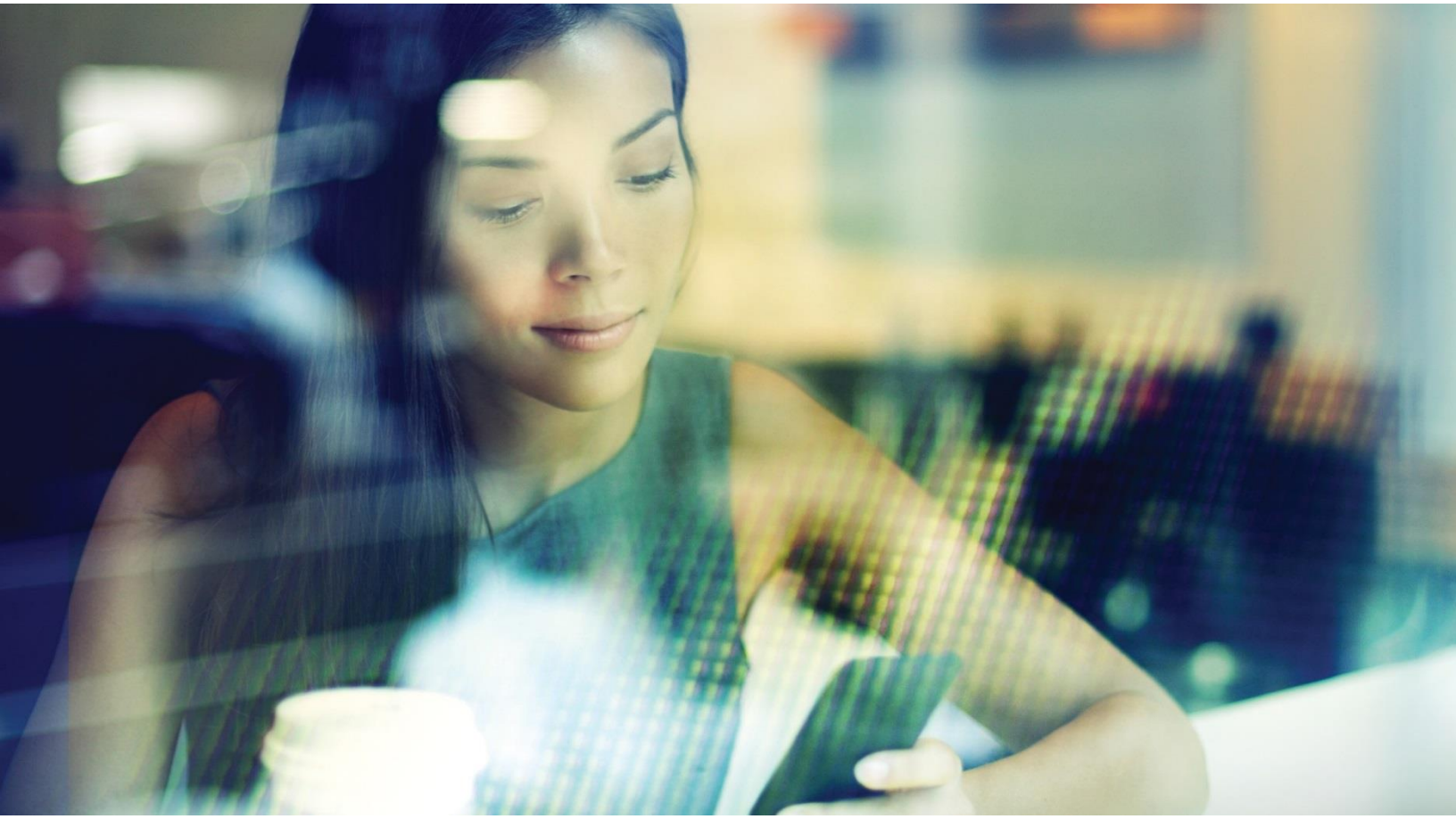


Citrix VDI Handbook and Best Practices

XenApp and XenDesktop 7.6
Long Term Service Release

Version 1.0

Table of Contents

Section 1: Overview 10
Introduction 10
Methodology 11
Section 2: Assess 12
Overview 12
Step 1: Define the Organization 12
Step 2: Define the User Groups 13
Step 4: Define the Applications 19
Step 5: Define the Project Team 21
Section 3: Design 29
Overview 29
Layer 1: The User Layer 29
Layer 2: The Access Layer 35
Layer 3: The Resource Layer 49
Layer 4: The Control Layer 70
Layer 5: The Hardware Layer 97
Section 4: Monitor 106
Overview 106
Process 1: Support 106
Process 2: Operations 118
Process 3: Monitoring 129
Acknowledgments 141
Authors 141
Subject Matter Experts 141
Revision History 142

About Citrix

Citrix (NASDAQ:CTXS) is leading the transition to software-defining the workplace, uniting virtualization, mobility management, networking and SaaS solutions to enable new ways for businesses and people to work better. Citrix solutions power business mobility through secure, mobile workspaces that provide people with instant access to apps, desktops, data and communications on any device, over any network and cloud. With annual revenue in 2015 of $3.28 billion, Citrix solutions are in use at more than 330,000 organizations and by over 100 million users globally. Learn more at www.citrix.com.

Section 1: Overview

Introduction

In traditional business environments, workers suffer from productivity loss in many ways, including downtime during PC refreshes, patches and updates, or simply when they are away from the office. Application and desktop virtualization centralizes apps and desktops in the datacenter, rather than on local devices. This allows IT to deliver apps and desktops to users on demand, to any device, anywhere.

Take the following response from a desktop virtualization user:

Experience from the Field

As a remote employee for [company], I struggled every time I needed to access the company's intranet, which forced me to VPN into the network. I also kept data on my local device because trying to access it over my broadband connection was too slow. Some coworkers did the same and lost data due to a virus, thankfully I was luckier.

Depending on my mood (and the weather), changing devices and locations was a challenge as I had to have my applications and data copied to many different endpoints. I know this was unsecure, but I didn't care because I was more concerned with flexibility.

Since moving to a virtual desktop, I'm able to use any device. I'm able to work from any location. And best of all, I don't have to worry about copying my data and applications onto all of my personal devices.

Unfortunately, organizations sometimes struggle to achieve this level of success. Why does one organization succeed while another organization struggles?

If we compare the factors between success and failure between desktop virtualization and other technology related projects, we see that there is little difference:

1. **Lack of justification –** Without a solid business reason, desktop virtualization is simply a new way to deliver a desktop. A business justification gives the project team a goal to strive towards.

2. **Lack of a methodology –** Many people who try and struggle to deploy a desktop virtualization solution do so because they jump right in without understanding or implementing the appropriate prerequisites. A structured methodology provides the path for the project.

3. **Lack of experience –** For many who embark on a desktop virtualization project, there is a lack of experience, which creates a lack of confidence in the design. Architects begin to second-guess themselves and the project stalls.

Our hope is that this handbook can alleviate the anxiety associated with desktop virtualization by showing how challenges can be resolved in a manner that is technically sound, but also feasible and effective for organizations facing deadlines and other organizational challenges.

Citrix has successfully employed the methodology, experience and best practices shared within this handbook across thousands of desktop virtualization projects.

Methodology

The Citrix VDI Handbook follows the Citrix Consulting methodology. A proven methodology that has been successfully employed across thousands of desktop virtualization projects. Each phase includes guidance on the important questions to ask, what tools to use and tips to help you succeed. The Citrix Consulting methodology consists of five phases:

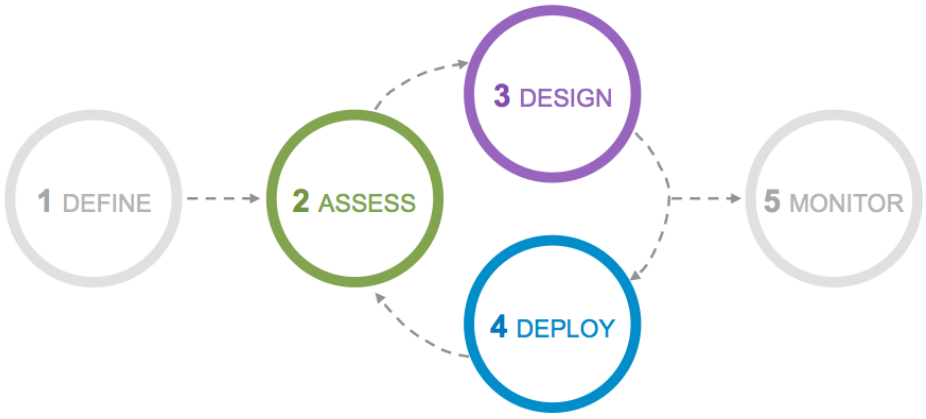


Figure 1: Citrix Consulting Methodology

1. **Define** – Builds the business case for desktop virtualization by creating a high-level project roadmap, prioritizing activities and estimating storage and hardware requirements.
2. **Assess –** Key business drivers are rated so that work effort can be prioritized accordingly. In addition, the current environment is reviewed for potential problems and to identify use cases for the project. This information will be used to set the direction of the Citrix deployment, upgrade, or expansion.
3. **Design –** Define architecture required to satisfy key business drivers and success criteria identified during the assess phase. Topics such as environment scalability, redundancy and high availability are addressed.
4. **Deploy –** During the deploy phase, the infrastructure is installed and configured as described in the design phase. All components of the infrastructure should be thoroughly unit and regression tested before users are provided with access to the environment.
5. **Monitor –** Define architectural and operational processes required to maintain the production environment.

The Citrix Consulting methodology follows an iterative Assess > Design > Deploy process for each major initiative of your project. In doing so, your organization is left with tangible improvements to the environment at the end of each engagement. For example, high priority user groups can progress through the assess, design and deploy phases earlier than other user groups

Note: *The VDI Handbook provides content on the Assess, Design and Monitor phases of the Citrix Consulting methodology.*

Section 2: Assess

Overview

Creating an app and desktop delivery solution begins with a proper assessment. Architects that fail to properly assess the current environment find that they require the assess information later on, forcing them to backtrack, which can potentially stall and put the project at risk.

By gathering all of the information from the outset, the architect will gain an appreciation for the current environment and be able to work from the beginning on properly aligning business and user requirements with the overall solution.

The assess phase is a four-step, simple to follow process:

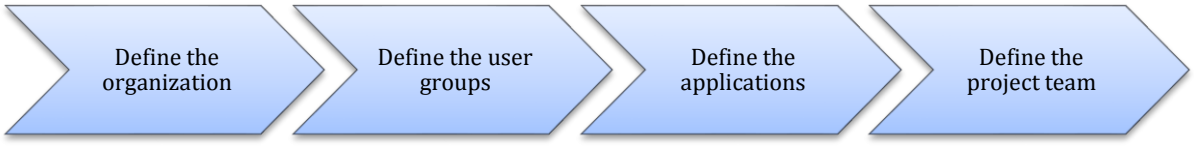


Figure 2: Assess Process

Step 1: Define the Organization

The first step in your virtual desktop project should be to understand and prioritize the strategic imperatives of the organization. This enables the project management team to define success criteria and allows the design team to create a tailored and optimized architecture.

Requirements can be captured during meetings or by distributing questionnaires. Meetings are more time consuming, but allow for follow-up questions to be asked and help to simplify the prioritization process. It is important that this exercise be completed jointly by both business managers and IT decision makers since both groups will have significantly different viewpoints. Take the following examples of what certain organizations faced, which drove their selection of desktop virtualization.

Experience from the Field

Finance – A large financial institution had a base of operations in the city designated as the host city for an upcoming G8 summit. As these types of meetings historically include riots, protests and other issues that can disrupt business and the safety of their employees, the financial organization needed an alternative allowing their users to work from the safety of their homes.

Agriculture – Due to thin margins, an agriculture organization wanted to save money by extending the life of desktop PCs while still being able to run the latest applications.

Healthcare – A large healthcare organization was in need of a solution to simplify application updates as the main application required updates on a weekly basis. Due to the distributed nature of the endpoint devices, the organization was in need of a better application delivery solution.

These are just a few examples, but they demonstrate how organizations think about their priorities. Most organization do not focus on technology, they focus on the needs of the user and of the organization. These needs can be met with technical solutions but it is imperative the team understands the "Why" of the project.

In addition to the three real-world examples, the following table identifies a few other priorities often stated from many organizations:

Requester	Requirement
Business managers	**Better IT agility and responsiveness –** Flexible desktop solution that is capable of accommodating periods of change such as rapid growth or downsizing. For example, enabling the business to setup project offices or temporary points of sale very rapidly without long delays or IT notification periods.
	Bring your own device – Empower employees to choose their own devices to improve productivity, collaboration and mobility.
	Collaboration – With an increase in both globalization and mobility, team members are often dispersed across multiple physical locations. Powerful collaboration capabilities are required to ensure high levels of productivity, efficiency and quality.
	Work from anywhere – The business needs to support home workers in order to attract and retain top talent, and / or travelling employees.
IT decision makers	**Better desktop management –** Simplify the management of desktop infrastructure. IT is not as proactive as they would like and spend too much time "fighting fires".
	Increase security – Data theft or the loss of devices containing sensitive data is a big risk and preventive measures are a top priority.
	Extend desktop hardware lifecycle – Replacing workstations every three to five years in order to keep up with the requirements of the operating system or the applications has been very costly.
	Reducing cost – Cut costs associated with supporting and maintaining traditional desktops.
	Improving user experience - Increasing performance or enabling features which would otherwise not be possible with a georgraphically dispersed user population

Table 1: Sample Business Drivers

The prioritization process should be completed in collaboration with the project team, business managers and IT managers so that all views are considered.

Step 2: Define the User Groups

Although there are multiple approaches towards defining user groups, it is often easiest to align user groups with departments as most users within the same department or organizational unit consumes the same set of applications.

User Segmentation

Depending on the size of the department, there might be a subset of users with unique requirements. Each defined user group should be evaluated against the following criteria to determine if the departmental user group needs to be further divided into more specialized user groups.

- **Primary datacenter** – Each user will have a primary datacenter assigned that will be used to host their virtual desktop, data, and application servers. Identify the datacenter that the user should be assigned to rather than the datacenter they are currently using. Users will be grouped based on their primary datacenter so that a unique design can be created for each one.
- **Personalization –** Personalization requirements are used to help determine the appropriate VDI model for each user group. For example, if a user group requires complete personalization, a personal desktop will be recommended as the optimal solution. There are three classifications available:

Personalization	Requirement
None	User cannot modify any user or application settings, for example - kiosk.
Basic	User can modify user-level settings of desktops and applications.
Complete	User can make any change, including installing applications.

Table 2: Personalization Characteristics

- **Security –** Security requirements are used to help determine the appropriate desktop and policy (or policies) for each user group. For example, if a user group requires high security, a hosted pooled desktop or a local VM desktop will be recommended as the optimal solution. There are three classifications available:

Security Level	Description
Low	Users are allowed to transfer data in and out of the virtualized environment.
Medium	All authentication and session traffic should be secured; users should not be able to install or modify their virtualized environment.
High	In addition to traffic encryption, no data should leave the data center (such as through printing or copy/paste); all user access to the environment should be audited.

Table 3: Security Characteristics

- **Mobility –** Mobility requirements are used to help determine the appropriate desktop model for each user group. For example, if a user group faces intermittent network connectivity, then any VDI model requiring an active network connection is not applicable. There are four classifications available:

Mobility	Requirement
Local	Always uses the same device, connected to an internal, high-speed and secured network.
Roaming Local	Connects from different locations on an internal, high-speed, secured network.
Remote	Sometimes connects from external variable-speed, unsecure networks.
Mobile	Often needs access when the network is intermittent or unavailable.

Table 4: Mobility Characteristics

- **Desktop Loss Criticality –** Desktop loss criticality is used to determine the level of high availability, load balancing and fault tolerance measures required. For example, if there is a high risk to the business if the user's resource is not available, the user should not be allocated a local desktop because if that local desktop fails, the user will not be able to access their resources. There are three classifications available:

Desktop loss criticality	Requirement
Low	No major risk to products, projects or revenue.
Medium	Potential risk to products, projects or revenue.
High	Severe risk to products, projects or revenue.

Table 5: Desktop loss criticality Characteristics

- **Workload –**Types and number of applications accessed by the user impacts overall density and the appropriate VDI model. Users requiring high-quality graphics will either need to utilize a local desktop implementation or a professional graphics desktop. There are three classifications available:

User Type	Characteristics
Light	1-2 office productivity apps or kiosk.
Medium	2-10 office productivity apps with light multimedia use.
Heavy	Intense multimedia, data processing or application development.

Table 6: Workload Characteristics

Note: *Performance thresholds are not identified based on processor, memory or disk utilization because these characteristics will change dramatically following the application rationalization and desktop optimization process. In addition, it is likely that the user's management tools and operating system will change during the migration process. Instead, workload is gauged based on the number and type of applications the user runs.*

Experience from the Field

Utility company – A large utility company collected data on every user in their organization. During the user segmentation process it was realized that the organization's existing role definitions were sufficiently well defined that all the users within a role shared the same requirements. This allowed a significant amount of time to be saved by reviewing a select number of users per group.

Government – A government organization discovered that there was significant deviation between user requirements within each role, particularly around security and desktop loss criticality. As such, each user needed to be carefully reviewed to ensure that they were grouped appropriately.

Assign VDI Models

As with physical desktops, it is not possible to meet every user requirement with a single type of VDI. Different types of users need different types of resources. Some users may require simplicity and standardization, while others may require high levels of performance and personalization. Implementing a single VDI model across an entire organization will inevitably lead to user frustration and reduced productivity.

Citrix offers a complete set of VDI technologies that have been combined into a single integrated solution. Because each model has different strengths, it is important that the right model is chosen for each user group within the organization.

The following list provides a brief explanation of each VDI model.

- **Windows Apps** – The Windows apps model utilizes a server-based or desktop-based Windows operating system, where only the application interface is seen by the user. This approach provides a seamless way for organizations to deliver a centrally managed and hosted application into the user's local PC. The Windows app model is often utilized when organizations must simplify management of a few line-of-business applications.
- **Browser Apps** – The browser apps model utilizes a server-based Windows operating system to deliver an app as a tab within the user's local, preferred browser. This approach provides a seamless way for organizations to overcome browser compatibility challenges when users have the ability to use their own preferred browser (Internet Explorer, Microsoft Edge, Google Chrome, Mozilla Firefox, etc.).
- **Shared Desktop –** With the shared desktop model, multiple user desktops are hosted from a single, server-based operating system (Windows 2008, 2012, 2016, Red Hat, SUSE, CentOS). The shared desktop model provides a low-cost, high-density solution; however, applications must be compatible with a multi-user server based operating system. In addition, because multiple users share a single operating system instance, users are restricted from performing actions that negatively impact other users, for example installing applications, changing system settings and restarting the operating system.
- **Pooled Desktop –** The pooled desktop model provides each user with a random, temporary desktop operating system. Because each user receives their own instance of an operating system, overall hypervisor density is lower when compared to the shared desktop model. However, pooled desktops remove the requirement that applications must be multi-user aware and support server based operating systems.
- **Personal Desktop –** The personal desktop model provides each user with a statically assigned, customizable, persistent desktop operating system. Because each user receives their own instance of an operating system, overall hypervisor density is lower when compared to the shared desktop model. However, personal desktops remove the requirement that applications must be multi-user aware and support server based operating systems.
- **Pro Graphics Desktop –** The pro graphics desktop model provides each user with a hardware-based graphics processing unit (GPU) allowing for higher-definition graphical content.
- **Local Streamed Desktop –** The local streamed desktop model provides each user with a centrally managed desktop, running on local PC hardware
- **Local VM Desktop –** The local VM desktop model provides each user with a centrally managed desktop, running on local PC hardware capable of functioning with no network connectivity.
- **Remote PC Access –** The remote PC access desktop model provides a user with secure remote access to their statically assigned, traditional PC. This is often the fastest and easiest VDI model to deploy as it utilizes already deployed desktop PCs.

Each user group should be compared against the following table to determine which VDI model best matches the overall user group requirements. In many environments, a single user might utilize a desktop VDI model and an app VDI model simultaneously.

Segmentation Characteristic	Hosted Windows Apps	Hosted Browser Apps	Hosted Shared Desktop	Hosted Pooled Desktop	Hosted Personal Desktop	Hosted Pro Graphics Desktop	Local Streamed Desktop	Local VM Desktop	Remote PC Access
Workload									
Light	✓	✓	o	o	o	✕	o	o	o
Medium	o	o	✓	✓	o	o	o	o	o
Heavy	✕	✕	✕	✕	o	✓	✓	✓	o
Mobility									
Local	✓	✓	✓	✓	o	o	✓	o	o
Roaming Local	✓	✓	✓	✓	o	o	✕	o	o
Remote	✓	✓	✓	✓	o	o	✕	o	✓
Mobile	✕	✕	✕	✕	✕	✕	✕	✓	✕
Personalization									
None	✓	✓	✓	✓	✕	o	o	o	o
Basic	✓	✓	✓	✓	✕	o	o	o	o
Complete	✕	✕	✕	✕	✓		✕	✓	✓
Security									
Low	o	o	o	o	o	o	o	o	o
Medium	✓	✓	✓	✓	o	o	o	o	o
High	o	o	o	✓	✕	o	✓	✓	✕
Desktop Loss Criticality									
Low	o	o	o	o	o	o	o	o	o
Medium	✓	✓	✓	✓	o	o	o	o	✕
High	✓	✓	✓	✓	✕	o	o	✕	✕

"✓": Recommended, "✕": Not Recommended, "o" Viable

Table 7: VDI Model Capability Comparison

Don't forget to follow these top recommendations from Citrix Consulting based on years of experience:

Citrix Consulting Tips for Success

1. **Start with Windows apps, shared and pooled desktops –** As you can see in the VDI capability table above, the Windows apps, hosted shared and pooled desktop models can be used in the majority of situations. The local streamed and local VM desktop models should only be used on an exception basis. By reducing the number of VDI models required, you will help to reduce deployment time and simplify management.

2. **Perfect match –** It may not be possible to select a VDI model that is a perfect match for the user group. For example, you can't provide users with a desktop that is highly secure and offers complete personalization at the same time. In these situations, select the VDI model which is the closest match to the organization's highest priorities for the user group.

3. **Desktop loss criticality –** There are only three VDI models that meet the needs of a high desktop loss criticality user group (backup desktops available) – none of which allow for complete personalization. If a high-desktop loss criticality user group also requires the ability to personalize their desktop they could be provided with a pool of backup desktops (hosted shared, pooled) in addition to their primary desktop. Although these desktops would not include customizations made to their primary desktop, they would allow users to access core applications such as mail, Internet and Microsoft Office.

4. **Consider Operations & Maintenance –** The ongoing support of each VDI model should be factored in when deciding on a VDI model. For example, pooled desktops can be rebooted to a known good state which often leads to reduced maintenance versus a personal desktop where each desktop is unique.

Step 4: Define the Applications

Once the users have been divided up in to groups the next step is to determine which applications they require. This is a two-step process:

1. **Application rationalization –** Help to simplify the application assessment by removing redundant applications from the inventory that were captured during the data capture.

2. **Link apps to users –** Use the results from the data capture process to map applications to user groups.

Application Rationalization

The number of applications identified during the inventory is often surprising, even for organizations that believe they have a high-level of control over applications. To help reduce complexity as well as overall time required, it's important to take the time to consolidate the list of applications.

The following guidelines will help ensure that your application list is consolidated appropriately:

- **Multiple versions –** Different versions of the same application may have been identified during the inventory. There are various reasons for this, including an inconsistent patching or upgrade process, decentralized application management, limited licenses and situations where users require specific application versions for compatibility with other applications, macros and document formats. Where possible, work with the application owners to reduce the number of versions required. The leading practice is to standardize on a single version of each application, typically the latest.

- **Non-business applications –** Applications that are not required by the business should be removed from the application inventory to reduce resource requirements and to help simplify the overall project. Non-business related applications are typically found in an application inventory when users have been provided with the ability to install their own applications and typically include games, communication clients, screen savers, peripheral software and media players.

- **Legacy applications –** The inventory may identify legacy applications that have since been retired or that are no longer required within the business. These applications may not have been removed from the desktops because there is no established process to do so or because there are always more high-priority activities to complete. These applications should be consolidated during the rationalization stage of the application assessment.

- **Management applications –** The antivirus, application delivery, monitoring, inventory, maintenance and backup applications will be completely re-designed across the organization during the desktop virtualization project. These applications should also be consolidated during this stage.

Experience from the Field

Government: A government organization identified that there were 2,660 applications installed across their desktop estate. Most of which were installed by users with local administrative rights. By following the application rationalization recommendations above, it was possible to reduce the number of applications required to 160.

Application Categorization

Each application included in the project should be categorized based on certain criteria, which will help determine the most appropriate way to host and integrate the app. Each application can be installed directly into the image, virtualized in an isolated container and streamed to the desktop (Microsoft App-V), captured in a unique layer and attached to the virtual machine (Citrix AppDisk) or installed locally on the user's endpoint device and seamlessly integrated into the user's virtual desktop (Citrix Local App Access). Due to the uniqueness of every application, many large-scale deployments simultaneously utilize multiple approaches.

Each application should be categorized as follows:

- **Common Apps** - Every organization includes a suite of applications utilized by almost every user, Microsoft Office for example. This suite of applications is often the most utilized application in a desktop VDI model.
- **Departmental Apps** - A certain set of applications are only relevant for a unique business unit or department. For example, an engineering department will often require software development applications.
- **User Apps** - Often making up the largest grouping of apps are the apps used by very few individual users. In a traditional PC implementation, these applications are installed by the user as a temporary requirement or a personal requirement, often not directly impacting the business.
- **Management Apps** – Many desktop deployments include a combination of antivirus, monitoring, inventory, maintenance and backup applications. Many of these applications have unique virtualization requirements and are often required across the entire organization.

Application Characterization

The following characteristics should be identified for each application so that the right application delivery model can be selected during the design phase of the project:

- **Complex –** An application should be classified as technically challenging if it is complex to set up, has extensive dependencies on other applications or requires a specialized configuration, for example an Electronic Medical Records (EMR) application. Technically challenging applications need to be identified during the application assessment because they are not generally appropriate for installation in to a base desktop image or delivery by application streaming. Delivering technically challenging applications as a hosted Windows app will help to reduce the complexity of the base desktop image.

- **Demanding –** Collecting application resource requirements allows the virtualization infrastructure to be sized and an appropriate application delivery model to be selected. For example, resource intensive applications will not be delivered via a hosted shared desktop because there is limited control over how the resources are shared between users. There are two classifications available in the user assessment worksheet:

Workload	Requirement
Resource Intensive	Application requires 1GB+ of RAM or averages 50%+ CPU utilization.
None	The application is not resource intensive.

Table 8: Application Workload Characteristics

- **Mobile –** Some user groups may require the ability to work while mobile, sometimes when offline. If so, it is important that the design can determine which applications will work without a network

connection and which ones will not. Applications that require backend infrastructure such as web and database servers are not typically available offline.

- **Peripherals –** If applications require connectivity with peripheral devices, identify the interface required so that it can be made available to the application when it is run from a virtual session.
- **Restrictions –** Application access may need to be restricted due to insufficient licenses / resources and to protect sensitive data / tools. For example, applications with a limited number of licenses should not be installed on a base image that is shared with unlicensed users. There are three restricted access categories in the application assessment workbook:

Restricted Access	Requirement
No	There are no restrictions for the application and it can be accessed by any user within the organization.
User group	The application may be installed on a multi-user operating system but only a specific group of users should be provided with an icon.
Virtual machine	Application should only be installed on a virtual machine that is accessible by authorized users.

Table 9: Restricted Access Characteristics

Step 5: Define the Project Team

Desktop virtualization is a fundamental change that requires close collaboration between various business and technical teams in order to be successful. For example, the virtualization and desktop teams need to work together to ensure that the virtual desktop image meets user needs while also being optimized for the datacenter. Failure to build a cohesive project team that consists of the right roles and skillsets can negatively impact performance, availability, user experience and supportability while also increasing costs and risk.

The following tables identify the business and technical roles required during an enterprise virtual desktop deployment. Although the list may seem quite large, many of these roles are only required for a short time and multiple roles may be performed by a single person. The project manager and Citrix architect are considered to be full time roles with other team members being brought in only when required. The project manager role is key to ensuring that the right people are involved in the project at the right time.

Business Roles

Role	Description	Example Responsibilities
Project sponsor	The project sponsor is a senior company executive who recognizes the benefits that desktop virtualization will bring to the business. The project sponsor role is often performed by the chief technology officer (CTO).	**Pre-project** • Promote desktop virtualization within business • Identify members of the steering committee **Secure funding** • Assess general costs associated with solution • Identify and prioritize key business drivers

Project manager	The project manager directs the project team and is responsible for ensuring that project objectives are completed on time and within budget.	**All steps** • Define key project milestones • Create and update project plan • Track progress against plan • Track expenditure against budget • Maintain issue and risk register • Manage scope changes • Create weekly project reports • Brief steering committee on progress • Organize project workshops and meetings • Ensure project teams are synchronized • Ensure pre-requisites are in place • Creates change control requests
Business manager	Depending on company structure and size, business managers oversee planning and performance at a department, region or company level. A business manager understands the requirements necessary for their employees to be successful.	**Assess** • Assist with application consolidation project • Provide details on connectivity requirements of user group, including offline usage • Provide details on risk tolerance of user group • Identify requirements for peripherals **Deploy** • Promote benefits of desktop virtualization • Assist with coordinating the rollout
Business continuity manager	The business continuity manager ensures that an organization can continue to function after a disruptive event such as natural disaster, crime or human/computer error.	**Assess** • Provide Citrix architect with detailed understanding of the current business continuity plan **Design** • Update business continuity plan to incorporate the new Citrix infrastructure **Deploy** • Test business continuity plan

Test manager	The test manager is responsible for ensuring that the test and user acceptance environments match the production environment as closely as possible. The test manager helps to reduce risk by ensuring that changes are fully tested before being implemented in production.	**Assess** • Provide Citrix architect with detailed understanding of current testing infrastructure and processes **Design** • Work with Citrix architect to design an appropriate testing infrastructure and test plan for new Citrix environment **Deploy** • Ensure that testing design is implemented correctly and new Citrix infrastructure is fully tested before rollout
Application owners	An application owner is a subject matter expert on specific applications deployed within the business. Application owners are responsible for ensuring that problems with the applications are resolved and that upgrades/updates are performed without issue. Application owners are also responsible for managing support agreements with the application vendors.	**Assess** • Assist with application consolidation project • Identify application licensing limitations • Provide details on security restrictions • Provide details on application dependencies • Provide location of backend resources **Deploy** • Provide installation pre-requisites and install guide • Assist Citrix team with installing and testing applications in VDI environment
Service desk manager	The service desk manager helps to improve productivity and end-user satisfaction by ensuring that production issues are logged, escalated and resolved in a timely manner. The service desk manager is also responsible for reporting on common issues, call volumes and service desk performance.	**Assess** • Identify common issues with existing environment • Provide details on support tools currently used **Design** • Assist Citrix architect with designing a delegated administration model • Participate in operations and support design workshops • Work with training manager to identify training requirements **Deploy** • Monitor helpdesk calls for rollout related issues

Training manager	The training manager ensures that support staff and end-users are proficient with new areas of technology. The training manager also has responsibility for ensuring that the training plan is up-to-date and followed appropriately.	**Assess** • Determine current skill set for support staff and end users **Design** • Create training plan for support staff and end users **Deploy** • Implement training plan for support staff and end users
Communications manager	The communication manager is responsible for disseminating key information throughout the organization.	**Design** • Work with project manager to create communications plan **Deploy** • Relay benefits of desktop virtualization • Inform users of key migration dates • Ensure expectations are set accordingly

Table 10: Business Roles

Technical Roles

Role	Description	Example Responsibilities

Citrix desktop architect	The Citrix architect acts as the design authority for all Citrix products and liaises with other architects to ensure that the Citrix infrastructure is successfully integrated into the organization.	**Assess** • Work with project sponsor and key stakeholders to identify and prioritize key business drivers • Oversee user segmentation and app. assessment • Map VDI models to user groups • Perform capabilities assessment to determine current state of readiness • Identify areas of risk and provides remedial actions **Design** • Create Citrix design that includes hardware and storage estimates • Coordinate with other architects to integrate Citrix infrastructure into organization • Work with monitoring architect to ensure that Citrix environment is monitored appropriately • Create operations and support design • Create implementation and rollout design • Create test plan **Deploy** • Ensure that the Citrix environment is implemented in accordance with design • Verify that implementation passes test plan • Ensure that the Citrix design is implemented correctly
Active directory architect	Design authority on Microsoft Active Directory, including Organizational Units (OU) and Group Policy Objects (GPOs).	**Assess** • Provide Citrix architect with detailed understanding of current Active Directory architecture **Design** • Work with the Citrix architect to design OU structure, group policies, permissions, service accounts, etc. for new Citrix environment • Update Active Directory infrastructure design to reflect centralization of user data and accounts **Deploy** • Ensure that Active Directory design is implemented correctly

Virtualization architect	Design authority on server and desktop virtualization using Citrix XenServer, Microsoft Hyper-V, Nutanix Acropolis or VMware vSphere.	**Assess** • Provide Citrix architect with detailed understanding of current virtualization architecture **Design** • Work with Citrix architect to design hardware, networking, storage, high availability, etc. for server and desktop virtualization • Work with monitoring architect to ensure that virtualization environment is monitored appropriately **Deploy** • Ensure that the virtualization design is implemented correctly
Network architect	Design authority on networking, including routing, VLANs, DHCP, DNS, VPN and firewalls.	**Assess** • Provide Citrix architect with detailed understanding of current networking architecture **Design** • Work with Citrix architect to design physical network, virtual networks, routing, firewalls, quality of service, remote access, network optimization, etc. for new Citrix environment • Work with monitoring architect to ensure that network is monitored appropriately **Deploy** • Ensure that network design is implemented correctly
Desktop architect	Design authority on Microsoft desktop operating systems, including Windows XP, Windows 7 and Windows 8.	**Assess** • Provide Citrix architect with detailed understanding of current desktop environment **Design** • Work with Citrix architect to design core desktop virtual image, core applications, desktop optimizations, etc. for new Citrix environment • Work with monitoring architect to ensure that the virtual desktops are monitored appropriately **Deploy** • Ensure that desktop design is implemented correctly

Storage architect	Design authority on storage solutions, including direct-attached storage, storage-attached networks and network attached storage.	**Assess** • Provide Citrix architect with detailed understanding of current shared storage environment **Design** • Work with Citrix architect to design storage architecture, tiers, sizing, connectivity, etc. for new Citrix environment • Work with monitoring architect to ensure that storage is monitored appropriately **Deploy** • Ensure that storage design is implemented correctly
Backup architect	Design authority on backup and recovery, including virtual machines, desktops, servers, user data and databases.	**Assess** • Provide Citrix architect with detailed understanding of current backup architecture and processes **Design** • Work with Citrix architect and disaster recovery architect to design backup architecture, process, schedule, retention, etc. for new Citrix environment **Deploy** • Ensure that backup design is implemented correctly
Application packaging architect	Design authority on packaging applications for deployment via the systems management team.	**Assess** • Provide Citrix architect with detailed understanding of current application packaging process and status **Deploy** • Ensure that all required applications are packaged according to design

Monitoring architect	Design authority on monitoring, including hardware, network, servers, storage and security appliances.	**Assess** • Provide Citrix architect with detailed understanding of current monitoring architecture and processes **Design** • Work with Citrix architect to design monitoring architecture, metrics, alerts, etc. for new Citrix environment and supporting infrastructure **Deploy** • Ensure that monitoring design is implemented correctly • Provide regular reports on capacity and trends during rollout
Systems management architect	Design authority on systems management, including server/desktop build process, patching and automated application installation.	**Assess** • Provide Citrix architect with a detailed understanding of the current systems management processes **Design** • Works with Citrix architect to define server/desktop build process, patching and application delivery strategy for new Citrix environment **Deploy** • Ensure that the systems management design is implemented correctly
Security architect	Design authority on security, including desktops, servers, networks and VPNs.	**Assess** • Provide Citrix architect with detailed understanding of current security policy **Design** • Work with Citrix architect to design security standards for new Citrix environment, including authentication, encryption, port numbers, firewall rules, etc. **Deploy** • Ensure that security design is implemented correctly

Table 11: Technical Roles

Section 3: Design

Overview

Designing a desktop virtualization solution is simply a matter of following a proven process and aligning technical decisions with organizational and user requirements. Without the standardized and proven process, architects tend to randomly jump from topic to topic, which leads to confusion and mistakes. The recommended approach focuses on working through five distinct layers:

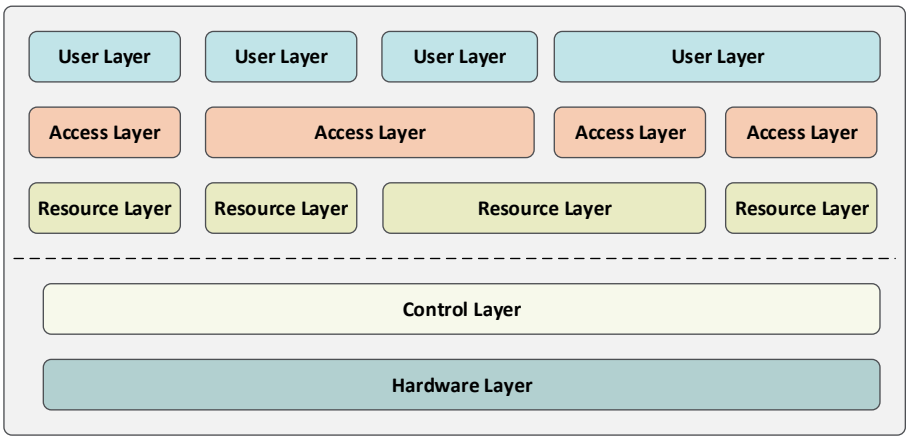


Figure 3: Five-Layer Design Model

The top three layers are designed for each user group independently, which were identified during the user segmentation section of the assess phase. These layers define the users' resources and how users access their resources. Upon completion of these three layers, the foundational layers (control and hardware) are designed for the entire solution.

