Paper 347-2009

Virtualization: What does it mean for SAS®?

Karl Fisher and Clarke Thacher, SAS Institute Inc., Cary, NC

ABSTRACT

SAS® groups virtualization into four categories: Hardware Virtualization, Presentation Virtualization, Application Virtualization, and Storage Virtualization. Hardware Virtualization (VMware, Xen, and Microsoft Hyper-V) and Presentation Virtualization (Windows Terminal Server, Citrix XenApp,) are technologies being used by SAS users today. Application Virtualization (Citrix XenApp, Microsoft App-V) is a newer technology for application deployment and maintenance that is generating interest in the SAS community. Storage Virtualization is something that we see as an important future technology.

This paper discusses how hardware, presentation, and application virtualization impact running SAS software.

INTRODUCTION

Companies and organizations are looking for ways to save money and make the best use of their information technology resources. The virtualization technologies can be used to:

- increase system utilization
- reduce datacenter costs
- enable central administration of servers and applications
- reduce software deployment costs
- reduce the effort to support legacy applications and processes
- speed system and application-provisioning requests

In today's computing environment, virtualization comes in several forms and is supported by a large number of software products from many software vendors. The virtualization technologies and their use can be confusing at first. Once understood, virtualization technologies can be used to address a variety of problems.

With the increasing use of virtualization, questions about running SAS software in a virtualized environment are common. Virtualization introduces some complexities in deployment, maintenance, performance, and support. This paper contains an overview of each virtualization category, some of the motivations for using the technology, and how it relates to running SAS software products.

This paper gives some examples of software products used to implement the virtualization solutions, but covers the virtualization topics in a general, product-neutral manner. The focus of the paper is primarily from a Microsoft Windows operating-systems perspective, but the concepts can be applied to other operating systems.

HARDWARE VIRTUALIZATION

Hardware virtualization is a technology to completely emulate a computer's physical hardware, such that multiple operating system instances can run on the same physical platform. Usually a software layer called a hypervisor is used to emulate the physical hardware. An instance of this emulated hardware is called a Virtual Machine (VM). With full hardware virtualization, an unmodified operating system can run in a VM. There are often operation system extensions that can be used when running in a VM to "enlighten" them to the virtualization. This enlightenment can improve performance for running in a virtual machine.

Another way to look at hardware virtualization is the separation of the operating system from the physical machine it is running on. This enables a single physical machine to run multiple operating systems simultaneously.

Hardware virtualization has been around for many years, and has a long history in the mainframe environment. Today hardware virtualization is gaining wide acceptance in the commodity and mid-range x86 and x64 server market. The original x86 architecture contained some features that could not be virtualized, so full hardware virtualization was very difficult to implement. By 2007 Intel and AMD had introduced hardware extensions, Intel VT-x and AMD-V, to their processors enabling hardware-assisted virtualization. This has made creating efficient hypervisors for x86 and x64 processors possible. When running on processors with Intel VT-x or AMD-V extensions,

modern hypervisors can run unmodified guest operating systems. VMware ESX, Xen, and Microsoft Hyper-V are some of the leading products being used today.

Most of the Virtualization vendors have programs and utilities to assist in the transition from physical hardware to virtual machines. This conversion and the tools are often called “P to V.” Virtual to physical conversion tools do not exist, so that is a more difficult conversion.

