standard virtual images much simpler as it becomes unnecessary to physically install applications or images on the server itself.

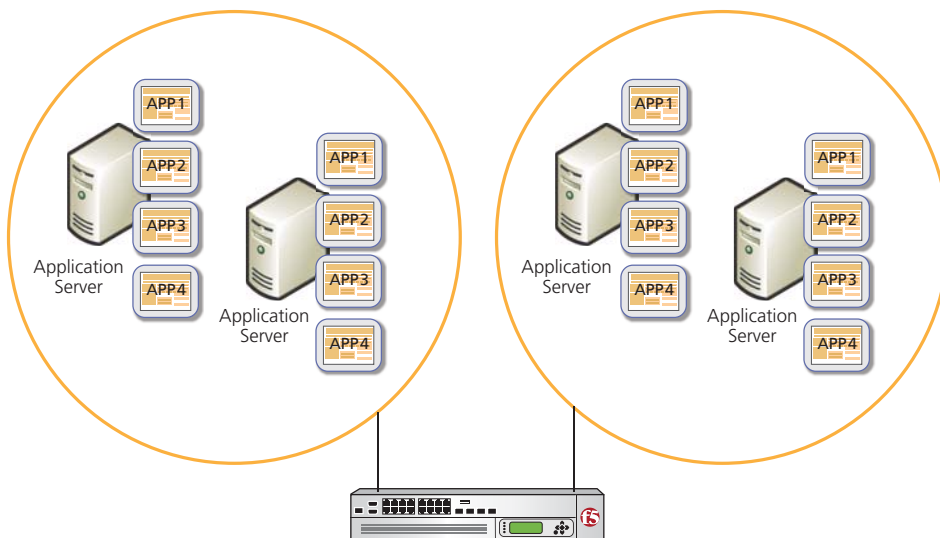


Figure 2: High availability architecture supporting server virtualization for automated provisioning

Application Delivery

The architecture must also include a solution that addresses the need for load balancing. It is nearly impossible to provision multiple instances of an application—regardless of the technology providing this functionality—without some way to load balance and deliver that application transparently to the client.



An Application Delivery Controller is a necessary component in both the operating system and server virtualization methods of automating the data center. In both cases the multiple servers need to appear as one. The responsibility of the ADC is to perform this task.

In the case that applications are being moved across resources based on real-time needs, the ADC must switch gears, and provide dynamic routing of application requests. It must mediate between the client and the server because “the server” may change from one moment to the next. An intelligent ADC, capable of inspecting application requests, can provide real-time, intelligent routing capability that is based on the current distribution of resources. Regardless of whether a utility computing-based model is based on OS or server virtualization, an ADC will be required to dynamically direct application requests to the appropriate application instance.

In the server virtualization scenario, all physical servers with the capability to serve a particular application are known to the ADC (Figure 2). Using an external automation solution or an internal execution engine, the ADC adds and removes application instances from its application pools based on a variety of parameters such as number of users, resources available on the server, and application performance. In this way the ADC is able to ensure SLAs are met while utilizing available resources in the most efficient manner possible.

Monitoring

In order to efficiently make use of resources while simultaneously ensuring capacity and performance needs are met, it is necessary to monitor the application, its environment, and application usage. This requires an intelligent solution that is capable of monitoring both the client and server side of an application, such as an ADC.

Without understanding where available resources may be within the infrastructure, it is impossible to provision resources such that all applications continue to perform within expected limits. Using advanced health monitoring techniques, an ADC can gain awareness of how all applications are performing while simultaneously determining which resources may be reprovisioned when the need arises.



An advanced ADC can provide the necessary monitoring capabilities using a variety of monitoring techniques including SNMP, agents, passive monitoring of transactions, and protocol specific (for example, HTTP, TCP, and IP monitors). External solutions may also be used, though such systems require integration into the overall automation framework.

Automation Engine

At the core of the automated data center is the ability to automatically provision and deprovision resources dynamically, based on conditions that exist in real-time within the data center. This automation should be accomplished via open, standards-based integration mechanisms such as a service-enabled API.

The benefit to automating the data center using server virtualization techniques is that it requires automation of only one system rather than two. In order to automate the data center using OS virtualization it is necessary to automate the server virtualization device (load balancer/ADC) as well as the creation and eventual disposal of virtual servers on back-end systems.

Unlike Business Process Management (BPM) solutions in the software world, there is no standard—or even a proposed standard—in the data center that provides a standards-based solution for automating the components necessary for dynamic provisioning. There are currently initiatives underway that may change this in the future, but at present there is no BPEL (Business Process Execution Language)-like counterpart in the data center automation realm.

There are several solutions to this issue. First, one can use the service-enabled APIs of components to develop a custom solution. This has advantages in that the solution will be what is needed for the organization. The disadvantage is that it requires development time and effort. Second, one can use a workflow automation solution capable of automation through service-enabled APIs. This reduces the amount of time and effort required to build an automation solution but increases the overall cost of acquisition, as it is yet another piece of software that must be purchased and managed. Lastly, many solutions providers offer professional services or consulting services that can be used to help implement such a solution. It may be the case that such a solution already exists that requires only a bit of tweaking and customization in order to serve an organization's specific needs.

Conclusion

Virtualization provides a tantalizing opportunity to automate the data center and reap the benefits of increased efficiency as well as decreased cost of management and power, all without sacrificing the need to meet SLAs.

There are many faces to virtualization. Both server virtualization and OS virtualization provide the foundation for building out an architecture that not only can be automated, but can be automated in a way that makes sense for each individual organization's needs and architectural preferences.

While there is currently much hype surrounding OS virtualization, it is not a necessary requirement for automating the data center and implementing a dynamic provisioning system. Server virtualization can provide the same benefits using similar techniques, without requiring the purchase of additional software and licenses and introducing another layer of technology that must be secured, managed, and upgraded.



F5 Networks, Inc.
Corporate Headquarters

401 Elliott Avenue West
Seattle, WA 98119
+1-206-272-5555 Phone
(888) 888GIP Toll-free
+1-206-272-5556 Fax
www.f5.com
info@f5.com

F5 Networks
Asia-Pacific

+65-6533-6103 Phone
+65-6533-6106 Fax
info.asia@f5.com

F5 Networks Ltd.
Europe/Middle-East/Africa

+44 (0) 1932 582 000 Phone
+44 (0) 1932 582 001 Fax
emeainfo@f5.com

F5 Networks
Japan K.K.

+81-3-5114-3200 Phone
+81-3-5114-3201 Fax
info@f5networks.co.jp