This process guides the design thinking in that decisions made higher up impact lower level design decisions.

Layer 1: The User Layer

The top layer of the design methodology is the user layer, which is defined for each unique user group.

The user layer appropriately sets the overall direction for each user group's environment. This layer incorporates the assessment criteria for business priorities and user group requirements in order to define effective strategies for endpoints and Citrix Receiver. These design decisions impact the flexibility and functionality for each user group.

Endpoint Selection

There are a variety of endpoints devices available, all with differing capabilities, including:

- Tablet based
- Laptop
- Desktop PC
- Thin client
- Smartphone

The user's primary endpoint device must align with the overall business objectives as well as each user's role and associated requirements. In many circumstances, multiple endpoints may be suitable, each offering differing capabilities.

Decision: Endpoint Ownership

In many organizations, endpoint devices are corporate owned and managed. However, more and more organizations are now introducing bring your own device (BYOD) programs to improve employee satisfaction, reduce costs and to simplify device management. Even if BYOD is a business priority, it does not mean that every user should be allowed to use a personal device in the corporate environment.

Certain user requirements, which were identified during the user segmentation, can greatly impact the suitability of personal devices:

- Security – Users requiring a high-level of security might not be able to bring a personal device into the secured environment for risk of data theft.
- Mobility – Users operating in a disconnected mode might not be able to use a personal device, as the local VM desktop VDI model associated with this type of requirement can have specific hardware requirements, or special maintenance requirements.
- Desktop loss criticality – Users with a high desktop loss criticality rating might require redundant endpoints in the event of failure. This would require the user to have an alternative means for connecting in the event their personal device fails, likely making these users poor candidates for a BYOD program.
- VDI models – A personal device should not be recommended for user groups utilizing a local VDI model like a local streamed desktop, local VM desktop or Remote PC Access. These VDI models typically require a specific hardware configuration or installation that will restrict device selection.

The following diagram provides guidance on when user owned devices could be used:

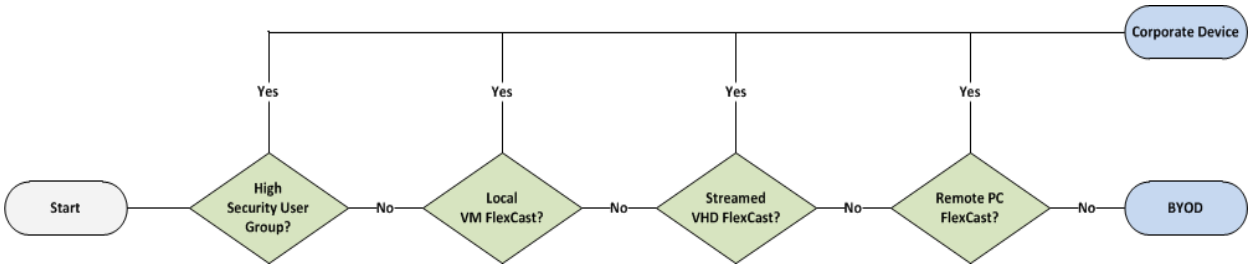


Figure 4: BYOD or Corporate Device

Decision: Endpoint Lifecycle

Organizations may choose to repurpose devices in order to extend refresh cycles or to provide overflow capacity for contract workers. Endpoints now offer more capabilities allowing them to have longer useful lifespans. In many cases, these hardware capabilities vastly outstrip the needs of a typical user. When coupled with the ability to virtualize application and desktop workloads, this provides new options to administrators such as repurposing existing workstations. These options go well beyond the simple three-year PC refresh cycle. However, the benefits of repurposing or reallocating a workstation should be balanced against the following considerations.

- Minimum standards – While cost factors of repurposing existing workstations may be compelling, certain minimum standards should be met to guarantee a good user experience. At a minimum, it is recommended that repurposed workstations have a 1GHz processor, 1GB of RAM, 16GB of free disk space and a GPU that is capable of supporting HDX features.
- Business drivers – Priorities underpin the success of any major project. Those organizations that have prioritized reducing capital expenditure by means of prolonging the hardware refresh cycle can benefit from repurposing hardware. Conversely, if an organization's business drivers include reducing power consumption as part of an overall green initiative, purchasing newer endpoints

may be beneficial in order to take advantage of the latest generation of power management capabilities available in the most modern devices.

- Workload – The type of work and VDI model for an end user can determine whether they are a good candidate for a repurposed endpoint, or may be better served with a new device. If the work performed by the individual involves locally installed applications, the individual may be best served by a new endpoint that offers the most powerful and recently updated processor and graphics architecture. However, if a user is largely performing tasks associated with virtualized applications that do not involve the latest multimedia capabilities such as webcams, VoIP and media redirection, then a repurposed workstation should be a viable alternative.

The following planning matrix outlines considerations when repurposing existing hardware:

Endpoint Provisioning Criteria	Repurpose Existing	Procure New
Capital restrained environment	✓	
High number of virtualized applications	✓	
Desire to prolong hardware refresh cycle	✓	
High failure rate among existing desktops		✓
Outmoded client-side feature set		✓
Power consumption or green initiative(s)		✓

Table 12: Endpoint Procurement Criteria

Decision: Endpoint Form Factor

The capabilities of endpoints have grown along with efficiencies offered in thin client form factors. Even mid-range thin clients now have graphics capabilities that allow utilization of HDX features such as multi-monitor support while offering management and power efficiency benefits. Citrix has developed a three-tiered classification for thin clients based on their HDX capabilities: HDX Ready, HDX Premium, and HDX 3D Pro, which can be used to help narrow the field of appropriate thin client devices based on the use case requirements. This expansion of capabilities has given IT administrators more options and flexibility than ever before.

Most organizations will likely deploy a mixture of fully featured clients as well as thin clients. However, certain endpoint devices are more appropriate when used in combination with certain VDI models as explained in the following table.

VDI Model	Thin Clients	Desktop PC	Laptop	Tablet	Smartphone
Hosted Windows Apps	✓	✓	✓	✓	✓
Hosted Browser Apps	✓	✓	✓	✓	✓
Hosted Shared Desktop	✓	✓	✓	○	○
Hosted Pooled Desktop	✓	✓	✓	○	○
Hosted Personal Desktop	✓	✓	✓	○	○
Hosted Pro Graphics Desktop	✓	✓	○	○	○
Local Streamed Desktop	✗	✓	✗	✗	✗
Local VM Desktop	✗	○	✓	✗	✗

Remote PC Access	✕	✓	✓	o	o

"✓": Recommended, "✕": Not Recommended, "o" Viable

Table 13: Primary Endpoint Selection

Experience from the Field

Large systems integrator – A large systems integrator recommended that a customer deploy a single type of low-end, limited capability endpoint for all users. Upon deployment to production, users immediately complained that they received a poor user experience when viewing multimedia content over the WAN. At great cost, the systems integrator and customer re-assessed the environment and chose to deploy endpoints that supported HDX MediaStream. The mistake caused a schism between systems integrator and the customer, resulting in lost time, capital and the end of a business relationship that was fostered over many years. It is critical that the endpoints assigned to each user group can support their requirements.

Receiver Selection

Citrix Receiver is an easy-to-install software client that provides access to applications, desktops and data easily and securely from any device, including smartphones, tablets, PCs and Macs.

The following section provides a series of design decisions that should be considered when deploying Citrix Receiver.

Decision: Receiver Type

While most organizations should simply deploy the latest Citrix Receiver compatible with their endpoint, it is important to recognize that there are certain differences between editions. The following table should be referenced to determine the most appropriate edition of Citrix Receiver for each user group. For the latest feature matrix, please refer to Receiver Feature Matrix.

Decision: Initial Deployment

There are several deployment options available for delivering Citrix Receiver to an endpoint. Although it is usually a best practice to have a full version of Citrix Receiver deployed to an endpoint to provide the greatest level of functionality, it is important to consider fallback options such as the HTML5 Receiver for those situations where the installation of Citrix Receiver is simply not possible. Note that although the HTML5 Receiver can be used as a fallback option, like the Java client was with Web Interface, it is not generally recommended as the primary Receiver for enterprises to standardize on due to the limited feature set and common browser restrictions around unsecured WebSockets connections (see CTX134123 for more information).

Experience from the Field

Furniture distributor – A furniture distributor maintains a configurator application for various furniture options. The configurator application is accessed via a limited functionality kiosk that is deployed at various furniture outlets, including small, independent retailers with little to no IT staff present. The kiosks are completely locked down in many situations, to the point where even the running of Java applications is limited. The kiosks do feature a modern browser (Google Chrome), and therefore, are able to utilize the HTML5 Receiver in order to provide access to the configurator application.

County government – A government IT organization provides services to all agencies operating in the county. A mixture of full desktops are applications are deployed to both Windows based desktops and iPads. Since the desktops are joined to the Active Directory domain, GPOs are utilized to deploy and configure Citrix Receiver. Mobile users accessing the Citrix environment via an iPad install and configure Receiver from the App Store. To allow for seamless provisioning, email based discovery was configured. This allows users to configure Receiver for both internal and external access through NetScaler Gateway by entering in their email address.

The following mechanisms are commonly used to deploy and update Citrix Receiver:

- **StoreFront –** If Citrix StoreFront is available, administrators can deploy Citrix Receiver via a Receiver for Web site by enabling the "Client Detection" feature. When deployed, a Receiver for Web site enables users to access StoreFront stores through a web page. If the Receiver for Web site detects that a user does not have a compatible version of Citrix Receiver, the user is prompted to download and install Citrix Receiver. The Receiver clients can be hosted on the StoreFront server, or users can be directed to citrix.com for the latest Receiver files.
- **Internal download site** – Users may be prevented from downloading software from the Internet, even if they have permission to install applications. Administrator can create an internal website for supported Citrix Receivers or host them on a common software distribution point for a more seamless user experience. This could be an alternative to enabling Client Detection on the StoreFront Receiver for Web site, which can result in an inconsistent user experience depending on browser's ActiveX settings.
- **Markets and stores** –Citrix Receiver is available on the Windows, Android and iOS stores..
- **Enterprise software deployment** – Many organizations employ an enterprise software deployment (ESD) or Mobile Application Management (MAM) solution. ESD/MAM solutions can be used to deploy Citrix Receiver to managed endpoint devices. Employee-owned devices can only be managed if the user successfully registered the device with the management tool.
- **Master image** – Most organizations have a group of master desktop images, which are deployed to each business owned desktop, laptop, server, or virtual desktop. A common mechanism to ensure access to virtual desktops and applications is to include a supported version of Citrix Receiver in the master image. Subsequent updates to Citrix Receiver are handled either by enterprise software deployment tools or manually.
- **Group policy** – For customers without a robust ESD solution, it is possible to deploy and configure Citrix Receiver via Microsoft Group Policy. Sample start-up scripts that deploy and remove Citrix Receiver are available on Citrix XenApp and XenDesktop media:

  Citrix Receiver and Plugins\Windows\Receiver\Startup_Logon_Scripts

- **Manual install** – All supported versions of Citrix Receiver are available from the Citrix Receiver Download site. Upon landing on this site, client detection is performed and a platform and operating system specific link is provided to allow users to download an appropriate edition of

Citrix Receiver. It is important to note that no configuration will be accomplished via this download, so users will receive the first time use prompt to enter a server URL or email address. This option is likely to be utilized in a BYOD environment.

Selecting the appropriate deployment method is based on the type of Citrix Receiver selected. The following table should be referenced to help identify the appropriate deployment options for Citrix Receiver.

Deployment Options	Thin clients	Desktop PC	Laptop	Tablet	Smartphone
Base image	✓	✓	✓	✕	✕
ESD / MAM	✕	✓	✓	✕	✕
Group Policy	✕	✓	✓	✕	✕
Receiver for Web Site	✕	✓	✓	✕	✕
Internal Download Site	✕	✓	✓	✕	✕
App Store	✕	✕	✕	✓	✓

"✓": Recommended, "✕": Not Recommended

Table 14: Receiver Deployment Options

Decision: Initial Configuration

Citrix Receiver must be configured in order to provide access to enterprise resources. The method of configuration varies by Citrix Receiver edition, the form factor of the device, and lastly the access method (local or remote) that is involved. Several methods may be viable for an organization. The method utilized is contingent on the resources (people, systems, time) available as well as larger organizational initiatives such as BYOD programs.

The following methods can be used to configure Citrix Receiver:

- **Email based discovery** – The latest releases of Citrix Receiver can be configured by entering an email address. Email based discovery requires Citrix StoreFront as well as an SRV DNS record which points to the FQDN of the StoreFront server.

  Note: *Any DNS platform should support email-based discovery, however only Windows DNS has been explicitly tested.*

  For remote access, NetScaler Gateway must be utilized with the corresponding SRV record in external DNS. A valid server certificate on the NetScaler Gateway appliance or StoreFront server must be present in order to enable email-based account discovery. This configuration assumes that the portion of the email address after the "@" is the DNS namespace that should be queried for this SRV record. This can be challenging for customers with different external and internal namespaces or email addresses that are different from DNS namespaces.

- **Group policy** – Microsoft Group Policy can be used to configure Citrix Receiver. This can be done via start up scripts used to deploy Receiver by ensuring there is a value for the *SERVER_LOCATION=Server_URL* parameter or by using the ADMX/ADML template files included with the installation of Citrix Receiver to set the StoreFront Account List option in conjunction with another Receiver deployment method. Provide the URL of the server running Citrix StoreFront in the format *https://baseurl/Citrix/storename/discovery*.

- **Provisioning file** – For environments running StoreFront, it is possible to provide users with a provisioning file that contains store information. Provisioning files are exported from the StoreFront console. The file is saved with a "*.cr" extension and can then be placed on a shared network resource, a Receiver for Web site, or other web based resource or emailed to users. The file can then be launched from an endpoint, which automatically configures Citrix Receiver to use

the store(s). If users browse to the Receiver for Web site and select the "Activate" option under their username, this also automatically downloads this same ".cr" file and configure the Receiver client for users.

- **Manually** – If allowed, it is usually possible to configure Citrix Receiver manually by entering the server URL. This method should be reserved for administrators or users that have advanced knowledge.
- **Studio** – In addition to the above methods, in order to configure Receiver deployed on a virtual desktop or server image (within a XenDesktop or XenApp environment), it is possible to set the StoreFront address via the properties of the Delivery Group.

Decision: Updates

Citrix Receiver is in active development. As such, periodic updates are released that provide enhanced functionality or address user issues. As with any actively developed product, the latest version of these products should be deployed to the endpoints so that users benefit from the latest functionality and to maintain compliance with product support lifecycles. There are multiple methods available to update Citrix Receiver and, if applicable, associated plug-ins.

- **Enterprise software deployment –** ESD tools provide an organization with direct control over the time/frequency of Receiver updates to managed devices. Additional thought must be given to updating unmanaged devices and endpoints outside of the corporate firewall.
- **Manual updates –** When no automated solution is available, manual methods can be used to update Citrix Receiver. Whether deployed on Receiver for Web site, StoreFront, an internal Citrix Receiver site, or an external site, these options will require user involvement in updating Citrix Receiver. Due to the involved nature of manual updates coupled with the opportunity for a user mistake, this option should only be considered as a last resort.

Layer 2: The Access Layer

The second layer of the design methodology is the access layer, which is defined for each user group.

Creating an appropriate design for the access layer is an important part of the desktop virtualization process. This layer handles user validation through authentication and orchestrates access to all components necessary to establish a secure virtual desktop connection.

The access layer design decisions are based on the mobility requirements of each user group as well as the endpoint devices used.

Authentication

Getting access to resources is based on the user's identity. Defining the authentication strategy takes into account the user's entry point into the environment as well as how the user will authenticate.

Decision: Authentication Point

Before a user connects to a virtual resource, they must first authenticate. The place of authentication is often determined by the user group's mobility requirements, which were defined during the user segmentation process. There are two authentication points available in XenDesktop 7.6:

- **StoreFront** – Citrix StoreFront provides authentication and resource delivery services for Citrix Receiver, enabling centralized enterprise stores to deliver desktops, applications and other resources.
- **NetScaler Gateway** – NetScaler Gateway is an appliance providing secure application access and granular application-level policy controls to applications and data while allowing users to work from anywhere.

The following table lists preferred authentication points according to user group mobility requirements:

User Group's Mobility Requirement	Preferred Authentication Point
Local	StoreFront
Roaming local	StoreFront
Remote	NetScaler Gateway
Mobile	NetScaler Gateway

Table 15: Preferred Authentication Point

Authentication for user groups with a mobility requirement of remote or mobile may occur directly on StoreFront where required. For example, DMZ security policies may prohibit access from the NetScaler Gateway to the domain, which is required to support SmartCard client certificate authentication. Access to StoreFront for authentication may then be delivered via a NetScaler SSL_BRIDGE virtual server, which provides a conduit for https traffic. Typically, the virtual server would be hosted alongside a NetScaler Gateway on the same NetScaler configured to provide HDX Proxy access to the virtual desktop environment. Although such a use case may sometimes be necessary, the recommended best practice is to authenticate external users via NetScaler Gateway.

Decision: Authentication Policy

Once the authentication point has been identified, the type of authentication must be determined. The following options are the primary methods available:

- **StoreFront** – Supports a number of different authentication methods, although not all are recommended depending on the user access method, security requirements and network location. Note that by default StoreFront authenticates users directly with Active Directory, not via XML as Web Interface did. StoreFront 3.0+ can be optionally configured to delegate authentication to XML if required (such as if the StoreFront servers are in a domain that does not trust the user domains).
    - **User name and password** – Requires users to logon directly to the site by entering a user name and password.
    - **Domain pass-through** – Allows pass-through of domain credentials from users' devices. Users authenticate to their domain-joined Windows computers and are automatically logged on when they access their stores.
    - **NetScaler Gateway pass-through** – Allows pass-through authentication from NetScaler Gateway. Users authenticate to NetScaler Gateway and are automatically logged on when they access their stores.
    - **Smart card** – Allows users to authenticate using smart cards and PINs through Citrix Receiver for Windows and NetScaler Gateway. To enable smart card authentication, user accounts must be configured either within the Microsoft Active Directory domain containing the StoreFront servers or within a domain that has a direct two-way trust relationship with the StoreFront server domain. Multi-forest deployments involving one-way trust or trust relationships of different types are not supported.
    - **Anonymous** – Allow users to access applications and desktops without presenting credentials to StoreFront or Citrix Receiver. Local anonymous accounts are created on demand on the Server VDA when sessions are launched. This requires a StoreFront store configured for authenticated access, a Server OS based VDA, and a XenApp 7.6 (or later) Delivery Group configured for unauthenticated users.

- **NetScaler Gateway** – The NetScaler Gateway supports several authentication methods. The list below includes those primarily used in virtual desktop environments. Each may be used individually, but are often combined to provide multi-factor authentication.
    - **LDAP** – The lightweight directory access protocol (LDAP) is used to access directory information services such Microsoft Active Directory. NetScaler Gateway uses LDAP to authenticate users and extract their group membership information.
    - **RADIUS (token)** – Remote authentication dial in user service (RADIUS) is a UDP based network security protocol that provides authentication, authorization and accounting. A network access server (NetScaler Gateway in this case) forwards credentials to a RADIUS server that can either check the credentials locally, or check them against a directory service. The RADIUS server could then accept the connection, reject the connection, or challenge and request a second form of authentication such as a token.
    - **Client certificate** – Users logging on to a NetScaler Gateway virtual server can also be authenticated based on the attributes of a client certificate presented to the virtual server. Client certificates are usually disseminated to users in the form of smartcards or common access cards (CACs) that are read by a reader attach to each user's device.

The authentication type for a user group is often determined based on security requirements as well as the authentication point used. The following table helps define the appropriate solution for each user group based on the level of security required:

Authentication Point	Security Requirement	Authentication Type
StoreFront	Low	• LDAP Username and Password • Pass-through
	Medium	• LDAP Username and Password • Pass-through
	High	• LDAP and/or Smartcard
NetScaler Gateway	Low	• LDAP Username and Password
	Medium	• LDAP Username and Password
	High	• LDAP and Token • LDAP and Smartcard • Token and Smartcard

Table 16: Authentication Policy Guidance

Experience from the Field

Retail – A small private retail company provides virtual desktop users with access to non-sensitive data such as marketing catalogs and email. They are not required to adhere to security regulations such as Sarbanes Oxley. Therefore, LDAP authentication has been implemented based on user name and password.

Financial – A medium financial enterprise provides their virtual desktop users with access to confidential data such as banking transaction records. They are governed by security regulations such as the Statement on Accounting Standards (SAS) 70 and are required to utilize multi-factor authentication for remote access users. LDAP authentication has been implemented based on user name and password along with RADIUS authentication using tokens.

Government – A large federal institution provides virtual desktop users with access to highly confidential data such as private citizens' personal records. They are subject to regulation by Department of Defense (DOD) security standards. LDAP authentication has been implemented based on user name and password, along with Client Certificate authentication using CAC cards.

Healthcare - A hospital is using XenApp to deliver their EMR application to users. ThinClient devices on stationary and mobile carts are being used by doctors and nurses to capture and retrieve patient data. Unauthenticated access has been configured to prevent medical staff from having to authenticate to the domain as well as the EMR application.

StoreFront

Citrix StoreFront authenticates users to XenApp and XenDesktop resources. StoreFront enumerates and aggregates available desktops and applications into a single interface that users access through Citrix Receiver for Windows, iOS, Android, or the StoreFront web site.

Decision: High Availability

If the server hosting StoreFront is unavailable, users will not be able to launch new virtual desktops, published applications or manage their subscriptions. Therefore at least two StoreFront servers should be deployed to prevent this component from becoming a single point of failure. By implementing a load balancing solution, users will not experience an interruption in their service. Options include:

- **Hardware load balancing** – An intelligent appliance, which is capable of verifying the availability of the StoreFront service and actively load balance user requests appropriately. Citrix NetScaler is a great example of a hardware load balancer. Citrix NetScaler is an ideal load balancer, coming pre-configured with StoreFront health checks.
- **DNS round robin** – Provides rudimentary load balancing across multiple servers without performing any checks on availability. If a StoreFront server becomes unavailable, DNS round robin would still route users to the failed server. Because of this, DNS round robin is not recommended by Citrix.
- **Windows network load balancing** – A Windows service capable of performing rudimentary checks to verify the server is available but cannot determine the status of individual services. This can cause users to be forwarded to StoreFront servers which are not able to process new requests. The user would then not be able to access applications or desktops.

Decision: Delivery Controller Reference

To provide users with desktops and applications, StoreFront must be configured with the IP address or DNS name of at least one Controller in each XenDesktop and XenApp site. For fault tolerance, multiple controllers should be entered for each site and/or farm specified. By default, StoreFront treats a list of servers in failover order (active/passive).

For large deployments or environments with a high logon load an active distribution of the user load (active/active) is recommended. This can be achieved by means of a load balancer with built-in XML monitors, such as Citrix NetScaler or by configuring StoreFront to load balance the list of Controllers instead of treating them as an ordered list.

Decision: Beacons

Citrix Receiver uses beacons (websites) to identify whether a user is connected to an internal or external network. Internal users are connected directly to StoreFront for authentication while external users are connected via Citrix NetScaler Gateway. It is possible to control what a user sees by restricting applications due to which beacon they have access to.

The internal beacon should be a site that is not resolvable externally. By default, the internal beacon is the StoreFront base URL. This will have to be adjusted if the same external and internal URL is configured. The external beacon can be any external site that produces an http response. Citrix Receiver continuously monitors the status of network connections (for example, link up, link down or change of the default gateway). When a status change is detected, Citrix Receiver first verifies that the internal beacon points can be accessed before moving on to check the accessibility of external beacon points. StoreFront provides Citrix Receiver with the http(s) addresses of the beacon points during the initial connection/configuration download process and provides updates as necessary.

It is necessary to specify at least two highly available external beacons that can be resolved from public networks.

Decision: Resource Presentation

By default, StoreFront allows users to choose (subscribe) to the resources they want to regularly use after they logon (favorites). This approach, deemed "Self-Service," allows users to restrict the resources that they see on their home screen to the ones that they use on a regular basis. The resources chosen by every user for each store are recorded by the subscription store service and stored locally on each StoreFront server (synced automatically between servers in the same server group) so that they can be displayed on the Citrix Receiver home screen from any device that the user connects from. Although by default subscriptions are per store and per server group, administrators can configure two stores within a server group to share a subscription database and/or sync subscriptions between two identically named stores in two separate server groups on a defined schedule if required.

Administrators should determine which applications should always be displayed to users on their home screen or the featured tab. In general, these applications are common applications such as the Microsoft Office Suite and any other applications that every user in an environment may need. StoreFront can filter/present these resources using Keywords defined within the published application properties Description field.

The following table explores the Keyword options:

Keyword	Description
Auto	Automatically subscribes all users of a store to an application. When users log on to the store, the application is automatically provisioned without users needing to manually subscribe to the application. Users can choose to subsequently remove this subscription if desired.
Mandatory	New in StoreFront 2.5, the Mandatory keyword will make applications automatically be subscribed to users of the store. However, users will not have the option to remove the application. This setting is useful when creating a core set of applications which must always be presented to all users.
Featured	Advertise applications to users or make commonly used applications easier to find by listing them in the Receiver Featured list.

Prefer	Specify a locally installed application should be used instead of an application available in Receiver. Receiver searches for the specified name/path to determine if the application is installed locally. If it is, Receiver subscribes the application and does not create a shortcut. When the user starts the application from the Receiver window, Receiver starts the locally installed (preferred) application. If a user uninstalls a preferred application outside of Receiver, the application is unsubscribed during the next Receiver refresh. If a user uninstalls a preferred application from the Receiver window, Receiver unsubscribes the application but does not uninstall it.
TreatAsApp	By default, XenDesktop VDI desktops and XenApp hosted shared desktops are treated like other desktops by Receiver for Web sites. By using the keyword "TreatAsApp," the desktop will be displayed in the application views of Receiver for Web sites rather than the desktop views. Users are required to subscribe before they can access the desktop.
Primary	When in a multi-site deployment, using this keyword ensures that an application is delivered from a designated site. If an application is available from multiple sites, with the same name, the application from the secondary site will only be displayed if the application is not available from the primary site.
Secondary	A same property as the "Primary" keyword, except it designates an application in the secondary site.

Table 17: Keywords for Application Delivery

Decision: Scalability

The number of Citrix Receiver users supported by a single StoreFront server depends on the resources assigned and level of user activity. Note that Receiver for Web users will consume more RAM on average than native Receiver users, but a minimum of 4 GB of RAM is recommended per StoreFront server in all cases as a baseline. Additionally, more sites/farms enumerated per store will increase both CPU utilization and server response time, with XenApp IMA farms having a greater scalability impact than XenApp/XenDesktop FMA site.

StoreFront deployment	CPU Usage	Simultaneous activities
• Standalone deployment • 4 CPUs • 4 GB RAM • Heavy Usage (logon, enumerate, subscribe, unsubscribe, logoff)	75%	• 291 per second
	90%	• 375 per second

• Cluster StoreFront deployment
• 2 Nodes each with:
  ○ 4 CPUs
  ○ 4 GB RAM
  ○ Heavy Usage (logon, enumerate, subscribe, unsubscribe, logoff)
	75%	• 529 per second

	90%	• 681 per second

Table 18: StoreFront Scalability

Tests have shown diminishing returns after a single StoreFront deployment grows beyond 3-4 StoreFront nodes with a maximum of 5-6 servers supported in a single server group.

NetScaler Gateway

Selection of the network topology is central to planning the remote access architecture to ensure that it can support the necessary functionality, performance, and security. The design of the remote access architecture should be completed in collaboration with the security team to ensure adherence to corporate security requirements. There are two primary topologies to consider, each of which provides increasing levels of security:

Decision: Topology

Selection of the network topology is central to planning the remote access architecture to ensure that it can support the necessary functionality, performance and security. The design of the remote access architecture should be completed in collaboration with the security team to ensure adherence to corporate security requirements. There are two primary topologies to consider, each of which provides increasing levels of security:

- **1-Arm (normal security)** – With a 1-arm topology, the NetScaler Gateway utilizes one physical or logical bonded interface, with associated VLAN and IP subnet, to transport both frontend traffic for users and backend traffic for the virtual desktop infrastructure servers and services.

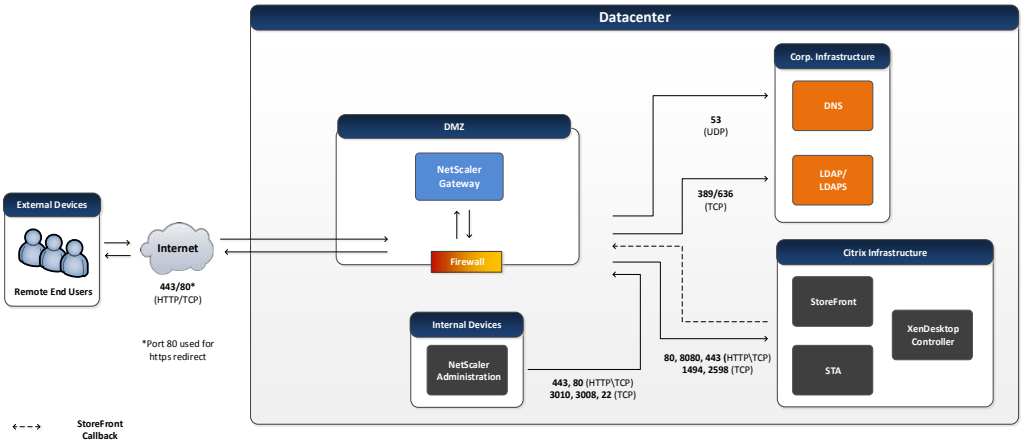


Figure 5: 1-Arm Topology

- **2-Arm (high security)** – With a 2-arm topology, the NetScaler Gateway utilizes two or more physically or logically bonded interfaces, with associated VLANS and IP subnets. Transport of the frontend traffic for users is directed to one of these interfaces. The frontend traffic is isolated from backend traffic, between the virtual desktop infrastructure servers and services, which is directed to a second interface. This allows the use of separate demilitarized zones (DMZs) to isolate frontend and backend traffic flows along with granular firewall control and monitoring.

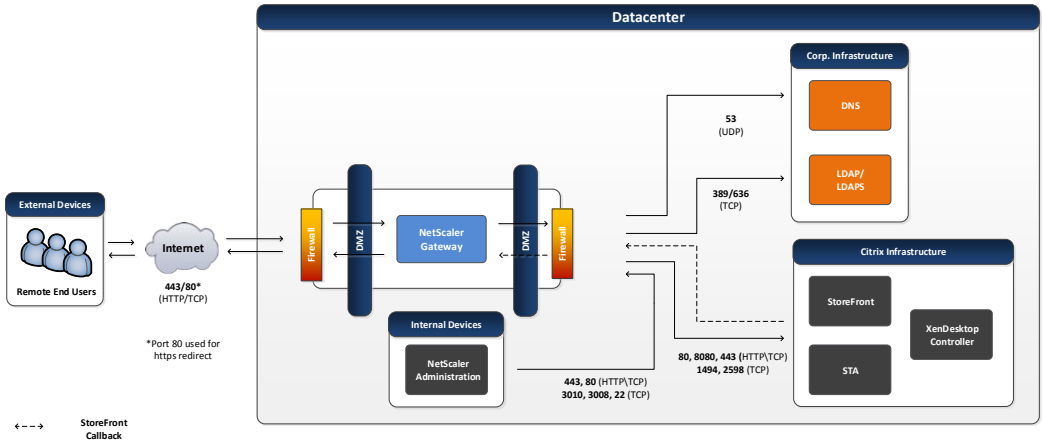


Figure 6: 2-Arm Topology

Decision: High Availability

If the NetScaler Gateway is unavailable, remote users will not be able to access the environment. Therefore at least two NetScaler Gateway hosts should be deployed to prevent this component from becoming a single point of failure.

When configuring NetScaler Gateway in a high availability (active/passive) pair, the secondary NetScaler Gateway monitors the first appliance by sending periodic messages, also called a heartbeat message or health check, to determine if the first appliance is accepting connections. If a health check fails, the secondary NetScaler Gateway tries the connection again for a specified amount of time until it determines that the primary appliance is not working. If the secondary appliance confirms the health check failure, the secondary NetScaler Gateway takes over for the primary NetScaler Gateway.

Note that in firmware 10.5 and above, clustering is also possible with multiple NetScaler Gateway instances to provide high availability, although support for spotted versus stripped configurations varies by firmware and Gateway configuration (full SSL VPN versus ICA proxy). (http://docs.citrix.com/en-us/netscaler/11-1/clustering/cluster-features-supported.html)

Decision: Platform

In order to identify an appropriate NetScaler platform to meet project requirements, the key resource constraints must be identified. Since all remote access traffic will be secured using the secure sockets layer (SSL), transported by Hypertext Transfer Protocol (HTTP) in the form of HTTPs, there are two resource metrics that should be targeted:

- **SSL throughput** – The SSL throughput is the gigabits of SSL traffic that may be processed per second (Gbps).
- **SSL transactions per second (TPS)** – The TPS metric identifies how many times per second an Application Delivery Controller (ADC) may execute an SSL transaction. The capacity varies primarily by the key length required. TPS capacity is primarily a consideration during the negotiation phase when SSL is first setup and it is less of a factor in the bulk encryption / decryption phase, which is the majority of the session life. While TPS is an important metric to monitor, field experience has shown that SSL throughput is the most significant factor in identifying the appropriate NetScaler Gateway.

The SSL bandwidth overhead average is often considered negligible relative to the volume of virtual desktop traffic and is not typically accounted for as part of required SSL throughput. However, making provisions for SSL bandwidth will help ensure the total throughput estimated is sufficient. The fixed bandwidth added to packet headers can vary according to the encryption algorithms used and the overall percentage of bandwidth may vary widely according to packet size. Ideally, the overhead should

be measured during a proof of concept or pilot. However, in the absence of such data incrementing the workload bandwidth by 2% is a reasonable rule of thumb. Therefore, to determine the SSL throughput required by a NetScaler platform, multiply the maximum concurrent bandwidth for a datacenter by 1.02:

$$SSL\ Throughput = Maximum\ Concurrent\ Bandwidth\ *\ 1.02$$

For example, assuming 128Mbps maximum concurrent bandwidth, the appropriate NetScaler model can be determined as follows:

$$\sim 130Mbps = 128Mbps\ *\ 1.02$$

The SSL throughput value should be compared to the throughput capabilities of various NetScaler platforms to determine the most appropriate one for the environment. There are three main platform groups available, each of which provides broad scalability options.

- **VPX** – A NetScaler VPX device provides the same full functionality as hardware NetScaler. However, NetScaler VPXs can leverage 'off the shelf' servers for hosting and are suitable for small to medium sized environments. Typically, organizations create a baseline cap for the VPX instances at 500 users.
- **MDX** – A NetScaler MDX is the hardware version of the NetScaler devices. The MDX device is more powerful than the virtual NetScaler and can support network optimizations for larger scale enterprise deployments, particularly when SSL offload will be configured as this is done in software on the VPX versus dedicated SSL chips on the MPX.
- **SDX** – A NetScaler SDX is a blend between the virtual and physical NetScaler devices. An SDX machine is a physical device capable of hosting multiple virtual NetScaler devices. This consolidation of devices aids with reducing required shelf space and device consolidation. NetScaler SDXs are suitable for handling network communications for large enterprise deployments and/or multi-tenant hosting providers.