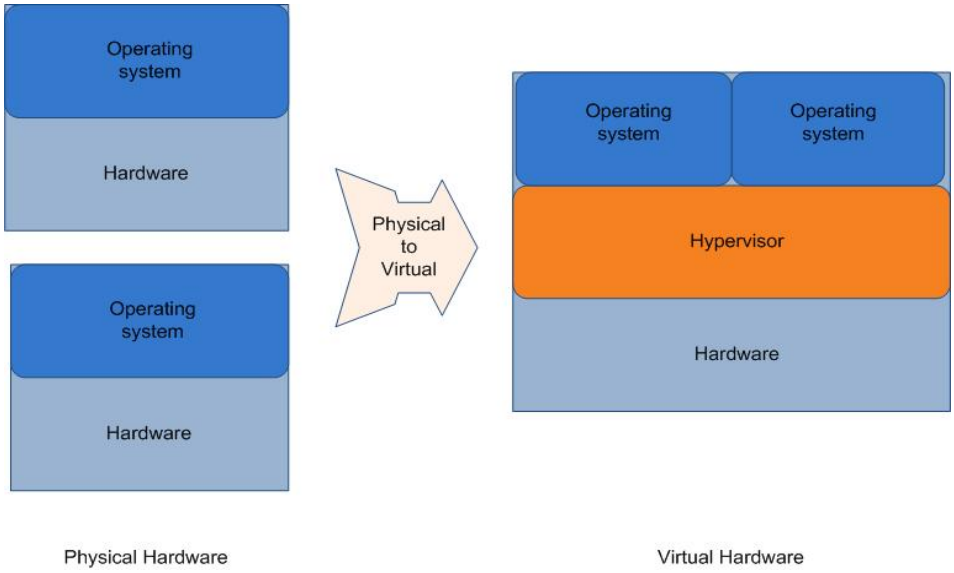

Figure 1. Physical versus Virtual: Two Physical Computers Being Combined on a Single System Using Hardware Virtualization

MOTIVATION FOR HARDWAE VIRTUALIZATOIN

Server Consolidation and Utilization: Hardware virtualization is often used to consolidate smaller physical servers for better server utilization. Many enterprises find themselves supporting and maintaining many small departmental, or workgroup, servers. As these systems need hardware replacement, they are transformed from physical to virtual servers, and consolidated onto newer, larger servers. This can lead to substantial savings by reducing the number of servers in corporate data centers, reducing the power consumption, and centralizing server management.

Legacy Application Servers: There are situations where a business can depend on applications that are running on older systems, and might not want to, or cannot, upgrade the systems. What happens when the server hardware needs to be replaced? Often the legacy system can be converted to a virtual machine, and can continue to run after that on any of the newer hardware servers in a virtual machine.

Efficient System Procurement: Once the VM infrastructure is in place, an organization can create standard VM images that are used to satisfy requests for new systems very quickly. Because new hardware does not need to be procured for the request, VMs can significantly reduce turnaround time for system deployment.

High Availability: Many of the hardware virtualization products have VM-management tools to support, monitor, and manage VMs. Since each virtual machine is run on virtual hardware, it can be moved from server to server as needed. Some tools support the moving of running virtual machines; others require the VM be stopped. Because the entire VM is encapsulated in a single physical file on the host physical system, moving a VM from one server to another is usually a very simple matter.

HARDWARE VIRTUALIZATOIN FUNCTIONAL AND PERFORMANCE CONSIDERATIONS

In general, user mode applications – like SAS software products – operate correctly in virtual machines. Applications that require device drivers to access low-level hardware do not tend to work well in virtual machines.

Server processes that have high CPU utilization might not be good candidates for virtual machines, particularly if the physical hardware that is hosting the VM is over committing CPU cores. This over committing occurs when the sum of the number of CPUs that are defined in virtual machines is greater than the number of physical CPU cores in the physical system.

Processes that are I/O bottlenecks might not work well in a hardware virtualized system. I/O operations present a performance challenge to virtual machines. Many I/O operations use direct memory access (DMA) to interface with devices for I/O operations. An operating system running in a VM does not have the address for DMA access, so the hypervisor needs to translate the address for the I/O request. This hypervisor intervention slows I/O operations for the VM.

Many SAS jobs are CPU- and I/O-intensive, so this should be carefully evaluated. Users might find their SAS jobs no longer meet performance requirements when they are run in a virtual environment.

LICENSING AND SUPPORT FOR HARDWARE VIRTUALIZATION

Pricing for SAS software products that run in a hardware virtualization environment depends on the server operating system.

For UNIX, Linux, or Windows server operating systems, pricing is based on the number of processor cores that are available to the physical machine, or in some cases, to the footprint that is assigned to SAS, and not the number of processors available to the virtual machine or machines. There are no license restrictions on the number of virtual machines that can be deployed on that physical machine.

