XenServer for XenApp - Better Together

Overview

For the last few years, IT professionals have questioned whether it makes sense to virtualize XenApp environments. The answer has always been “it depends.” In short, that answer has not changed. Most organizations realize that having an environment that is 100% physical or 100% virtual may not be the best solution. Adding server virtualization impacts single server scalability, requires management and technical expertise, and adds environmental complexity. However, the benefits can include higher utilization, greater availability and increased flexibility. Before deciding the best solution for your environment, the challenges must be considered, the goals must be understood and the environment must be conducive for server virtualization.

Objectives

IT organizations everywhere are either currently evaluating or in the early phases of investigating server virtualization solutions. While there are many compelling arguments for the adoption of virtualization technologies, the main drivers from both a business and a technical perspective can be categorized into three distinct groups:

- Higher utilization
- Greater availability
- Increased flexibility

Higher Utilization

Organizations have built XenApp server farms to deliver varied types of applications, ranging from mission-critical line-of-business applications to the outdated, resource-intensive applications that are only occasionally used. Depending on the application and server resources, a single XenApp server has the potential to deliver applications to a handful of users or several hundred users simultaneously. Increased scalability can have a large impact on reducing the overall environment’s server footprint if overall resource utilization can be increased. Trying to get more efficiency from of a XenApp environment without requiring huge investments in hardware is a challenge organizations face. This challenge focuses on the following questions:

- How to best fully utilize servers without sacrificing the user experience?
- How to allow systems to perform multiple functions without requiring significant integration testing?

Greater Availability

Organizations are looking for ways to provide greater availability and responsiveness to applications. This means being able to overcome faults, failures and surges in application usage without significantly increasing the hardware footprint of the XenApp infrastructure, raising the questions:

- How to provide adequate availability taking into account cyclical surges peaks in usage?
- How to overcome faults in the environment without significantly increasing hardware resources?
- How to make applications available quickly in lieu of a disaster?

Increased Flexibility

There are more options than ever for delivering applications to users, and there are likewise more requirements than ever for providing the optimal working environment. Organizations must be able to develop and test system changes before rolling out into production, as well as find ways to integrate technical solutions into the XenApp environment. Providing a greater level of flexibility requires one to look into the following aspects:

- How to provide a development/test environment that closely mimics production?
- How to assimilate a mixed 32/64 bit environment?
- How to deliver a XenApp environment faster?
- How to provide the optimal IT environment allowing the business to grow?

Challenges

Currently, most XenApp deployments are functioning in a purely physical environment. Each server in the infrastructure has a single purpose either to function as Web Interface, licensing, zone data collector, data store or member server resulting in a 1:1 relationship between role and physical device. These environments typically resemble something like Figure 1.