SSL throughput capabilities of the NetScaler platforms may be found in the Citrix NetScaler data sheet. Therefore, based on the example calculation above, a NetScaler MPX 5550 appliance would be sufficient to handle the required load. However, actually scalability will depend on security requirements. NetScaler SSL throughput decreases with the use of increasingly complex encryption algorithms and longer key lengths. Also, this calculation represents a single primary NetScaler. At a minimum, N+1 redundancy is recommended which would call for an additional NetScaler of the identical platform and model.

Note: The Citrix NetScaler data sheet typically represents throughput capabilities under optimal conditions for performance. However, performance is directly affected by security requirements. For example, if the RC4 encryption algorithm and a 1k key length are used, a VPX platform may be able to handle more than 500 HDX proxy connections. However, if a 3DES encryption algorithm and 2k key length are used (which are becoming more common), the throughput may be halved.

Decision: Pre-Authentication Policy

An optional pre-authentication policy may be applied to user groups with NetScaler Gateway as their authentication point. Pre-authentication policies limit access to the environment based on whether the endpoint meets certain criteria through Endpoint Analysis (EPA) Scans.

Pre-authentication access policies can be configured to test antivirus, firewall, operating system, or even registry settings. These policies can be used to prevent access entirely or can be used by XenDesktop to control session features such as clipboard mapping, printer mapping and even the availability of specific applications and desktops. For example, if a user device does not have antivirus installed, a filter can be set to hide sensitive applications.

The following figure provides an overview of how multiple policies can be used to customize the features of a virtualization resource:

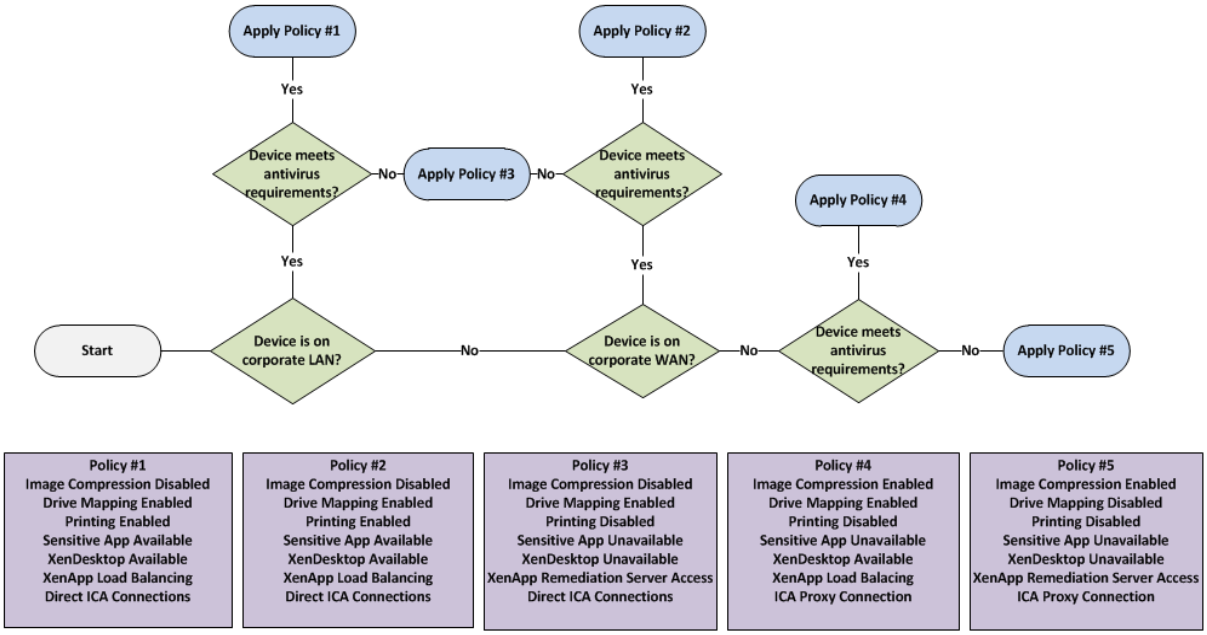


Figure 7: Simplified SmartAccess Decision Logic

Experience from the Field

Retail – A small private retail company use EPA to scan for the presence of updated antivirus definitions prior to allowing access.

Financial – A medium financial enterprise use EPA scans of the Domain SID to verify that users are members of the enterprise domain prior to allowing access.

Government – A large federal institution use EPA to scan endpoint devices to ensure that a specific certificate (or set of certificates) has been installed on the device prior to allowing access.

Decision: Session Policy

User groups with NetScaler Gateway as their authentication point must have corresponding session policies defined. Session policies are used to define the overall user experience post-authentication.

Organizations create sessions policies based on the type of Citrix Receiver used. For the purpose of session policy assignment, devices are commonly grouped as either non-mobile (such as Windows, Mac and Linux OS based), or mobile (such as iOS or Android). Therefore a decision on whether to provide support for mobile devices, non-mobile devices, or both should be made based on client device requirements identified during the assess phase.

To identify devices session policies, include expressions such as (http://docs.citrix.com/en-us/netscaler-gateway/11-1/storefront-integration/ng-clg-session-policies-overview-con.html):

- **Mobile devices** – The expression is set to REQ.HTTP.HEADER User-Agent CONTAINS CitrixReceiver which is given a higher priority than the non-mobile device policy to ensure mobile devices are matched while non-mobile devices are not.

- **Non-mobile devices** – The expression is set to ns_true which signifies that it should apply to all traffic that is sent to it.

An alternative use of session policies is to apply endpoint analysis expressions. These session policies are applied post authentication yet mimic the previously mentioned pre-authentication policies. Use of session policies is an option to provide a fallback scenario to endpoints that do not meet full security requirements such read-only access to specific applications.

Decision: Session Profile

Each session policy must have a corresponding session profile defined. The session profile defines details required for the user group to gain access to the environment. There are two primary forms of session profiles that determine the access method to the virtual desktop environment:

- **SSLVPN** – Users create a virtual private network and tunnel all traffic configured by IP addresses through the internal network. The user's client device is able to access permitted intranet resources as if it were on the internal network. This includes XenDesktop sites and any other internal traffic such as file shares or intranet websites. This is considered a potentially less secure access method since network ports and routes to services outside of the virtual desktop infrastructure may be opened leaving the enterprise susceptible to risks that may come with full VPN access. These risks may include denial of service attacks, attempts at hacking internal servers, or any other form of malicious activity that may be launched from malware, trojan horses, or other viruses via an Internet based client against vulnerable enterprise services via routes and ports.

  Another decision to consider when SSLVPN is required is whether to enable split tunneling for client network traffic. By enabling split tunneling, client network traffic directed to the intranet by Citrix Receiver may be limited to routes and ports associated with specific services. By disabling split tunneling, all client network traffic is directed to the intranet, therefore both traffic destined for internal services as well as traffic destined for the external services (Internet) traverses the corporate network. The advantage of enabling split tunneling is that exposure of the corporate network is limited and network bandwidth is conserved. The advantage of disabling split tunneling is that client traffic may be monitored or controlled through systems such as web filters or intrusion detection systems.

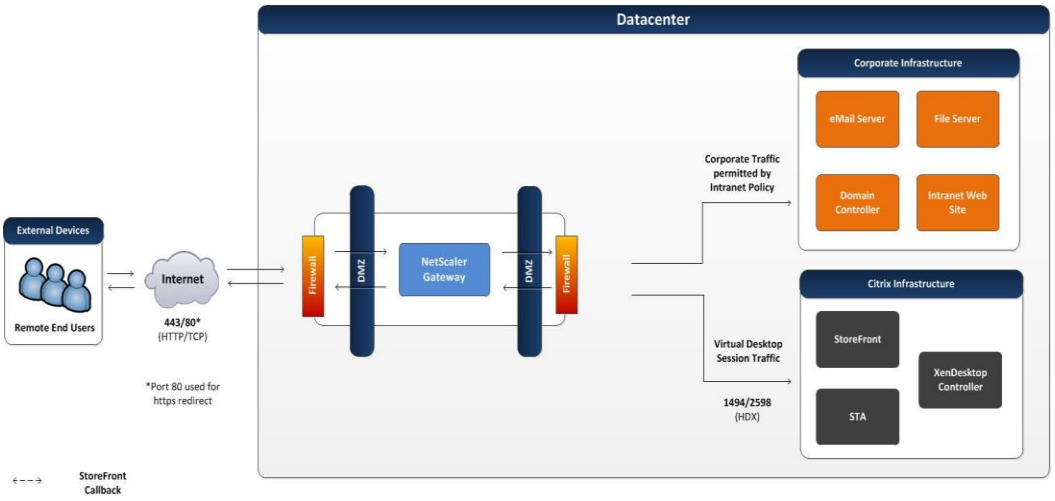


Figure 8: SSL VPN

- **HDX proxy** – With HDX Proxy, users connect to their virtual desktops and applications through the NetScaler Gateway without exposing internal addresses externally. In this configuration, the NetScaler Gateway acts as a micro VPN and only handles HDX traffic. Other

types of traffic on the client's endpoint device, such as private mail or personal Internet traffic do not use the NetScaler Gateway.

Based on the endpoint and Citrix Receiver used, a decision must be made as to whether this method is supported for each user group. HDX Proxy is considered a secure access method for remote virtual desktop access since only traffic specific to the desktop session is allowed to pass through to the corporate infrastructure. Most Citrix Receivers support HDX Proxy and it is the preferred method:

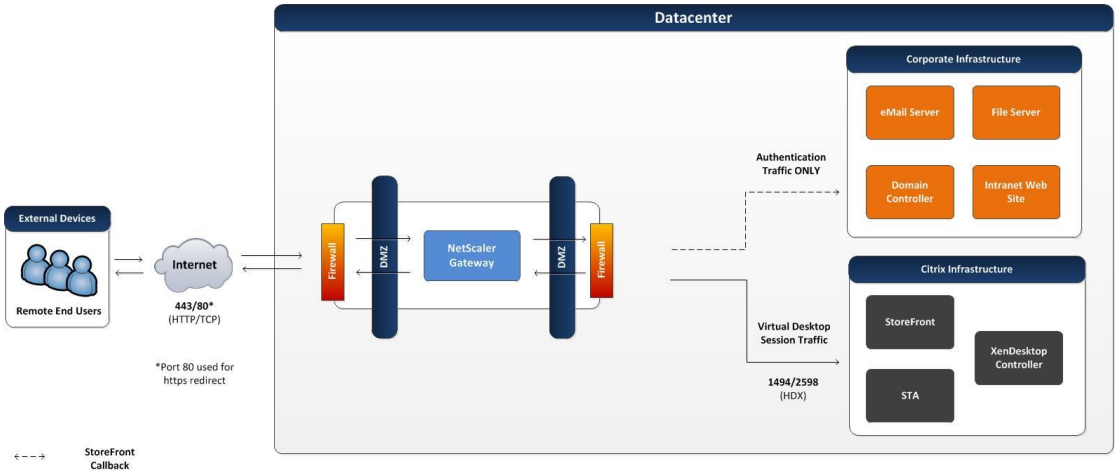


Figure 9: HDX Proxy

Decision: Preferred Datacenter

Enterprises often have multiple active datacenters providing high availability for mission critical applications. Some virtual desktops or applications may fall into that category while others may only be accessed from a specific preferred datacenter. Therefore, the initial NetScaler Gateway that a user authenticates to in a multi-active datacenter environment may not be within the preferred datacenter corresponding to the user's VDI resources. StoreFront is able to determine the location of the user's assigned resources and direct the HDX session to those resources; however, the resulting path may be sub-optional (WAN connection from the NetScaler Gateway to the virtual desktop/application resources as opposed to LAN connection).

There are static and dynamic methods available to direct HDX sessions to their virtual desktop resources in their primary datacenter. The decision regarding which method to select should be based on the availability of technology to dynamically assign sites links such as Global Server Load Balancing (GSLB) along with the network assessment of intranet and Internet bandwidth as well as Quality of Service (QoS) capabilities.

Note: *For more information on configuring the static and dynamic methods of GSLB, please refer to Citrix Product Documentation -* *Configuring GSLB for Proximity*.

- Static
    - Direct – The user may be given a FQDN mapped to an A record that is dedicated to the primary datacenter NetScaler Gateway(s) allowing them to access their virtual desktop directly wherever they are in the world. This approach eliminates a layer of complexity added with dynamic allocation. However, it also eliminates fault tolerance options such as the ability to access the virtual desktop through an alternative intranet path when a primary datacenter outage is limited to the access infrastructure.
- Dynamic

- Intranet – For most dynamic environments, the initial datacenter selected for authentication is the one closest to the user. Protocols such as GSLB dynamic proximity calculate the least latency between the user's local DNS server and the NetScaler Gateway. Thereafter, by default, the HDX session is routed through the same NetScaler Gateway to whichever datacenter is hosting the user's virtual desktops and applications. The advantage of this approach is that the majority of the HDX session would traverse the corporate WAN where quality of service may be used.

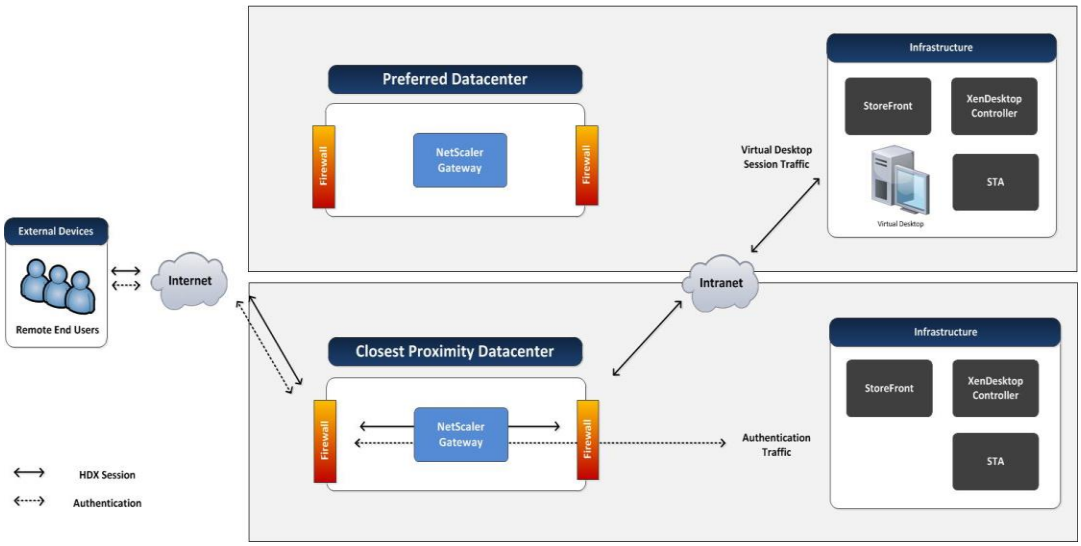


Figure 10: Intranet Connection

- Internet - Alternatively, the HDX session can be re-routed through an alternate NetScaler Gateway proximate to the backend VDI desktop / XenApp server, resulting in most of the HDX session travelling over the Internet. For example, a user with a dedicated desktop in the United Stated, traveling in Europe may be directed to a NetScaler Gateway hosted in a European datacenter based on proximity. However, when the user launches their desktop, an HDX connection will be established to the virtual desktop via a NetScaler Gateway hosted in the preferred datacenter in the United States.

  This conserves WAN network usage (at the cost of QoS) and is recommended in cases where the user's Internet connection may provide a more reliable experience than the corporate WAN.

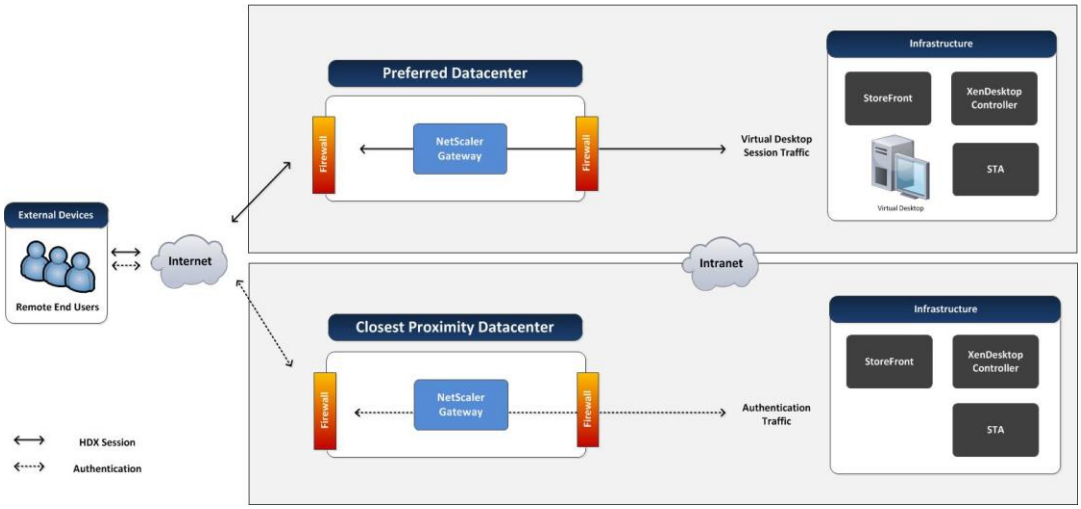


Figure 11: Internet Connection

Some customers will used a combination of these methods, such as geo-specific dymanic URLs such that fault tolerance is provided within a geographic area (such as North America) without incurring the complexity of global GSLB.

Layer 3: The Resource Layer

The resource layer is the third layer of the design methodology and the final layer focused specifically on the user groups.

The overall user acceptance of the solution is defined by the decisions made within the resource layer. Profiles, printing, applications and overall desktop image design play a pivotal role in how well the desktop is aligned with the user group's requirements, which were identified within the assess phase.

User Profiles

A user's profile plays a critical role in delivering a consistently positive experience within a virtual desktop or virtual application scenario. Even a well-designed virtual desktop solution can fail if users are frustrated due to lengthy logon times or lost settings.

The user profile solution chosen must align with the personalization characteristics of the user group captured during the assess phase as well as the VDI model selected.

Decision: Profile Type

This section provides an overview on the different profile types available and provides guidance on the optimal user profile for each VDI model.

- **Local profiles** – Local profiles are stored on each server or desktop operating system and are initially created based on the default user profile. Therefore, a user accessing these resources would create an independent profile on each system. Users are able to retain changes to their local profile on each individual system, but changes are only accessible for future sessions on that system. Local profiles require no configuration; if a user logging into a server or desktop operating system does not have a profile path administratively defined, a local profile is created by default.
- **Roaming profiles** – Roaming profiles are stored in a centralized network repository for each user. Roaming profiles differ from local profiles in that the information in the profile (whether it is a printer, a registry setting, or a file stored in the documents folder) can be made available to user sessions accessed from all systems in the environment. Configuring a user for a roaming profile requires an administrator to designate the user's profile path (for virtual desktops) or terminal server profile path to a particular network share. The first time the user logs on to a server or desktop operating system, the default user profile is used to create the user's roaming profile. During logoff, the profile is copied to the administrator-specified network location.
- **Mandatory profiles** – Mandatory profiles are typically stored in a central location for many users. However, the user's changes are not retained at logoff. Configuring a user for a mandatory profile requires an administrator to create a mandatory profile file (NTUSER.MAN) from an existing roaming or local profile and assign users' with a terminal services profile path. This can be achieved by means of Microsoft Group Policy, customizing the user properties in Active Directory or Citrix Profile Management.
- **Hybrid profiles** – Hybrid profiles combine a robust profile core (a mandatory profile or a local default profile) with user specific registry keys or files that are merged during logon. This technique enables administrators to tightly control which changes are retained and to keep the user profiles small in size. Furthermore, hybrid profiles address the last write wins issue using mature queuing techniques that automatically detect and prevent simultaneous writes that could potentially overwrite changes made in another session. Thus minimizing, user frustration resulting from lost profile changes when accessing multiple servers or virtual desktops simultaneously. In addition, they capture and record only the changes within the

profile, rather than writing the entire profile at logoff. A good example of a hybrid profile solution is Citrix Profile Management, which will be discussed in detail within this chapter.

The following table compares the capabilities of each profile type:

Feature	Local	Roaming	Mandatory	Hybrid
Central management / roams with user	✕	✓	o	✓
User settings are stored persistently	✓	✓	✕	✓
Granular capture of user settings	✕	✕	✕	✓

"✓": Functionality available "o": Optional "✕": Functionality not available

Table 19: Profile Type Capability Comparison

In order to select the optimal profile type for each user group it is important to understand their personalization requirements in addition to the FlexCast model assigned.

The following table provides recommendations on selecting the appropriate user profile type based on VDI resource:

	Local	Roaming	Mandatory	Hybrid
User setting persistence required (personalization characteristic: basic / complete)				
Hosted Windows App	✕	✓	✕	✓
Hosted Browser App	✕	✓	✕	✓
Hosted Shared Desktop	✕	✓	✕	✓
Hosted Pooled Desktop	✕	✓	✕	✓
Hosted Personal Desktop	o	✓	✕	✓
Hosted Pro Graphics Desktop	o	✓	✕	✓
Local Streamed Desktop	✕	✓	✕	✓
Local VM Desktop	✓	o	✕	o
Remote PC Access	✓	o	✕	o
User setting persistence not required or not desired (personalization characteristic: none)				
Hosted Windows App	✕	✕	✓	✕
Hosted Browser App	✕	✕	✓	✕
Hosted Shared Desktop	✕	✕	✓	✕
Hosted Pooled Desktop	✓	✕	✓	✕
Hosted Personal Desktop	✕	✕	✓	✕
Hosted Pro Graphics Desktop	o	✕	✓	✕
Local Streamed Desktop	✓	✕	✓	✕
Local VM Desktop	o	✕	✓	✕
Remote PC Access	o	✕	✓	✕

"✓": Recommended "o": Viable "✕": Not Recommended

Table 20: Profile Type Selection

Decision: Folder Redirection

Redirecting special folders can supplement any of the described profile types. While redirecting profile folders, such as user documents and favorites, to a network share is a good practice to minimize profile size, architects need to be aware that applications may frequently read and write data to profile folders such as AppData, causing potential issues with file server utilization and responsiveness. It is important to thoroughly test profile redirection before implementation in production to avoid these issues. Therefore, it is important to research profile read / write activities and to perform a pilot before moving to production. Microsoft Outlook is an example of an application that regularly performs profile read activities, as the user signature is read from the user profile every time an email is created.

The following table provides general recommendations to help identify the appropriate folders to redirect:

Folder	Local	Roaming	Mandatory	Hybrid
Application Data	✕	○	✕	○
Contacts	✕	✓	✕	○
Desktop	✕	✓	✕	○
Downloads	✕	○	✕	○
Favorites	○	✓	○	✓
Links	✕	✓	✕	○
My Documents	○	✓	○	✓
My Music	○	✓	○	○
My Pictures	○	✓	○	○
My Videos	○	✓	○	○
Saved Games	✕	○	✕	○
Searches	✕	✓	✕	○
Start Menu	✕	✕	✕	✕

"✓": Recommended "○": Optional "✕": Not recommended

Table 21: Folder Redirection Matrix

Decision: Folder Exclusion

Excluding folders from being persistently stored as part of a roaming of hybrid profile can help to reduce profile size and logon times. By default Windows excludes the AppData\Local and AppData\LocalLow folders, including all subfolders, such as History, Temp and Temporary Internet Files. In addition, the downloads and saved games folders should also be excluded. All folders that are redirected should be excluded from the profile solution.

Decision: Profile Caching

Local caching of roaming or hybrid user profiles on a server or virtual desktop is default Windows behavior and can reduce login times and file server utilization / network traffic. With profile caching, the system only has to download changes made to the profile. The downside of profile caching is that it can consume significant amounts of local disk storage on multi-user systems, such as a hosted shared desktop hosts.

In certain VDI models and configurations, the VDI resource is reset to a pristine state. Having locally cached profiles be deleted upon logoff is an unnecessary consumption of resources. Based on this, the leading recommendation is to **not** deleting locally cached profiles for the following VDI models:

- Hosted Personal Desktops
- Hosted Pooled Desktops – only in situations where a reboot occurs after logoff.
- Local VM Desktops
- Remote PC Access

Configuring the "Delay before deleting cached profiles" Citrix policy sets an optional extension to the delay before locally cached profiles are deleted at logoff. Extending the delay is useful if a process keeps files or the user registry hive open during logoff. This can also reduce logoff times for large profiles.

Decision: Profile Permissions

For security reasons, administrators, by default, cannot access user profiles. While this level of security may be required for organizations that deal with very sensitive data, it is unnecessary for most environments and can complicate operations and maintenance. Therefore, consider enabling the "Add the Administrators security group to roaming user profiles" policy setting. The configuration of this policy should be aligned with the security characteristics of the user groups captured during the assess phase. For more information on the permissions required for the file share hosting user profiles and data, please refer to Microsoft TechNet - Deploying Roaming Profiles.

Decision: Profile Path

Determining the network path for the user profiles is one of the most significant decisions during a user profile design process. In general it is strongly recommended to leverage a redundant and high performance file server or NAS device.

There are four topics that must be considered for the profile share:

- **Performance** – File server performance will affect logon and logoff times, and depending on other decisions such as redirected folders and profile streaming, can impact the user's experience within the session. For large virtual desktop infrastructures, a single file server cluster may not be sufficient to handle periods of peak activity. In order to distribute the load across multiple file servers, the file server address and share name will need to be adjusted.
- **Location** – User profiles are transferred over the network by means of the SMB protocol, which does not perform well on high-latency network connections. Furthermore, WAN connections are typically bandwidth constrained, which can add additional delay to the profile load process. Therefore, the file server should be located in close proximity to the servers and virtual desktops to minimize logon times.
- **Operating system platforms** – User profiles have a tight integration with the underlying operating system and it is not supported to reuse a single user profile on different operating systems or different platforms like 64-Bit (x64) and 32-Bit (x86). For more information, please refer to the Microsoft knowledge base article KB2384951 – Sharing 32 and 64-bit User Profiles. Windows 2008 and Windows Vista introduced a new user profile structure, which can be identified by .V2 profile directory suffix, which makes older user profiles incompatible with newer operating systems such as Windows 2012, 7 and 8. In order to ensure that a separate profile is used per platform, the profile directory has to be adapted.
- **Indexing capabilities** – To take full advantage of Windows Search functionality on a user's redirected data, Windows file servers that index the user's data must be used, as opposed to a

share on a NAS appliance. This is important for use cases that are heavily dependent on Windows Search or are especially sensitive to perception of slowness or latency.

There are two methods that can be used to address these challenges that are based on Windows built-in technology:

- **User object** – For every user object in Active Directory, an individual profile path, which contains file server name and profile directory, can be specified. Since only a single profile path can be specified per user object, it is not possible to ensure that a separate profile is loaded for each operating system platform.
- **Computer group policy or system variables** – The user profile path can also be configured by means of computer specific group policies or system variables. This enables administrators to ensure that a user profile is dedicated to the platform. Since computer specific configurations affect all users of a system, all user profiles will be written to the same file server. To load balance user profiles across multiple servers dedicated XenDesktop delivery groups have to be created per file server.

Note: Microsoft does not support DFS-N combined with DFS-R for actively used user profiles. For more information, please refer to the Microsoft articles:

- Information about Microsoft support policy for a DFS-R and DFS-N deployment scenario
- Microsoft's Support Statement Around Replicated User Profile Data

When using Citrix Profile Management, a third option is available to address these challenges:

User object attributes and variables – Citrix Profile Management enables the administrator to configure the profile path by means of a computer group policy using attributes of the user object in Active Directory to specify the file server dynamically. In order to achieve this, three steps are required:

1. Create a DNS alias (for example, fileserver1) that refers to the actual file server
2. Populate an empty LDAP attribute of the user object (for example, l or UID) with the DNS Alias
3. Configure Citrix Profile Management by means of GPO to use a profile path that refers to the LDAP attribute (for example, If attribute UID is used the profile path becomes \\#UlD#\Profiles\profiledirectory)

In addition, Citrix Profile Management automatically populates variables to specify the profile path dynamically based on the operating system platform. Valid profile management variables are:

- **!CTX_PROFILEVER!** – Expands to v1 or v2 depending on the profile version.
- **!CTX_OSBITNESS!** – Expands to x86 or x64 depending on the bit-level of the operating system.
- **!CTX_OSNAME!** – Expands to the short name of the operating system, for example Win7

By combining both capabilities of Citrix Profile Management, a fully dynamic user profile path can be created, which can be load balanced across multiple file servers and ensure profiles of different operating system platforms are not mixed. An example of a fully dynamic user profile path is shown below:

\\#UID#\profiles$\%USERNAME%.%USERDOMAIN%\!CTX_OSNAME!!CTX_OSBITNESS!

Decision: Profile Streaming

***Note**: The following design decision only applies to those environments that use Citrix Profile Management.*

With user profile streaming, files and folders contained in a profile are fetched from the user store (file server) to the local computer when a user accesses them. During the logon process, Citrix Profile Management immediately reports that the profile load process has completed reducing profile load time to almost zero.

Citrix recommends enabling profile streaming for all scenarios. If it is desired to keep a local cached copy of the user profile for performance reasons, it is recommended to enable the "Always Cache" setting and configure a size of 0. This ensures that the user profile is downloaded in the background and enables the system to use this cached copy going forward.

Note: *Profile streaming is not required and does not work with the personal vDisk feature of Citrix XenDesktop. Even if explicitly enabled by means of Group Policy, the profile streaming setting is automatically disabled.*

Experience from the Field

General – Some poorly written applications might load faster if their AppData has already been streamed to the VDI resource. Enabling the "Always Cache" option for profile streaming can help improve performance when the AppData folder is not redirected.

Decision: Active Write Back

Note*: The following design decision only applies to those environments that use Citrix Profile Management.*

By enabling the active write back feature, Citrix Profile Manager detects when an application has written and closed a file and copies the file back to the network copy of the profile during idle periods. In scenarios where a single user leverages multiple virtual desktops or hosted shared desktops simultaneously, this feature can be tremendously beneficial. However, Citrix Profile Management does not copy any registry changes back to the network, except during an ordered logoff. As such, there is a risk that the registry and files may get out of alignment on non-persistent systems, where locally cached profile information is wiped upon reboot. Therefore, it is recommended to disable active write back functionality for non-persistent scenarios.

Decision: Configuration Approach

Note*: The following design decision only applies to those environments that use Citrix Profile Management.*

Citrix Profile Management can be configured by means of an ".ini" file, Microsoft Group Policy and Citrix Policy (Citrix Profile Management 5.0 and newer). While each option offers the same configuration settings, Group Policy is recommended because it allows administrators to perform Windows and Citrix profile configurations from a single point, minimizing the tools necessary for profile management.

Note: *With Citrix Profile Management 5.0 and newer, the desktop type is automatically detected and Citrix Profile Management policies set accordingly. For more information, please refer to Citrix eDocs –* [*How automatic configuration works*].

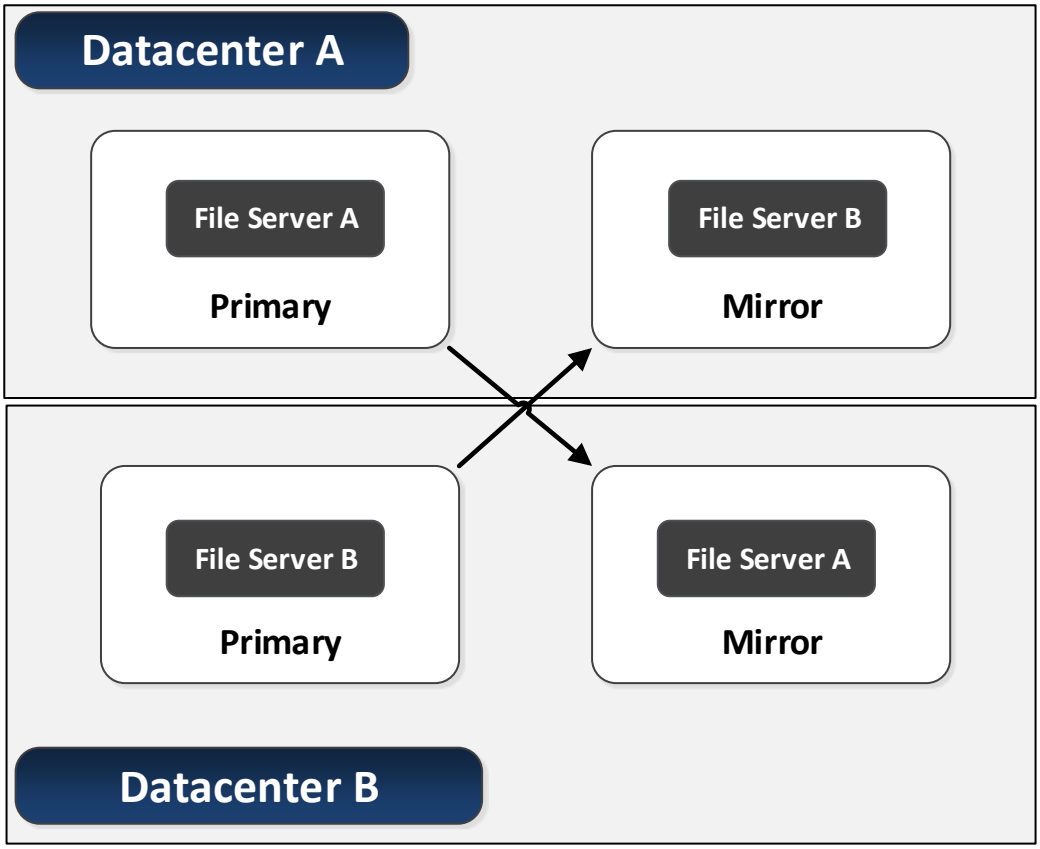
Decision: Replication

While having an active/active datacenter on a network level is easily accomplished with GSLB, the replication of user data makes having a fully active/active deployment complex in most situations. To have an active/active configuration where users are not statically assigned to a specific datacenter, will require users to have no form of personalization requirements. This will limit the user's ability to make any configuration changes and will not allow them to create any documents or persistent data. The exception to this is when a high-spee,d low latency connection such as dark fibre is available between datacenters. This will let resources in both locations can point to the same file server allowing for a true

active/active solution. Also, an active/active configuration can be accomplished when applications are used that rely solely on a backend database that is actively replicated between datacenters and do not store any data in the user profile.

For redundancy and failover purposes, user data such as Windows profiles and documents should be synchronized between datacenters. Although it is recommended to replicate user data between datacenters, the replication would be an active/passive configuration. This means the data can only be actively consumed from a single datacenter. The reason for this limitation is the distributed file locking method inside Windows that only allows a single user to actively write to a file. Therefore, active/active replication of user data is not supported. Any supported configuration consists of a one-way replication of data that is active in a single datacenter at any point in time.

For example, the figure below describes a scenario where user data is passively replicated from Datacenter A to Datacenter B. In this example, File Server A is the primary location for user data in Datacenter A and File Server B is the primary location in Datacenter B. One-way replication of the user data occurs for each fileserver to allow for the user data to be available in the opposite datacenter if a failover occurs. Replication technologies such as Microsoft DFS can be configured to mirror user profiles and documents to a file server in another datacenter. DFS Namespaces can also be used to have a seamless path for the location of the user data. However, implementing a replication solution like this requires an administrator familiar with Microsoft DFS and user profiles.



Policies

Policies provide the basis to configure and fine tune XenApp and XenDesktop environments, allowing organizations to control connection, security and bandwidth settings based on various combinations of users, devices or connection types.

When making policy decisions it is important to consider both Microsoft and Citrix policies to ensure that all user experience, security and optimization settings are considered. For a list of all Citrix-related policies, please refer to the Citrix Policy Settings Reference.

Decision: Preferred Policy Engine

Organizations have the option to configure Citrix policies via Citrix Studio or through Active Directory group policy using Citrix ADMX files, which extend group policy and provide advanced filtering mechanisms.

Using Active Directory group policy allows organizations to manage both Windows policies and Citrix policies in the same location, and minimizes the administrative tools required for policy management. Group policies are automatically replicated across domain controllers, protecting the information and simplifying policy application.

Citrix administrative consoles should be used if Citrix administrators do not have access to Active Directory policies. Architects should select one of the above two methods as appropriate for their organization's needs and use that method consistently to avoid confusion with multiple Citrix policy locations.

It is important to understand how the aggregation of policies, known as policy precedence flows in order to understand how a resultant set of policies is created. With Active Directory and Citrix policies, the precedence is as follows:

Policy Precedence	Policy Type
Processed first (lowest precedence)	Local server policies
Processed second	Citrix policies created using the Citrix administrative consoles
Processed third	Site level AD policies
Processed fourth	Domain level AD policies
Processed fifth	Highest level OU in domain
Processed sixth and subsequent	Next level OU in domain
Processed last (highest precedence)	Lowest level OU containing object

Table 22: Policy Precedence

Policies from each level are aggregated into a final policy that is applied to the user or computer. In most enterprise deployments, Citrix administrators do not have rights to change policies outside their specific OUs, which will typically be the highest level for precedence. In cases where exceptions are required, the application of policy settings from higher up the OU tree can be managed using "block inheritance" and "no override" settings. Block inheritance stops settings from higher-level OUs (lower precedence) from being incorporated into the policy. However, if a higher-level OU policy is configured with no override, the block inheritance setting will not be applied. Given this, care must be taken in policy planning, and available tools such as the "Active Directory Resultant Set of Policy" tool or the "Citrix Group Policy Modeling" wizard should be used to validate the observed outcomes with the expected outcomes.

