



TOP CHALLENGES AND BEST PRACTICES VIRTUALIZING **CITRIX** XenApp / XenDesktop ON **vmware** vSphere

An eG Innovations Technical White Paper

Top Challenges and Best Practices for Virtualizing Citrix XenApp and XenDesktop on VMware vSphere

Avoid Performance Issues Before They Occur

A growing number of companies are virtualizing Citrix XenApp and XenDesktop on VMware vSphere and other virtualization platforms to reduce costs and boost flexibility. While Citrix virtualization seems like an easy, straightforward project, many of these initiatives experience performance problems – delayed deployments, cost overruns and frustrated users, among others.

In this white paper, you will learn how to:

- Prevent common problems in the storage, network, compute and memory tiers
- Configure VMware vSphere and Citrix XenApp and XenDesktop for maximum reliability, performance and scalability
- Identify and monitor the right performance metrics
- Get complete visibility into your virtualized environment to deliver performance, end user satisfaction and cost control

In a recent poll, more than 800 Citrix professionals were asked for their biggest challenge when virtualizing Citrix XenApp and XenDesktop on VMware vSphere. The results are shown in Figure 1.

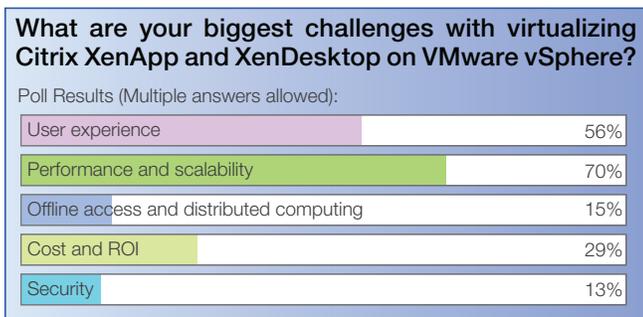


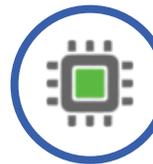
Figure 1: The biggest challenges when virtualizing Citrix XenApp and XenDesktop on VMware vSphere

Not surprisingly, performance and scalability issues top the list of challenges at 70%, followed by user experience with 56%, and coming in third, cost and ROI at 29%.

Fine-Tuning the Virtualization Platform

Most organizations virtualizing Citrix and VMware solutions are very familiar with these technologies, but one thing they often neglect is that the platform and the hardware on which these software components are deployed also has to be fine-tuned. These software applications have to be able to leverage the power of the platform and the hardware in order to get the best performance and scalability possible.

Top 5 Infrastructure Performance Challenges and Recommendations



#1 Don't Discard Compute

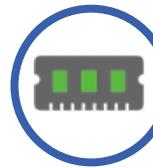
Many organizations deploy compute out of default configurations. Compute can have a large impact on how easily you can scale and what the performance is going to look like for your infrastructure. Key recommendations for improved Citrix/virtual desktop performance include:

- **Use CPUs with hardware assists:** CPUs with MMUs (e.g., Nehalem+ for Intel, Shanghai+ for AMD) will give you a significant performance boost as compared to ones without.
- **BIOS settings matter:** Aside from having the right advanced CPU that will do hardware-assisted tasks, the Basic Input/Output System (BIOS) settings are also going to be critical to help scalability. For instance, by default, Non-Uniform Memory Access (NUMA) should be enabled on all the blade chassis and rack-mount servers; if it is not, then you will want to enable it. What NUMA does is group a subset of physical memory and physical CPUs into a node. Then it will assign specific workloads to these nodes. The idea

here is that if you can put a specific workload on a very specific set of hardware – memory and CPU – then it is going to perform better as opposed to spreading it across all the memory and all the CPU that is available in the system, because that will introduce some form of latency.

So if you can contain specific workloads on a particular NUMA node, you're going to get significantly better performance out of it. You want to make sure the applications you are using are NUMA-aware. For example, VMware vSphere is NUMA-aware, as are Microsoft SQL Server and Exchange. This way, the application or virtualization platform itself knows where to place its workloads on the most efficient memory and CPU to gain the best performance.