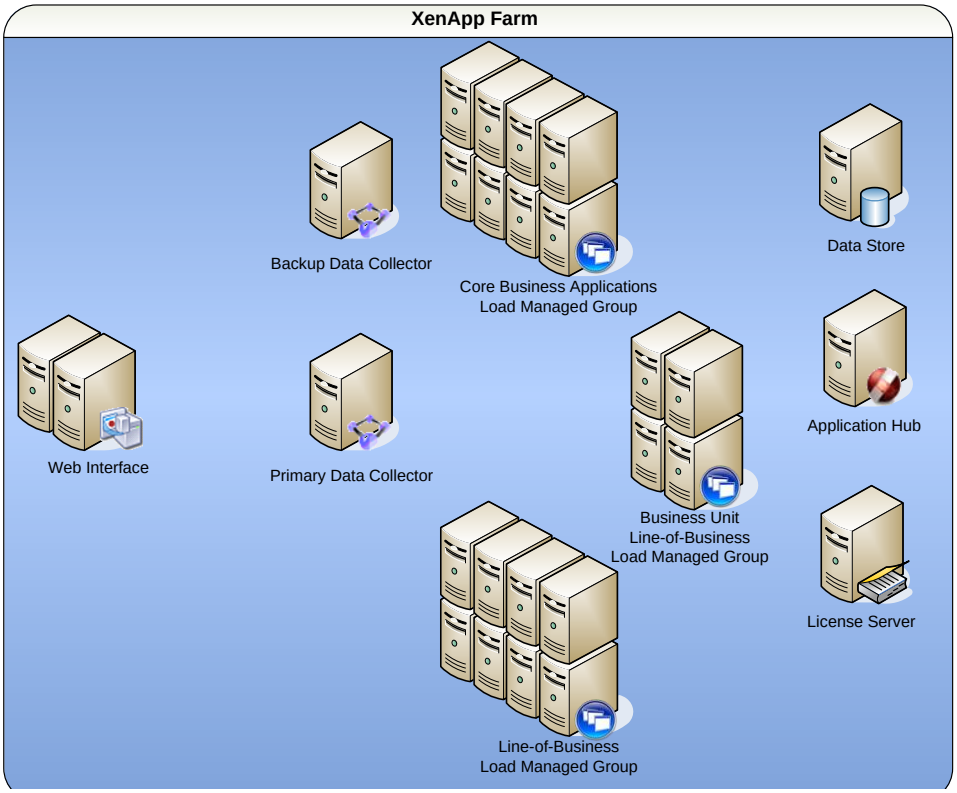

Figure 1: Physical XenApp Architecture

Improving availability and flexibility often results in lower utilization rates because more physical servers are required. For example, look at the following aspects of this architecture:

- Web Interface: The Web Interface servers are responsible for delivering applications to the users by acquiring their authentication information, enumerating and displaying available applications and forwarding the application launch request to the XenApp server. Because Web Interface is a critical component, redundant servers must be available to provide fault tolerance. As the XenApp farm increases in size, additional Web Interface servers will be added to handle the greater potential workload. Providing fault tolerance and designing for the maximum possible workload results in an even greater number of underutilized servers.
- Data Collector: The data collector is responsible for authenticating users, identifying accessible applications, and identifying which XenApp server a user should connect. The data collector is the brokering mechanism for requests coming from the end user and Web Interface destined to the XenApp farm. As the size of the XenApp farm increase, the data collector moves from becoming a shared server, responsible for delivering applications, to a dedicated server. If the primary data collector were to fail, a backup, with the same hardware and software configuration, should also be available. Similar to Web Interface, providing fault tolerance through the use of additional allocated hardware and providing optimal responses through the use of dedicated hardware requires more physical servers resulting in an even greater number of underutilized servers
- Load Managed Groups: Whether applications are installed or delivered via streaming to the XenApp servers, organizations might create load managed groups based on business requirements. Load managed groups are created to focus a set of XenApp servers on a particular set of applications. This is done for numerous business and technical reasons including application update frequency, business unit server ownership, application criticality, regional applications, and application and server language requirements.

When creating a load managed group, each group must provide enough redundancy to be capable of supporting all users in the event of a server failure. This results in an N+1 scenario where there is at least one additional XenApp server per load managed group. In many situations, organizations implement an N+10% strategy where an additional 10% of XenApp servers per load managed group are allocated in order to allow for multiple server failures or maintenance.

In this architecture, there are three load managed groups focused on the following:

- Line-of-Business: Every organization has a set of applications that are critical to the proper functioning of the business, often called Line-of-Business applications. In this particular example, the Line-of-Business servers experience high utilization just from the day-to-day activities of the business. These systems are optimal in their physical world as utilization is high, but if a few servers fail, the remaining systems will not be able to support the business.
- Business Unit Line-of-Business: Driven by organizational structure, this scenario includes a load managed group for another line-of-business application meant for a single business unit. The business unit purchased the hardware and had it integrated into the corporate XenApp environment. This group of servers is lightly loaded and is meant for only a single business unit.
- Core Applications: Many XenApp environments contain a load managed group focused on the delivery of the remaining core business applications like Microsoft Office, Adobe Acrobat, Internet Explorer, etc. This load managed group is not mission critical, but this group of servers experiences the highest user loads of all load managed groups. Although many users are connected to the servers, most applications are idle. For example, most users keep their email client open all day, but only interact with the application when a new message appears. However, during certain periods throughout the year, this load managed group experiences a huge surge in utilization due to month-end and year-end report generation. Because it is essential that there are enough servers available his peak demand, the environment was designed for maximum usage, resulting in underutilized servers the rest of the year.