For Windows workstation operating systems, like Windows XP and Windows Vista, pricing is based on a PC-Use metric. *PC Use* is defined as the greater of the total number of virtual machines or the total number of users that access any of the virtual machines at any time (not concurrent).

This information is current as of February 2009. These pricing policies are under review and will likely change in the near future. Under consideration is pricing based on the total number of processor cores that are assigned to a virtual machine.

All SAS products beginning with 9.1.3 are supported in a hardware virtualization environment. Please see the support statement link in the "References" section below.

PRESENTATION VIRTUALIZATION

Presentation virtualization separates a program's execution from the user's interaction. The possessing executes on one system, while the display, mouse, and keyboard are presented on a different machine.

Citrix XenApp (former names include MetaFrame Server and Presentation Server), Microsoft Windows Terminal Services and Remote Desktop, and VNC are common ways to use presentation virtualization on Microsoft Windows systems. For Linux and UNIX systems, methods include exporting the SAS display with X-Windows X11 or remote logons with a terminal or terminal emulation such as an xterm.

Presentation virtualization is very common, and many users do not even think about it as a form of virtualization. For these users, it is just the way they access servers in a remote location or systems located in corporate data centers. This illustrates how well presentation virtualization works for some people.

Presentation virtualization is often use along with hardware virtualization. A virtual machine can be run on a large server physically located in a secure data center. Users of applications that are running in the VM can access the server with remote connections from their office computers.

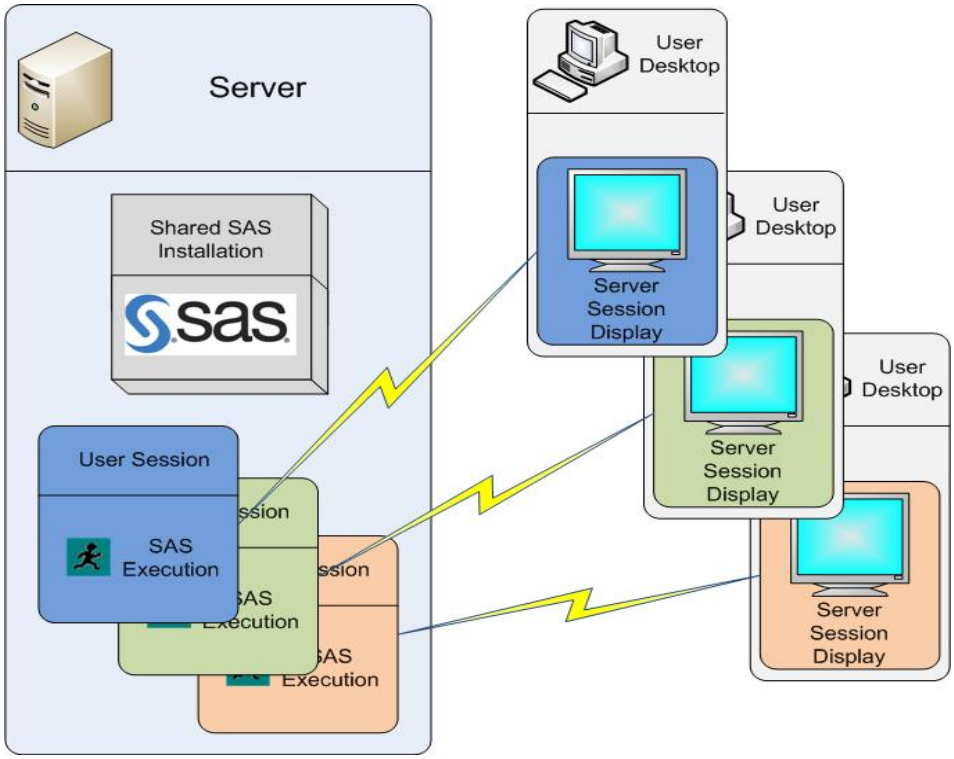

Figure 2. Presentation Virtualization: A Server System with SAS Software Installed, and Three User Desktop Systems Running Server Sessions, Executing SAS Software on the Server

MOTIVATION FOR PRESENTATION VIRTUALIZATION

Using presentation virtualization enables multiple users to run software applications on a shared server. This can reduce the software deployment and maintenance costs.