Note: *some Citrix policy settings, if used, need to be configured through Active Directory group policy, such as Controllers and Controller registration port, as these settings are required for VDAs to register.*

Decision: Policy Integration

When configuring policies, organizations often require both Active Directory policies and Citrix policies to create a completely configured environment. With the use of both policy sets, the resultant set of policies can become confusing to determine. In some cases, particularly with respect to Windows Remote Desktop Services (RDS) and Citrix policies, similar functionality can be configured in two different locations. For example, it is possible to enable client drive mapping in a Citrix policy and disable client drive mapping in a RDS policy. The ability to use the desired feature may be dependent upon the combination of RDS and Citrix policy. It is important to understand that Citrix policies build upon functionality available in Remote Desktop Services. If the required feature is explicitly disabled in RDS policy, Citrix policy will not be able to affect a configuration as the underlying functionality has been disabled.

In order to avoid this confusion, it is recommended that RDS policies only be configured where required and there is no corresponding policy in the XenApp and XenDesktop configuration, or the configuration is specifically needed for RDS use within the organization. Configuring policies at the highest common denominator will simplify the process of understanding resultant set of policies and troubleshooting policy configurations.

Decision: Policy Scope

Once policies have been created, they need to be applied to groups of users and/or computers based on the required outcome. Policy filtering provides the ability to apply policies against the requisite user or computer groups. With Active Directory based policies, a key decision is whether to apply a policy to computers or users within site, domain or organizational unit (OU) objects. Active Directory policies are broken down into user configuration and computer configuration. By default, the settings within the user configuration apply to users who reside within the OU at logon, and settings within the computer configuration are applied to the computer at system startup, and will affect all users who logon to the system. One challenge of policy association with Active Directory and Citrix deployments revolves around three core areas:

- **Citrix environment specific computer policies –** Citrix servers and virtual desktops often have computer policies that are created and deployed specifically for the environment. Applying these policies is easily accomplished by creating separate OU structures for the servers and the virtual desktops. Specific policies can then be created and confidently applied to only the computers within the OU and below and nothing else. Based upon requirements, virtual desktops and servers may be further subdivided within the OU structure based on server roles, geographical locations or business units.
- **Citrix specific user policies –** When creating policies for XenApp and XenDesktop there are a number of policies specific to user experience and security that are applied based on the user's connection. However, the user's account could be located anywhere within the Active Directory structure, creating difficulty with simply applying user configuration based policies. It is not desirable to apply the Citrix specific configurations at the domain level as the settings would be applied to every system any user logs on to. Simply applying the user configuration settings at the OU where the Citrix servers or virtual desktops are located will also not work, as the user accounts are not located within that OU. The solution is to apply a loopback policy, which is a computer configuration policy that forces the computer to apply the assigned user configuration policy of the OU to any user who logs onto the server or virtual desktop, regardless of the user's location within Active Directory. Loopback processing can be applied with either merge or replace settings. Using replace overwrites the entire user GPO with the policy from the Citrix server or virtual desktop OU. Merge will combine the user GPO with the GPO from the Citrix server or desktop OU. As the computer GPOs are processed after the user GPOs when merge is used, the Citrix related OU settings will have precedence and be applied in the event of a conflict. For more information, please refer to the Microsoft TechNet article - Understand User Group Policy Loopback Mode.

- **Active Directory policy filtering –** In more advanced cases, there may be a need to apply a policy setting to a small subset of users such as Citrix administrators. In this case, loopback processing will not work, as the policy should only be applied to a subset of users, not all users who logon to the system. Active Directory policy filtering can be used to specify specific users or groups of users to which the policy is applied. A policy can be created for a specific function, and then a policy filter can be set to apply that policy only to a group of users such as Citrix administrators. Policy filtering is accomplished using the security properties of each target policy.

Citrix policies created using Citrix Studio have specific filter settings available, which may be used to address policy-filtering situations that cannot be handled using group policy. Citrix policies may be applied using any combination of the following filters:

Filter Name	Filter Description	Scope
Access control	Applies a policy based on access control conditions through which a client is connecting. For example, users connecting through a Citrix NetScaler Gateway can have specific policies applied.	User settings
Citrix CloudBridge	Applies a policy based on whether or not a user session was launched through Citrix CloudBridge.	User settings
Client IP address	Applies a policy based on the IPv4 or IPv6 address of the user device used to connect the session. Care must be taken with this filter if IPv4 address ranges are used in order to avoid unexpected results.	User settings
Client name	Applies a policy based on the name of the user device used to connect the session.	User settings
Delivery group	Applies a policy based on the delivery group membership of the desktop running the session	User and computer settings
Delivery group type	Applies a policy based on the type of machine running the session. For example, different policies can be set depending upon whether a desktop is pooled, dedicated or streamed.	User and computer settings
Organizational unit	Applies a policy based on the OU of the desktop or server running the session.	User and computer settings
Tag	Applies a policy based on any tags applying to the desktop running the session. Tags are strings that can be added to virtual desktops in XenDesktop environments that can be used to search for or limit access to desktops.	User and computer settings
User or group	Applies a policy based on the Active Directory group membership of the user connecting to the session.	User settings

Table 23: Citrix Policy Filters

***Note:** Citrix policies in XenDesktop 7.x provide a merged view of settings that apply at the user and computer level. In table 24, the Scope column identifies whether the specified filter applies to user settings, computer settings, or both.*

Decision: Baseline Policy

A baseline policy should contain all common elements required to deliver a high-definition experience to the majority of users within the organization. A baseline policy creates the foundation for user access, and any exceptions that may need to be created to address specific access requirements for

groups of users. It should be comprehensive to cover as many use cases as possible and should have the lowest priority, for example 99 (a priority number of "1" is the highest priority), in order to create the simplest policy structure possible and avoid difficulties in determining the resultant set of policies. The unfiltered policy set provided by Citrix as the default policy may be used to create the baseline policy as it is applied to all users and connections. In the baseline configuration, all Citrix policy settings should be enabled, even those that will be configured with the default value, in order to explicitly define desired/expected behavior, and to avoid confusion should default settings change over time.

Citrix Policy templates can be used to configure Citrix policies to effectively manage the end-user experience within an environment and can serve as an initial starting point for a baseline policy. Templates consist of pre-configured settings that optimize performance for specific environments or network conditions. The built-in templates included in XenDesktop are shown below:

Built-in Templates	
High definition user experience	Includes settings for providing high quality audio, graphics, and video to users.
High server scalability	Includes settings for providing an optimized user experience while hosting more users on a single server.
Optimized bandwidth for WAN	Includes settings for providing an optimized experience to users with low bandwidth or high latency connections.
Security and control	Includes settings for disabling access to peripheral devices, drive mapping, port redirection, and Flash acceleration on user devices.

Table 24: XenDesktop 7 Built-in Policy Templates

For more information on Citrix policy templates, please refer to Citrix eDocs - Manage Citrix Policy Templates.

A baseline policy configuration should also include Windows policies. Windows policies reflect user specific settings that optimize the user experience and remove features that are not required or desired in a XenDesktop environment. For example, one common feature turned off in these environments is Windows update. In virtualized environments, particularly where desktops and servers may be streamed and non-persistent, Windows update creates processing and network overhead, and changes made by the update process will not persist a restart of the virtual desktop or application server. Also in many cases, organizations use Windows software update service (WSUS) to control Windows updates. In these cases, updates are applied to the master disk and made available by the IT department on a scheduled basis.

In addition to the above considerations, an organization's final baseline policy may include settings specifically created to address security requirements, common network conditions, or to manage user device or user profile requirements:

Printing

Citrix XenApp and Citrix XenDesktop support a variety of different printing solutions. In order to plan and successfully implement the proper printing solution it is important to understand the available technologies as well as their benefits and limitations.

Decision: Printer Provisioning

The process of creating printers at the start of a XenApp or XenDesktop session is called printer provisioning. There are multiple approaches available:

- User Added – Allowing users to manually add printers gives them the flexibility to select printers by convenience. The drawback to manually adding network-based printers is that it requires the users to know the network name or path of the printers. There is also a chance that the native print driver is not installed in the operating system and the Citrix Universal Print Driver is not compatible, thereby requiring the user to seek administrative assistance. Manually adding printers is best suited in the following situations:
    - Users roam between different locations using the same client device (i.e. laptop, tablet).
    - Users work at assigned stations or areas whose printer assignments will rarely change.
    - Users have personal desktops with sufficient rights to install necessary printer drivers.
- Auto Created – Auto-creation is a form of dynamic provisioning that attempts to create some or all of the available printers on the client device at the start of a user session. This includes locally attached printers as well as network-based printers. Auto-creating all client printers can increase the session logon time as each printer is enumerated during the logon process.
- Session Based – Session printers are a set of network-based printers assigned to users through a Citrix policy at the start of each session.
    - Proximity Based: Session printers filtred by IP subnet. The network printers created under this policy may vary based on where the user's endpoint device is located. Proximity printing is recommended in situations where: Users roam between different locations using the same endpoint device (i.e. laptop, tablet) and where thin clients are used, which do not have the ability to connect to network-based printers directly.
    - Session printers may be assigned using the "Session Printer" policy or the "Printer Assignments" policy. The "Session printer" policy is intended to be used to set default printers for a farm, site, large group, or OU. The "Printer Assignments" policy is used to assign a large group of printers to multiple users. If both policies are enabled and configured, the session printers will be merged into a single list.
- Universal Printer – The Citrix Universal Printer is a generic printer object that is auto-created at the start of a session and is not linked to a printing device. When using the Citrix Universal Printer it is not required to enumerate the available client printers during logon, which can greatly reduce resource usage and decrease user logon times. By default the Citrix Universal Printer will print to the client's default printer, however the behavior can be modified to allow the user to select any of their compatible local or network-based printers.

  The Citrix Universal Printer is best suited for the following scenarios:

    - The user requires access to multiple printers both local and network-based which may vary with each session.
    - The user's logon performance is a priority and the Citrix policy "Wait for printers to be created" must be enabled due to application compatibility.
    - The user is working from a Windows based device or thin client.

Note: *Other options for provisioning printers, such as Active Directory group policy, "follow-me" centralized print queue solutions, and other 3rd party print management solutions can be used to provision printers into a Citrix session.*

Decision: Printer Drivers

Managing print drivers in XenApp and XenDesktop can be a tedious task, especially in large environments with hundreds of printers. In XenApp and XenDesktop there are several methods available to assist with print driver management.

- User Installed – When adding a printer within a XenApp or XenDesktop session and the native print driver is not available, the drivers can be installed manually, by the user. Many different print drivers can potentially be installed on different resources creating inconsistencies within the environment. Troubleshooting printing problems and maintenance of print drivers can become very challenging since every hosted resource may have different sets of print drivers installed. To ensure consistency and simplify support and troubleshooting, user installed drivers is not recommended.
- Automatic Installation – When connecting a printer within a XenApp or XenDesktop session, a check is made to see if the required print driver is already installed in the operating system. If the print driver is not already installed, the native print driver, if one exists, will be installed automatically. If users roam between multiple endpoints and locations, this can create inconsistencies across sessions since users may access a different hosted resource every time they connect. When this type of scenario occurs, troubleshooting printing problems and maintenance of print drivers can become very challenging since every hosted resource may have different sets of print drivers installed. To ensure consistency and simplify support and troubleshooting, automatic installed drivers is not recommended.
- Universal Print Driver – The Citrix Universal Printer Driver (UPD) is a device independent print driver, which has been designed to work with most printers. The Citrix Universal Printer Driver (UPD) simplifies administration by reducing the number of drivers required on the master image. For autocreated client printers, the driver records the output of the application and sends it, without any modification, to the end-point device. The endpoint uses local, device-specific drivers to finish printing the job to the printer. The UPD can be used in conjuction with the Citrix Universal Print Server (UPServer) to extend this functionality to network printers.

Decision: Printer Routing

Print jobs can be routed along different paths: through a client device or through a print server.

- Client Device Routing – Client devices with locally attached printers (printers attached through USB, LPT, COM, TCP, etc.) will route print jobs directly from the client device to the printer.
- Windows Print Server Routing – By default, print jobs sent to auto-created network-based printers will be routed from the user's session to the print server. However, the print job will take a fallback route through the client device when any of the following conditions are true:
    - The session cannot contact the print server
    - The print server is on a different domain without a trust established
    - The native print driver is not available within the user's session
- Citrix Universal Print Server Routing – Print job routing follows the same process as Windows Print Server Routing except that the Universal Print Driver is used between the user's session and the Citrix Universal Print Server.

The specifics with print job routing are based on the printer provisioning method. Auto-created and user-added printers can route print jobs based on the folloing diagrams:

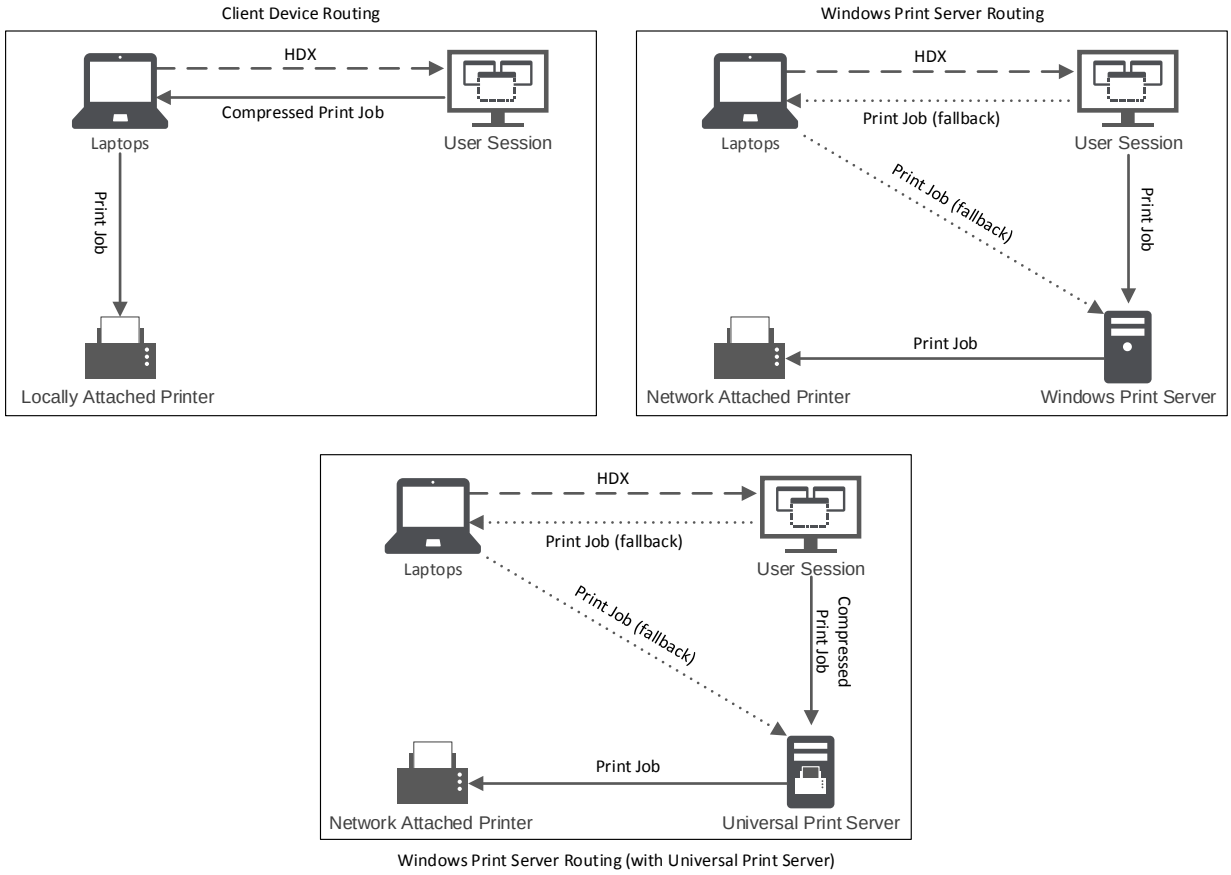


Figure 12: Auto-created and User-Added Print Job Routing

However, if the printers are provisioned as session printers, the print job routing options changes slightly. The jobs are no longer able to route through the user's endpoint device.

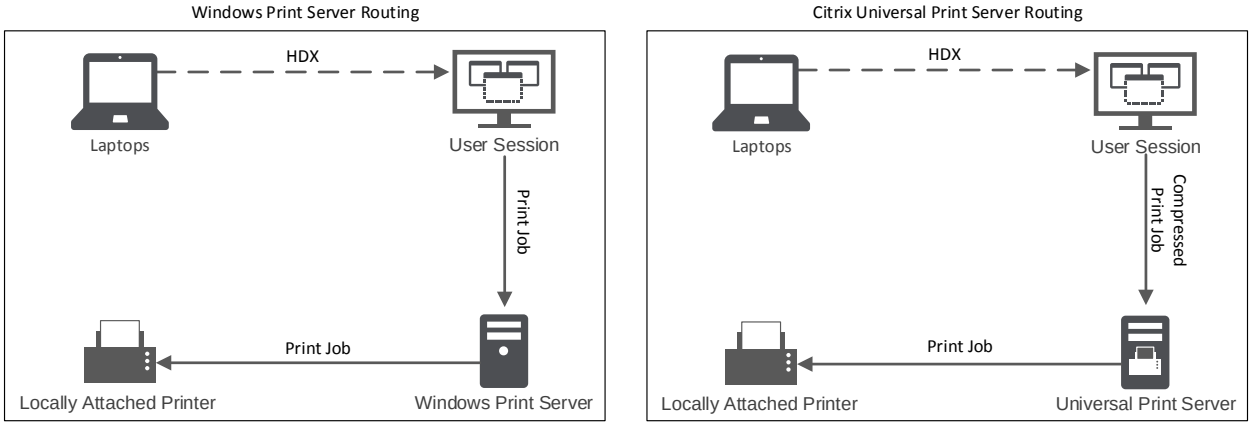


Figure 13: Session Printers Print Job Routing

The recommended option is based on the network location of the endpoint device, the user's session and the print server.

- Client Device Routing

    - Use for locally attached printer implementations.
    - Use if a Windows endpoint device and printer are on the same high-speed, low-latency network as the Windows Print Server.
- Windows Print Server Routing
    - Use if the printer is on the same high-speed, low-latency network as the Windows Print Server and user session.
- Windows Print Server Routing (with Universal Print Server)
    - Use if non-Windows endpoint device and printer are on the same high-speed, low-latency network as the Windows Print Server.

Decision: Print Server Redundancy

Network-based printers, managed with a Microsoft print server or the Citrix Universal Print Server should be configured with redundancy in order to eliminate a single point of failure. The Citrix Universal Print Server should be defined within a Citrix Policy.

Experience from the Field

A print media company leverages Thin Clients and Windows-based workstations at the company headquarters. Network based printers are placed throughout the building (one per floor). Windows print servers reside in the datacenter and manage the network printers. XenDesktop and XenApp servers also reside in the datacenter.

A regional office has numerous Windows, Linux and Mac endpoints with network attached printers.

A remote branch office has a few Windows workstations with locally attached printers.

Three different print strategies are applied:

Headquarters

A Citrix Universal Print Server is used for printing within the XenApp and XenDesktop session. Native print drivers are not required on the Windows based workstations. A session printer policy is configured per floor which connects the floor printer as the default printer. The policies are filtered based on the subnet of the thin client for proximity printing.

Quality of Service (QoS) policies are implemented. Inbound and outbound network traffic on ports TCP 1494 and TCP 2598 are prioritized over all other network traffic. This will prevent HDX user sessions from being impacted by large print jobs.

Regional Office

A Universal Print Server is deployed within the regional office. The print job uses the Universal Print Driver and is compressed and delivered from the user's session to the Universal Print Server, across the WAN. The job is then sent to the network-attached printer in the office.

Branch Office

Since all branch users work on Windows based workstations, auto-created client printers in conjunction with the Citrix Universal Printer Driver are used. Since the print job is delivered over ICA, the print data is compressed which saves bandwidth. The Citrix Universal Printer Driver ensures all printers connected to the client can be used within the XenApp or XenDesktop session without concern of the printer model used.

Applications

Properly integrating an application requires understanding compatilibty and how the user/business requirements influences the appropriate delivery method.

Decision: Compatibility

VDI typically requires significant changes to be made to an organization's application delivery and management strategy. For example, many organizations will take the opportunity to upgrade their desktop operating system and to simplify management by reducing the number of applications installed into the base image using techniques such as application streaming and application layering. These are significant changes that require comprehensive compatibility testing. Important compatibility requirements that may need to be verified include:

- Operating system –the application must be compatible with the preferred operating system.
- Multi-User – Some applications may be more appropriate for delivery via a hosted shared desktop or a hosted Windows App. In these situations, the compatibility of the application must be verified against the multi-user capabilities of a server operating system like Windows Server 2012R2.

- Application architecture – It is important to understand whether the application includes 16-bit, 32-bit or 64-bit code so that an appropriate operating system can be selected. 16-bit code cannot be executed on a 64-bit operating system. However, a 16-bit application can be delivered to users as a Hosted Windows App from a 32-bit desktop-based operating system like x86 editions of Windows 7, 8 or 10.
- Interoperability – Some applications may experience complications if they coexist on the same operating system. Possible causes include shared registry hives, dll files or INI files as well as incompatible dependencies. Application interoperability issues should be identified so that appropriate remediation steps can be taken or an alternative delivery model selected.
- Dependency – Applications may need to interact with each other to provide the users with a seamless experience. For example, applications that present information in a PDF format require a suitable PDF viewer to be available. Many times, the dependent (child) applications are version specific to the parent application. .
- Application virtualization – The use of application virtualization techniques, like streaming and layering, helps to simplify image management by reducing the number of applications installed into the base image. However, not all applications are suitable for streaming and layering because they may install device drivers, use COM+ or form part of the operating system.

Application compatibility can be achieved by doing a combination of manual, user testing, utilizing pre-verified lists maintained by the software vendor, or using an automated application compatibility solution, like Citrix AppDNA which runs through thousands of tests to verify compatibility.

Decision: Application Delivery Method

It is unlikely that a single delivery method will meet all requirements. Based on the outcome of the application categorization assessment process, several application delivery methods can be considered.

Choosing one of the appropriate application delivery method helps improve scalability, management and user experience.

- Installed app – The application is part of the base desktop image. The install process involves dll, exe and other files being copied to the image drive as well as registry modifications.
- Streamed App (Microsoft App-V) – The application is profiled and delivered to the desktops across the network on-demand. Application files and registry settings are placed in a container on the virtual desktop and are isolated from the base operating system and each other, which helps to address compatibility issues.
- Hosted Windows App - The application is installed on a multi-userXenApp host and deployed as an application and not a desktop. The hosted Widnwos app is accessed seamless from the user's VDI desktop or endpoint device, hiding the fact that the app is executing remotely.
- Local App – The application is deployed on the endpoint device. The application interface appears within the user's hosted VDI session even though it executes on the endpoint.

The following table provides recommendations on the preferred approaches for integrating applications into the overall solution,

App Category	Installed App	Streamed App	Hosted Windows App	Local App
Common	✓	°	°	✕
Departmental	°	✓	✓	✕

User	✕	∘	∘	✓
Management	✓	✕	∘	✕

"✓": Recommended, "✕": Not Recommended, " " Viable

Table 25: App Deployment Recommendations

Experience from the Field

Energy – An energy company installs applications on the base image for the majority of users and streams departmental applications as required.

Financial – A banking customer maintains and deploys multiple desktop images containing user group focused applications as required by various departments.

Virtual Machines

Virtual resources require proper allocation of the processor, memory and disk. These decisions have a direct impact on the the amount of hardware required as well as the user experience.

The key to successful resource allocation is to ensure that virtual desktops and applications offer similar levels of performance to physical desktops. Otherwise, productivity and overall user satisfaction will be affected. Allocating resources to the virtual machines above their requirements however is inefficient and expensive for the business.

The resources allocated should be based on the workload characteristic of each user group, identified during the assess phase.

Decision: Virtual Processor (vCPU)

For hosted desktop-based VDI models (hosted pooled desktops and hosted personal desktops), the general recommendation is two or more vCPUs per virtual machine so that multiple threads can be executed simultaneously. Although a single vCPU could be assigned for extremely light workloads, users are more likely to experience session hangs.

For hosted server-based VDI models (hosted Windows apps, hosted browser apps, hosted shared desktops), the proper vCPU allocation is based on the Non-Uniform Memory Access (NUMA) architecture of the processors.

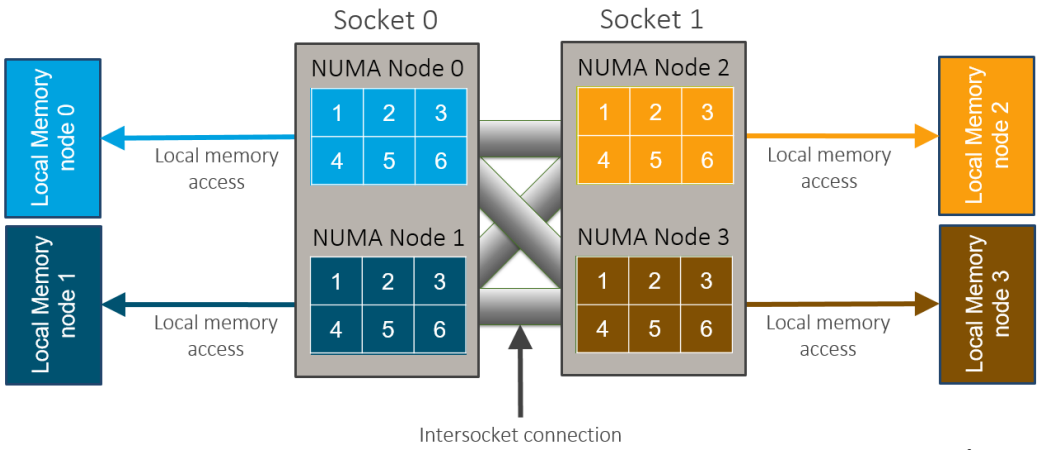


Figure 14: NUMA Architecture

Each socket is divided into one or more NUMA nodes. Hosted server-based VDI models will often utilize 4 or more processors. Allocating more vCPU than the NUMA node contains results in a performance hit. Allocating a portion of a NUMA node to a virtual machine results in a perofmrance hit if the portion allocated is not easily divisible by the size of the NUMA node. It is often ideal to allocate the number of cores within a NUMA node to a virtual machine or allocate ½ of the cores to a virtual machine, while doubling the number of virtual machiens.

User Workload	Operating System	vCPU Configured for Scale	vCPU Configured for Experience
Light	Windows 7	2 vCPU	2 vCPU
	Windows 10	2 vCPU	2 vCPU
	Windows 2012R2	NUMA or ½ of NUMA	NUMA or ½ of NUMA
Medium	Windows 7	2 vCPU	3 vCPU
	Windows 10	2 vCPU	3 vCPU
	Windows 2012R2	NUMA or ½ of NUMA	NUMA or ½ of NUMA
Heavy	Windows 7	3 vCPU	4 vCPU
	Windows 10	3 vCPU	4 vCPU
	Windows 2012R2	NUMA or ½ of NUMA	NUMA or ½ of NUMA

Table 26: vCPU Allocation

Note: Windows 2012R2 recommendations are based on the hosted Windows app, hosted browser app and hosted shared desktop VDI model.

Decision: Virtual Memory (vRAM)

The amount of memory allocated to each resource is a function of the user's expected workload and application footprint. Assigning insufficient memory to the virtual machines will cause excessive paging to disk, resulting in a poor user experience; allocating too much RAM increases the overall cost of the solution.

The following table provides guidance on the virtual RAM that should be assigned based on workload.

User Workload	Operating System	vRAM Configured for Scale	vRAM Configured for Experience
Light	Windows 7	2 GB	3 GB
	Windows 10	2 GB	3 GB
	Windows 2012R2	256 MB per user	
Medium	Windows 7	3 GB	4 GB
	Windows 10	3 GB	4 GB
	Windows 2012R2	512 MB per user	
Heavy	Windows 7	6 GB	8 GB
	Windows 10	6 GB	8 GB
	Windows 2012R2	1024 MB per user	

Table 27: vRAM Allocation

Note: Windows 2012R2 recommendations are based on the hosted Windows app, hosted browser app and hosted shared desktop VDI model.

Note: Memory allocation above 4GB requires a 64-bit operating system

Note: If used, the Machine Creation Sevices and Provisioning Sevices cache in RAM amount should be added onto the virtual machine RAM specifications.

Decision: Disk Cache

The amount of storage that each VM requires will vary based on the workload and the image type. If creating hosted personal desktop without leveraging an image management solution, each VM will require enough storage for the entire OS and locally installed applications.

Deploying machines through Machine Creation Services or Provisioning Services can substantially reduce the storage requirements for each virtual machine. Disk space requirements for the write cache and difference disk will depend on application usage and user behavior. However, the following table provides a starting point for estimating disk space requirements based on machine sized with vCPU and vRAM as per the guidelines above:

User Workload	Operating System	Storage Space (Differencing Disk / Write Cache Disk)
Light	Windows 7	10 GB
	Windows 10	10 GB
	Windows 2012R2	40 GB
Medium	Windows 7	15 GB
	Windows 10	15 GB
	Windows 2012R2	40 GB
Heavy	Windows 7	20 GB
	Windows 10	20 GB
	Windows 2012R2	40 GB

Table 28: Disk Cache Allocation

Decision: RAM Cache

Provisioning Services and Machine Creation Services have the capability to utilize a portion of the virtual machine's RAM as a buffer for the storage cache. The RAM cache is used to improve the performance of traditional storage by sharing the virtual machine's non-paged pool memory

User Workload	Operating System	RAM Cache Configured for Scale	RAM Cache Configured for Experience
Light	Windows 7	128 MB	256 MB
	Windows 10	128 MB	256 MB
	Windows 2012R2	2 GB	
Medium	Windows 7	256 MB	512 MB
	Windows 10	256 MB	512 MB
	Windows 2012R2	2 GB	
Heavy	Windows 7	512 MB	1024 MB
	Windows 10	512 MB	1024 MB
	Windows 2012R2	2 GB	

Table 29: RAM Cache Allocation

Note: If used, the Machine Creation Sevices and Provisioning Sevices cache in RAM amount should be added onto the virtual machine RAM specifications.

Note: If additional RAM is available on the host, the RAM Cache amounts can be increased to provide even greater levels of performance.

Decision: Storage IOPS

Storage performance is limited by the number of operations it can handle per second, referred to as IOPS. Underallocating storage IOPS results in a VDI desktop where apps, web pages and data are slow to load.

The following table provides guidance on the number of storage IOPS generated per user based on workload and operating system. Storage IO activity will be higher during user logon/logoff.

User Workload	Operating System	Storage IOPS (without RAM-Based Cache)	Storage IOPS (with RAM-Based Cache)
Light	Windows 7	10 IOPS	1 IOPS
	Windows 10	12 IOPS	1 IOPS
	Windows 2012R2	3 IOPS	0.5 IOPS
Medium	Windows 7	15 IOPS	1 IOPS
	Windows 10	20 IOPS	1.5 IOPS
	Windows 2012R2	4 IOPS	0.5 IOPS
Heavy	Windows 7	25 IOPS	2 IOPS
	Windows 10	35 IOPS	3 IOPS
	Windows 2012R2	5 IOPS	½ IOPS

Table 30: IOPS Allocation

Decision: Graphics (GPU)

Without a graphical processing unit (GPU), graphical processing is rendered with software by the CPU. A graphical processing unit (GPU) can be leveraged to improve server scalability and user experience or enable the use of graphically intensive applications. During the desktop design it is important to decide how the GPU (if used) will be mapped to the virtual machines. There are three methods available.

- Pass-Through GPU – Each physical GPU is passed through to a single virtual machine (hosted apps or hosted desktops).
- Hardware Virtualized GPU – Using a hypervisor's vGPU technology, an NVIDIA GRID or Intel Iris Pro is virtualized and shared between multiple machines. Each virtual machine has the full functionality of GPU drivers and direct access to the GPU.
- Software Virtualized GPU – The GPU is managed by the hypervisor and intercepts requests made by the VDI desktops. This process is used if a GPU is not installed within the host.

	Pass-Through GPU	Hardware Virtualized GPU (NVidia)	Hardware Virtualized GPU (Intel)	Software Emulated GPU
Citrix XenServer				
XenDesktop	✓	✓	✓	✓

XenApp	✓	✓	✓	✓
Microsoft Hyper-V				
XenDesktop	✓	✕	✕	✓
XenApp	✓	✕	✕	✓
VMware vSphere				
XenDesktop	✓	✓	✕	✓
XenApp	✓	✓	✕	✓

"✓": Available "✕": Not Supported

Table 31 GPU Allocation Options

User groups with a heavy use of graphical applications will often require the use of a NVidia hardware virtualized GPU. User groups who rely on office-based applications can have an observable benefit with the use of a hardware virtualized GPU from Intel.

Layer 4: The Control Layer

Active Directory

Decision: Forest Design

Multi-forest deployments, by default, do not have inter-domain trust relationships between the forests. An AD administrator can establish trust relationships between the multiple forests, allowing the users and computers from one forest to authenticate and access resources in another forest.

For forests that have inter-domain trusts, it is recommended that the appropriate settings be configured to allow the Delivery Controllers to communicate with both domains. When the appropriate trusts are not configured, multiple XenDesktop sites for each forest must be configured. This section outlines the storage requirements on a per product basis and provides sizing calculations. For more information, please refer to Citrix article: CTX134971 – Successfully Deploying XenDesktop in a Complex Active Directory Environment

Decision: Organizational Unit Structure

Infrastructure components for a XenApp and XenDestkop deployment should reside within their own dedicated organizational units (OUs); separating workers and controllers for management purposes. By having their own OUs, the objects inside will have greater flexibility with their management while allowing Citrix administrators to be granted delegated control.

A sample Citrix OU structure can be seen below.

- Citrix
  - <Datacenter 1>
    - Citrix Infrastructure
      - Delivery Controllers
      - Provisioning Services Servers
      - StoreFront Servers
    - Citrix VDAs
      - Desktop OS
      - Server OS

Figure 15: Example Citrix OU Structure

Decision: User Groups

Whenever possible, permissions and authorization should be assigned to user groups rather than individual users, thereby eliminating the need to edit a large amount of resource permissions and user rights when creating, modifying, or deleting user accounts.

Permission application example:

- An application published to one group of 1,000 users requires the validation of only one object for all 1,000 users.
- The same application published to 1,000 individual user accounts requires the validation of all 1,000 objects.

Database

The majority of Citrix products discussed within this document require a database. The following table outlines the usage on a per product basis:

Product	Configuration Data	Runtime Data	Audit / Change Log Data	Monitoring Data
XenDesktop	✓	✓	✓	✓
Provisioning Services	✓		①	
XenClient	✓	✓	✓	

"①": Optional

Table 32: Database usage

Decision: Edition

There are multiple editions of Microsoft SQL Server 2012: Express, Web, Standard, Business Intelligence, and Enterprise. Based on the capabilities of the various SQL Server editions available, the Standard edition is often used for hosting the XenApp and XenDesktop databases in production environments.

The Standard edition provides an adequate amount of features to meet the needs of most environments. For more information on the databases supported with Citrix products please refer to the Citrix Database Support Matrix. Different versions of Citrix products support different versions of the SQL server; therefore it is important to check the support matrix to ensure the version of SQL server used is compatible with the Citrix product being deployed.

Decision: Database Server Sizing

The SQL server must be sized correctly to ensure the performance and stability of an environment. Since every Citrix product uses SQL server in a different way, no generic all-encompassing sizing recommendations can be provided. Instead, per-product SQL server sizing recommendations are provided below.

XenApp and XenDesktop

XenApp and XenDesktop Brokers use the database as a message bus for broker communications, storing configuration data and storing monitoring and configuration log data. The databases are constantly in use and the performance impact on the SQL server can be considered as high.

Based on results from Citrix internal scalability testing the following SQL server specification for a server hosting all XenDesktop databases are recommended:

- 2 Cores / 4 GB RAM for environments up to 5,000 users

- 4 Cores / 8 GB RAM for environments up to 15,000 users
- 8 Cores / 16 GB RAM for environments with 15,000+ users

The database files and transaction logs should be hosted on separate hard disk subsystems in order to cope with a high number of transactions. For example, registering 20,000 virtual desktops during a 15 minute boot storm causes ~500 transactions / second and 20,000 users logging on during a 30 minute logon storm causes ~800 transactions / second on the XenDesktop Site Database.