- License Server: The license server receives license check-in and license check-out requests from XenApp server. This service is fairly lightweight, but often hosted on its own physical server. Most servers include dual or quad processors as standard, but as the license server is a single threaded application these additional processors remain unused. If the license server is unavailable for more than 30 days, users will be denied access. This has led some organizations to implement either a hot or cold standby. In either case, a new physical server is required but unused unless the primary license server fails. This server takes up data center space and potentially consumes power and cooling resources just to mitigate against the possibility that the primary server fails and cannot be restored in 30 days.
- Data Store: The data store is a critical component as all static farm information is stored within the data base. Due to the criticality of this database, some organizations dedicate a server for this purpose or at least dedicate the server for additional XenApp farm databases like Configuration Logging, SmartAuditor, or EdgeSight. In this scenario, an entire physical server is being used to host multiple databases that could range in size from a few megabytes to gigabytes. Regardless of what the server is hosting, the important concern for the Data Store is that it is backed up and can be restored quickly in the event of a hardware failure.
- Dev/Test Environment: A large percentage of XenApp environments have implemented some form of a development and test environment to validate operating system changes and application changes before being rolled out into production. The Dev/Test environment, to be of value, should mimic production as closely as possible. Dev/Test environments require their own dedicated physical infrastructure and must be built, managed and maintained in an identical fashion to production systems.
- Hardware: Due to technical limitations within the operating system, organizations have been forced into relatively small server standards, which take up a considerable amount of real estate in the data center. Because of a 4GB limit on the Windows 32bit operating system, organizations have had to purchase many servers just to support the users. Moving towards a 64bit platform would go a long way in overcoming the memory bottleneck, but due to application or driver inconsistencies, many organizations are not ready to make the lead into 64bit computing. In a physical environment, this leaves organizations with few options.

In addition, organizations are also interested in becoming more flexible by moving with changing customer needs. To achieve this, the IT organization must also be able to change or expand the infrastructure easily. In a truly physical environment, this is difficult as servers require installation and configuration. There are deployments tools available allowing for these processes to be automated, but it can still take hours to build a single server.

Ultimately, a purely physical XenApp environment has the potential to introduce a number of underutilized servers into the architecture. Conversely, if resources are fully utilized, then there is an inherent chance of limiting fault tolerance within the environment. Finding ways to provide higher utilization, while improving availability and flexibility is a one of the key drivers towards to the adoption of server virtualization solutions like Citrix XenServer.

Architecture

Server virtualization is touted as the ultimate solution for providing better utilization, availability and flexibility. But 100% server virtualization is not the ultimate solution in many cases. In fact, most environments will achieve the best utilization, availability and flexibility with a blend of virtual and physical servers. Trying to integrate physical and virtual environments can be challenging because the processes needed to manage these disparate environments are distinctly different. To better illustrate how a mixture of both physical and virtual provides the best solution for utilization, availability and flexibility and how the physical/virtual environments can be merged into a single, cohesive solution, two different architectures will be explored:

1. XenServer Enterprise
2. XenServer Platinum

XenServer Enterprise

Introducing server virtualization into XenApp environments can improve ROI by improving utilization, availability and flexibility. However, incorporating server virtualization within a XenApp environment without proper planning can be counterproductive as the solution chosen might negatively impact the user's environment. The environment must be analyzed before deciding which servers to virtualize and which to keep physical, which is explained in the Citrix article "***XenServer and XenApp – Better Together: Design Considerations***".