- **Hyper-Threading is a performance booster:** Hyper-threading is technology that is used to improve parallelization of computations (doing multiple tasks at once). Earlier versions of VMware vSphere (4.0 and earlier) had issues with hyper-threading. This issue has been addressed in VMware vSphere 4.1 and above. So hyper-threading should be enabled when you are virtualizing Citrix XenApp, XenDesktop and Provisioning server.
- **Use Turbo Boost where possible:** Turbo Boost is a technology available in certain processors that enables the processor to run above its base operating frequency via dynamic control of the processor's clock rate. If your hardware supports Turbo Boost, make sure it is turned on.
- **Power Management is a performance drain:** The biggest gotcha with respect to performance of a blade server relates to Power Management and load balancing of the power to the CPU. For Citrix XenApp and XenDesktop environments in particular, if the blade chassis is regulating the power to the CPU, which means it is trying to load balance the power and not sending the maximum power to it, this automatically introduces a latency into the environment and the user experience suffers and degrades right away. Therefore, for best performance, power management should be off. Processor C-state which refer to the idle power saving state of the processor should also be off.



#2 Memory Matters

When you use VMware vSphere as a hypervisor for deploying Citrix XenApp and XenDesktop, there are several memory management technologies built into VMware vSphere that come in handy. There may be situations when you need to run your vSphere servers in a memory overcommit mode – for example, if a few servers go down or you are facing a disaster recovery situation when you are working with fewer servers than usual.

- One of these technologies is transparent page sharing - a method by which redundant copies of pages are eliminated in memory, which is particularly useful for virtual desktop environments where multiple VMs are running the same operating system and there is a lot of potential to save on duplicate copies of the same data.
- Other memory management technologies of VMware vSphere that are useful include memory ballooning and compression.
- Using the Swap to host cache is another useful memory management technique. This will severely reduce the performance impact of VMkernel swapping. It is recommended to use a local SSD drive for the cache to eliminate any network latency and to optimize performance.

Other important memory strategies include:

- Balance physical memory on the right NUMA nodes: Check the manufacturer documentation to make sure you have the right amount of memory per NUMA node so you are taking the highest advantage of the hardware possible.
- Full memory reservations at the resource pool level for Citrix XenApp, XenDesktop and Provisioning server is recommended.
- Verify the monitor mode: In some instances, the VMware vSphere hypervisor thinks it is faster than the hardware itself – whether it's memory or CPU – and as a result it reverts back to using software to run some of these tasks as opposed

to leveraging the hardware that comes with the CPU and the memory. As you are checking your virtual machines, make sure you are checking the VMware log and make sure the MMU is set to hardware and not software so you are leveraging the abilities of the actual hardware.

Often administrators consider whether they should use faster memory rather than configuring more memory. Given a choice, choose more memory over memory speed. This is the case for desktop virtualization but this is also the case when you are virtualizing any type of tier 1 application (e.g., Citrix XenApp, Exchange, Microsoft SQL, or SharePoint). In doing so, you will see a significant performance enhancement from having more memory rather than faster memory.

Finally, when virtualizing Citrix XenApp, you have to decide how to size your virtual machines. If you decide to go with large VMs, you have to make sure that your hardware is sized correctly. For instance, if you have a VM with 8 vCPUs, you should not be running this on a two socket blade. In this case, you will need a system with a quad socket or higher.



#3 Tweak the Network

A common question regarding the network is what type of an uplink teaming policy should be used for the vSwitches? In most cases, the default load balancing policies are efficient (i.e., routing based on MAC addresses). If you are virtualizing a Citrix provisioning server, you may want to aggregate some of the network links and you can use the policy of IP-based hashing on the virtual switch. If you couple this with an EtherChannel on the physical switch, you will now be able to aggregate these links for the Provisioning server.

Another feature that is disabled by default with VMware vSphere is the Network I/O Control (NetIOC), which allows you to support QoS and to aggregate the network links. It is strongly recommended that you enable the NetIOC for best performance.

It is also recommended that if you are deploying a large XenApp or XenDesktop environment, to have a separate isolated network for your IP storage. Storage traffic is very sensitive and a single glitch or a single re-transmit could have a negative effect on performance, and on the stability of the Citrix environment. By separating storage and other user traffic, you are isolating the different traffic types and increasing performance.