Provisioning Services

In addition to static configuration data provisioning servers store runtime and auditing information in the database. Depending on the boot and management pattern, the performance impact of the database can be considered as low to medium.

Based on this categorization, a SQL server specification of 4 Cores and 4 GB RAM is recommended as a good starting point. The SQL server should be carefully monitored during the testing and pilot phase in order to determine the optimal configuration of the SQL server. .

Decision: Instance Sizing

When sizing a SQL database, two aspects are important:

- **Database file** – Contains the data and objects such as tables, indexes, stored procedures and views stored in the database.
- **Transaction log file** – Contains a record of all transactions and database modifications made by each transaction. The transaction log is a critical component of the database and, if there is a system failure, the transaction log might be required to bring the database back to a consistent state. The usage of the transaction log varies depending on which database recovery model is used:
    - **Simple recovery** – No log backups required. Log space is automatically reclaimed, to keep space requirements small, essentially eliminating the need to manage the transaction log space. Changes to the database since the most recent backup are unprotected. In the event of a disaster, those changes must be redone.
    - **Full recovery** – Requires log backups. No work is lost due to a lost or damaged database data file. Data of any arbitrary point in time can be recovered (for example, prior to application or user error). Full recovery is required for database mirroring.
    - **Bulk-logged** – Requires log backups. This is an adjunct of the full recovery model that permits high-performance bulk copy operations. It is typically not used for Citrix databases.

    For further information, please refer to the Microsoft Developer Network – SQL Server Recovery Models.

In order to estimate storage requirements, it is important to understand the disk space consumption for common database entries. This section outlines the storage requirements on a per product basis and provides sizing calculations. For more information, please refer to Citrix article: CTX139508 – XenDesktop 7.x Database Sizing.

XenDesktop General

XenApp 7.x and XenDesktop 7.x use three distinct databases:

- **Site Configuration database –** Contains static configuration and dynamic runtime data
- **Monitoring database –** Contains monitoring data which is accessible via Director

- **Configuration logging database –** Contains a record for each administrative change performed within the site (accessible via Studio)

Site Database

Since the database of a XenApp or XenDesktop site contains static configuration data and dynamic runtime data, the size of the database file depends not only on the physical size of the environment but also user patterns. The following factors all impact the size of the database file:

- The number of connected sessions
- The number of configured and registered VDAs
- The number of transactions occurring during logon
- VDA heartbeat transactions

The size of the Site Database is based on the number of VDAs and active sessions. The following table shows the typical maximum database size Citrix observed when scale testing XenApp and XenDesktop with a sample number of users, applications, and desktop delivery methods.

Users	Applications	Desktop Types	Expected Maximum Size (MB)
1,000	50	Hosted Shared	30
10,000	100	Hosted Shared	60
100,000	200	Hosted Shared	330
1,000	N/A	Hosted Pooled	30
10,000	N/A	Hosted Pooled	115
40,000	N/A	Hosted Pooled	390

Table 33: XenDesktop Site DB sample size calculations

Note: This sizing information is a guide only. Actual database sizes may differ slightly by deployment due to how databases are maintained.

Determining the size of the transaction log for the Site database is difficult due to factors that can influence the log including:

- The SQL Database recovery model
- Launch rate at peak times
- The number of desktops being delivered

During XenDesktop scalability testing, Citrix observed the transaction log growth rate at 3.5MB an hour when the system is idle, and a per user per day growth rate of ~32KB. In a large environment, transaction log usage requires careful management and a regular backup, to prevent excessive growth. This can be achieved by means of scheduled jobs or maintenance plans

Monitoring Database

Of the three databases, the Monitoring database is expected to be the largest since it contains historical information for the site. Its size is dependent on many factors including:

- Number of Users
- Number of sessions and connections
- Number of workers

- Retention period configuration – Platinum customers can keep data for over a year (default 90 days). Non-platinum customers can keep data for up to 7 days (default 7 days).
- Number of transaction per second. Monitoring service tends to execute updates in batches. It is rare to have the number of transactions per second go above 20.
- Background transaction caused by regular consolidation calls from the Monitoring service.
- Overnight processing carried out to remove data outside the configured retention period.

The following table shows the estimated size of the Monitoring database over a period of time under different scenarios. This data is an estimate based on data seen within scale testing XenApp and XenDesktop (assuming a 5 day working week).

Estimates with 1 connection and 1 session per user with a 5 day work week					
Users	Type	1 week (MB)	1 month (MB)	3 months (MB)	1 year (MB)
1,000	HSD	20	70	230	900
10,000	HSD	160	600	1,950	7,700
100,000	HSD	1,500	5,900	19,000	76,000
1,000	VDI	15	55	170	670
10,000	VDI	120	440	1,400	5,500
40,000	VDI	464	1,700	5,400	21,500
Estimates with 2 connections and 1 session per user with a 5 day work week					
Users	Type	1 week (MB)	1 month (MB)	3 months (MB)	1 year (MB)
1,000	HSD	30	100	330	1,300
10,000	HSD	240	925	3,000	12,000
100,000	HSD	2,400	9,200	30,000	119,000
1,000	VDI	25	85	280	1,100
10,000	VDI	200	750	2,500	9,800
40,000	VDI	800	3,000	9,700	38,600

Table 34: Monitoring DB size estimations

Note: *The 100,000 HSD tests are based on a test environment consisting of:*

- *2 Delivery Controllers*
- *43 Hosted Shared Desktop workers*
- *3 SQL servers, configured with databases held within one Always On Availability Group.*

For more information please see the Citrix Support article – XenDesktop 7.x Database Sizing.

The size of the transaction log for the Monitoring Database is very hard to estimate, but XenApp and XenDesktop scalability testing showed a growth rate of about 30.5 MB an hour when the system is idle, and a per user per day growth rate of ~9 KB.

Configuration Logging Database

The Configuration Logging Database is typically the smallest of the three databases. Its size and the size of the related transaction log depends on the daily administrative activities initiated from Studio, Director or PowerShell scripts, therefore its size is difficult to estimate. The more configuration changes

are performed, the larger the database will grow. Some factors that can affect the size of the database include:

- The number of actions performed in Studio, Director and PowerShell.
- Minimal transactions which occur on the database when no configuration changes are taking place.
- The transaction rate during updates. Updates are batched whenever possible.
- Data manually removed from the database. Data within the Configuration Logging Database is not subject to any retention policy, therefore it is not removed unless done so manually by an administrator.
- Activities that have an impact on sessions or users, for example, session logoff and reset.
- The mechanism used for deploying desktops.

In XenApp environments not using MCS, the database size tends to fall between 30 and 40MB. For MCS environments, database size can easily exceed 200MB due to the logging of all VM build data.

Temporary Database

In addition to the Site, Monitoring, and Configuration Logging databases, a system-wide temporary database (tempdb) is provided by SQL Server. This temporary database is used to store Read-Committed Snapshot Isolation data. XenApp 7.x and XenDesktop 7.x uses this SQL Server feature to reduce lock contention on the XenApp and XenDesktop databases. Citrix recommends that all XenApp 7.x and XenDesktop 7.x databases use Read-Committed Snapshot Isolation. For more information please see How to Enable Read-Committed Snapshot in XenDesktop.

The size of the tempdb database will depend on the number of active transactions, but in general it is not expected to grow more than a few MBs. The performance of the tempdb database does not impact the performance of XenApp and XenDesktop brokering, as any transactions that generate new data require tempdb space. XenApp and XenDesktop tend to have short-lived transactions, which help keep the size of the tempdb small.

The tempdb is also used when queries generate large intermediate result sets. Guidance and sizing the tempdb can be found on the Microsoft TechNet article Optimizing tempdb Performance.

Provisioning Services

The Provisioning Services farm database contains static configuration and configuration logging (audit trail) data. The record size requirements outlined below can be used to help size the database:

Configuration Item	DB Space Required (KB)	Number of Items (Example)	Total (KB)
Base farm configuration	112	-	112
User group w/ farm access	50	10	250
Site	4	5	20
Device collection	10	50	500
Farm view	4	10	40
Farm view to device relationship	5	1	5,000
Site View	4	5	20
Site view to device relationship	5	1	5,000
Device	2	5,000	10,000

Device bootstrap	10	-	-
Device to disk relationship	35	1	175,000
Device printer relationship	1	-	-
Device personality data	1	-	-
Device status (when booted)	1	5,000	5,000
Device custom property	2	-	-
vDisk	1	20	20
vDisk version	3	5	300
Disk locator	10	1	200
Disk locator custom property	2	-	-
Server	5	10	50
Server IP	2	1	20
Server status (when booted)	1	20	20
Server custom property	2	-	-
vDisk store	8	5	40
vDisk store to server relationship	4	1	40
Connection to XenServer (VirtualHostingPool)	4	-	-
vDisk update task	10	10	100
Administrative change (auditing enabled)	1	10,000	10,000
Total			**211,732KB (~212MB)**

Table 35: Provisioning Services Farm DB sample size calculations

During the PVS farm setup, a database with an initial file size of 20MB is created. Due to the nature of the data in the PVS farm database the transaction log is not expected to grow very quickly, unless a large amount of configuration is performed.

In contrast to XenApp, which also offers the ability to track administrative changes, the related information is not written to a dedicated database but directly to the Provisioning Services farm database. In order to limit the size of the Provisioning Services database it is recommended to archive the audit trail data on a regular schedule.

Decision: Database Location

By default, the Configuration Logging and Monitoring databases are located within the Site Configuration database. Citrix recommends changing the location of these secondary databases as soon as the configuration of the site has been completed, in order to simplify sizing, maintenance and monitoring. All three databases can be hosted on the same server or on different servers. An ideal configuration would be to host the Monitoring database on a different server from the Site Configuration and Configuration Logging databases since it records more data, changes occur more frequently and the data is not considered to be as critical as the other databases. For more information, please refer to Citrix eDocs – Change secondary database locations.

Note: *The location of the Configuration Logging database cannot be changed when mandatory logging is enabled.*

Decision: High-Availability

The following table highlights the impact to XenApp, XenDesktop and Provisioning Services when there is a database outage:

Component	Impact of Database Outage
Site configuration database	Users will be unable to connect or reconnect to a virtual desktop. *Note: Connection leasing in XenApp and XenDesktop 7.6 allows users with Hosted Shared Desktops, Hosted Windows and Browser Applications, and Personal Desktops to reconnect to their most recently used applications and desktops even when the site database is unavailable.*
Monitoring database	Director will not display any historical data and Studio cannot be started. Brokering of incoming user requests and existing user sessions will not be affected.
Configuration logging database	If allow changes when the database is disconnected has been enabled within XenApp and XenDesktop logging preferences, an outage of the configuration logging database will have no impact (other than configuration changes not being logged). Otherwise, administrators will be unable to make any changes to the XenApp and XenDesktop site configuration. Users are not impacted.
Provisioning Services farm database	When offline database support is enabled and the database becomes unavailable, the stream process uses a local copy of the database to retrieve information about the provisioning server and the target devices supported by the server. This allows provisioning servers and the target devices to remain operational. However, when the database is offline, the console and the management functions listed below become unavailable: • AutoAdd target devices • vDisk creation and updates • Active Directory password changes • Stream process startup • Image update service • PowerShell and MCLI based management If offline database support has not been enabled, all management functions become unavailable and the boot and failover of target devices will fail.

Table 36: Impact of a database outage

Note: *Please review HA options for 3rd party databases (for example, App-V, SCVMM or vCenter) with the respective software vendor.*

In addition to the built-in database redundancy options, Microsoft SQL Server, as well as the underlying hypervisor (in virtual environments), offer a number of high availability features. These enable administrators to ensure single server outages will have a minimal impact (if any) on the XenApp and XenDesktop infrastructure. The following the SQL / hypervisor high availability features are available:

- **VM-level HA –** This high availability option is available for virtual SQL servers only, which need to be marked for High Availability at the hypervisor layer. In case of an unexpected shutdown of the virtual machine or the underlying hypervisor host, the hypervisor will try to restart the VM immediately on a different host. While VM-level HA can minimize downtimes in power-outage scenarios, it cannot protect from operating system level corruption. This solution is less expensive than mirroring or clustering because it uses a built-in hypervisor feature. However,

the automatic failover process is slower, as it can take time detect an outage and start the virtual SQL server on another host. This may interrupt the service to users.

- **Mirroring –** Database mirroring increases database availability with almost instantaneous failover. Database mirroring can be used to maintain a single standby or mirror database, for a corresponding principal or production database. Database mirroring runs with either synchronous operation in high-safety mode, or asynchronous operation in high- performance mode. In high-safety mode with automatic failover (recommended for XenDesktop) a third server instance, known as a witness, is required, which enables the mirror server to act as a hot standby server. Failover from the principal database to the mirror database happens automatically and is typically completed within a few seconds. It is a good practice to enable VM-level HA (or a similar automatic restart functionality) for at least the witness to ensure SQL service availability in case of a multi-server outage.

  Note: Microsoft is planning to remove mirroring as a high availability option in a future release of SQL Server and is discouraging its use in new network development. Please refer to the Microsoft article – Database Mirroring (SQL Server) for more information.

- **AlwaysOn Failover Cluster Instances –** Failover clustering provides high-availability support for an entire instance of Microsoft SQL Server. A failover cluster is a combination of two or more nodes, or servers, using a shared storage. A Microsoft SQL Server AlwaysOn Failover Cluster Instance, introduced in SQL Server 2012, appears on the network as a single computer, but has functionality that provides failover from one node to another if the current node becomes unavailable. The transition from one node to the other node is seamless for the clients connected to the cluster. AlwaysOn Failover cluster Instances require a Windows Server Failover Clustering (WSFC) resource group. The number of nodes supported in the WSFC resource group will depend on the SQL Server edition. (Please refer to the table in the Decision: Edition earlier in this chapter.) For more information please refer to MSDN – AlwaysOn Failover Cluster Instances (SQL Server).

- **AlwaysOn Availability Groups –** AlwaysOn Availability Groups is an enterprise-level high-availability and disaster recovery solution introduced in Microsoft SQL Server 2012, which enables administrators to maximize availability for one or more user databases. AlwaysOn Availability Groups require that the Microsoft SQL Server instances reside on Windows Server failover clustering (WSFC) nodes. Similar to failover clustering a single virtual IP / network name is exposed to the database users. In contrast to failover clustering, shared storage is not required since the data is transferred using a network connection. Both synchronous and asynchronous replication to one or more secondary servers is supported. As opposed to mirroring or clustering secondary servers can be actively used for processing incoming read-only requests, backups or integrity checks. This feature can be used to offload user resource enumeration requests to a secondary SQL server in XenDesktop environments to essentially scale-out a SQL server infrastructure. Since the data on active secondary servers can lag multiple seconds behind the primary server, the read-only routing feature cannot be used for other XenDesktop database requests at this point in time. For more information, please refer to MSDN – AlwaysOn Availability Groups (SQL Server).

The following table outlines the recommended high availability features for Citrix databases:

Component	VM-Level HA	Mirroring	AlwaysOn Failover Cluster Instances	AlwaysOn Availability Groups
Site database	①	✓	°	°
Configuration logging database	①	°	°	°

Monitoring database	①	✓	○	○
Provisioning Services farm database	①	✓	○	×
XenClient database	①	×	○	○

"✓": Recommended "○": Viable "×": Not Supported "①": Recommended for test environments only

Table 37: Recommended SQL high availability options

Decision: Connection Leasing

Connection leasing is a new XenApp and XenDesktop 7.6 feature that allows Hosted Shared, Hosted Windows and Browser Apps and Personal VDI users to connect and reconnect to their most recently used applications and desktops, even when the site database is unavailable. Connection Leasing is not available for users with a Pooled VDI desktop.

The lease information along with the application, desktop, icon, and worker information is stored on the controller's local disk and synchronized between controllers in the site. If the site database becomes unavailable, the controllers enter a "leased connection mode" and replay cached operations from an XML file on the local disk to connect or reconnect users to a recently used application or desktop.

Administrators familiar with the local host cache in XenApp 6.5 and earlier should understand the similarities and differences with connection leasing because it can have an impact on the design and scalability of the XenApp and XenDesktop 7.6 solution. In XenApp 6.5 and earlier, the IMA service is responsible for synchronizing the local host cache with the data store. In XenApp and XenDesktop 7.6, the FMA service caches the brokering operations (leases) to an XML file containing the address of the VDA, application path, and other details required for the session to launch. The FMA also caches dynamic information such as user sessions, VDA registrations, and load. These files are uploaded to the SQL database and synchronized between all controllers in the site. The controllers will download the files on a regular basis so that any other controller in the site can connect a user to their session.

Each controller needs additional disk space for the cached lease files. At a minimum, 4KB is required for each lease file. Each resource entry in the enumeration lease will take anywhere from 200 bytes to a few KBs depending on the number of entries and resources published. Citrix testing has shown that 200,000 leased connections for server hosted applications and desktops required approximately 3GB of disk space. 40,000 leased connections for assigned desktops required approximately 156MB of disk space.

By default, connection leases have an expiration period of two weeks. Applications and desktops must have been launched within the two last weeks to still be accessible when the database is unavailable. The expiration period is configurable using PowerShell cmdlets or editing the registry and can be set from 0 minutes to several years. Setting the expiration period too short will prevent users from connecting to their virtual desktops and applications in the event of an outage. Setting the expiration period too long will increase storage requirement on the controllers.

By default, connection leasing affects the entire site, however, leases can be revoked for specific users, which prevents them from accessing any applications or desktops when the site database is unavailable.

For more information on connection leasing considerations and configuration, please refer to eDocs – Connection Leasing.

Citrix Licensing

Citrix offers organizations the flexibility of multiple licensing models that align with common usage scenarios. The different licensing models vary based on the Citrix product used, but can include per user/device and per concurrent user. Several Citrix products use the license server, while other products require a license to be installed on the product itself.

Product	License Location
XenDesktop	Citrix License Server
XenApp	Citrix License Server
Provisioning Services	Citrix License Server
XenServer	Citrix License Server
NetScaler	On the product
NetScaler Gateway	On the product

For more information on XenDesktop 7.x licensing, please refer to CTX128013 - XenDesktop Licensing.

For more information on Microsoft Licensing, please refer to the Microsoft document – Licensing Microsoft's Virtual Desktop Infrastructure Technology.

Decision: Sizing

Internal scalability testing has shown that a single virtual license server with two cores and 2GB of RAM can issue approximately 170 licenses per second or 306,000 licenses per half hour. If necessary, the specification of the license server can be scaled out to support a higher number of license requests per second.

Decision: High Availability

For a typical environment, a single license server is sufficient. Should the license server become unavailable, dependent Citrix products will enter a 30-day grace period, which provides more than enough time to resolve connectivity issues and/or restore or rebuild the license server.

Note: If the license server and the Citrix product do not communicate within 2 heartbeats (5-10 min), the Citrix product will enter a grace period and will allow connections for up to 30 days. Once communication with the license server is re-established, the license server will reconcile the temporary and actual licenses.

Note: A CNAME record in DNS is a convenient way to reference the license server. Using CNAMEs allows the license server name to be changed without updating the Citrix products.

If additional redundancy is required, Citrix supports the following high availability solutions for the license server.

- **Windows Clustering** – Cluster servers are groups of computers that work together in order to increase availability. Clustering allows the license server role to automatically failover in the event of a failure. For more information on clustering, please see the Citrix eDocs article – Clustered License Servers.
- **Duplication of license server** – Create a VM level backup of the license server. This backup should not be stored on the same host as the license server. Instead, it should be stored in a safe location, such as a highly available storage solution, or backed up to tape or disk. The duplicate server is not active, and will remain on standby until the need arises to restore the active license server. Should the license server be restored using this backup, any new licenses must be re-downloaded to the server.

For more information, please refer to Citrix eDocs – Licensing Architecture Overview.

Each method allows an administrator to exchange a single license server for another without an interruption in service; assuming that the change occurs during the grace period and that the following limitations are considered.

- License files will reference the server specified during the allocation process. This means that the license files can only be used on a server with the same binding information (Hostname) as the server that was previously specified.
- Two Windows-based, domain joined license servers cannot share the same name and be active in the environment at the same time.
- Because license servers do not communicate with each other, any additional licenses must be placed on both the active and backup license server.

Decision: Optimization

License server performance can be optimized by tuning the number of "receive" and "processing" threads. If the thread count is set too low, requests will be queued until a thread becomes available. Conversely, if the thread count is set too high, the license server will become overloaded.

The optimal values are dependent on the server hardware, site configuration, and license request volume. Citrix recommends testing and evaluating different values to determine the proper configuration. Setting the maximum number of processing threads to 30 and the maximum number of receiving threads to 15 is a good starting point for large scale deployments.

This optimization will improve the Citrix License Server 's ability to provide licenses by increasing its ability to receive and process license requests.

For more information, please refer to the Citrix eDocs – Improving Performance by Specifying Thread Use

Delivery Controllers

Decision: Topology

A XenApp and XenDesktop site groups desktops and applications together to form a single architectural and management entity. All persistent and dynamic data for the site, including site configuration, desktop assignments, and session state, is stored in a site's database.

A single site can be subdivided into zones, often associated with geographical locations. There are several factors that must be considered when determining the overall topology of the XenApp and XenDesktop solution:

- **Risk Tolerance –** Multiple XenDesktop sites can be created to minimize the impact from a site-wide outage. For example, corruption of the XenDesktop site database could affect site-wide availability. For many organizations, the decreased risk from implementing multiple sites outweighs the additional management overhead and supporting infrastructure required.

  Experience from the Field

  Finance – A large financial institution hosts 10,000 desktops from a single datacenter. To reduce risk, it was decided that no site should exceed 5,000 desktops. Therefore, despite the desktops being connected by a fast and redundant network, two sites were created.

- **Security –** Although delegated administration is available, high-security organizations may require complete separation between environments to demonstrate compliance with specific service level agreements.

Experience from the Field

Retail – A retail organization required complete separation for employees responsible for managing financial data. To meet this requirement, two separate sites were created within the same datacenter – one for the financial employees and a second for all other employees.

- **Geographical Connectivity –** Although the implementation of zones does allow a single site to span geographical locations, there must be enough bandwidth between the satellite zone and primary zone for session information to be captured within the site database.

Session count	Max concurrent session launches	Min site-to-site bandwidth	Max site-to-site round trip latency
Less than 50	20	1 Mpbs	250 ms
50 to 500	25	1.5 Mbps	100 ms
500 to 1,000	30	2 Mbps	50 ms
1,000 to 3,000	60	8 Mbps	10 ms
Over 3,000	60	8 Mbps	5 ms

Table 38: Zone Networking Requirements

In general, the number of XenDesktop sites and zonesshould be kept to a minimum to reduce architectural complexity and administrative effort.

Decision: Server Sizing

Delivery Controller scalability is based on CPU utilization. The more processor cores available, the more virtual desktops a controller can support. Each desktop startup, registration, enumeration and launch request impacts the controller's processor. As the storm increases in intensity, the CPU utilization of the controller will increase. If the CPU reaches a critical threshold, roughly 80%, the site will need to either scale up or scale out.

Adding additional CPU cores to a Delivery Controller will lower the overall CPU utilization, thus allowing for greater numbers of desktops supported by a single controller. This is really only feasible when dealing with virtualized controllers as adding virtual CPUs is fairly easy and straightforward. The other alternative is to add another controller into the site configuration. The controller would have the same configuration as other controllers, and the load would be evenly distributed across all controllers, thus helping to reduce the overall load on each single controller.

Testing has shown that a single Delivery Controller, using the following configuration, can support more than 5,000 desktops.

Component	Specification
Processor	4 vCPU
Memory	4GB RAM
Network	Bonded virtual NIC
Host Storage	40GB shared storage
Operating System	Windows Server 2012
XenDesktop	7

Table 39: Delivery Controller Specification for 5K Sessions

The following formula can be used to calculate the number of Delivery Controllers required for a Citrix site.

$$Number\ of\ Delivery\ Controllers = \frac{Number\ of\ Active\ Sessions\ per\ Site}{5{,}000} + 1$$

Decision: High Availability

If the server hosting the Delivery Controller is unavailable, users will not be able to access their virtual desktops or published applications. Therefore at least two Delivery Controllers (N+1 redundancy) should be deployed per zone on different physical servers to prevent this component from becoming a single point of failure. If one controller fails, the others can manage connections and administer the site.

The locations of all Delivery Controllers are specified on the VDA, allowing it to automatically failover if communication with one Delivery Controller is unavailable. The VDA checks the following locations, in order, stopping at the first place it finds the Delivery Controller:

1. A persistent storage location maintained for the auto-update feature. This location contains controller information when auto-update is enabled and after the VDA successfully registers for the first time after installation.

For its initial registration after installation, or when auto-update is disabled, the VDA checks the following locations.

2. Policy settings (Delivery Controllers, Delivery Controller SIDs).
3. The Delivery Controller information under the VDA ListofDDCs registry key. The VDA installer initially populates these values, based on the information specified when installing the VDA.
4. OU-based discovery. This is a legacy method maintained for backward compatibility.
5. The Personality.ini file created by Machine Creation Services.

Citrix Consulting recommends utilizing the auto-update feature (enabled by default). This feature will simplify management of the environment by keeping VDA's updated when adding and removing Delivery Controllers.

Decision: XML Service Encryption

In a typical session, the StoreFront server passes credentials to the Citrix XML Service on a Delivery Controller. The Citrix XML protocol uses clear text to exchange all data, with the exception of passwords, which are transmitted using obfuscation.

If the traffic between the Storefront servers and the XenDesktop Controllers can be intercepted it will be vulnerable to the following attacks:

- Attackers can intercept the XML traffic and steal resource set information and tickets.
- Attackers with the ability to crack the obfuscation can obtain user credentials.
- Attackers can impersonate the XenDesktop Controller and intercept authentication requests.

For most organizations, the Citrix XML traffic will be isolated on a dedicated physical or virtual datacenter network making interception unlikely. However, for safety consider using SSL encryption to send StoreFront data over a secure HTTP connection.

Decision: Server OS Load Management

Default Load Management policies are applied to all Server OS delivery groups. The default settings specify the maximum number of sessions a server can host at 250 and do not consider CPU and Memory usage. Capping session count does not provide a true indication of load, which can lead to an

overburdening of Server OS delivery groups resulting in a degradation of performance or an underutilization of Server OS delivery groups resulting in an inefficient usage of resources.

Citrix Consulting recommends creating unique "custom" Load Management policies for each Delivery Group based on performance and scalability testing. Different rules and thresholds can be applied to each Delivery Group depending on the different resource bottlenecks identified during testing. For more information on the available Load Management policy configurations refer to Citrix eDocs – Load Management policy settings.

If adequate testing cannot be performed prior to production, Citrix Consulting recommends implementing the following "custom" Load Management policy which can be applied to all servers as a baseline:

- **CPU Usage** - Full Load: 80%
- **CPU usage excluded process priority –** Below Normal or Low
- **Memory Usage -** Full Load: 80%
- **Memory Usage base load –** Report zero load (MBs): 786
- **Maximum number of sessions** – X

The "Maximum number of sessions" policy is included for capping purposes – this is considered a best practice for resiliency. Organizations can choose an initial value of 250 (denoted by "X" above). **It is highly recommended that this value and others be customized based on the results from scalability testing**.

Provisioning Services

Citrix Provisioning Services (PVS) uses streaming technology to simplify the deployment of virtual and physical machines. Computers are provisioned and re-provisioned in real-time from a single shared-disk image. In doing so, administrators can completely eliminate the need to manage and patch individual systems. Instead, all image management is performed on the master image.

Decision: Topology

A Provisioning Services farm represents the top level of the Provisioning Services infrastructure, which can be further broken down into sites. All provisioning servers in a farm share the same SQL database and Citrix license server.

Each site is a logical entity containing perovisioning servers, vDisk pools and target device collections. Although all sites within a farm share the same database, target devices can only fail over to other pvosioning servers within the same site.

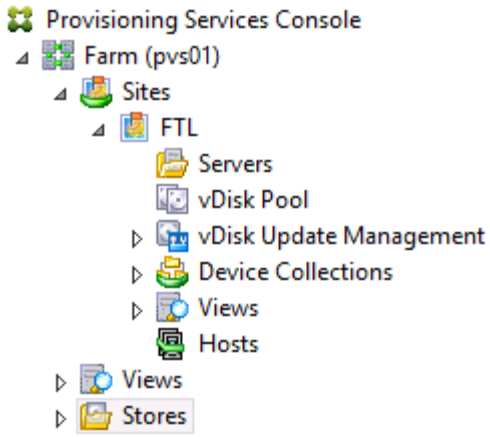


Figure 16: PVS Site structure

There are factors that must be considered when determining the overall Provisioning Services topolgy:

- **Network –** Provisioning servers are constantly communicating with the farm database to retrieve system configuration settings. Therefore, separate farms should be created for each physical location where target devices reside, unless they are connected to the database server by a fast and robust connection. .
- **Administration –** Organizations may need to maintain the separation of administrative duties at a departmental, regional or countrywide basis. Additional Provisioning Services farms will add some complexity to the management of the environment. However, this overhead is typically limited to initial configuration, desktop creation and image updates.
- **Organization –** A practical reason for building multiple sites is due to organizational changes. For example, two companies may have recently merged through acquisition, but need to keep resources separate while integration takes place. Configuring the organization to use separate sites is one way to keep the businesses separate but managed centrally through the Provisioning Services console.

Only create additional sites if the business requirements warrant it. A single site per farm is easier to manage and requires no additional configuration.

Decision: Device Collections

Device collections provide the ability to create and manage logical groups of target devices. Creating device collections simplifies device management by allowing actions to be performed at the collection level rather than the target device level.

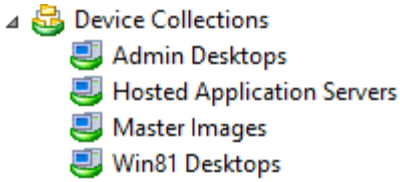


Figure 17: Device Collection structure

Device collections can represent physical locations, subnet ranges, chassis or different departments within an organization. Collections can also be used to logically separate production target devices from test and maintenance ones.

Consider creating device collections based on vDisk assignment so that the status of all target devices assigned to a particular vDisk can be quickly identified.

Decision: High Availability

Provisioning Services is a critical component of the virtual desktop infrastructure. The following recommendations should be followed to eliminate single points of failure:

- **Provisioning Server –** A minimum of two provisioning servers should always be implemented per site. Sufficient redundancy should be incorporated into the design so that a single server failure does not reduce the total number of target devices that can be supported per site.

  The Provisioning Services boot file should be configured for high availability. Up to four Provisioning Servers may be listed in the boot file. Target devices will try to contact the servers in the order that they are listed. The server that responds may not necessarily be the server that will provide streaming services to the target device. If Load Balancing is enabled, the target device may be reassigned to another server in the site that is less loaded than the others.

- **vDisks and Storage –** For vDisk stores hosted on local, Direct Attached Storage (DAS) or Storage Area Network (SAN), replication should be used to synchronize the vDisks. If using Network Attached Storage (NAS), ensure that the vDisks are hosted on a highly available network share.

- **Networking –** The provisioning server's should have redundant NICs. If the provisioning server is deployed as a physical server, redundant NIcs should be teamed and if the provisioning server is deployed as a virtual server, the underlying hypervisor should incorporate redundant NICs.

  Note: *The target devices will only failover to NICs that are in the same subnet as the PXE boot NIC.*

Trivial File Transfer Protocol (TFTP) is a communications protocol used for transferring configuration or boot files between machines. Provisioning services can use TFTP to deliver the bootstrap file to target devices. There are several options available to make the TFTP service highly available. Some of the more commonly used options are:

- **DNS Round Robin –** A DNS entry is created for the TFTP service with multiple A records corresponding to the TFTP services running on the provisioning servers in the farm. This method is not recommended since the state of the TFTP service is not monitored. Clients could potentially be sent to a non-functioning server.

- **Hardware load balancer –** Use a hardware load balancer, such as Citrix NetScaler, to create virtual IPs that corresponds to the provisioning servers. The NetScaler can intelligently route traffic between the provisioning servers. In the event that one of the servers becomes unavailable, NetScaler will automatically stop routing TFTP requests to that server. This is the best method for making TFTP highly available, but can be complicated to setup.

- **Multiple DHCP Option 66 entries** – This method is easy to implement but requires a DHCP service that supports entering multiple entries in option 66. Microsoft DHCP server allows one option 66 entry so this method would not be feasible in environments with Microsoft DHCP services. If using a non-Microsoft DHCP server or appliance, check with the manufacturer to verify that multiple option 66 entries is supported.

There are other options available that can achieve the same result without having to use TFTP:

- **Proxy DHCP –** Use the provisioning servers PXE service to provide the bootstrap information. If one of the servers is down, the next available server in the farm can provide the bootstrap information. This method requires the provisioning servers to be on the same broadcast domain as the target devices. If there are other PXE services running on the network (Altiris,

SCCM, etc.) then multiple VLANs may be required to keep the PXE services from interfering with each other.

- **Boot Device Manager –** Use the Boot Device Manager to create a bootstrap file that is either placed on the local hard drive, or used as a bootable ISO file. If the ISO file is used, configure the target devices to boot from the CD/DVD-ROM drive, and place the ISO file on a highly available shared network location or local storage of each target device. When either method is utilized, the TFTP service is not used at all.

High availability should always be incorporated into the Provisioning Services design. Although high availability may require additional resources and increased costs, it will provide a highly stable environment so that users experience minimal impact due to service outages.

Decision: Bootstrap Delivery

A target device initiates the boot process by first loading a bootstrap program which initializes the streaming session between the target device and the provisioning server. There are three methods in which the target device can receive the bootstrap program:

- **Using DHCP Options –**

    1. When the target device boots, the target device sends a broadcast for IP address and boot information. DHCP will process this request and provide an IP as well as scope option settings 66 (the name or IP address of the Provisioning Services TFTP server) and 67 (the name of the bootstrap file).

        Note: *If using a load balancer for the TFTP service then the address of the load balancer is entered in option 66.*

    2. Using TFTP, a request for the bootstrap file is sent from the target device to the provisioning server. The target device downloads the boot file from the provisioning server.

    3. The target device boots the assigned vDisk image.

    Note: *Requires UDP/DHCP Helper to be configured when targets are not on the same subnet as the DHCP servers in order to receive PXE broadcasts.*

- **Using PXE Broadcasts –**

    1. When a target device boots from the network, the target device sends a broadcast for an IP address and boot information. DHCP will process this request and provide an IP address. In addition, all provisioning servers that receive the broadcast will return boot server and boot file name information. The target device will merge the information received and start the boot process.

    2. Using TFTP, a request for the bootstrap file is sent from the target device to the provisioning server which responded first. The target device downloads the boot file from the provisioning server.

    Note: *Make sure no other PXE services are in use on the same subnet, such as the Altiris PXE service, or isolate using VLANs otherwise conflicts may occur with Provisioning Services.*

    Note: *Requires UDP/DHCP Helper to be configured when targets are not on the same subnet as the DHCP and PVS servers in order to receive PXE broadcasts.*

- **Using Boot Device Manager –** The Boot Device Manager (BDM) creates a boot file that target devices obtain through an ISO image mounted from a network share, a physical CD/DVD drive placed on the server, or the boot file information is written to a hard drive partition local to the target devices.

A summary of the advantages and disadvantages for each delivery method is listed in the following table.

Delivery Method	Advantages	Disadvantages
DHCP Options	Easy to implement.	Requires changes to production DHCP service. DHCP service may only allow one option 66 entry. Requires UDP/DHCP helper for targets on different subnets.
PXE	Easy to implement	Can interfere with other running PXE services on the same subnet. Requires UDP/DHCP helper for targets on different subnets.
BDM	Does not require PXE or TFTP services	Extra effort required to boot physical target devices.

Table 40: Bootstrap delivery options and advantages/disadvantages.

Note: *When configuring the bootstrap file, up to four provisioning servers may be listed. The order in which the provisioning servers appear in the list determines the order which the provisioning servers are accessed. If the first server does not respond, the next server in the list is contacted.*

Decision: vDisk Format

Provisioning Services supports the use of fixed-size or dynamic vDisks:

- **Fixed-size disk –** For vDisks in private mode, fixed-size prevents disk fragmentation, and offers improved write performance over dynamic disks.
- **Dynamic disk –** Dynamic disks require less storage space than fixed-size disks, but offer significantly lower write performance. Although vDisks in Shared mode do not perform writes to the vDisk, the time required to complete vDisk merge operations will increase with dynamic disks. This is not a common occurrence as more environments choose to create new vDisks when updating.

Since most reads will be to the System Cache in RAM, there is no significant change in performance when utilizing fixed-size or dynamic disks. In addition, dynamic disks require significantly less storage space. Therefore dynamic disks are recommended.