Part of the analysis is to determine which systems to virtualize. It is true that just about any server can undergo server virtualization, but what is the best virtualization solution for a particular workload? Virtualizing XenApp is not the same as virtualizing a web server or an email server. Many XenApp administrators are familiar with different optimization settings to make on a XenApp server, which are not required on non-XenApp servers. This is because a XenApp server is a unique application delivery system that must be understood before it can be optimized. For example, XenApp can support hundreds of simultaneous user sessions, each with multiple processes resulting in thousands of concurrent processes. Due to the sheer number of running processes, the number of memory page table entries and context switches also increases significantly. As an example of the tight integration between XenServer and XenApp, the XenServer memory algorithms have been enhanced to accommodate the uniqueness of the virtual server. These enhancements are easily implemented with a flip of a switch on the virtual machine. Without the memory optimizations, scalability of the virtualized XenApp server would be compromised. The impact these enhancements have will be demonstrated in a forthcoming scalability document.

Once the analysis is complete and the workload is understood, integrating XenServer Enterprise within a XenApp environment will most likely result in an architecture similar to that shown in Figure 2:

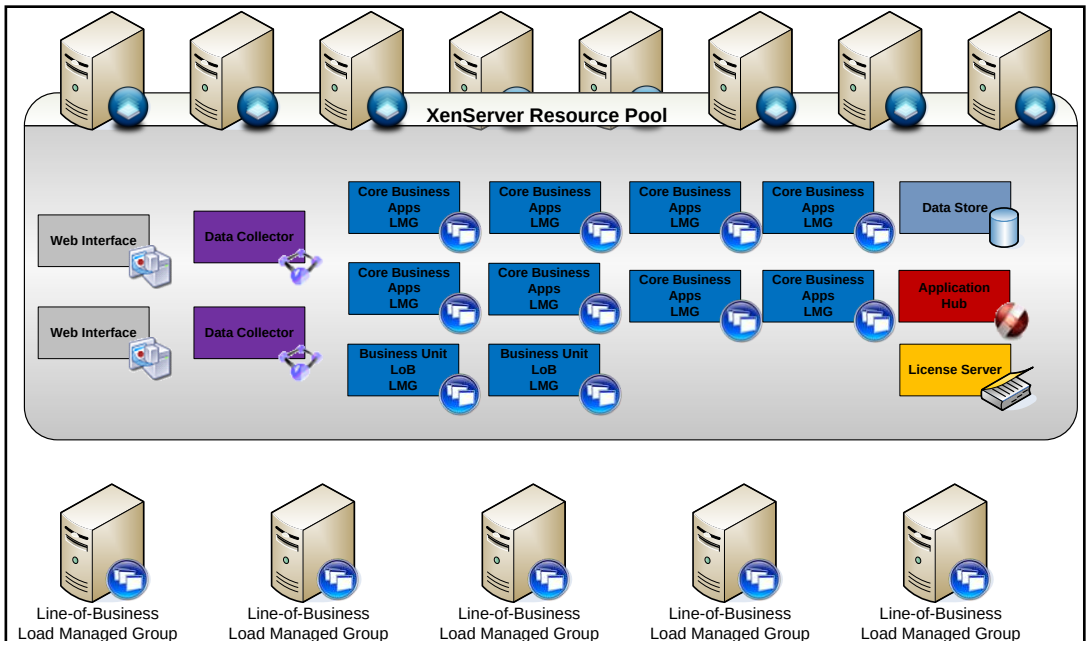

Figure 2: XenServer Enterprise Architecture

All components, except the Line-of-Business load managed group, will be virtualized inside of a XenServer Resource Pool. This architecture helps us to improve upon the 1:1 server to role ratio in the physical world, delivering a 1:Many ratio in the virtual world. Many different systems and roles can be executed on a single physical server. In the simplest form, moving a physical XenApp server to a XenServer solution is achieved by running physical-to-virtual migration tools where each physical server is placed inside of a virtual server instance. This migration has the following impact on the overall architecture and components:

- The Line-of-Business load managed group is already fully utilizing the physical hardware resources. Virtualizing this set of servers will not improve utilization. In fact, as another layer is added, the overall user concurrency of the system may be reduced slightly, requiring more virtual servers. Virtualizing this load managed group does not mitigate the original risk that these servers are already fully utilized and any server downtime would result in degraded performance. Based on the risk, an additional 10% could easily be added to the environment as virtual servers.
- More Effective Resource Allocation: All virtualized components benefit because they are allocated an appropriate number of resources based on their requirements. There are still redundant systems to provide better availability, but these additional virtual servers are also allocated the appropriate level of resources. For example, multiple physical Web Interface servers were deployed to provide fault tolerance and better availability. As these physical servers were mostly unused, the extra CPU cycles were wasted. In a virtualized environment, multiple Web Interface servers are still deployed; unused CPU cycles are used by other virtual machines within the resource pool.
- Fault Tolerance: Providing fault tolerance for some components, like the license server or data store, required the allocation of an additional physical server. The redundant systems would either be idle or powered off.
    - Virtualizing the data store allows for fault tolerance across hardware as a virtualized data store could be migrated to another physical server in the event of a potential hardware failure.
    - Virtualizing the license server allows a cold standby to be easily created. Once the license server is configured and ready for production, a copy or clone is created and kept powered off. No CPU or memory resources are allocated while in a shut down state. The powered off system is only consuming hard drive space. If the active license server fails and cannot be restarted, the cold standby can be powered on and immediately take over the license server responsibilities.

- Dev/Test: Utilizing XenServer Enterprise in the Dev/Test environment allows for easier synchronization with production at a great cost savings. A production-level virtual server could be copied into the Dev/Test environment, allowing for identical environments with little administrative overhead, albeit minor changes would be required to link a dev/test server with a dev/test backend. As a Dev/Test environment must remain flexible to execute different test scenarios, the virtual Dev/Test environment could quickly take on numerous roles as different virtual servers could be started.
- Hardware: Even if organizations are unable to move to the 64bit computing model, virtualization allows for more options on the hardware selected for the XenApp architecture. Organizations can now utilize servers with 16 or even 32GB of memory by slicing the physical server into multiple virtual servers each consisting of 4GB of memory. This allows the organization to "virtually" overcome the 4GB memory limitation of the 32bit Windows operating system.
- Storage Optimization: Physical systems rely on local storage that is shared with an operating system not dedicated to file sharing. Each XenApp server relies on the correct components, installation and configuration of the local storage in order to perform at the most optimal level. However, when XenApp servers are moved into a virtual environment, local storage is no longer an issue because XenServer utilizes shared network storage. The shared network storage should be a dedicated device that is dedicated to file sharing and whose operating system is optimized for file sharing.

One of the greatest benefits from combining XenServer Enterprise with a XenApp environment comes from helping to improve overall utilization because the resources from underutilized servers are appropriately allocated and shared with the rest of the resource pool. This integration adds a significant amount of flexibility and availability to the environment while allowing for better utilization of the hardware. However, this integrated solution still poses a few noticeable challenges that must be overcome.

The first challenge is focused on virtual images. Every virtual server is, in all respects, a separate physical system encapsulated within a file. Although a physical server can now host multiple virtual servers (1:Many), at the virtual level, there is still a 1:1 relationship between the virtual server and its respective physical system encapsulated file. The ease with which new servers can be provisioned and brought online can lead to 'virtual server sprawl' and actually decrease the benefits of virtualization. This brings about a few concerns:

- How much shared storage is required to support multiple distinct virtual servers?
- How are the virtual servers maintained from an application and operating system patching perspective?