Another performance tip is to disable physical NIC (pNIC) interrupt coalescing. This change reduces latency, trading it off for a slightly higher bandwidth. For performance sensitive use cases like Citrix XenApp and XenDesktop, where good user experience is critical, it is recommended that pNIC interrupt coalescing be turned off.



#4 Storage is Critical

While most of the innovation today is happening at the storage level, that is also where most of the challenges are.

If you are still using Windows Server 2003 for your XenApp environments, then you have to do storage alignment between the VM's virtual disk and the server's physical disk/LUN. Without this, there will be a significant performance strain at the storage array level. If you are on Windows 2008 or above, you do not have to worry about this.

Snapshots are not backups. If you are not using a snapshot, do not let it linger – delete it. Snapshots have a tendency to auto-grow and will cause performance issues.

From a vSphere host perspective, multi-pathing (i.e., having more than one physical path between host and its storage tier) is a must. Storage I/O Control (SIOC) is used to control the I/O usage of a virtual machine and to gradually enforce the predefined I/O share levels. This must also be enabled to enhance performance.

Also, make sure your storage array supports vSphere APIs for Array Integration (VAAI), which allows you to offload a lot of the SCSI reservations and other copy tasks that happen within the hypervisor to the storage array and onto the storage processor so you can get a significant performance boost.

With regards to spindle counts and array configuration, do not assume that using RAID-10 is going to provide the best performance. If you have enough spindles and put it in a RAID-5 configuration, you can still get better performance and increased capacity.

Overall, as you're starting to go down the route of desktop virtualization or virtualizing Citrix XenApp, make sure to get your storage administrator involved in the project. They have to understand what the profile of this application is and what the workload is going to be like on the storage tier. They can then make the determination together with you on what the RAID configuration should be like and what the spindle count on the back end should look like. For instance, you can choose to put the gold image on a RAID-10 and then write cache for a Citrix provisioning server on a RAID-5. Bring the storage admin into the conversation early on to start breaking down some of the walls that have become established between the server, desktop and storage teams.



Often administrators do not look at the virtual hardware closely. There are a lot of options that you can configure at the virtual machine virtual hardware layer and these can significantly improve performance:

- **Use paravirtualized virtual NICs (like vmxnet3):** This will give you a significant performance increase. By using this on a Citrix provisioning server, you can reduce boot times of VMs streamed by this server by 30 to 40 seconds.
- **Take advantage of paravirtualized SCSI (PVSCSI) drivers:** A para-virtualized SCSI controller provides a bigger pipe to access the virtual disk (vmdk). This is ideal for VMs with

high IOPS requirements (1500 iops or more) like ones running Microsoft SQL, XenApp or Citrix provisioning server.

- **Disable ISO CD-ROMs:** Many administrators forget to check if the CD-ROM is connected or not. When a CD-ROM is connected, it is consuming bandwidth to the ESX host. This can have a negative impact on performance. vMotion is also disabled if a VM has a CD-ROM. So all in all, disable the CD-ROM as soon as you are done using it.
- **Setting NIC as first boot device:** In the VM BIOS make sure you are setting your NIC as the first boot device.
- **vNUMA sizing:** VMware vSphere and ESX have been NUMA-aware for a long time. In the past, the virtualization platform made a choice of which NUMA node should run a VM. Now, with vNUMA, the virtualization platform exposes the hardware landscape all the way to the application that is running in the VM. With this, the application can make the decision on where to place the workload. This can result in significant performance gains.

When a VM has more than 8 vCPUs, vNUMA is enabled by default. If you want to enable vNUMA on a VM with less than 8 vCPUs, you have to explicitly enable it.
- **Disable interrupt coalescing at the virtual NIC level as well.** This will help lower latency, trading off for a little bit more bandwidth.

Fine-Tuning the Citrix XenApp/ XenDesktop Infrastructure

A lot of times, customers ask if they need to have a separate vCenter server for their VMware environment that is hosting Citrix XenDesktop. The best practice for a large Citrix deployment is to dedicate a vCenter for the Citrix XenDesktop infrastructure. The reason for this is sometimes Citrix XenDesktop tasks (power operations in particular) can overwhelm vCenter. Hence, for best performance, ensure that there is a dedicated vCenter for the Citrix XenDesktop infrastructure.