Decision: vDisk Replication

vDisks hosted on a local, Direct Attached Storage or a SAN must be replicated between vDisk stores whenever a vDisk is created or changed. Provisioning Services supports the replication of vDisks from stores that are local to the provisioning server as well as replication across multiple sites that use shared storage. The replication of vDisks can be performed manually or automatically:

- **Manual –** Manual replication is simple, but can be time consuming, depending on the number of vDisks and vDisk stores. If an error occurs during the replication process, administrators can catch them straight away and take the appropriate steps to resolve them. The risk of manual replication is vDisk inconsistency across the provisioning servers which will result in load balancing and failover to not work properly. For example, if a vDisk is replicated across three servers and then one of the vDisks is updated, that vDisk is no longer identical and will not be

considered if a server failover occurs. Even if the same update is made to the other two vDisks, the timestamps on each will differ, and therefore the vDisks are no longer identical.

- **Automated –** For large environments, automated replication is faster than the manual method due to the number of vDisks and vDisk Stores required. Some automated tools, such as Microsoft DFS-R, support bandwidth throttling and Cross File Remote Differential Compression (CF-RDC), which use heuristics to determine whether destination files are similar to the file being replicated. If so, CF-RDC will use blocks from these files to minimize the amount of data transferred over the network. The risk of automated replication is that administrator do not typically monitor replication events in real-time and do not respond quickly when errors occur, unless the automation tool has an alerting feature. Some tools can be configured to automatically restart the copy process in the event of a failure. For example, Robocopy supports "resume copying" in the event that the network connection is interrupted.

For medium and large projects, use a tool to automate vDisk replication. Select a tool that is capable of resuming from network interruptions, copying file attributes and preserving the original timestamp.

Note: *Load balancing and high availability will not work unless the vDisks have identical timestamps.*

Decision: Server Sizing

Generally, a Provisioning Server is defined with the following specifications:

Component	Specification
Model	Virtual
Processor	4 to 8 vCPU
Memory	2GB + (# of vDisks * 2GB)
Network	10 Gbps NIC
Host Storage	40 GB shared storage
Operating System	Windows Server 2012R2

Table 41: General Provisioning Services specifications

Model

Citrix Provisioning Services can be installed on virtual or physical servers:

- **Virtual –** Offers rapid server provisioning, snapshots for quick recovery or rollback scenarios and the ability to adjust server resources on the fly. Virtual provisioning servers allow target devices to be distributed across more servers helping to reduce the impact from server failure. Virtualization makes more efficient use of system resources.
- **Physical –** Offers higher levels of scalability per server than virtual servers. Physical provisioning servers mitigate the risks associated with virtual machines competing for underlying hypervisor resources.

In general, virtual provisioning servers are preferred when sufficient processor, memory, disk and networking resources can be made available and guaranteed to be available.

Note: *For high availability, ensure that virtual Provisioning Servers are distributed across multiple virtualization hosts. Distributing the virtual servers across multiple hosts will eliminate a single point of failure and not bring down the entire Provisioning Services farm in the event of a host failure.*

CPU

Provisioning Services is not CPU intensive. However, underallocating the number of CPUs does impact the optimization of the network streams. The number of streams that a Provisioning Services server can run concurrently can be determined by the following formula:

$$Max\ Number\ of\ Streams = \#\ of\ Ports * \#\ of\ Threads/Port$$

By default the Streaming Service is configured with 20 sequential network ports, and 8 threads per port. Therefore, by default, a provisioning server can support 160 concurrent targets. If more than 160 streams are required, Provisioning Services continuously switches between streaming different target devices

Ideally, if the environment needs to support more than 160 concurrent targets, the number of ports, and threads per port can be adjusted in the Provisioning Services console. Best performance is attained when the threads per port is not greater than the number of cores available on the provisioning server. If the provisioning server does not have sufficient cores, the server will show a higher CPU utilization, and target devices waiting for requests to be processed will have a higher read latency.

Even though Provisioning Services is not CPU intensive, allocating 2 CPUs will require a larger continguous network port range.

- Small environments (up to approximately 500 virtual machines) 4 vCPUs are recommended.
- Larger environments 8 vCPUs are recommended.

RAM

The Windows operating system hosting Provisioning Services partially caches the vDisks in memory (system cache) reducing the number of reads required from storage. Reading from storage is significantly slower than reading from memory. Therefore, Provisioning Servers should be allocated sufficient memory to maximize the benefit from this caching process.

The following formula can be used to determine the optimal amount of memory that should be allocated to a provisioning server:

$$Total\ Server\ RAM = 2GB + (\#\ of\ vDisks * 2GB)$$

Network

Unlike most other XenApp and XenDesktop components, Provisioning Services does not bottleneck the CPU. Provisioning Services scalability is based on network throughput.

The following table shows the approximate amount of data that Provisioning Services requires to boot different operating systems:

Operating System	Avg Boot Data Usage (MB)
Windows 10 x64	240
Windows 8 x86	178
Windows 8 x64	227
Windows 7 x86	166
Windows 7 x64	210
Windows 2012	225
Windows 2012 R2	232
Windows 2008 R2	251
Windows Vista x86	190

Windows Vista x64	240

Table 42: Approximate boot data usage by OS

Determining how much time will be required to boot the target devices can be estimated using the following formula:

$$Seconds\ to\ Boot = \frac{(Number\ of\ Targets\ *\ MB\ Usage)}{\text{Network Throughput}}$$

Operating System	Number of VMs	Network Throughput	Time to Boot
Windows 10 x64	500	1 Gbps	960 Seconds (16 minutes)
Windows 10 x64	500	10 Gbps	96 Seconds (1 minute, 36 seconds)

Table 43: Boot time estimate

A 10Gbps network is recommended for use with Provisioning Services. If a 10Gbps network is not available, consider link aggregation to provide additional bandwidth to the provisioning servers, or a dedicated physical streaming network.

Tip: *Firewalls can add latency and create bandwidth bottlenecks in Provisioning Services environments. If the use of firewalls cannot be avoided, refer to the Citrix whitepaper CTX101810 – Communication Ports Used By Citrix Technologies, for the list of ports that should be enabled for full functionality.*

Growth

As the farm grows, administrators will need to decide whether to add more resources to the provisioning servers or to add more provisioning servers to the farm.

There are a number of environmental factors that need to be considered when determining whether the Provisioning Servers should be scaled up or scaled out:

- **Redundancy –** Spreading user load across additional less-powerful servers helps reduce the number of users affected from a single provisioning server failure. If the business is unable to accept the loss of a single high-specification server, consider scaling out.
- **Failover times –** The more target devices connected to a single provisioning server, the longer it will take for them to failover in the event that the server fails. Consider scaling out to reduce the time required for target devices to failover to another server.
- **Data center capacity –** The data center may have limited space, power and/or cooling available. In this situation, consider scaling up.
- **Hardware costs –** Initially, it may be more cost effective to scale up. However, there will be a point where scaling out actually becomes more cost effective. A cost analysis should be performed to make that determination.
- **Hosting costs –** There may be hosting and/or maintenance costs based on the number of physical servers used. If so, consider scaling up to reduce the long-term cost of these overheads.

Decision: Network Configuration

As mentioned before it is essential that the network is sized correctly to prevent network bottlenecks causing high disk access times and directly affecting virtual desktop performance. The following diagram outlines a common Provisioning Services network infrastructure:

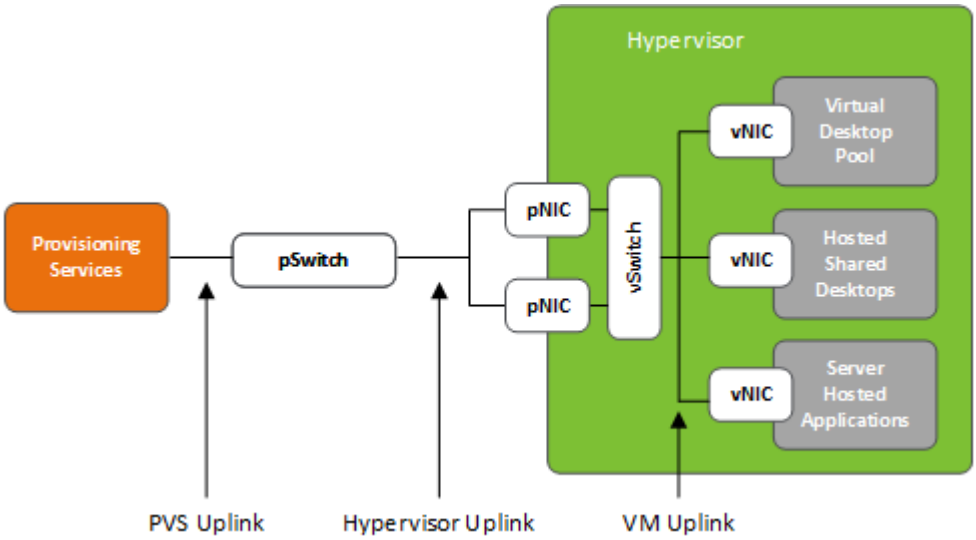


Figure 18: Sample PVS Network Configuration

The following network configuration is recommended for the network sections outlined within the diagram:

- **PVS Uplink –** All disk access from the target devices will be transferred via the PVS network uplink. This means hundreds or even thousands of devices will use this network connection. Therefore it is vital that this connection is redundant and can failover without any downtime. Furthermore Citrix recommends a minimum bandwidth of 1Gbps per 500 target devices. For virtual provisioning servers a respective QoS quota or a dedicated physical network uplink should be configured to ensure best performance.
- **Hypervisor Uplink –** Used by all PVS target devices hosted on a particular hypervisor host. Therefore redundancy with transparent failover is strongly recommended. Unless the target devices run a very I/O intensive workload or perform I/O intensive tasks (e.g. booting) simultaneously, a bandwidth of 1Gbps is sufficient for this uplink.
- **VM Uplink –** All network traffic for a virtual machine, including PVS streaming traffic, will traverse this virtual network connection. Unless the workload is extremely I/O intensive a bandwidth of 100 Mbps is sufficient to handle even peak loads during I/O intensive tasks, such as booting from vDisk. For example a Windows 2012 R2 Server will read approximately 232MB during a period of 90 seconds from the vDisk until the Windows Logon Screen is shown. During this period an average data rate of 20.5 Mbps with peaks up to 90 Mbps can be observed.

The following switch settings are recommended for Provisioning Services:

- **Disable Spanning Tree or Enable PortFast** – In a switching environment the Spanning Tree Protocol (STP) places ports into a blocked state while it transmits Bridged Protocol Data Units (BPDUs) and listens to ensure the BPDUs are not in a loopback configuration. The port is not placed in a forwarding state until the network converges, which depending on the size of the network, may incur enough time to cause Preboot Execution Environment (PXE) timeouts. To eliminate this issue, disable STP on edge-ports connected to clients or enable PortFast.
- **Storm Control** - Storm Control is a feature available on Cisco switches that allows a threshold to be set whereby, multicast, broadcast, or unicast traffic may be suppressed. Its purpose is to prevent malicious or erroneous senders from flooding a LAN and affecting network

performance. PVS Servers may send a large amount of traffic by design that falls within a storm control threshold, therefore the feature should be configured accordingly.

- **Broadcast Helper** – The broadcast helper is required to direct broadcasts from clients to servers that would otherwise not be routed. In a PVS environment it is necessary to forward PXE boot requests when clients are not on the same subnet as the servers. If possible the recommended network design is to have PVS servers residing on the same subnet as the target devices. This mitigates the risk of any service degradation due to other networking infrastructure components.

The following network interface features should be taken into consideration when selecting a network interface for Provisioning Services:

- **TCP Offloading** – Offloading I/O tasks to the network interface reduces CPU usage and improves overall system performance, however, PVS Streaming Services can be negatively impacted when Large Send Offload is enabled due to the extra work placed on the network adapter. Many network adapters will have Large Send Offload and TCP checksum offload enabled by default.

  Note: *If Large Send Offload is enabled and the switch that the traffic is passing through does not support the frame size sent by the Large Send Offload engine, the switch will drop the frame causing data retransmission. When retransmitting, the operating system will segment the frames instead of the network adapter, which can lead to severe performance degradation.*

- **Receive Side Scaling** (RSS) – Receive side scaling enables packets received from a network adapter to be balanced across multiple CPUs which allows incoming TCP connections to be load balanced, preventing bottlenecks from occurring to a single CPU. In Windows Server 2008 R2 and Windows Server 2012/2012 R2, RSS is enabled by default..

Note: *For more information on PVS networking best practices please refer to Best Practices for Configuring Provisioning Services Server on a Network.*

Note: *For Provisioning Services implementations on low bandwidth networks (1Gbps or slower), performance may be improved by isolating streaming traffic from other network traffic on the LAN.*

Note: *Microsoft does not support NIC teaming with Hyper-V on Windows Server 2008 R2; however, third party solutions are available. Microsoft does support NIC teaming with Hyper-V on Windows Server 2012/2012 R2. All support queries regarding teaming with Hyper-V should be directed to the NIC OEM.*

Decision: Subnet Affinity

The Provisioning Services Subnet Affinity is a load balancing algorithm that helps to ensure target devices are connected to the most appropriate provisioning server. When configuring subnet affinity the following options are available:

- **None** – Ignore subnets; uses the least busy server.
- **Best Effort** – Uses the least busy server/NIC combination from within the same subnet. If no server/NIC combination is available within the subnet, select the least busy server from outside the subnet. If more than one server is available within the selected subnet, perform load balancing between those servers. This is the default setting.
- **Fixed** – Use the least busy server/NIC combination from within the same subnet. Perform load balancing between servers within that subnet. If no server/NIC combination exists in the same subnet, do not boot target devices assigned to this vDisk.

The following examples show common network configurations for physical provisioning servers. Similar configurations can be implemented for virtual provisioning servers without compromising on performance or functionality.

Blade Design

The provisioning servers and the target devices that they support reside within the same chassis. In most cases, the chassis will have a dedicated 10Gbps switch shared among all blade servers within the chassis.

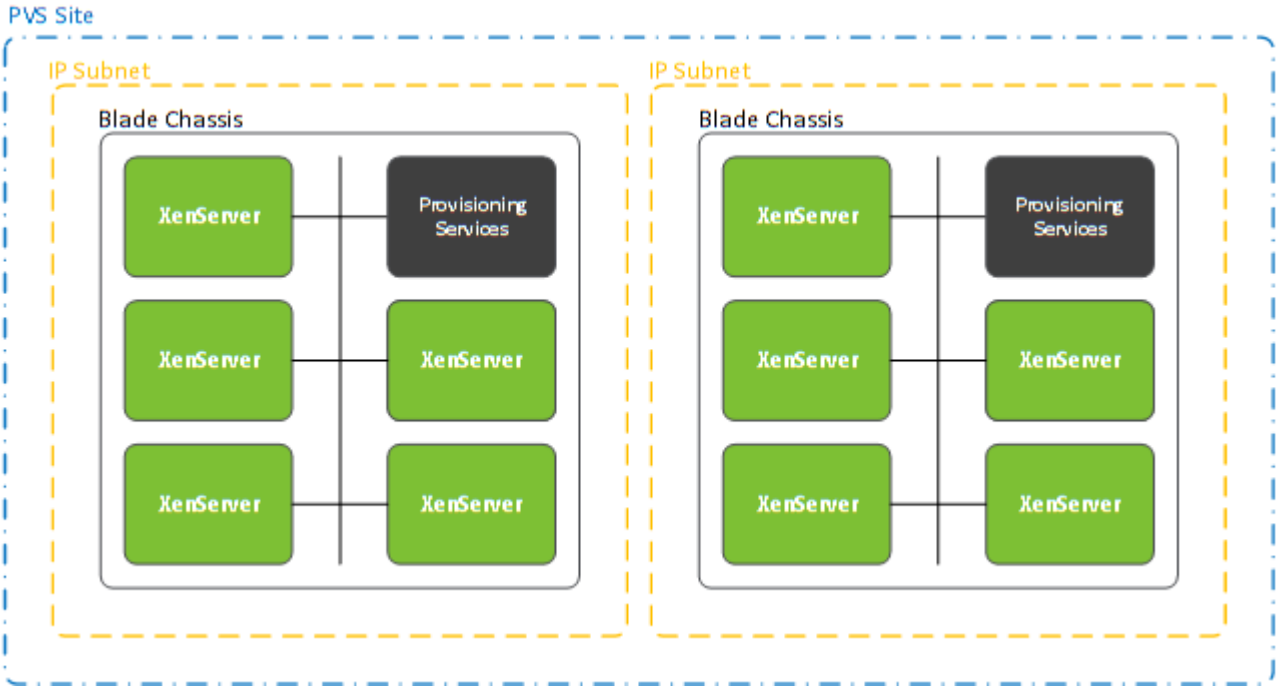


Figure 19: PVS Blade Enclosure Design

The "Best Effort" subnet affinity option is used to keep Provisioning Services traffic within the same chassis. Should the provisioning server become unavailable, the targets will failover to the second provisioning server in the second chassis, but same Provisioning Services site.

Rack Design

The second example is based on a rack design that uses rack switches to keep the provisioning traffic within the rack.

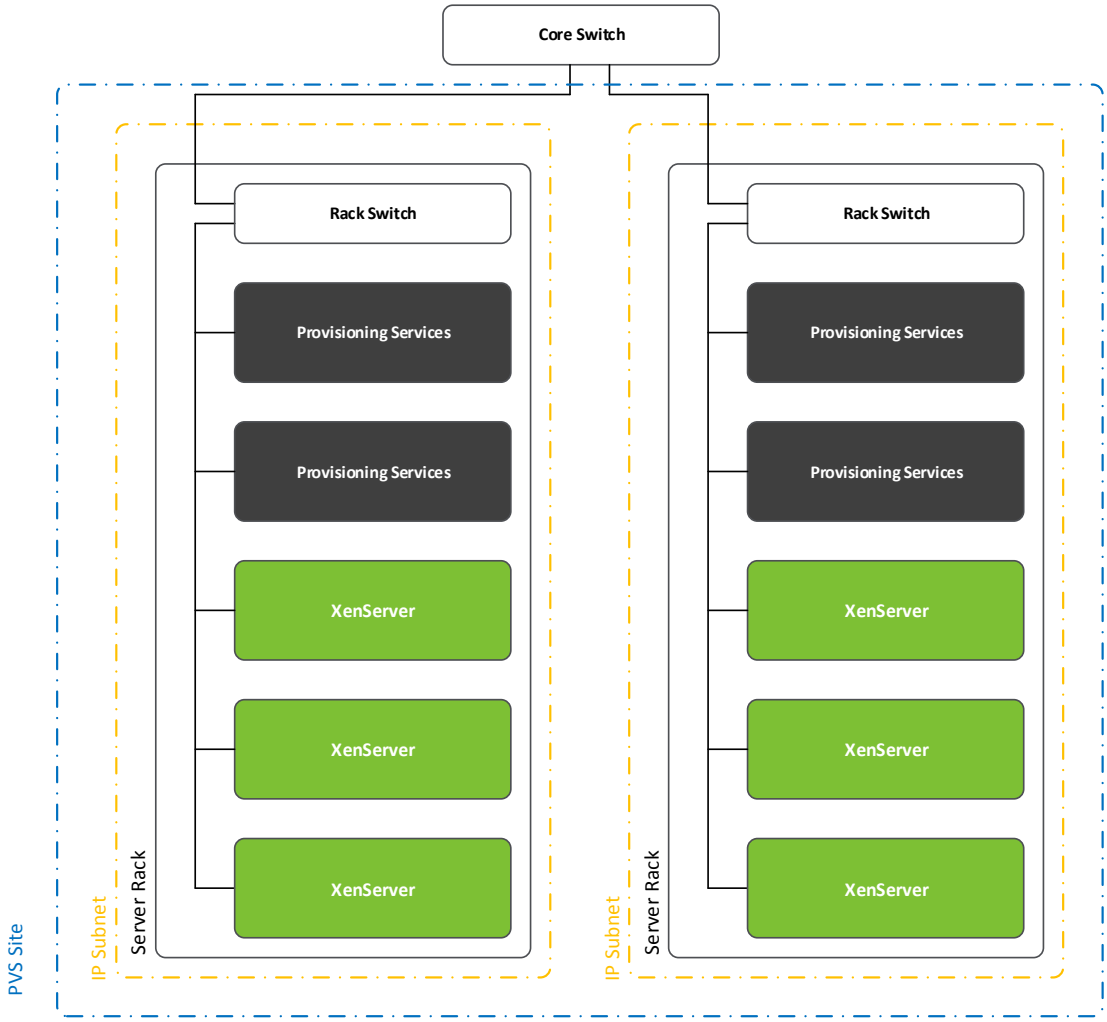


Figure 20: PVS Rack Design

As opposed to the blade chassis design, the subnet affinity feature is not used. Instead a Provisioning Services site with two provisioning servers will be configured per server rack. This will ensure that the target devices are streamed from provisioning servers within the same rack.

Experience from the Field

Manufacturing – A manufacturing company is designing a Provisioning Services solution to support five thousand virtual desktops. The company has concerns that Provisioning Services streaming traffic will create a bottleneck on the network affecting other applications. The company chose to build the environment on blade servers so that provisioning traffic is contained within the blade enclosure and will not impact other traffic on the network.

Decision: Antivirus

By default, most antivirus products scan all files and processes, which has a significant impact on Provisioning Services performance. For details on how antivirus software can be optimized for Provisioning Services, please refer to CTX124185 – Provisioning Services Antivirus Best Practices.

Antivirus software can cause file-locking issues on provisioning servers. The vDisk Store and write cache should be excluded from antivirus scans in order to prevent file contention issues.

When a virtual disk is running in standard mode and needs to be restarted, it downloads all of the previously loaded virus definitions. This can cause performance degradation when restarting several target devices at a time, often causing network congestion while the operation persists. In extreme cases, the target device and provisioning server can become sluggish and consume more resources than necessary. If the antivirus software supports it, definition files should be redirected to the write cache drive so that they are preserved between reboots.

Security

Depending on the requirements of the organization, different security standards should be implemented within the solution. It is advisable refer to the following papers:

- Getting Started Guide for Security
- End-to-End Encryption
- System Hardening Guide
- Common Criteria

Layer 5: The Hardware Layer

Hardware Sizing

This section covers hardware sizing for the virtual infrastructure servers, virtual desktops, and virtual application hosts. The sizing of these servers is typically done in two ways.

- The first and preferred way is to plan ahead and purchase hardware based on the workload requirements.
- The second way is to use existing hardware in the best configuration to support the different workload requirements.

This section will discuss decisions related to both methods.

Decision: Workload Separation

When implementing a XenApp and XenDesktop deployment, the workloads for the infrastructure, XenDesktop, and XenApp workloads can be separated into dedicated resource clusters or mixed on the same physical hosts. Citrix recommends using resource clusters to separate the workloads, especially in an enterprise deployment. This allows better host sizing as each workload has unique requirements such as overcommit ratios and memory usage.

In smaller environments where resource clusters are cost prohibitive, the workloads may be mixed in a manner which still allows for a highly available environment. Citrix leading practice is to separate the workloads however mixed workloads is a cost based business decision.

Decision: Physical Processor (pCPU)

The following table provides guidance on the number of virtual desktops that can be supported for light, medium and heavy workloads per physical core. Each desktop correlates to a single concurrent user.

User Workload	Operating System	Users per Physical Core
Light	Windows 7	13
	Windows 8	12
	Windows 10	12
	Windows 2008R2	18
	Windows 2012R2	21
Medium	Windows 7	10
	Windows 8	9
	Windows 10	9
	Windows 2008R2	12
	Windows 2012R2	14
Heavy	Windows 7	5
	Windows 8	4
	Windows 10	4
	Windows 2008R2	6
	Windows 2012R2	7

Table 44: Processor Requirements by Workload

The estimate for "Users per Physical Core" is a baseline number running Microsoft Office 2010. The baseline number must be adjusted based on specific infrastructure requirements.

Characteristic	Server Density Impact
Antivirus	25% decrease
Monitoring	15% decrease
Office 2013	20% decrease
Office 2016	25% decrease
Hyperthreading	20% increase

Table 45: Server Density Adjustments

To estimate the total number of physical cores required for the XenApp and XenDesktop workload, use the following formula for each user group:

$$\text{Total XenDesktop pCPU} = \sum_i \frac{Users_i}{UsersPerCore_i} * \left(1 + (AV + Mon + Off13 + Off16 - HV)\right)$$

$$\text{Total XenApp pCPU} = \sum_i \frac{Users_i}{UsersPerCore_i} * \left(1 + (AV + Mon + Off13 + Off16 - HV)\right)$$

∑ represents the sum of all user group combinations "i".
$Users_i$ = Number of concurrent users per user groups
$UsersPerCore_i$ = Number of users per physical core
AV = Antivirus impact (default = 0.25)
Mon = Monitoring tools impact (default = 0.15)
Off13 = Office 2013 impact (default = .2)
Off16 = Office 2016 impact (default = .25)
HT = Hyperthreading impact (default = .2)

If workloads will be separated (XenApp and XenDesktop workloads), the formula should be calculated twice, once for all XenDesktop users and the second for all XenApp users in order

Decision: Physical Memory (pRAM)

The recommended method for sizing memory to a physical host is to size based on the total memory required to support the machines and the CPU capacity of the host. In order to calculate the total memory required for XenApp and XenDesktop, simply multiply the number of users in a user group by the amount of memory allocated to the desktop. The sum of all of the user groups will be the total RAM required for XenApp and XenDesktop hosts. This is shown in the formula below.

$$Total\ XenDesktop\ pRAM = \sum_i Users_i * vRAM_i$$

$$Total\ XenApp\ pRAM = \sum_i Users_i * vRAM_i$$

∑ represents the sum of all user group combinations "i".
$Users_i$ = Number of concurrent users per user groups
$vRAM_i$ = Amount of RAM assigned to each virtual machine

If workloads will be separated onto different hosts (XenApp and XenDesktop workloads), the formula should be calculated twice, once for all XenDesktop users and the second for all XenApp users.

Decision: Physical Host (pHost)

In most situations, the number of physical hosts (pHost) to support the XenApp and XenDesktop workloads will be limited on the number of processor cores available.

The following formula provides an estimate for the number of hosts required for the user workloads. The formula is based on the best practice of separating the XenApp and XenDesktop workloads due to the different recommended CPU overcommit ratios for each.

$$XenDesktop\ pHosts\ = \left(\frac{Total\ XenDesktop\ pCPU}{Cores\ per\ pHost} + 1\right)$$

$$XenApp\ pHosts\ = \left(\frac{Total\ XenApp\ pCPU}{Cores\ per\ pHost} + 1\right)$$

Once the number of physical hosts has been determined based on processor cores, the amount of RAM for each host is caulated.

$$XenDesktop\ pRAM\ per\ pHost\ = HypervisorRAM + \left(\frac{Total\ XenDesktop\ pRAM}{XenDesktop\ pHosts - 1}\right)$$

$$XenApp\ pRAM\ per\ pHost\ = HypervisorRAM + \left(\frac{Total\ XenApp\ pRAM}{XenApp\ pHosts - 1}\right)$$

Decision: GPU

Hosts used to deliver graphical workloads require graphics processors to deliver a high end user experience. Specific hardware hosts and graphics cards are required to support high end graphics using HDX 3D Pro. An updated list of tested hardware is available in a knowledge base article. Sizing of the desktop and application hosts of high end graphics users should be based on the GPU requirements ensuring that the host then has adequate CPU and memory resource to support the workload.

NVIDIA GRID cards can be leveraged with vGPU profiles to support multiple users. Sizing guidelines are provided from NVIDIA in the table below.

NVIDIA GRID Graphics Board	Virtual GPU Profile	Application Certifications	Graphics Memory	Max Displays Per User	Max Resolution Per Display	Max Users Per Graphics Board	Use Case
GRID K2	K260Q	✔	2,048 MB	4	2560x1600	4	Designer/Power User
	K240Q	✔	1,024 MB	2	2560x1600	8	Designer/Power User
	K220Q	✔	512 MB	2	2560x1600	16	Designer/Power User
	K200		256 MB	2	1900x1200	16	Knowledge Worker
GRID K1	K140Q	✔	1,024 MB	2	2560x1600	16	Power User
	K120Q	✔	512 MB	2	2560x1600	32	Power User
	K100		256 MB	2	1900x1200	32	Knowledge Worker

Storage Sizing

Decision: Storage Architecture

The primary storage architectures are as follows:

- **Local Storage -** Uses hard disks directly attached to the computer system. The disks cannot be shared with other computer systems, but if the computer is hosting pooled or hosted shared desktops, a shared storage solution is not necessary. In many cases local storage can perform as well as shared storage. Scalability is limited to the number of drive bays available in the computer system. Many blade servers for example have just two drive bays, so using local storage to support a XenDesktop deployment may not be optimal.
- **DAS -** Storage sub-system directly attached to a server or workstation using a cable. It uses block-level storage and can be a hard disk local to the computer system or a disk shelf with multiple disks attached by means of external cabling. Unlike local disks, disk shelves require separate management. Storage shelves can be connected to multiple servers so the data or disks can be shared.
- **NAS -** Provides file-level storage to computer systems through network file shares. The NAS operates as a file server, and NAS systems are networked appliances which contain one or more hard drives, often arranged into logical, redundant storage containers or RAID arrays. Access is typically provided using standard Ethernet and network file sharing protocols such as NFS, SMB/CIFS, or AFP.

  Note: *NAS can become a single point of failure. If the network share becomes unavailable, all target devices streamed from the disk will be unavailable as well.*
- **SAN -** Dedicated storage network that provides access to consolidated, block-level storage. SANs allow computers to connect to different storage devices, so no server has ownership of the storage subsystem enabling data to be shared among multiple computers. A SAN will typically have its own dedicated network of storage devices that are generally not accessible through the network by standard means. In order to connect a device to the SAN network a specialized adapter called the Host Bus Adapter (HBA) is required. SANs are highly scalable with no noticeable change in performance as more storage and devices are connected. SANs can be a costly investment both in terms of capital and the time required to learn, deploy and manage the technology.
- **Hybrid -** A NAS head refers to a NAS which does not have any on-board storage, but instead connects to a SAN. In effect, it acts as a translator between the file-level NAS protocols (NFS, CIFS, etc.) and the block-level SAN protocols (Fibre Channel and iSCSI). Thus it can combine the advantages of both technologies and allows computers without Host Bus Adapters (HBA) to connect to centralized storage.

The following table summarizes the storage options available and rates their suitability for XenDesktop deployments.

Storage Properties	Local	DAS	NAS	SAN
Implementation costs	Low	Medium	Medium	High
Administration	Low	Medium	Medium	High
Performance	High[1]	Med - High	Med - High	High
Redundancy	Low - Med	Medium	Med – High	High

Scalability	Low	Low - Med	Med - High	High
Typical use case	Small to medium production and test environments	Small to medium production environments.	Small to medium production environments.	Medium to large production environments.

Table 46: Storage feature comparison

Note: *Hyper-V 2008 R2 does not support NAS technology. Hyper-V 2012/2012 R2 only supports NAS solutions that support the SMB 3.0 protocol. For more information please refer to the Hyper-V 2008 R2 and Hyper-V 2012 R2 sections of the handbook.*

Local storage is best suited for storing virtual machines which do not have high availability requirements or persistent data attached such as random (pooled) desktops or hosted shared desktops. Local and DAS is suited for storing user data and home directory files. If using Machine Creation Services, master images as well as any updates must be replicated to each server.

NAS and SAN storage is best suited for infrastructure servers supporting the XenDesktop environment, and virtual machines with persistent data such as static (dedicated) desktops, and random (pooled) desktops with Personal vDisks.

Decision: RAID Level

To choose the optimal RAID level, it is necessary to consider the IOPS and read/write ratio generated by a given application or workload in combination with the individual capabilities of a RAID level. For hosting read intensive workloads, such as the Provisioning Services vDisk store, RAID levels that are optimized for read operations such as RAID 1, 5, 6, 10 are optimal. This is because these RAID levels allow read operations to be spread across all disks within the RAID set simultaneously.

For hosting write intensive workloads, such as Provisioning Services write cache and Machine Creation Services differencing disks, RAID levels such as RAID 1 or 10 are optimal, as these are optimized for writes and have a low write penalty.

The following table outlines the key quantitative attributes of the most commonly used RAID levels:

RAID	Capacity	Fault Tolerance	Read Performance	Write Performance	Minimum # of Disks
0	100%	None	Very High	High (Write Penalty 1)	2
1	50%	Single-drive failure	Very High	Medium (Write Penalty 2)	2
5	67%-94%	Single-drive failure	High	Low (Write Penalty 4)	3
6	50%-88%	Dual-drive failure	High	Low (Write Penalty 6)	4
10	50%	Single-drive failure in each sub array	Very High	Medium (Write Penalty 2)	4

Table 47: RAID levels

Note*: The write penalty is inherent in RAID data protection techniques, which require multiple disk I/O requests for each application write request, and ranges from minimal (mirrored arrays) to substantial (RAID levels 5 and 6).*

Decision: Number of Disks

To determine the number of disks required it is important to understand the performance characteristics of each disk, the characteristics of the RAID level and the performance requirements of the given workload. The basic calculation for determining the total number of disks needed is:

$$Total\ \#\ of\ Disks = \frac{\left(Total\ Read\ IOPS + (Total\ Write\ IOPS * RAID\ Penalty)\right)}{Disk\ Speed\ IOPS}$$

For example, a disk manufacturer is reporting that a particular disk array which they have developed has a total workload IOPS of 2000. The raw IOPS per disk is 175. To determine how many disks are required to support a workload with 20% read operations and 80% write operations on RAID 10:

$$Total\ \#\ of\ Disks = \frac{\left((20\% * 2000) + (80\% * 2000) * 2\right)}{175} = 20.57\ or\ 21\ Disks$$

Based on the previous example, the following table shows how the disk count will vary based on the RAID level and the read/write ratio.

RAID	RAW IOPS (per disk)	Workload IOPS	Read %	Write %	Disk count
0	175	2000	20%	80%	12
	175	2000	80%	20%	12
1 / 10	175	2000	20%	80%	21
	175	2000	80%	20%	14
5	175	2000	20%	80%	39
	175	2000	80%	20%	19

Table 48: Example of how disk count changes per RAID level and R/W ratio

Decision: Disk Type

Hard disk drives (HDDs) are the traditional variation of disk drives. These kinds of disks consist of rotating platters on a motor-driven spindle within a protective enclosure. The data is magnetically written to and read from the platter by read/write heads.

Different implementations of this technology are available on the market, which differ in terms of performance, cost and reliability.

- Serial ATA (SATA) disk transmit data serially over two pairs of conductors. One pair is for differential transmission of data, and the other pair is for differential receiving of data. SATA drives are widely found in consumer desktop and laptop computers. Typical SATA drives have transfer speeds ranging from 1500 – 6000Mbps and support hot-swapping by design.
- Small Computer Systems Interface (SCSI) disks use a buffered, peer to peer interface that uses handshake signals between devices. Many SCSI devices require a SCSI initiator to initiate SCSI transactions between the host and SCSI target. SCSI disks are common in workstations and servers and have throughputs ranging from 40 – 5120Mbps. iSCSI (Internet Small Computer System Interface) is a mapping of the regular SCSI protocol over TCP/IP, more commonly over Gigabit Ethernet.
- Fibre Channel (FC) disk is the successor to the parallel SCSI disk and is common in SAN storage devices. Fibre Channel signals can run on an electrical interface or fibre-optic cables. Throughput can range from 1 – 20Gbps, and connections are hot-pluggable.

- Serial Attached SCSI (SAS) disk uses a new generation serial communication protocol to allow for higher speed data transfers than SATA disks. Throughput can range from 2400 – 9600Mbps.

In contrast to traditional hard disks, Solid State Disks (SSDs) use microchips to retain data in either NAND non-volatile memory chips (flash) or DRAM and contain no moving parts. SSDs are less susceptible to physical shock, have lower access times and latency and have higher I/O rates. SSDs have significantly higher random read performance. An SSD drive can attain anywhere from 5,000 to 20,000 random reads per second. SSDs are also more expensive per gigabyte (GB) and typically support a limited number of writes over the life of the disk.

Flash memory-based SSDs can be either based on multi-level cells (MLC) or single-level cells (SLC). SLC devices only store one bit of information in each cell. MLC devices can store multiple bits of information with each cell. Flash based SSDs cost lower than DRAM based SSDs but perform slower. DRAM based SSD devices are used primarily to accelerate applications that would otherwise be held back by the latency of flash SSDs or traditional HDDs.

SSDs were previously not viable for enterprise storage solutions because of the high cost, low capacity and fast wear of the drives. Improvements in SSD technology and lowering costs are making them more favorable over HDDs. Solid state hybrid drives (SSHD) combine the features of SSDs and HDDs, by containing a large HDD drive with an SSD cache to improve performance of frequently accessed data.

Comparing SSDs and HDDs is difficult since HDD benchmarks are focused on finding the performance aspects such as rotational latency time and seek time. As SSDs do not spin, or seek, they may show huge superiority in such tests. However SSDs have challenges with mixed reads and writes and their performance may degrade over time.