Running SAS on a remote server can allow SAS to run "closer to the data." If the data that SAS is consuming resides on the server, it can be significantly faster to run the SAS process on the same system. Access to data can also be a security issue, and keeping the data on the server might be desirable.

Presentation virtualization can also give users access to SAS from multiple locations. When running on a server system, users can remotely access products from their office, home, or while traveling.

PRESENTATION VIRTUALIZATION FUNCTIONAL AND PERFORMANCE CONSIDERATIONS

First let's consider the performance of the presentation virtualization product. The performance of the presentation virtualization products vary a bit, but the biggest factor is the speed and latency of the network connection between the local and remote computers. On a fast corporate intranet, these connections approach the feel of being on the physical machine. With slower connections, the interaction with the desktop can be sluggish and lag behind user interaction.

With presentation virtualization, the software applications are running on the server system. This can lead to contention for server-system resources, such as physical memory, CPU-instruction cycles and I/O-bandwidth contention. This can be a significant issue when running SAS products. Many SAS jobs are memory-, CPU-, and I/O-intensive, so care must be taken to limit the use of system resources by the SAS system. SAS options (such as MEMSIZE, SORTSIZE, CPUCOUNT, and THREADS) can be used to control and limit the resources consumed by SAS.

There is also a finite limit to the number of user sessions that the server operating system can support. When the limit is reached, additional users cannot connect to the server system.

LICENSING AND SUPPORT FOR PRESENTATION VIRTUALIZATION

In a presentation-virtualization environment, SAS products are installed and processing on the server and not on the remote computers. Pricing for SAS software products that are running in a presentation virtualization environment is based on the number of processor cores that are available to the entire physical machine, or in some cases, to the footprint that is assigned to SAS. This applies for UNIX, Linux, and Windows server-operating systems. There are

no license restrictions on the number of virtual users interfacing with SAS on this server, nor limitations on the number of remote desktops.

Many of the software products used for remote access are freely available or part of the operating system. Use of the software can result in additional licensing fees for the presentation virtualization, and the use of additional client access licenses.

The SAS virtualization support site does not currently have a specific support statement for presentation virtualization, but a support statement is being developed. Please see the support statement link in the "References" section below for the current, official support statement from SAS.

APPLICATION VIRTUALIZATION

Application virtualization is a newer technology for application deployment and maintenance that is generating interest in the SAS community. The technology enables running software products on multiple desktop systems, without the overhead of installing on each desktop computer.

Microsoft App-V (formerly SoftGrid), VMware ThinApp (formerly Thinstall), and Citrix XenApp client-side virtualization are just a few of the products that are used for application virtualization.

When the virtualized application is launched on the user's computer system, the application virtualization software downloads the application's modules to the system. The modules can be cached on the user's system, so on subsequent invocations, the download is avoided.

The virtualized applications are then run in an isolated sandbox environment. This means that the virtualized applications cannot see other local or virtualized applications. This is achieved by using virtual file systems, virtual registries, and virtual services. Application virtualization is very useful in resolving application conflicts due to different required versions of common software components.

To configure application virtualization, the product administrator obtains the product media. The administrator then runs the product installation while the application virtualization product "monitors" or "sequences" the installation. This installation creates a virtualization package that is then placed on an application virtualization server. This package is not an executable image of the product. You can think of it as a proprietary blob of data, only understood by the application-virtualization server product. The software product cannot be run directly on the server system from this package. The package contains all of the information about the product installation. This includes the program modules that are delivered during the installation and information about other updates to the system – things such as registry updates, component registrations, menus, and desktop icons.

Most of the application-virtualization software products can be configured to customize the use of desktop cache, license counting, and off-line accessibility of virtualized products.