The second major concern is around physical/virtual integration. Because some physical systems would not show business benefits to being virtualized, two different systems must be maintained: the physical and the virtual. They are maintained differently, thus requiring the administrator to remember what is physical and what is virtual, while managing them appropriately and with the correct tools. For example, to provide greater levels of fault tolerance, additional virtual servers were added for the Line-of-Business servers, even though the core servers are physical. Patching, imaging, etc., for the Line-of-Business servers would follow different processes, not to mention all other servers within the environment. This introduces greater complexity into the environment, resulting in a higher probability of manual errors.

XenServer Platinum, when integrated with XenApp, can help alleviate these challenges as the next section will explain.

XenServer Platinum

With server virtualization, we are able to provide for better utilization by sharing resources between virtual servers. However, this still posed a challenge around flexibility and availability as physical and virtual servers are managed differently and the maintenance of virtual servers can be even more complex and tedious as the physical. Even though a physical server can now host multiple virtual servers, each virtual server is still a separate entity that must be managed and maintained individually. By integrating XenServer Platinum within a XenApp environment, greater flexibility and availability are achieved while still fully utilizing the physical servers.

XenServer Platinum introduces a new solution called Provisioning Server for Data Centers. Instead of managing and maintaining hundreds of virtual machine images, Provisioning Server for Data Centers only requires one virtual disk

image per server role. The architectural advantages introduced by leveraging the Provisioning Server component of XenServer Platinum can be seen in Figure 3.

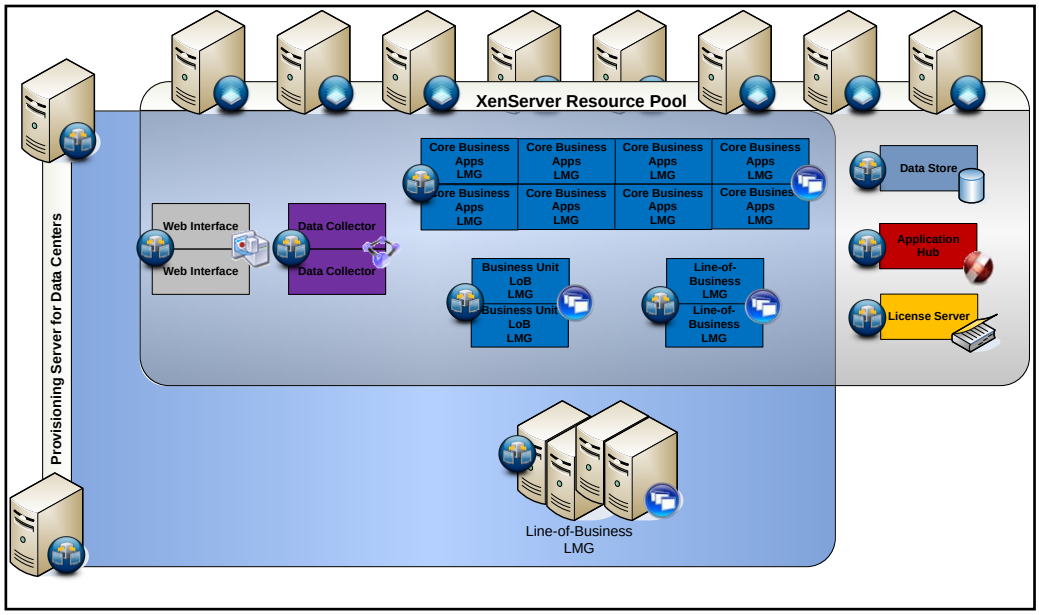

Figure 3: XenServer Platinum Architecture

With the defined XenApp architecture, there are five distinct roles:

1. Web Interface
2. Data Collector
3. Core Business Apps
4. Business Unit Line-of-Business
5. Line-of-Business

Each role would have its own base server image. The image would be created and maintained as a single entity. For common/shared images, this one image would then be delivered (streamed) to any number of physical or virtual servers, which can be easily changed through the management console. The common or shared images are read only. Changes made to the system are lost upon reboot. This guarantees that unforeseen changes will not impact a server.