Often administrators clone VMs when setting up an environment. In many environments, this causes issues with Citrix XenApp, XenDesktop and the Provisioning server if you have not changed the SID or properly sysprepped your environment. Hence, if you plan to use templates, make sure you have sysprepped your templates, and if you plan to clone a VM, make sure you have changed the SID.

At the same time, your Citrix infrastructure services – the connection broker, storefront servers, license server, provisioning servers, etc. can be installed on the general VMware vSphere servers or clusters in the infrastructure. These servers need not be administered by the vCenter server dedicated for Citrix XenDesktop. There are reasons for this – the general vSphere servers may be configured with DRS and high availability and other advanced enterprise features that may be disabled or toned down in the XenDesktop vSphere environment.

VM affinity rules and host affinity rules are often overlooked, but are very important for eliminating single points of failure in your server farm. Most environments are moving to blade servers and a blade chassis has 7 to 18 servers. If the key servers in the Citrix infrastructure are on VMs that are on servers on the same chassis, this creates a single point of failure. Hence, it is important to spread the load so that critical servers do not reside on the same blade chassis. VM affinity rules are useful for hosting XenApp/XenDesktop controller, provisioning servers, etc. where you do not want them to reside on the same blade chassis. Host affinity rules can ensure that specific VMs only reside on a specified set of blades. This comes in handy for example when virtualizing vCenter. Without this, vCenter could be vmotioned and you will struggle to find out which ESX server the vCenter has moved to. This can also be used to place XenApp servers so they are all not on the same host or the same chassis.

You should also want to make sure that some of the performance intensive applications do not participate in DRS. For instance, you may not want your Citrix provisioning server to be moved around across hosts.

Having multiple XenApp/XenDesktop controllers is important for load balancing. With the Citrix FMA architecture in XenApp/XenDesktop 7.x, Citrix servers no longer have a local host cache. So database backend failures are service impacting. You must plan for high availability of your SQL datastore to eliminate a single point of failure in the backend. Power on/off operations can overwhelm vCenter if you are running virtual desktops. There are Citrix XenDesktop controller host parameters that you can use to control how many power operations are sent concurrently to vCenter. Tune this setting for optimal performance.

VERY IMPORTANT

Anti-Virus and Angry VMs

For virtual machines and virtual desktops, in particular if you include an anti-virus agent inside each VM, you are going to take your IOPs profile from 10 to 30 IOPs depending on the application to 100s of IOPs per VM. Each anti-virus agent will download the latest definitions and all the scans can happen in parallel on the VMs. This will significantly increase the load on the storage tier and in turn, negatively affect the user experience. So, it is very important to have a strategy for anti-virus deployment especially for Citrix XenDesktop.

The best practice for anti-virus deployment is to avoid having an agent on every virtual desktop. With VMware vSphere, you can leverage vShield Endpoint. This installs a SCSI driver within the VM that redirects the IO traffic from within the VM into an anti-virus appliance. With this approach, you are just doing the scan once and you are updating anti-virus definitions only once.

Another option is cloud-based anti-virus agents – in this case, you still need an agent on the VM, but the agent has a very small footprint. All of the definition files get updated in the cloud and it saves a subset of them locally, simply to support you offline (similar to a Local Host cache approach for anti-virus).

You've Followed Best Practices. Now What Should You Monitor?

Citrix XenApp and XenDesktop environments have a high visibility in your infrastructure because they touch the user. Therefore, monitoring must be proactive. The last thing you want to do is to wait for users to experience issues and complain.

Most organizations use the in-built tools in Citrix and VMware platforms to monitor the infrastructure. The challenge that most administrators face is that these tools are all diverse and not tied to one another. So when a problem happens, an administrator has to look at a multitude of tools – one for the network, another for Citrix, yet another for storage, and so on.