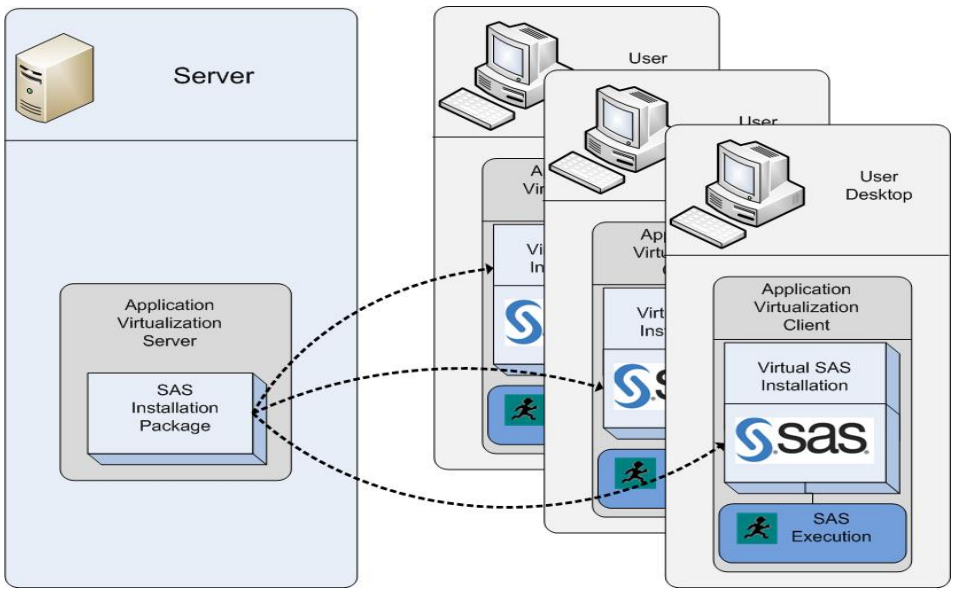

Figure 3. Application Virtualization: A Server Machine Running Application-Virtualization Server Software Containing a SAS Software Installation Package, and Three User-Desktop Machines Running the Application-Virtualization Client.

As Figure 3 shows, within the virtualization client, there is a virtual installation of SAS software and an execution environment.

MOTIVATION FOR APPLICATION VIRTUALIZATION

Application virtualization centralizes product management, deployment, and maintenance. Standard applications can be made readily available to users. Applications can be deployed to many users without running the installation program on each user's system.

The application-virtualization products can run the virtualized products in a protected sandbox, therefore, vitalizing the registry and other local resources used. These sandboxes can also enable multiple levels of products to run on a single system without unwanted interaction.

APPLICATION VIRTUALIZATION FUNCTIONAL AND PERFORMANCE CONSIDERATIONS

In application virtualization, there is a performance penalty the first time an application is run. Additional slow downs occur when the application is updated for maintenance or when the version is upgraded. The performance penalty is caused by the need to download the application modules from the server to the desktop. Usually this time is spent at installation time, not application use.

Once the application has been downloaded to the client, there is only a slight delay on each invocation as the server is checked for application updates and application-use counts.

Because each user runs the application on their computer, application virtualization offers a scalable solution for centralized, standard-enterprise product use.

LICENSING AND SUPPORT FOR APPLICATION VIRTUALIZATION

At this time, application virtualization is not supported by SAS for SAS software products. There are customers that are using SAS products in this way. For those that choose to run in this unsupported environment, each end-user system that is running SAS software products needs to have a SAS software license because SAS software is computing on each end user's system. Pricing is based on the number of processor cores available to each end-user system.

This information is current as of February 2009. These pricing policies are under review and will likely change in the near future. Under consideration is pricing that is based on the total number of processor cores assigned to each virtual client.

In addition to the SAS software licensing cost, the application virtualization software product needs to be purchased.

The SAS virtualization support site does not currently have a specific support statement for application virtualization, but a support statement is being developed. Please see the support statement link in the "References" section below for the current, official support statement from SAS.

DESKTOP VIRTUALIZATION

Desktop virtualization is much like application virtualization, but taken one step further. The entire desktop operating system, with installed applications, is virtualized for the desktop computer. This enables an organization to have one or more standard desktop images – a combination of operating system and installed applications that can quickly be deployed to users. It can also aid in supporting legacy applications that might not operate correctly on the current level of the operating system.