The power of this solution is based on the ease of management and maintenance of the images. If there are 100 core application servers, this solution only requires changes to be made to the common image. All 100 core application servers will receive the changes without administrative intervention. This solution also makes it easier to add new capacity to deal with surges in usage. If capacity needs to be made to the Line-of-Business group of servers, then the base image for that role would be applied to other server objects within the console, regardless of the server object being a physical or virtual server. When that server starts, it will become a Line-of-Business server.

As there is one image for hundreds of servers, all servers within the role share the same characteristics (operating system, application, configuration settings, computer name and SID). Every server belonging to the role is identical except for a MAC address. This one image must be capable of taking on different identities in order to function properly in a XenApp environment. This is achieved through the Preparation, Startup and XenApp Integration phases:

Preparation Phase: Before an image is delivered to multiple servers, the image must be prepared appropriately.

- To be integrated into Active Directory, the defined server, within the Provisioning Management console, must be added to Active Directory. Provisioning Server uses standard Windows API calls to add the server into the correct domain and organizational unit.
- Provisioning Server is also set to manage computer account password with Active Directory so the provisioned servers will be able to authenticate with Active Directory.
- Before a XenApp server is delivered, an integration utility must be executed to prepare the server for XenApp provisioning. XenApp services are stopped, local host cache instances are cleared and computer names are

modified. These modifications allow the server to become part of the XenApp farm during server delivery, as explained in the XenApp Integration step.

Startup Phase: The server is no longer using its localized operating system present on its local storage. Instead, the server must receive its role across the network. This process takes the following steps during startup:

- A physical/virtual server starts and requests an IP address from the DHCP server. The DHCP response includes two additional configuration options telling the server what the boot program is called and what server can deliver the boot program.
- The server makes a TFTP request for the boot program from the specified server
- The Provisioning Server sends the boot program to the requesting server
- The server executes the boot program
- The boot program determines, based on the Provisioning Server configuration, which image the server should receive based on the server's MAC address
- The image is streamed to the server so the server can launch he operating system. Additional data is continually streamed to the server throughout the duration of the session.
- During startup, the server uses the installed Provisioning Server drivers to change the computer name to the specified value within the Provisioning Server Console

XenApp Integration: XenApp environments store member server information within the data store. This information is created when a new server is added. The information is modified based on changes made within the XenApp management console. Because a group of XenApp servers share the same role, they originate from the same base Provisioning Server image. Each server must have its presence established within the data store. This is accomplished by the following:

- When the server starts and its computer name has been changed based on the previous processes, a XenApp integration utility is automatically executed
- The utility adds/syncs the server with the XenApp data store by restarting the XenApp services
- The XenApp server receives its instructions for published applications and starts accepting connections

Results

Based on the three architectures identified within this document, it should be clear that a “one size fits all solution” is not the best approach. The best approach should be able to integrate the best utilization, availability and flexibility for physical servers as well as virtual servers. Recapping the three architectures, we can observe the following:

	Physical	XenServer Enterprise	XenServer Platinum
Utilization	**Standard Solution** • Server resources are not shared between other servers	**Good Solution** • Optimized for XenApp’s unique workload • Low latency, small/efficient codebase and specialized hardware component utilization results in nominal impact to the user’s experience • Physical server resources are shared between virtualized servers	**Best Solution** • Includes all XenServer Enterprise benefits • Storage is better utilized as a single image can be used across multiple physical/virtual servers • The N+10% or N+1 algorithms are not as critical because the server role can be updated, allocated, and delivered without removing servers for maintenance windows
Availability	**Standard Solution** • Adding redundancy results in wasted utilization because system resources are minimally used or idle	**Good Solution** • Unused redundant systems can be powered down while resources are used by other virtual servers • Hardware issues are overcome by moving active virtual servers to new hosts or by starting new systems on available virtual or physical systems	**Best Solution** • Includes all XenServer Enterprise benefits • Stable base image not modifiable, unless put into modification mode • Faults (blue screens) to virtual and physical servers because of a software issue are overcome with a reboot • Disaster recovery easily accomplished by pointing images to new server objects, regardless of them being physical or virtual
Flexibility	**Standard Solution** • Each physical server must be built, installed and configured before it becomes usable, a process that can take hours or days. • Each physical server is a separate entity and managed distinctly • Each new server model requires new drivers and configurations, making physical server images difficult to maintain	**Good Solution** • New servers can be brought online in a few seconds as virtual servers can be cloned. • Virtual servers can be started and moved to different physical servers without requiring downtime • Dev/Test environments can be built using the same images as production	**Best Solution** • Includes all XenServer Enterprise benefits • Allows a single image to be used for multiple servers simultaneously • Images can be easily managed and delivered across physical and virtual servers • A physical or virtual server’s role can be changed quickly allowing environment to change as demands change

Conclusion

Although server virtualization is in high demand, it should be used to address business requirements. Some IT organizations are considering implementing a server virtualization solution for many aspects of their infrastructure believing it will solve all of their environmental challenges. However, server virtualization alone is not the answer. In fact, within certain environments, server virtualization might not even make sense. This reference design focuses on where and when server virtualization fits within Citrix XenApp environments by focusing on the three major reasons why organizations are looking at server virtualization solutions: utilization, availability and flexibility.

Virtualizing XenApp environments requires more than adding server virtualization. The solution must consider the workload, how images will need to be managed, how much storage is required and the how scalability will be affected. Looking at three potential architectures, the following can be seen:

- Physical Environment: Although some systems are running at full utilization, many of the other systems are not optimized and are wasting resources. This environment is completely static, making it difficult for organizations to quickly and dynamically meet their rapidly changing business requirements.
- XenServer Enterprise: Optimized for XenApp workloads by changing the way memory is managed and utilized between virtual servers. Integrating server virtualization alleviates utilization challenges by allowing multiple systems to share resources; however, server virtualization does not solving the management and maintenance concerns of a growing infrastructure. Although the environment is dynamic in that it allows systems to move from different physical servers, it can only provide a basic step into virtualization.
- XenServer Platinum: The virtual boundaries are broken, allowing this solution to allocate server roles to any number of servers including the virtual and physical. Consolidating hundreds or thousands of server images into less than ten greatly simplifies costs, management and maintenance activities. This time and cost savings allows organizations to appreciate how their environment can provide greater value to business rather than spending money to support the infrastructure.

The best virtualization solution for your organization incorporates an architecture that delivers resources to users in the most cost efficient manner available.

Notice

The information in this publication is subject to change without notice.

THIS PUBLICATION IS PROVIDED "AS IS" WITHOUT WARRANTIES OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT. CITRIX SYSTEMS, INC. ("CITRIX"), SHALL NOT BE LIABLE FOR TECHNICAL OR EDITORIAL ERRORS OR OMISSIONS CONTAINED HEREIN, NOR FOR DIRECT, INCIDENTAL, CONSEQUENTIAL OR ANY OTHER DAMAGES RESULTING FROM THE FURNISHING, PERFORMANCE, OR USE OF THIS PUBLICATION, EVEN IF CITRIX HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES IN ADVANCE.

The exclusive warranty for Citrix products, if any, is stated in the product documentation accompanying such products. Citrix does not warrant products other than its own.

Version History			
Author	Version	Change Log	Date
Daniel Feller	0.1	Document created	April 1, 2008
Daniel Feller	0.2	Updates made to XS-E section based on feedback	April 8, 2008
Daniel Feller	1.0	Document Finalized	April 11, 2008
Daniel Feller	1.1	Document diagram updated	April 23, 2008

851 West Cypress Creek Road Fort Lauderdale, FL 33309 954-267-3000 http://www.citrix.com