A multi-tool monitoring approach requires a lot of manual effort, takes time and requires expertise. The experts are often required for troubleshooting problems. They have to look at a multitude of metrics at different tiers including the kernel latency and device latency on the ESX servers, network latency to the user terminal, retransmit patterns at the TCP layer, network bandwidth usage, IOPs on the LUNs, logon times for Citrix users, HDX latency seen by users, concurrent application usage, etc.

Every minute of slowness costs time and money for your organization. And if Citrix applications or virtual desktops are slow, your entire application/desktop transformation project may be in jeopardy – users will want their physical desktops back.

It's All About Correlated Visibility

To minimize your operations costs and to help you diagnose problems quickly and proactively, you need to take a holistic and unified approach to monitoring your Citrix XenApp and XenDesktop infrastructure. Here are some best practices you should follow:

- **Use a unified monitoring approach** to Citrix XenApp/XenDesktop monitoring: Citrix issues can be caused by problems in the network, storage, virtualization or the client tiers. With a single unified console from where you can monitor the health of every tier of the Citrix infrastructure, your administrators will be able to more efficiently understand, diagnose and resolve potential issues.

- **Monitor your Citrix infrastructure as a service** – not silo by silo: When users call your helpdesk, they are not saying that the storage is slow or that VMware ESX is slow. Their complaints are that “Citrix is slow” or “Citrix is not working”. You need to understand that the user is referring to the Citrix service – not the Citrix servers. Hence, take a service-oriented view of Citrix performance.
- **Look for monitoring tools with automated analysis capabilities:** There are hundreds of key metrics to look at from every tier of the infrastructure. Administrators do not have the time to manually review and analyze each metric. Therefore, look for monitoring tools that can automatically determine baselines for each metric and alert you proactively when any metric is deviant from its norm.
- **Correlated analysis is critical for root-cause identification** in Citrix infrastructures. Because of the inter-dependencies between different tiers, a problem in one tier – say the profile server – can affect other tiers of the infrastructure and ultimately the user experience. Look for monitoring tools that can analyze the inter-dependencies in the infrastructure and automatically correlate alerts from different tiers to identify where the real cause of a problem lies.
- **Use historical trends and reports to right-size your Citrix infrastructure:** Performance monitoring is not about problem diagnosis alone. Ideally, you want all the metrics collected by the monitoring tool to be subjected to trend analysis. Historical analysis and predictions based on past history can help you right-size your infrastructure and to plan for future growth.

eG Enterprise: Monitoring, Diagnosis and Reporting for Virtualized Citrix Infrastructures

The eG Enterprise solution addresses all of the above key requirements for effective Citrix performance monitoring, diagnosis and reporting. Its unified monitoring and correlative intelligence technologies are tailored for virtualized Citrix XenApp and XenDesktop infrastructures. With this solution in place, IT administrators will be alerted to problems proactively, before users complain. With a single click, they will be able to diagnose where

The Solution: eG Enterprise



Figure 2: eG Enterprise provides a total performance management solution for virtualized Citrix infrastructures

the real root-cause of a problem lies and fix it quickly, thereby ensuring high service uptime and improved user satisfaction and productivity.

eG Enterprise is also a force multiplier for IT productivity. Citrix and other domain experts will not be required to solve routine problems. IT helpdesk staff will be able to resolve routine issues by themselves. This leads to significant cost savings for the organization.

About eG Innovations

eG Innovations is dedicated to helping businesses across the globe transform IT service delivery into a competitive advantage and a center for productivity, growth and profit through award-winning performance monitoring and management solutions. Many of the world's largest businesses and mission-critical environments trust eG Enterprise Universal Insight technology to enhance IT service performance in physical, virtual and cloud environments, increase operational efficiency, ensure IT effectiveness and deliver on the ROI promise of transformational IT investments

For a live demonstration of eG Enterprise and how it can transform the performance and productivity of your Citrix infrastructure, [click here](#).

Restricted Rights

The information contained in this document is confidential and subject to change without notice. No part of this document may be reproduced or disclosed to others without the prior permission of eG Innovations, Inc. eG Innovations, Inc. makes no warranty of any kind with regard to the software and documentation, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose.

© Copyright eG Innovations, Inc. All rights reserved.

All trademarks, marked and not marked, are the property of their respective owners. Specifications subject to change without notice.