Citrix Provisioning Server for Desktops and Microsoft Enterprise Desktop Virtualization (MED-V) are two examples of products that provide this virtualization. Provisioning Server for Desktops is currently available. MED-V is in beta test at this time and is part of the Microsoft Desktop Optimization Pack.

Implementation details vary for desktop virtualization. One method is to bootstrap the client computer, so that it boots the operating system from a network storage location. Once booted from the network image, the desktop is running just like a natively installed operating system. Another approach is to run the virtualized desktop as a virtual machine on the desktop system. In this case, the virtual desktop acts much like the virtual machine described in the "Hardware Virtualization" section above.

DESKTOP VIRTUALIZATION FUNCTIONAL AND PERFORMANCE CONSIDERATIONS

SAS products are not tested in this environment, but a functional equivalence between a physical and virtual environment is expected.

Performance characteristics will depend on the implementation of the desktop virtualization. The system and application startup performance of bootstrapped systems can be less than natively booted systems. Performance characteristics for virtualized desktops that are running in a virtual machine are discussed in the Hardware Virtualization section of this paper.

LICENSING AND SUPPORT FOR DESKTOP VIRTUALIZATION

At this time, desktop virtualization is not supported by SAS for SAS software products. There are SAS users that are using SAS products in this way. For those that choose to run in this unsupported environment, each virtualized desktop that is running SAS software products needs to have a SAS software license because SAS software is computing on each desktop.

When the virtual desktops are running UNIX, Linux, or Windows server-operating systems, pricing is based on the number of processor cores that are available to each virtual desktop.

When the virtual desktops are running a Windows workstation-operating system like XP or Vista, pricing is based on a PC-Use metric. PC Use is defined as the greater of the total number of virtual desktops or the total number of users that access any of the virtual desktops at any time (not concurrent).

This information is current as of February 2009. These pricing policies are under review and will likely change in the near future. Under consideration is pricing that is based on the total number of processor cores assigned to each virtual desktop.

The SAS virtualization support site does not currently have a specific support statement for desktop virtualization, but a support statement is being developed. Please check the support statement in the "References" section below for the current, official support statement from SAS.

CONCLUSION

Virtualization technology is here to stay, and SAS users want and need to run SAS products in virtual environments. Virtual environments can be more cost effective and more flexible than physical systems. But with this flexibility comes increased complexity and potentially difficult problem resolution.

SAS cannot test all SAS software products in all the virtualization environments that are available from all the virtualization vendors. The virtualization products claim to provide a functional equivalent to the physical environment. SAS requires this functional equivalence to operate correctly. Until SAS gains more experience with virtualization, SAS might request users to reproduce any functional issues in a physical environment.

Performance of SAS software in virtual environments is not usually equivalent to a physical environment. It is up to the SAS users to configure and tune their virtual environment to meet their performance requirements.

REFERENCES

- SAS Institute Inc. (2009). SAS® Product Support for Hardware Virtualization Environments. SAS online. Available at http://support.sas.com/techsup/pcn/virtualization.html.
- Citrix Systems, Inc. (2009). Citrix Online. Available at http://www.citrix.com.
- Microsoft Virtualization. (2009). How far will you take virtual?Microsoft Virtualization. Microsoft Online. Available at http://www.microsoft.com/virtualization.
- VMware, Inc. (2009). VMware Online. Available at http://www.vmware.com.
- 1104 Enterprise Computing Group (ECG). (2009). *Virtualization Review. Available at* http://virtualizationreview.com.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Karl Fisher
SAS Institute Inc.
SAS Campus Drive
Cary, NC 27513
Work Phone: (919) 531-4160
Fax: (919) 677-4444)
Karl.Fisher@sas.com

Clarke Thacher
SAS Institute Inc.
SAS Campus Drive
Cary, NC 27513
Work Phone: (919) 531-7786
Fax: (919) 677-4444
Clarke.Thacher@sas.com