The following table compares the transfer rates of some of the more common storage types available on the market today.

Technology	Rate (bit/s)
iSCSI over Fast Ethernet	100Mbps
Ultra-2 wide SCSI (16 bits/40 MHz)	640Mbps
iSCSI over Gigabit Ethernet	1,000Mbps
SATA rev 3	6,000Mbps
SAS 3	9,600Mbps
FCoE over 10GbE	10,000Mbps
SATA rev 3.2 – SATA Express	16,000Mbps
iSCSI over Infiniband	32,000Mbps

Table 49: Common disk types and transfer rates

SCSI and SATA disks are best suited for storing data that does not have high performance requirements like the PVS vDisk store. SAS, Fibre Channel, or SSD drives are best suited for storing data that have high performance requirements like the PVS write cache.

Decision: Storage Bandwidth

Storage bandwidth is the connectivity between servers and the storage subsystem. Understanding bandwidth requirements can help determine the proper hardware for delivering data and applications at speeds for a positive end user experience. For most datacenters 10Gbps Ethernet or 10Gbps FCoE is sufficient for storage connections. Smaller environments however may only need 1Gbps bandwidth. In

virtualized environments it is not just important to look at the bandwidth requirements of the physical host and storage subsystem, but determining how much bandwidth is required for each virtual machine plays a factor too.

In order to plan for the required bandwidth, it is necessary to determine the throughputs for every individual system that uses a shared component or network path. For example, the following information is provided for an environment with 100 similar virtual machines (hosted on 10 virtualization hosts and connected to one NAS head).

	Average	Peak
Throughput per VM	10Mbps	30Mbps
Throughput per host	100Mbps (10 VMs * 10Mbps)	300Mbps (10 VMs * 30Mbps)
Throughput per storage	1Gbps (10 hosts * 100Mbps)	3Gbps (10 hosts * 300Mbps)

Table 50: Throughput

The NIC used for storage communication needs to be a 1Gbps adapter in order to handle the peak load. The NAS head as well as its network connection need to support 3Gbps worth of data traffic in order to support the peak load of all systems.

Decision: Tiered Storage

A one-size-fits-all storage solution is unlikely to meet the requirements of most virtual desktop implementations. The use of storage tiers provides an effective mechanism for offering a range of different storage options differentiated by performance, scalability, redundancy and cost. In this way, different virtual workloads with similar storage requirements can be grouped together and a similar cost model applied.

For example, a XenDesktop implementation using tiered storage may look like the following:

- Tier 1 storage group - Write intensive files such as the write cache and differencing disks are placed in a storage group consisting of SSDs.
- Tier 2 storage group - Mission critical data, or data that requires high availability such as Personal vDisks, are placed in a storage group consisting of less expensive high performing drives.
- Tier 3 storage group - Seldom used data files, read-only files, or other non-mission critical data placed in a storage group consisting of low cost and lower performing drives.

Decision: Thin Provisioning

Thin provisioning allows more storage space to be presented to the virtual machines than is actually available on the storage repository. This lowers storage costs by allowing virtual machines access to disk space that is often unused. This is particularly beneficial to Machine Creation Services which uses a linked-clone approach to provisioning virtual machines. Thin provisioning minimizes the storage space required for the master image copies used to build virtual machines. Thin provisioning is possible at the physical storage layer, a feature usually available with most SAN solutions, and at the virtual layer. NFS based storage solutions will usually have thin provisioning enabled by default.

At the physical storage layer it is important to ensure that sufficient storage is available to prevent the risk of virtual machines not being available in a storage "overcommit" scenario when available disk space is exhausted. Organizations should decide if the cost savings thin provisioning provides outweighs the associated risk and consider enabling if the storage solution supports it.

Note: *Virtual machines may not function if disk space is exhausted so it is important to have a process in place, either through alerts or notifications that will give administrators enough time to add more disks to the storage solution so that the XenDesktop environment is not impacted.*

Decision: Data De-Duplication

Data de-duplication is a data compression technique whereby duplicate data is replaced with pointers to a single copy of the original item. This reduces storage requirements and costs by improving storage utilization, however it can impact storage performance. There are two implementations of de-duplication available:

- Post-process de-duplication – The de-duplication is performed after the data has been written to disk. Post-process de-duplication should be scheduled outside business hours to ensure that it does not impact system performance. Post Process de-duplication offers minimal advantages for random desktops as the write-cache/difference disk is typically reset on a daily basis.
- In-line de-duplication – Examines data before it is written to disk so that duplicate blocks are not stored. The additional checks performed before the data is written to disk can sometimes cause slow performance. If enabled, in-line duplication should be carefully monitored to ensure that it is not affecting the performance of the XenDesktop environment.

If the storage solution supports it, enabling post-process data de-duplication is recommended for minimal impact to XenDesktop performance.

Section 4: Monitor

Overview

Like any integrated system, the overall health status of the solution must be monitored and maintained. Without proper support, operations and health monitoring systems in place, the user experience will slowly start to degrade.

Process 1: Support

When problems arise, technical support is the first point of contact. This section addresses the proper staffing, organization, training, delegated administration and tools that should be used to maintain the Citrix deployment.

Decision: Support Structure

Multiple levels of support has been found to be the most effective ways of addressing support issues. Low criticality, low complexity or frequently occurring issues should be managed and resolved at the lower support levels. High criticality and complex issues are escalated to more experienced architects or infrastructure owners. The diagram below outlines a common multi-level support structure.

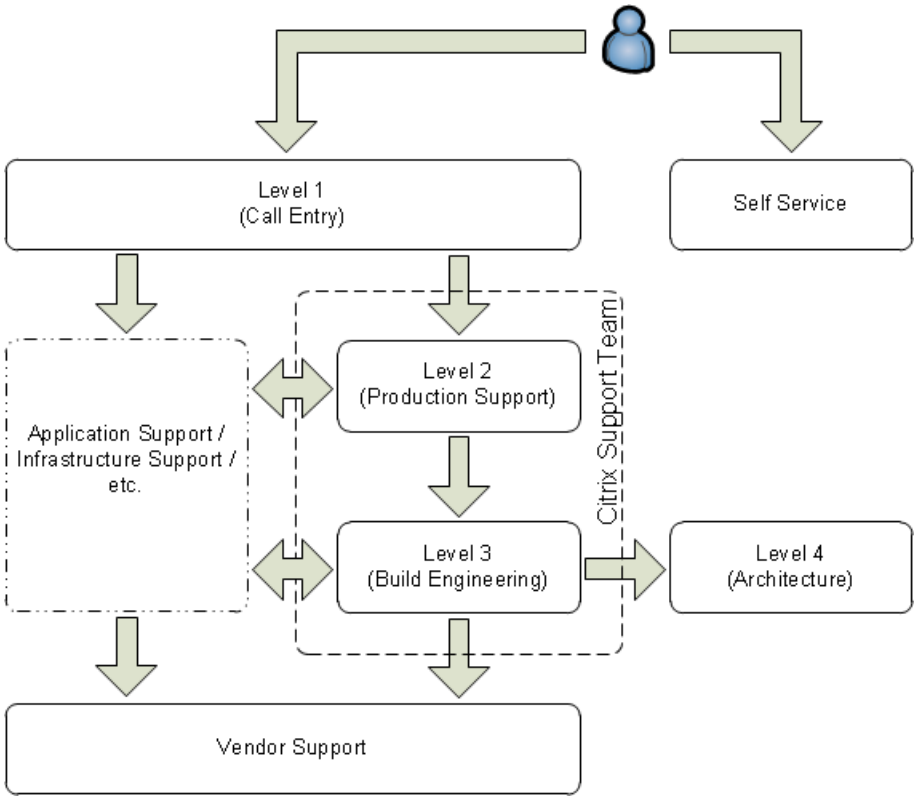


Figure 21: Support Structure

If a user encounters an issue, Level-1 support (help desk) is the entry point to the support system. Level-1 should resolve 75% of all issues encountered, of which a majority will be routine problems that only require a limited knowledge of the Citrix environment. At this level, issues are quickly resolved and some may be automated (self-service), for example password resets and resource provisioning.

Non-routine problems that exceed Level-1's abilities are escalated to Level-2 (Operators). This support level is generally comprised of administrators supporting the production Citrix environment. Information

on the end user's problem and attempted troubleshooting steps are documented at the first level allowing Level-2 technicians to immediately begin addressing the problem. Level-2 technicians should handle only 20% of the support tickets and are highly knowledgeable on the Citrix environment.

Complex issues that exceed Level-2's abilities should be escalated to Level-3 (Implementers). Level-2 and Level-3 support may often both be members of the Citrix Support Team, with Level-3 comprising the senior staff maintaining the Citrix environment. Level-3 issues are complicated and often mission critical requiring expert knowledge of the virtual desktop and application environment. Level-3 support tickets should amount to no more than 5% of all support issues.

The final level, Level-4 (Architects), is focused on the strategic improvements for the solution, testing new technologies, planning migrations, and other high level changes. Generally, Level-4 is not involved in active support of a production environment.

Should support discover an issue that is related to an application or underlying infrastructure, the ticket is handed to the appropriate team for troubleshooting. If a program bug is discovered, the issue is then re-escalated and a ticket is established with the appropriate vendor.

Decision: Support Responsibilities and Skill Set

The following table highlights the recommended characteristics of each support level.

Support Level	Description	Responsibilities	Skill Set
Level-1 (Help Desk)	Provide first-line support of reported issues. Initially, servicing support messages and phone calls. This level needs to perform initial issue analysis, problem definition, ticket routing, and simple issue resolution. Often processes requests for application access or support with configuring plugins.	• Perform issue definition, initial analysis and basic issue resolution • Perform initial troubleshooting to determine the nature of the issue • Create ticket, collect user information, and log all troubleshooting steps performed • Resolve basic Citrix related issues, connectivity problems and application related issues using existing knowledge base articles • Escalate issue to Level-2 if advanced skills or elevated permissions are required • Ability to isolatethe issue to be Citrix related, Microsoft related or third party Application related • If it affects the production environment or is potentially causing a system level ourtage, escalate directly to Level-3 • Generate requests for additional issue resolution guides as necessary • Follow up with end users when a support ticket is closed to ensure the problem has been satisfactorily resolved	• General Citrix XenApp/XenDesktop knowledge (CCA, CCA-V) • General Windows client OS/server OS knowledge (MCP) • General Active Directory knowledge • General Networking knowledge (CCNA)

Level-2
(Operators)
	Primarily supporting day-to-day operations of the Citrix environment; may include proactive monitoring and management. In addition, this role should also perform intermediate level troubleshooting and utilize available monitoring or troubleshooting tools.
	Assist with resolving issues escalated by Level-1 support.
		• Perform intermediate issue analysis and resolution.
		• Identify root cause of issues.
		• Respond to server alerts and system outages.
		• Create weekly report on number of issues, close rate, open issues, etc.
		• Review vendor knowledge base articles.
		• Respond to out-of-hours helpdesk calls.
		• Respond to critical monitoring alerts.
		• Generate internal knowledge base articles and issue resolution scripts and maintain Level-1 troubleshooting workflows.
		• Perform basic server maintenance and operational procedures.
		• Manage user profiles and data.
		• Escalate ticket to Level-3 or appropriate technology owner if advanced skills or elevated permissions are required.
		• Generate requests for additional issue resolution scripts and knowledge base articles as necessary.
		• Able to read built-in event logs for Windows and Citrix to do basic troubleshooting following public information via Google/Bing.
			• Experience with Microsoft Windows Server including but not limited to:
			  ○ Configuring operating system options
			  ○ Understanding Remote Desktop Services policies and profiles
			  ○ Using Active Directory
			  ○ Creating users/managing permissions and administrator rights
			  ○ Creating and modifying Active Directory group policies
			• Basic administration skills, including:
			  ○ An understanding of protocols (TCP)
			  ○ An understanding of firewall concepts
			  ○ An understanding of email administration and account creation
			  ○ An understanding of Remote Desktop Services policies and profiles
			  ○ The ability to create shares and give access to shared folders/files
			• Experience performing the following:
			  ○ Managing, maintaining, monitoring and troubleshooting Citrix solutions
			  ○ Backing up components in Citrix environments
			  ○ Updating components in Citrix environments
			  ○ Creating reports for trend analysis

Level-3
(Implementer)
	Central point for implementing, administering and maintaining Citrix desktop and application virtualization infrastructure. This role focuses on deploying new use cases and leading lifecycle management initiatives. Generally, one Implementer could focus on one use-case at a time. For example, three new concurrent use cases would require three Implementers. Escalates issues to software vendor specific technical support and notifies Level-4 about this issue.	• Perform advanced issue analysis and resolution.
		• Perform maintenance and environment upgrades.
		• Addresses high severity issues and service outages.
		• Manage the Citrix environment.
		• Oversee and lead administrative tasks performed by Level-2.
		• Manage network and storage infrastructure as it relates to the Citrix environment (depending on size of company or Citrix environment).
		• Review periodic reports of server health, resource usage, user experience, and overall environment performance.
		• Review vendor knowledge base articles and newly released updates.
		• Perform policy-level changes and make Active Directory updates.
		• Review change control requests that impact the Citrix environment.
		• Perform advanced server and infrastructure maintenance.
		• Review knowledge base articles and issue resolution scripts for accuracy, compliance, and feasibility.
		• Create knowledge base articles and issue resolution scripts to address Level-2 requests.
		• Escalate ticket to vendor specific technical support, when necessary, and notify Level-4 of the issue.
			• Knowledge of how the following Windows components integrate with Citrix technologies:
			  ○ Active Directory Domain Services
			  ○ Active Directory Certificate Services
			  ○ Policies
			  ○ Domain Name System (DNS)
			  ○ Dynamic Host Configuration Protocol (DHCP)
			  ○ Group Policy Objects (GPOs)
			  ○ NTFS Permissions
			  ○ Authentication and Authorization
			  ○ Knowledge of IIS
			  ○ Microsoft Windows Operating Systems
			  ▪ Windows 10
			  ▪ Windows 8.1
			  ▪ Windows 7
			  ▪ Windows Server 2012 R2
			  ▪ Windows Server 2008 R2
			• Roles and features of Windows Server
			• Knowledge of SQL 2008 R2 and newer
			• Knowledge of SQL Clustering, Mirroring and AlwaysOn Availability Groups.
			• General networking skills (i.e. routing, switching)
			• Knowledge of hypervisors.
			• Knowledge of shared storage configuration and management.

Level-4 (Architect)	The Level-4 team has minimal exposure to administrative tasks but focuses on scoping, planning and executing Citrix-specific service and project requests. An architect translates business requirements into a technical design.	• Provide technical leadership for upcoming projects.
		• Lead design updates and architecture revisions.
		• Address high severity issues and service outages
		• Oversee technology integration workflows.
		• Review periodic reports of server health, resource usage, user experience, and overall environment performance to determine next steps and upgrade paths.
		• Initiate load testing to determine capacity of environment.
		• Review frequently recurring helpdesk issues
		• Ensure technical specifications continue to meet business needs.
		• Update design documentation.
			• Advanced architectural assessment and design skills for:
			  ○ Citrix XenApp
			  ○ Citrix XenDesktop
			  ○ Citrix XenServer / VMware vSphere / Microsoft Hyper-V
			  ○ Citrix Provisioning Services
			  ○ Citrix NetScaler
			  ○ Citrix StoreFront
			  ○ Active Directory
			  ○ Storage solutions
			  ○ Networking
			  ○ Application delivery
			  ○ Disaster recovery
			  ○ Policies/policy structures and security restrictions
			  ○ Licensing
			  ○ Methodology
			• Intermediate knowledge of:
			  ○ General networking skills
			  ○ Change control process
			  ○ Project management
			  ○ Risk assessment

Vendor Support	Vendor assistance may be necessary should defects in a program be discovered. At this stage, Level-3 engineers need to establish a support ticket with the appropriate vendor to assist with finding a solution.		
Self-Service	A self-service portal should be utilized for noncritical tasks such as application access, permissions, password resets, etc. The portal can range from a simple FAQ page to a fully automated process requiring no human interaction. The purpose of the self-service portal is to add an additional touch point for end users to address basic issues, preventing the creation of new support tickets.		

Table 51: Support Levels

Decision: Certifications and Training

The following table details the recommended training, certifications and experience for each support level.

Role	Recommended Training	Recommended Course(s)	Recommended Certification	Relevant Experience

Help Desk (Level 1)	Level-1 support personnel should be provided with basic training on Citrix XenApp, Citrix XenDesktop and supporting technologies. This can include internal training from subject matter experts or from a Citrix Authorized Learning Center. The training provided should focus on the following topics: • High level overview of the XenApp and XenDesktop implementation • Using Citrix Director to manage user sessions • Troubleshooting Citrix XenApp and XenDesktop sessions • Troubleshooting methodology In addition, regular training should be provided to the Tier-1 team members on the latest troubleshooting recommendations from the Level-2 and Level-3 teams as well as details on any relevant changes to the environment. This will help to ensure a good baseline knowledge level across the team and consistent customer service.	CXD-105: Citrix XenApp and XenDesktop Help Desk Support	N/A	1+ years (Entry level also acceptable)
Operator (Level-2)	Level-2 personnel should conduct regular team training sessions to refine administrative skills and ensure a baseline knowledge level across the team. Formalized trainings are also essential when there are architectural updates to the environment and the Level-2 team is working with unfamiliar technologies. All members of the Level-2 team should achieve the Citrix Certified Associate (CCA) certification for Citrix XenApp and XenDesktop. Advanced training on Windows concepts will also be essential for Level-2 team members who do not have desktop or server support experience. Finally, on-the-job training along with close integration with Level-3 administrators is essential as the Level-2 roles are formalized and responsibilities are handed over from Level-3 to Level-2.	CXD-203: Managing App and Desktop Solutions with Citrix XenApp and XenDesktop 7.6	Citrix Certified Associate - Virtualization	2-3 years
Implementer (Level-3)	Level-3 support team members hold a minimum of three years of enterprise experience implementing and supporting XenApp, XenDesktop, Provisioning Services and Windows operating systems. Level-3 staff should also complete the Citrix Certified Professional (CCP) certification track as this will prepare them to proactively manage the user community and implement Citrix solutions according to Citrix leading practices.	CXD-300: Deploying App and Desktop Solutions with Citrix XenApp and XenDesktop 7.6	Citrix Certified Professional - Virtualization	3-4 years

Architect (Level 4)	Experience is essential for Level-4 staff. A qualified Level-4 resource should have a minimum of five of experience implementing, supporting, and serving in a technology architect role for a XenApp and/or XenDesktop environment as well as additional administrative experience with integrated technologies such as application and profile management solutions. The ideal candidate will have served in such a capacity at two or more environments for purposes of product exposure and in at least one environment of over 1,200 concurrent users. A Citrix Certified Expert (CCE) certification or comparable training and experience should be a prerequisite of the role.	CXD-400: Designing App and Desktop Solutions with Citrix XenApp and XenDesktop	Citrix Certified Expert - Virtualization	5+ years

Table 52: Training Recommendations

Decision: Support Staffing

The following table provides guidance on the recommended number of support staff.

Role	**Small Environment** Sites: 1 Users: <500 Images: 1-2	**Mid-size Environment** Sites: 1-2 Users: 1000-5000 Images: 3-5	**Large Environment** Sites: 2+ Users: >5000 Images: 5+
Help Desk (Level-1)	3	5-10	15-20
Operator (Level-2)	1-2	2-3	4-5
Implementer (Level-3)	1	1-2	2-3
Architect (Level-4)	1	1	1-2

***Note**: This table should only be used as a baseline. Support staffing decisions should be evaluated against the defined requirements, projected workloads, and operational procedures of an organization. Multiple levels can be combined, for example there may be insufficient design projects to have a dedicated architect role or a more senior member of the Citrix team can act as an Operator and Implementer.*

Decision: Job Aids

General Support Tools

The following table details tools that should be made available to all support levels.

Tools	Details
Ticket Management System	Used to document customer information and issues. A typical ticket management system provides the following functionality: • Monitoring the queue of tickets. • Setting a limit on the number of open tickets. • Establishing thresholds such as how long a certain type of ticket should take to be answered. • Identifying a group of users or individuals who require higher priority assistance. • Informing the user when their ticket is open, updated, or closed. • Provide an internal knowledge base for the support professionals to search for known resolved issues.
Call Scripts	The first contact help desk personnel should have documented scripts to ensure that all relevant data is captured while the user is on the phone. This practice also assists in proper triage and allows the next support level to perform research prior to customer contact. A sample call script is provided for reference.
Remote Assistance Tools	Remote assistance tools are useful when troubleshooting user issues. Support technicians and administrators can remotely observe a user's actions.
Knowledge Base	Documentation should be created and maintained in a knowledge base or library of known issues. Articles should be searchable for quick recovery. Knowledge bases help support staff to quickly resolve known issues and reduce the need to perform time consuming research.

Table 53: Recommended General Helpdesk Tools

Citrix Support Tools

The following table provides recommendations on the Citrix support tools that should be made available to each support level.

Tool	Description	Products				Support Level			
		XD	XA	PVS	XS	L1	L2	L3	L4
Citrix Director	Citrix Director provides an overview of hosted desktops and application sessions. It enables support teams to perform basic maintenance tasks and to monitor and troubleshoot system issues.	X	X			X	X	X	X
Citrix Studio	Citrix Studio enables administrators to perform configuration as well as maintenance tasks for a XenApp/XenDesktop site and associated virtual desktops or hosted applications.	X	X			X	X	X	X
Citrix Insight Services	Run from a single Citrix Delivery Controller to capture key data points and CDF traces for selected computers followed by a secure and reliable upload of the data package to Citrix Technical Support for escalation.	X	X	X	X			X	X
HDX Monitor	HDX Monitor is a tool to validate the operation of the Citrix ICA/HDX stack of a user session. HDX Monitor provides information about client capabilities, network performance / activity, session settings and many more items.	X	X			X	X	X	X
Provisioning Services Console	The Provisioning Services Console enables administrators to perform configuration and maintenance tasks for a Provisioning Services farm.			X				X	X
XenCenter	XenCenter enables administrators to perform configuration and maintenance tasks for a XenServer Resource Pool.				X			X	X

Table 54: Support Citrix Tool Assignment

Citrix Insight Services

Administrators can utilize Citrix Insight Services to simplify the support and troubleshooting of the Citrix environment. Citrix Insight Services is run locally to collect environment information. Online analysis capabilities analyze that information and provide administrators recommendations based on their Citrix environment and configuration. Additional information regarding Citrix Insight Services can be referenced in the Citrix Support article CTX131233 - FAQ: Citrix Insight Services.

A full list of the available tools provided by Citrix Support to assist with troubleshooting can be referenced in Citrix Supportability Pack.

Call Script

The following call script can be used as an initial baseline for a Citrix Help Desk team. Citrix Consulting recommends reviewing this sample call guide and adding any environment specific details that may need to be collected.

Step	Details

1.	**What is the name and location of the user?** This question will identify if the user is accessing the environment from an external or internal network location.
2.	**Is the problem always reproducible? If it is a Yes, get the exact reproduce steps.** This question is very important for the support team to troubleshooting an issue.
3.	**Do any other users at the site/location experience the same issue? Can they have a colleague logon from same and/or different workstation?** These questions help to determine if this is a workstation issue or a user issue.
4.	**What type of endpoint device is the user utilizing? (Corporate device, BYOD, thin client, pc, laptop, etc.)** This question will help determine if the issue is related to the user's endpoint.
5.	**What is the Citrix Receiver version and connection information?** This question will verify if the user is using the right version of Receiver (the latest Receiver version or the version standardized by the company).
6.	**Can the user see the StoreFront authentication page?** This question helps to identify network issues.
7.	**What is the name of the application (or virtual desktop) the user is attempting to use? Does the user see the appropriate application or desktop icon on the StoreFront site?** These questions help to determine if there is an issue with user access and/or group membership.
8.	**Does the application (or desktop) launch when the icon is selected? Does the application logon screen appear (if applicable)?** These questions help to determine if a connection is made into the Citrix XenDesktop infrastructure.
9.	**Can the user authenticate into the application (if applicable)? Does the issue occur after the application is launched?** This question helps to determine if the issue is with the application rather than the application delivery infrastructure.
10.	**What is the specific error seen (if applicable)? Get screen captures** This question identifies the specific error. The user should be requested to provide a screenshot, if available.

Decision: Delegated Administration

Each support level must be provided with sufficient rights to effectively perform their role. The following tables provide guidance on the recommended privileges per support level.

XenApp/XenDesktop Delegated Rights

Admin Role	Support Level
Help Desk Administrator	Level-1
Full Administrator	Level-2
Full Administrator	Level-3
Full Administrator	Level-4

Table 55: XenApp/XenDesktop Delegated Rights

For further information about delegated rights within a XenApp/XenDesktop Site, please refer to Citrix Product Documentation - XenApp and XenDesktop Delegated Administration.

Provisioning Services Delegated Rights

Admin Role	Support Level
N/A	Level-1
Site Administrator	Level-2
Farm Administrator	Level-3
Full Administrator	Level-4

Table 56: Provisioning Services Delegated Rights

For further information about delegated rights within a Provisioning Services Site, please refer to Citrix eDocs - Provisioning Services Managing Administrative Roles.

StoreFront Delegated Rights

Admin Role	Support Level
N/A	Level-1
N/A	Level-2
Local Administrator on SF server	Level-3
Full Administrator	Level-4

Table 57: StoreFront Delegated Rights

Users with local administrator rights have access to view and manage all objects within StoreFront or Web Interface. These users can create new sites and modify existing ones.

Citrix License Server Delegated Rights

Admin Role	Support Level
N/A	Level-1
N/A	Level-2
Administrator	Level-3
Administrator	Level-4

Table 58: Citrix License Server Delegated Rights

By default, the account used during the installation of the license server becomes the administrator for the console. Often the accounts used for the installation are not the intended accounts for the regular administration tasks. For the steps on how to change the default administrator, please reference CTX135841 - How to Change the Default Administrator for the Citrix Licensing Server Version 11.10. All users created through this process are full administrators of the Citrix License Server.

XenServer Delegated Rights

Admin Role	Support Level
N/A	Level-1
Virtual Machine Operator	Level-2
Pool Administrator	Level-3
Full Administrator	Level-4

Table 59: XenServer Delegated Rights

For further information about delegated rights within a XenServer Resource Pool, please refer to XenServer 7.0 Administrators Guide (see chapter Role Based Access Control).

Process 2: Operations

This section defines routine operations for the Citrix environment that help to improve stability and performance.

Decision: Administrative Tasks

The Citrix Support Team should perform regular operations and maintenance tasks to ensure a stable, scalable Citrix environment.

Each operation is categorized by the associated component of the solution as well as the frequency of the operation (ongoing, daily, weekly and yearly). Tasks have been aligned to the roles described within Decision: Support Responsibilities and Skill Set.

If the administrators performaning operations are the same the support team, then the designations are linked as follows:

- Level 2 Support = Operators
- Level 3 Support = Implementers

Daily Periodic Tasks

The following table outlines the tasks that should be performed by the Citrix Support Team on a daily basis.

Component	Task	Description	Responsible
Generic	Review Citrix Director, Windows Performance Monitor, Event Log, and other monitoring software alerts	Check for warnings or alerts within Citrix Director, event logs, or other monitoring software. Investigate the root cause of the alert if any. **Note**: A computer and monitor can be set up to display the Citrix Director dashboard to create a Heads up Display for the Citrix department. This ensures the status of the environment is clearly visible. Monitoring recommendations for XenDesktop and XenApp 7.x are included in the Monitoring section of the VDI Handbook.	Operators
Generic	Verify backups completed successfully	Verify all scheduled backups have been completed successfully. This can include but is not limited to: • User data (user profiles / home folders) • Application data • Citrix databases • StoreFront configuration • Web Interface configuration • Provisioning Services vDisks (virtual desktops and application servers) • XenServer VM/Pool metadata (or equivalent for other hypervisors) • Dedicated virtual desktops • License files	Operators

Generic	Test environment access	Simulate a connection both internally and externally to ensure desktop and application resources are available before most users log on for the day. This should be tested throughout the day and may even be automated.	Operators
XenApp/ XenDesktop	Virtual machine power checking	Verify that the appropriate number of idle desktops and application servers are powered on and registered with the Delivery Controllers to ensure availability for user workloads.	Operators
XenApp/ XenDesktop	Perform incremental backup of Citrix related databases	Perform incremental-data backups of the following Citrix databases: • Site Database • Configuration Logging Database • Monitoring Database	Operators, Database team (if Citrix environment is using a shared SQL)
Provisioning Services	Check Citrix Provisioning Server utilization	Check the number of target devices connected to the Citrix Provisioning Servers and balance the load across servers, if required.	Operators
Provisioning Services	Perform incremental backup of Citrix PVS database	Incremental backup of Citrix Provisioning Server database hosted on SQL Server infrastructure.	Operators, Database team (if Citrix environment is using a shared SQL)

Table 60: Daily Operations

Weekly Periodic Tasks

The following table outlines the tasks that should be performed by the Citrix Support Team on a weekly basis.

Component	Task	Description	Responsible
Generic	Review latest hotfixes and patches	Review, test, and deploy the latest Citrix hotfixes and ascertain whether the Delivery Controllers and Server-Based OS / Desktop-Based OS virtual machines require them. **Note**: Any required hotfixes should be tested using the recommended testing process prior to implementation in production.	Operators, Implementers (review process)
Generic	Create Citrix environment status report	Create report on overall environment performance (server health, resource usage, user experience) and number of Citrix issues (close rate, open issues, and so on).	Operators
Generic	Review status report	Review Citrix status report to identify any trends or common issues.	Implementers, Architect
Generic	Maintain internal support knowledge base	Create knowledge base articles and issue resolution scripts to address Level-1 and Level-2 support requests. Review knowledge base articles and issue resolution scripts for accuracy, compliance, and feasibility.	Operators (Level-2 requests), Implementers (Level-3 requests, and review process)

XenApp/ XenDesktop	Check Configuration Logging reports	Confirm that Citrix site-wide changes implemented during the previous week were approved through change control.	Auditors
XenApp/ XenDesktop	Perform full backup of Citrix related databases	Perform full-data backups of the following Citrix databases: • Site Database • Configuration Logging Database • Monitoring Database	Operators, Database team (if Citrix environment is using a shared SQL)
Provisioning Services	Check storage capacity (only prior to updating a vDisk)	Review storage utilization, used and free storage space, for vDisk store and each vDisk. **Note:** Lack of space within the vDisk repository will be an issue only when the vDisks are updated using versioning or when a vDisk is placed in private mode during an update procedure. Storage utilization within vDisk should also be investigated. For example a 20GB vDisk may only have 200MB of free storage. If the vDisk itself is limited for storage then it needs to be extended. Citrix does not support resizing of a VHD file. Refer to the Microsoft link Resize-VHD for information on resizing a VHD file.	Operators
Provisioning Services	Perform vDisk updates (as necessary)	Perform a **full backup** of the vDisk before implementing any updates. Update the master vDisk image files and apply the following: • Windows software updates and patches • Operating system and application changes • Anti-virus pattern and definitions updates **Note**: Updates should be tested using the recommended testing process prior to implementation in production.	Operators
Provisioning Services	Check auditing reports	Review the Citrix Provisioning Services auditing Logs. **Note:** Provisioning Server auditing is off by default and can be enabled to record configuration actions on components within the Provisioning Services farm. To enable auditing refer to the Citrix eDocs article Enabling Auditing Information	Auditors
Provisioning Services	Perform full backup of Citrix PVS database	Backup of Citrix Provisioning Server database hosted on SQL Server infrastructure.	Operators, Database team (if Citrix environment is using a shared SQL)

Table 61: Weekly Operations

Monthly Periodic Tasks

The following table outlines the tasks that should be performed by the Citrix Support Team on a monthly basis.

Component	Task	Description	Responsible
Generic	Perform capacity assessment	Perform capacity assessment of the Citrix environment to determine environment utilization and any scalability requirements. **Note**: Recommendations for performing a capacity assessment are included in Decision: Capacity Management in the Monitoring section of the Virtual Desktop Handbook.	Architect

Table 62: Monthly Operations

Yearly Periodic Tasks

The following table outlines the tasks that should be performed by the Citrix Support Team on a yearly basis.

Component	Task	Description	Responsible
Generic	Conduct Citrix policy assessment	Review Citrix policies and determine whether new policies are required and existing policies need to be updated.	Implementers
Generic	Review software upgrades	Review and assess the requirement for new Citrix software releases or versions.	Architect
Generic	Perform Business Continuity Plan (BCP)/ Disaster Recovery (DR) test	Conduct functional BCP/DR test to confirm DR readiness. This plan should include a yearly restore test to validate the actual restore process from backup data is functioning correctly.	Implementers
Generic	Perform application assessment	Review the usage of applications outside and within the Citrix environment. Assess the validity of adding additional applications to the Citrix site, removing applications that are no longer required, or upgrading the applications to the latest version.	Architect
Provisioning Services	Archive audit reports	Perform an archive of the Citrix Provisioning Server Audit Trail Information for compliance requirements.	Auditors

Decision: Backup Location

The location of backups directly effects the recovery time and reliability of the Citrix environment. It is recommended to store backups of critical data both onsite and at an offsite location. If offsite backups are not possible due to costs associated or sensitivity of the data, backups should be placed at separate physical locations within the same datacenter.

Each backup option is discussed further below.

- **Onsite Backups** – Onsite backups should be located on a storage device in the datacenter that will allow the data to be recovered quickly in the event of a failure. Onsite backups are ideal for issues that only affect a small subnet of hardware in the datacenter. Backups can also be stored on a cold storage solution such as tape. While this medium is slower to recover from, it provides additional protection since it is only active during the backup process.
- **Offsite Backups** – Although the time to recover is much higher, offsite backups provide additional protection in the event of a disaster. Offsite backups may require transferring data over the Internet to a third party provider or they are created onsite and then transported to a remote

location on storage mediums such as tape. It is typical to put a limited number of backups offsite. For example one backup a week or month.

Decision: Testing Process

Regular updates and maintenance are an everyday part of IT operations. Standard processes must be followed to ensure updates do not negatively impact the production environment. This includes maintaining a dedicated testing infrastructure where modifications can be validated prior to being implemented in production.

Since changes to Citrix infrastructure can impact thousands of virtual desktop and application users, multi-phase testing is critical for the reliability and performance of the environment. As such, the process for testing should resemble the following:

Figure 22: Testing Process

- **Development -** The development infrastructure exists outside of the production network. Typically, it consists of short-lived virtual machines whose configuration matches production as closely as possible. The purpose of the development phase is to provide change requestors a non-production environment to perform proof of concepts, determine integration requirements and perform iterative testing as part of a discovery phase. Proposed changes should be documented so they can be applied in the test phase.
- **Testing -** The test environment is a standalone 1:1 copy of the production infrastructure and is used to confirm that the proposed changes can be easily repeated prior to the pre-production staging environment. The changes made should follow documentation from the development stage. If testing fails within the testing stage, the architect must determine the severity of failure and determine whether minor updates to documentation is sufficient or a full development cycle is needed.
- **Pre-production -** The per-production environment should mimic the current production environment. The goal of staging is to implement the proposed changes with little risk or uncertainty. It is expected that any changes made to the staging infrastructure have been tested and documented for repeatability. There should not be any iterations or adjustments required within this phase. During this phase and within this environment User Acceptance Testing (UAT) should be performed.
- **Production -** The production environment is a fully redundant and scalable solution designed for normal usage by end users. There should be minimal changes to the environment. If possible, all approved changes should be rolled out in stages to the production environment. This process is known as a staged rollout and mitigates risk by allowing changes to be rolled back, if necessary, without impacting the entire environment.

Decision: Change Control

Standardized processes that manage changes throughout a system's lifecycle are necessary to ensure consistent and accountable performance. The following change control leading practices should be considered.

- Use a change control window so that all applicable parties know when there might be downtime.

- Make sure that all teams are represented in the Change Advisory Board (CAB).
- Every change should have a roll back plan.
- If a change fails have a "hot wash" to determine what went wrong.
- Always use an automated change control system so that support staff can quickly and easily identify changes.
- When available, ensure configuration logging is enabled to track any changes made to the Citrix environment.

The change control process should be closely followed starting with a change request. A change request form should be filled out detailing changes requested, reasons for the change, and intended timeframes for the action. This is then reviewed and edited if required by a Change Manager and advisory board. When the change request has gone through the entire change approval process it is given to a change implementer who stages the change for testing, and finally conducts the implementation in production.

A sample change control process, including detailed steps, is provided in the diagram below:

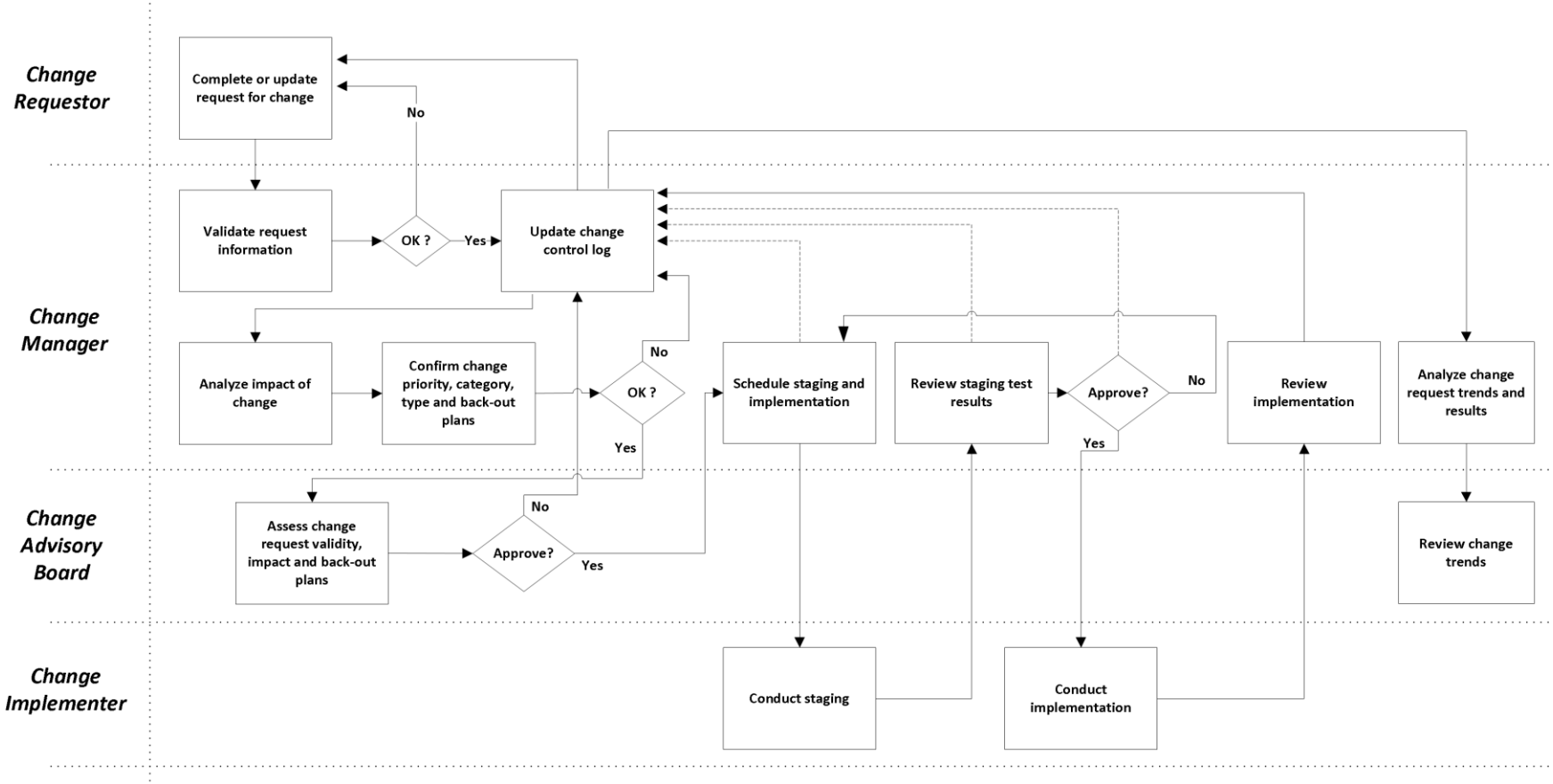


Figure 23: Change Control Process

The process is as follows:

1. The Change Request (CR) form is completed by any person requesting a change.
2. After appropriate manager approvals have been acquired, the CR is forwarded to the appropriate Change Manager(s).
3. The Change Manager validates the CR for completeness and logs the CR information into the Change Control Log for tracking. Incomplete change requests are returned to the requestor for update and re-submission.
4. The Change Manager assesses the impact of the change in conjunction with subject matter experts and/or managers of the teams associated/affected by this change.
5. The Change Manager works with the associated/affected teams as well as the change requestor in order to confirm the priority, category and type of the change as well as the proposed rollback plan.
6. If the change is approved by the Change Manager, the CR is forwarded to the CAB for approval. If the change is rejected, the Change Control Log is updated with the current status as well as the reason of the rejection and the CR is send back to the requestor.
7. The CAB reviews and validates the change in detail, and discusses and evaluates purpose, reasons, impact, cost and benefits. Each board member represents their department and provides guidance on the change requests. The CAB also reviews multiple requests to coordinate implementations and "package" requests into a single release schedule.
8. Upon approval the change is sent back to the Change Manager to schedule the change for implementation into the staging environment.
9. The change is implemented and tests are conducted. The results are sent back to the Change Manager.
10. If the staging implementation and testing are successful, the change is scheduled for production implementation. In case the staging phase was not successful another staging iteration will be conducted.
11. If possible, the change is rolled out in stages to the production environment. This process is known as a staged rollout and mitigates risk by allowing changes to be rolled back, if necessary, without impacting the entire environment. A rollback plan should be in place if there is an issue implementing a change in the production environment.
12. The Change Manager reviews the implementation and finally updates the Change Control Log.
13. On a periodic basis, the Change Manager reviews the Change Control Log to identify trends on type, frequency and size of changes and forwards the results to the CAB for review.

In an emergency, the processes may be expedited. Should an issue be declared an emergency, a change request form is still filled out and delivered to the appropriate change management representative. When approved, the requested change is immediately implemented and the advisory board notified.

Decision: Availability Testing

Availability testing is focused on ensuring resources are still available in the instance of a component failure. These tests are essential to ensuring users always have access to business critical resources. The testing should be conducted during nonbusiness hours or during a scheduled maintenance weekend when appropriate notice has been given to end users to make them aware if any unforeseen issues arise.

The following is a list of the key components that should be tested on a regular basis.

- **StoreFront** – StoreFront should be load balanced and health checked by a NetScaler or other load balancing device. To validate its configuration, all but one of the StoreFront servers should be shutdown. This will validate that the load balancing device is detecting the failure and directing users to the functioning server.
- **SQL** – SQL Server should be in a high availability configuration. To validate the configuration, the primary SQL server should be taken offline and then the Citrix Studio console should be opened. Since Citrix Studio will not be accessible without a functioning SQL server, it will validate that the SQL server failover mechanisms are functioning properly.
- **Delivery Controllers** - Resources deployed should be configured with a list of multiple Delivery Controllers. If one is made unavailable, desktops and application hosts will automatically establish a connection to another server in the list. To validate this, shutdown one of the Delivery Controller hosts and determine if the resources initially connected to it automatically register to another server. This can be determined by viewing the registration status of the resources inside Citrix Studio.

Sample Testing Workflow: Citrix Provisioning Services

Prerequisites and configuration requirements:

- Hypervisor, XenApp, and XenDesktop services are up and running.
- At least two PVS servers are installed and configured, providing the streamed disk image.
- Resilient networking and storage infrastructure with multiple links to each server.
- Test users are active on the XenApp or XenDesktop machines.

Steps	Expected Results
PVS Server Outage • Shutdown one of the Provisioning Servers. • Validate PVS continues to function. • Restart PVS Server. • Validate connections rebalance between PVS Servers. • Try the other(rest) PVS server(s) one by one.	• Existing XenApp/XenDesktop machines connect to another PVS server. • There is limited to no impact to the users utilizing that server. • New XenApp/XenDesktop machines can be booted and start correctly. • SCOM reports that the PVS server is down / not available. • Live connections are rebalanced between both PVS servers once both PVS servers are made available again.

PVS Bond Disruption • Disable / unplug a NIC in the PVS Streaming Bond on the PVS server.	• Provisioning Server continues to stream over remaining NICS in PVS Streaming Bond.
SQL Server PVS Database Mirror Failover • Admin logs on to Principle SQL Server. • Initiate failover of PVS database. • Validate PVS continues to function. • Initiate failback of PVS database. • Validate PVS continues to function.	• PVS continues to function.
SQL Service Outage • Admin reboots both Principle & Mirror SQL Servers simultaneously. • Validate PVS continues to function, but that administration is not possible. • Wait for the SQL Server to come back online. • Validate PVS administrative functions are once again possible.	• PVS continues to function. • PVS administrative functions are no longer available. • PVS administrative functions are available once the SQL services are restored.

Sample Testing Workflow: Citrix XenDesktop and XenApp Services

Prerequisites and configuration requirements:

- Hypervisor, XenDesktop, and StoreFront services are up and running.
- Network and storage services available.
- Provisioning Services is providing the streamed disk images.
- Test users are active on the virtual machines.
- SQL (Mirroring) and XenDesktop servers are up and running.
- Ensure multiple StoreFront servers are running.
- NetScaler load balancing services.

Steps	Expected Results
XenApp/XenDesktop 7.x Delivery Controller Citrix Broker Service Outage: • Stop the Citrix Broker Service on one of the Delivery Controller servers. • Validate virtual desktops or applications can still be enumerated and launched. • Start the Citrix Broker Service on the Delivery Controller server. • Shutdown one of the Desktop Controllers. • Validate virtual desktops or applications can still be enumerated and launched. • With a desktop launched, determine which Controller owns the host connection. Shut the Controller down and verify that another Controller takes over the session. ***Note**: This should be done during the maintenance window. Once complete, the VDI resources should be rebooted so the VDAs are evenly distributed across all controllers.*	• StoreFront correctly identifies service as being unavailable and redirects connections to remaining Delivery Controller. • Desktops continue to be enumerated and launch successfully. • Launched desktop can be supported if a hosting Controller goes down.
SQL Server Database Mirror Failover: • Admin logs on to principle SQL Server. • Initiate failover of XenApp/XenDesktop database. • Validate XenApp/XenDesktop continues to function.	• The database should failover and the Citrix Studio should pick up the failover database with no issues. • Existing sessions are not impacted. • New sessions are possible. • Administrative functions are possible.
SQL Service Outage: • Admin restarts both principle & mirror SQL Servers simultaneously. • Validate XenApp/XenDesktop continues to function, but that administration is not possible. • Wait for the SQL Service to come back online. • Validate administrative functions are once again possible.	• Existing XenDesktop sessions are not impacted • Recently used applications, hosted shared desktops and assigned VDI can be accessed due to connection leasing. New sessions are not possible if they do not meet the criteria for connection leasing. For more information on connection leasing, reference Citrix eDocs Connection leasing. • XenDesktop Administrative functions are not possible • XenDesktop Administrative functions are possible once SQL service is available.

Sample Testing Workflow: Citrix Licensing Services

Prerequisites and configuration requirements:

- Citrix Licensing Server up and running (with valid licenses installed).
- Hypervisor, XenApp/XenDesktop and StoreFront services are up and running.
- Users are active on the Server OS or Desktop OS machines.

Steps	Expected Results
Service continuity during complete failure of the Citrix Licensing Server: • Shutdown the Citrix Licensing server. • Reboot an existing Server OS machine. • Logon to the Citrix StoreFront and launch a published application. • Reboot an existing Desktop OS machine. • Logon to the Citrix StoreFront and launch a virtual desktop.	• License Server connectivity error posted in Event Log. • Provisioned Server OS boots successfully. • Users are able to launch published applications. • Provisioned Desktop OS boots successfully. • User is able to launch a virtual desktop. • Administrators will have 30 days grace to recover the Citrix Licensing Server.

Process 3: Monitoring

By having an in-depth understanding of current and expected behavior of the Citrix environment and its components, administrators are better equipped to discover an issue before it impacts the user community. Furthermore the data tracked during normal operations can be used for trending and capacity planning. This section defines how a Citrix environment should be monitored, as well as some common tools that can be used.

Decision: Performance Monitor Metrics

Monitoring the performance of the overall environment is crucial towards making sure all components are available and performing effectively to ensure users have a high quality experience.

Different components within the overall solution require monitoring of unique metrics with appropriately set thresholds. The metrics and thresholds presented are based on real world experience but may not apply to all environments. Organizations will need to perform their own baselining, validity testing and validation before implementing within a production environment.

***Note**: Some hypervisors, such as VMware vSphere and Hyper-V, provide specific performance counters for tracking CPU and Memory utilization within virtual machines (i.e. "VM Processor \ % Processor Time"). These performance counters should be used in addition to the general counters listed below.*

General

These performance counters should be used to monitor the key performance metrics of the Citrix infrastructure, application servers, and virtual desktops.

Metric	Description	Warning (Yellow)	Critical (Red)	Troubleshooting / Remediation
Processor - % Processor Time	% Processor Time is the percentage of elapsed time that the processor spends to execute a non-Idle thread. It is calculated by measuring the duration of the idle thread is active in the sample interval, and subtracting that time from interval duration. (Each processor has an idle thread that consumes cycles when no other threads are ready to run). This counter is the primary indicator of processor activity, and displays the average percentage of busy time observed during the sample interval. It is calculated by monitoring the time that the service is inactive and subtracting that value from 100%.	80% for 15 minutes	95% for 15 minutes	Identify the processes/services consuming processor time using Task Manager or Resource Monitor. If all processes/services work within normal parameters and the level of CPU consumption is an expected behavior it should be considered to add additional CPU resources to this system in the future. If a process/service can be identified which works outside normal parameters, the process should be killed. Please note that killing a process can cause unsaved data to be lost.
System - Processor Queue Length	Processor queue length is the number of threads in the processor queue. Unlike the disk counters, this counter shows ready threads only, not threads that are running. There is a single queue for processor time even on computers with multiple processors. Therefore, if a computer has multiple processors, you need to divide this value by the number of processors servicing the workload. A sustained processor queue of less than ten threads per processor is normally acceptable, dependent of the workload.	5 (per core) for 5 minutes **or** 6 (per core) for 15 minutes	10 (per Core) for 10 minutes **or** 12 (per core) for 30 minutes	A long CPU queue is a clear symptom of a CPU bottleneck. Please follow the steps outlined for counter "**Processor** - % Processor Time".
Memory – Available Bytes	Available memory indicates the amount of memory that is left after nonpaged pool allocations, paged pool allocations, process' working sets, and the file system cache have all taken their piece.	<30% of total RAM **or** 20% of physical memory over 6 minutes	<15% of total RAM **or** 5% of physical memory over 6 minutes	Identify the processes/services consuming memory using Task Manager or Resource Monitor. If all processes/services work within normal parameters and the level of memory consumption is an expected behavior it should be considered to add additional memory to this system in the future. If a process/service can be identified which works outside normal parameters, the process should be killed. Please note that killing a process can cause unsaved data to be lost.

Memory – Pages/sec	Pages/sec is the rate at which pages are read from or written to disk to resolve hard page faults.	>10	>20	A high value reported for this counter typically indicates a memory bottleneck, except if "**Memory** – Available Bytes" reports a **high** value at the same time. In this case most likely an application is sequentially reading a file from memory. Please refer to Microsoft Knowledge Base article KB139609 – High Number of Pages/Sec Not Necessarily Low Memory for further information.
Paging File - %Usage	This is the percentage amount of the Page File instance in use.	>40% **or** 80% over 60 minutes	>70% **or** 95% over 60 minutes	Review this value in conjunction with "**Memory** - Available Bytes" and "**Memory** - Pages/sec" to understand paging activity on the affected system.
LogicalDisk/PhysicalDisk - % Free Space	% Free Space is the percentage of total usable space on the selected logical disk drive that is free.	<20% of physical disk **or** 20% reported after 2 minutes	<10% of physical disk **or** 15% reported after 1 minute	Identify which files or folders consume disk space and delete obsolete files if possible. In case no files can be deleted, consider increasing the size of the affected partition or add additional disks.
LogicalDisk/PhysicalDisk - % Disk Time	% Disk Time marks how busy the disk is.	>70% consistently **or** 90% over 15 minutes (_Total)	>90% consistently **or** 95% over 15 minutes (_Total)	Identify the processes / services consuming disk time using Task Manager or Resource Monitor. If all processes/services work within normal parameters and the level of disk consumption is an expected behavior it should be considered to move the affected partition to a more capable disk subsystem in the future. If a process/service can be identified which works outside normal parameters, the process should be killed. Please note that killing a process can cause unsaved data to be lost.
LogicalDisk/PhysicalDisk – Current Disk Queue Length	Current disk queue length provides a primary measure of disk congestion. It is an indication of the number of transactions that are waiting to be processed.	>=1 (per spindle) consistently **or** 3 over 15 minutes (_Total)	>=2 (per spindle) consistently **or** 10 over 30 minutes (_Total)	A long disk queue length typically indicated a disk performance bottleneck. This can be caused by either processes/services causing a high number of I/Os or a shortage of physical memory. Please follow the steps outlined for counter "**LogicalDisk/PhysicalDisk** - % Disk Time" and counter "**Memory** – Available Bytes"

LogicalDisk/PhysicalDisk – Avg. Disk Sec/Read – Avg. Disk Sec/Write – Avg. Disk Sec/Transfer	The Average Disk Second counters show the average time in seconds of a read/write/transfer from or to a disk.	>=15ms consistently	>=20ms consistently	High disk read or write latency indicates a disk performance bottleneck. Systems affected will become slow, unresponsive and application or services may fail. Please follow the steps outlined for counter "**LogicalDisk/PhysicalDisk** - % Disk Time"
Network Interface – Bytes Total/sec	Bytes Total/sec shows the rate at which the network adaptor is processing data bytes. This counter includes all application and file data, in addition to protocol information, such as packet headers.	< 8 MB/s for 100 Mbit/s adaptor <80 MB/s for 1000 Mbit/s adaptor **or** 60% of NIC speed inbound and outbound traffic for 1 min.	70% of NIC speed inbound and outbound traffic for 1 min.	Identify the processes / services consuming network using Task Manager or Resource Monitor. If all processes/services work within normal parameters and the level of bandwidth consumption is an expected behavior it should be considered to move the respective process/service to a dedicated NIC (or team of NICs). If a process/service can be identified which works outside normal parameters, the process should be killed. Please note that killing a process can cause unsaved data to be lost.

Table 63: Recommended Metrics to Monitor for all Virtual Machines

XenApp/XenDesktop

These performance counters are specific to the Delivery Controllers.

Metric	Description	Warning (Yellow)	Critical (Red)	Troubleshooting / Remediation
Database Avg. Transaction Time	The time on average, in seconds, taken to execute a database transaction. A baseline needs to be established in the environment in order to accurately establish threshold values.	Based on baseline values	Based on baseline values	In case the reported values exceed the baseline response time constantly, a potential performance issue needs to be investigated at the SQL server level.
Database Connected	Indicates whether this service is in contact with its database. (1 is connected; 0 is not connected).	0	0 (for over 30 minutes)	Both values report connectivity issues of the XenDesktop Broker service with the database. In case issues are reported, SQL server and network availability needs to be verified.

Database Transaction Errors/sec	The rate at which database transactions are failing.	None	>0	Both values report connectivity issues of the XenDesktop Broker service with the database. In case issues are reported, SQL server and network availability needs to be verified.

Table 64: Recommended XenApp/XenDesktop Metrics

StoreFront

These performance counters are specific to the StoreFront servers.

Metric	Description	Warning (Yellow)	Critical (Red)	Troubleshooting / Remediation
ASP.NET – Request Queued	The number of requests waiting to be processed by ASP. A baseline needs to be established in the environment in order to accurately establish threshold values.	Based on baseline values	Based on baseline values	In case the queue length exceeds the critical limit requests may be rejected. In this case it should be considered to add additional StoreFront or Web Interface servers to the load balancing team in order to distribute the load across more nodes.
ASP.NET – Requests Rejected	The number of requests rejected because the request queue was full.	None	>=1	When this limit is exceeded, requests will be rejected with a 503 status code and the message "Server is too busy." Please follow the steps outlined for counter "ASP.NET – Request Queued"

Table 65: Recommended StoreFront/Web Interface Metrics

Citrix License Server

These performance counters are specific to the Citrix License Server.

Metric	Description	Warning (Yellow)	Critical (Red)	Troubleshooting / Remediation
Citrix Licensing – Last Recorded License Check-Out Response Time	Displays the last recorded license check-out response time in milliseconds.	>2000 ms	> 5000 ms	If the reported values exceed the 5000 ms response time, a potential performance issue needs to be investigated in the Citrix License Server.
Citrix Licensing – License Server Connection Failure	Displays the number of minutes that XenDesktop has been disconnected from the License Server.	> 1 minute	> 1440 minutes	Both values report connectivity issues with the License Server. In case issues are reported, License Server and network availability needs to be verified.

Table 66: Recommended Citrix License Server Metrics

Decision: Services Monitoring

Windows services that are critical to basic server functionality should be automatically monitored to ensure that they are running properly. The following table provides a list of the common Windows services that should be monitored. When any of these services are restarted or stopped a warning (Yellow) or critical (Red) alert should be assigned respectively. The recommended recovery actions for the services listed below are as follows:

- First failure: Restart the Service
- Second Failure: Restart the Service
- Subsequent Failures: Put the server in maintenance mode and investigate the root cause

XenApp/XenDesktop

Service	Functionality	Administration Risk
Citrix AD Identity Service	Manages Active Directory computer accounts. Dependencies: • WMI Service	Machine Creation Service relies on this service to create virtual machines. Administrators will be unable to create new or modify existing Machine Catalogs. Administrators will be unable to establish new connections to Citrix Studio.
Citrix Broker Service	Manages connections to virtual machines and applications.	If this service is stopped administrators will be unable to make changes to the environment or establish new connections to Citrix Studio. Any existing administrator connections to Citrix Studio can also be terminated. If this service is stopped existing user connections are not affected. No new connections can be established. Users logging into StoreFront will be unable to see any resources available for selection. Once the service is restarted users will need to re-login to StoreFront to establish connections.
Citrix Configuration Logging Service	Logs administrator activity and configuration changes in a XenDesktop deployment.	If this service is stopped XenApp/XenDesktop will be unable to communicate with the Configuration Logging Database. Administrators will be unable make changes to the environment or establish new connections to Citrix Studio.
Citrix Configuration Service	Stores service configuration information. Dependencies: • WMI Service	If this service is stopped administrators will be unable to make changes to the environment or establish new connections to Citrix Studio.
Citrix Delegated Administration Service	Manages configuration of delegated administration permissions.	If this service is stopped XenApp/XenDesktop cannot assign administrative permissions. Administrators will be unable to make changes to the environment or establish new connections to Citrix Studio. Administrators will be unable to establish new connections to Citrix Director and existing sessions within Citrix Director will be interrupted.

Citrix Diagnostic Facility COM Server Service	Manages and controls Citrix diagnostic trace sessions on the system. Dependencies: • RPC Service	This service has no impact on the production environment. It is used to generate CDF trace files which aid in troubleshooting issues.
Citrix Environment Test Service	Manages tests for evaluating the state of a XenDesktop Site.	If this service is stopped administrators will be unable to establish new connections to Citrix Studio. Administrators will also be unable to check the status of the Citrix site configuration, machine catalogs, and delivery groups by running the tests under "Common Tasks" in the Citrix Studio administration console.
Citrix Host Services	Manages host and hypervisor connections. Dependencies: • WMI Service	Administrators will be unable to create new Machine Catalogs or control virtual machine power settings via Citrix Studio. Administrators will be unable to establish new connections to Citrix Studio. Users may experience issues connecting to virtual desktops when this service is not available. If this service is stopped existing connections are not affected.
Citrix Machine Creation Service	Creates new virtual machines. Dependencies: • WMI Service	Administrators will be unable to create new or modify existing Machine Catalogs or establish new connections to Citrix Studio. Administrators will be unable to establish new connections to Citrix Studio.
Citrix Monitor Service	Monitors the FlexCast system.	If this service is stopped XenApp/XenDesktop will be unable to communicate with the Monitoring Database. Citrix Director will be unable to retrieve any data on the environment. Administrators will be unable to establish new connections to Citrix Studio.
Citrix StoreFront Service	Manages deployment of StoreFront.	Administrators will be unable to establish new connections to Citrix Studio.

Table 67: XenApp/XenDesktop 7.x Services

Delivery Controller Services Monitoring in Citrix Director

The **Infrastructure** pane within the Citrix Director dashboard provides status of the services running on the Delivery Controllers and will provide warning indications if a service or Controller is unavailable. These alerts can be accessed by clicking the **Alert** hyperlink within the **Infrastructure** pane.

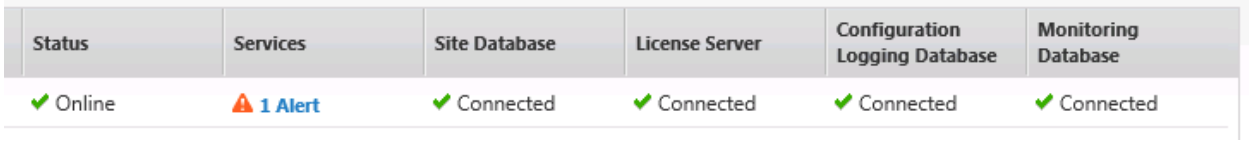


Figure 24: Citrix Director Infrastructure Pane

Provisioning Services

Service	Functionality	Risk
Citrix PVS PXE Service	Provides the PVS PXE Boot Server functionality. ***Note:*** *Only applicable when PXE boot is used.*	On failure of this service target devices may not be able to boot successfully if PXE booting is leveraged.
Citrix PVS Stream Service	Streams contents of the vDisk to the target device on demand.	If this service stopped it will not be possible to stream vDisk images.
Citrix PVS SOAP Service	Provides framework for external or existing solutions to interface with Provisioning services. ***Note:*** *Only impacts console operations. User is unaffected*	If this service fails PVS Server to PVS Server communication as well as PVS Console to PVS Server communication is not possible.
Citrix PVS TFTP Service	Provides the TFTP Server functionality. ***Note:*** *Only applicable when TFTP is used.*	On failure of this service target devices may not be able to boot if this server is used as TFTP server for the bootstrap.
Citrix PVS Two-Stage Boot Service	Provides the bootstrap functionality for devices booting by means of a BDM ISO file. ***Note:*** *Only when BDM boot partitions are used.*	On failure of this service target devices may not be able to boot if a BDM ISO file is used.

Table 68: Provisioning Server Services

StoreFront

Service	Functionality	Risk
Citrix Cluster Join Service	Provides Server Group join services.	This service is started when adding additional StoreFront servers to a Server Group. If this service does not start or is interrupted when this process is initiated the additional server will be unable to join the indicated Server Group and the process will result in an error.
Citrix Configuration Replication	Provides access to Delivery Services configuration information.	This service only exists on the primary StoreFront server of a Server Group. If this service is stopped additional StoreFront servers will be unable to join the Server Group and any changes made to the primary StoreFront server will not be replicated to other servers. This can result in servers within the Server Group being out of sync.
Citrix Credential Wallet	Provides a secure store of credentials. Dependencies: • Citrix Peer Resolution Service	If this service is stopped users will be unable to login to access their desktops or applications. Users logged into StoreFront will be unable to launch new application or desktop sessions. Existing application or desktop sessions are unaffected.
Citrix Default Domain Services	Provides authentication, change password, and other domain services.	If this service is stopped users will be unable to login to access their desktops or applications. Users currently logged in will not be affected.

Citrix Peer Resolution Service	Resolves peer names within peer-to-peer meshes.	On failure of this service both the Citrix Credential Wallet and Citrix Subscriptions store are stopped generating the risks associated with those services.
Citrix Subscriptions Store	Provides a store and replication of user subscriptions. Dependencies: • Citrix Peer Resolution Service	If this service is stopped Citrix Receiver cannot add, remove, and reposition applications within StoreFront. Users will need to re-add applications and all changes made to their selection of applications within the StoreFront store will not be saved or replicated to other sessions. Original user configuration will be restored once the service is restarted.
World Wide Web Publishing Service	Provides web connectivity and administration through the Internet Information Services Manager. Dependencies: • HTTP • RPC Service	Access to published applications or published desktops will not be available through StoreFront. Users will be unable to resolve the Receiver for Web login page. Users logged into StoreFront will be unable to launch new application or desktop sessions and will need to reenter credentials when the service is restarted. Existing application or desktop sessions are unaffected.

Table 69: StoreFront Services

Web Interface

Service	Functionality	Risk
World Wide Web Publishing Service	Provides web connectivity and administration through the Internet Information Services Manager. Dependencies: • HTTP • RPC Service	Access to published applications or published desktops will not be available through Web Interface if the WWW service is not available.

Table 70: Web Interface Services

Citrix License Server

Service	Functionality	Risk
Citrix Licensing Service	Provides licensing services for Citrix products.	Licensing mode changes to grace period when service is stopped or License Server cannot be contacted. If not monitored, functionality of Citrix products will cease after grace period expires.
Citrix Licensing Support Service	This account controls reading the license files and updating strings with license trailers (data dictionary functionality).	None
Citrix Licensing WMI	The Citrix License Management Console collects license data information using the WMI service.	None

Table 71: Citrix License Server Service

Decision: Events Monitoring

Monitoring the Windows Event Log for unknown or critical events can help to proactively discover issues and allow administrators to understand event patterns:

- **Licensing -** Errors in the Event Log dealing with Remote Desktop licensing should be investigated. This might be a result of the installed Citrix product not being able to contact the Remote Desktop Licensing Server or the Citrix Licensing Server. If errors in the Event Log are not reviewed, users might eventually be denied access because they cannot acquire a valid license.
- **Hardware Failure -** Any event notification that relates to a hardware failure should be looked at immediately. Any device that has failed will have an impact on the performance of the system. At a minimum, a hardware failure will remove the redundancy of the component.
- **Security Warnings -** Customers should investigate security warnings or audit failure events regarding failed logons in the security log. This could be an indication that someone is attempting to compromise the servers.
- **Disk Capacity -** As the drives of a Windows system reach 90% of capacity, an event error message will be generated. To ensure continuous service, customers should poll these event errors. As the system runs out of hard disk space, the system is put at severe risk. The server might not have enough space left to service the requests of users for temporary file storage.
- **Application / Service errors -** Any event notification that relates to application or services errors should be investigated.
- **Citrix errors -** All Citrix software components will leverage the Windows Event Log for error logging. A list of the known Event Log warnings and errors issued by Citrix components can be found at the following links:
    - Event Codes Generated by PVS
    - XenDesktop 7 - Event Log Messages

  It is important to periodically check the Event Viewer for Citrix related warnings or errors. Warnings or errors that repeatedly appear in the logs should be investigated immediately, because it may indicate a problem that could severely impact the Citrix environment if not properly resolved.

In multi-server environments it becomes easier to administer the servers when logs can be collected and reviewed from a central location. Most enterprise grade monitoring solutions provide this functionality. More sophisticated monitoring solutions enable an administrator to correlate event information with other data points such as performance metrics or availability statistics. In case the selected monitoring solution does not provide this functionality the Windows Server 2008 R2 or Windows Server 2012/2012 R2 Event Log subscription feature can be used. This feature allows administrators to receive events from multiple servers and view them from a designated collector computer. Please see Microsoft TechNet article Manage Subscriptions for more information.

XenServer is also capable of sending its logs to a central syslog server. The administrator sets the IP address of the syslog daemon server in the properties of each XenServer in the pool. This configuration allows administrators to capture real-time activity across multiple XenServer hosts. Further information can be found within the XenServer Admin Guide.

Decision: Capacity Management

In addition to the day-to-day monitoring of system-level metrics, performance metrics should be tracked from a historical perspective to help plan for future growth as more users access the environment.

A baseline of the environment performance should be taken so that it can be compared against performance over time. For example, if a user complains of poor performance, this baseline can be used for comparison purposes to identify if the issues are related to the user load exceeding the capacity of the environment.

An example of baseline performance metrics for capacity management would include historical data for CPU, Memory, and network utilization on the Delivery Controller and application servers or desktops.

Citrix Director

Administrators can utilize the **Trends** view within Citrix Director to track different parameters of the Citrix XenApp/XenDesktop deployment over time. These parameters can be leveraged for capacity planning of the Citrix environment.

From the Trends view, administrators can see historical data that is broken up into several categories including:

- **Sessions** - Provides the concurrent session usage over time enabling the ability to size the environment appropriately.
- **Connection Failures** - Gives an overview of the different types of connection failures that have occurred across different Delivery Groups.
- **Failed Desktop OS Machines** – Gives an overview of the different problems associated with failures in desktop machines.
- **Failed Server OS Machines - G**ives an overview of the different problems associated with failures in server machines.
- **Logon Performance** – Shows how long it takes for users to log on to their applications and desktops.
- **Load Evaluator Index** – Provides various performance counter-based metrics, including CPU, Memory, and Disk Usage for Server OS machines.
- **Hosted Application Usage** – Details all applications published in the site and can provide usage information about each individual applications in detail (concurrent instances, launches, usage duration, and so on).

  ***Note**: Requires XenApp or XenDesktop Platinum licensing*
- **Network** – Network analytics provided through NetScaler HDX Insight.

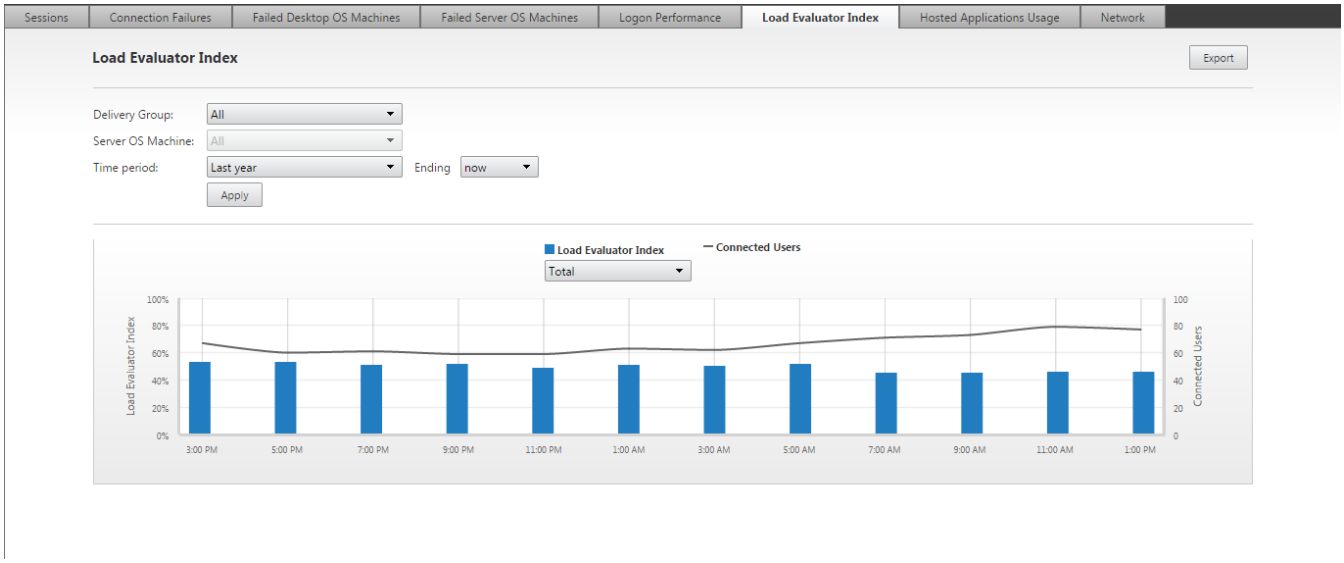


Figure 25: Citrix Director Trends View

For more information on Citrix Director Trends, please refer to the following.

- Citrix Blogs – Citrix Director: Trends Explained
- Citrix Support – CTX139382 Best Practices for Citrix Director

Acknowledgments

The creation of the handbook is a time consuming process and requires real deployment experience across many scenarios. Citrix would like to thank the authors and subject matter experts who contributed to the Citrix VDI Handbook.

Authors

Name	Title
Daniel Feller	Lead Architect
Nicholas Rintalan	Lead Architect
Andy Baker	Senior Architect
Thomas Berger	Sr. Product Marketing Manager
Amit Ben-Chanoch	Technical Product Manager
Rich Meesters	Architect
Matthew Brooks	Sr. Product Marketing Manager
Roger LaMarca	Principal Consultant
Adeel Arshed	Sr. Consultant
Rafael Jose Gomez	Consultant
Ed Duncan	Consultant
Kevin Nardone	Principal Consultant

Subject Matter Experts

Name	Title
Brendan Lin	Sr. Architect
Sarah Steinhoff	Sr. Architect
Dan Morgan	Principal Consultant
Diego Madiedo	Architect
Cid Neves	Sr. Architect
Jeff Qui	Principal Consultant
Michael Havens	Principal Consultant
Ryan Robott	
Josh Fu	
Maria Chang	Senior Consultant
Pablo Legorreta	Architect
Uzair Ali	Principal Consultant
Steven Kruger	Lead Systems Engineer

Revision History

Revision	Change Description	Updated By	Date
1.0	XD7.6 LTSR Handbook Released	Daniel Feller (Lead Architect)	September 30, 2016