High Availability for Citrix XenApp

Enhancing XenApp Availability with NetScaler Reference Architecture
Contents

Contents ........................................................................................................................................... 2

Introduction ..................................................................................................................................... 3

Desktop Availability ......................................................................................................................... 3

  Hardware ....................................................................................................................................... 4
  Operating System .......................................................................................................................... 4
  Applications ................................................................................................................................... 5

High Availability ............................................................................................................................. 5

  XenApp Server Virtualization ......................................................................................................... 6
  XenApp Server Delivery .................................................................................................................. 8
    Web Interface .............................................................................................................................. 8
    Application Delivery .................................................................................................................... 8
    XML Service ............................................................................................................................... 9
  SQL Server ...................................................................................................................................... 10
  Active Directory Infrastructure ....................................................................................................... 10

Business Continuity/Disaster Recovery ............................................................................................ 11

  Global Server Load Balancing ....................................................................................................... 11
  Roaming User Support ................................................................................................................... 13
  Disaster Recovery .......................................................................................................................... 15

Conclusion ........................................................................................................................................ 16

Revision History ............................................................................................................................... 17
Introduction

As organizations simplify desktop management activities through the use of desktop virtualization, creating a highly available architecture becomes ever more important. Historically, users were minimally impacted by desktop outages. In many situations, an outage typically impacted a single user, as one desktop device belonged to a single user. In a virtualized desktop operating model, there is the potential for an outage to impact hundreds or thousands of users if the environment is not designed appropriately.

This white paper focuses on the fault tolerant and high-availability options integrated with Citrix XenApp 6.5 as they align with three different focus areas:

- **Layer 1 – Application Availability:** Users should always be able to work with an available hosted shared desktop or application.

- **Layer 2 – High Availability:** Failures within a site should not impact desktop or application availability.

- **Layer 3 – Business Continuity/Disaster Recovery:** The loss of an entire site should not prevent users from accessing their desktops or applications.

By focusing on three distinct layers, an organization can feel confident in the operational availability of the architecture, even in the event of a catastrophic failure.

Desktop Availability

When delivered via XenDesktop or XenApp, a user’s desktop environment is built around four different layers:

- Hardware
- Operating System
- Applications
- Personalization

In order to provide a user with the correct desktop, each one of the desktop layers must be highly-available. The focus of this whitepaper is on delivering high availability for published applications and hosted shared desktops. Information on delivering high availability for hosted virtual desktops can be found in the whitepaper [CTX131255](http://www.citrix.com).


Hardware

With XenApp, the end point hardware becomes less of an issue for providing availability to the user’s applications. When the user’s physical desktop contains all operating system, application and data functionality, the loss of the desktop meant the loss of work until the device was repaired.

In a virtualized desktop and application model, the user’s operating system, applications and personal settings are abstracted from the end point hardware. Essentially, the user is able to jump across different machines without impacting the usability of their desktop. This provides fault tolerance in the following ways:

- **Endpoint Failure:** If the user’s physical endpoint fails, any new endpoint can be used to gain access to the virtual desktop or to the user’s applications.

- **Hosted Shared Desktop Failure:** If the XenApp server delivering the hosted shared desktop or user applications fails, the user can immediately initiate a new connection with another server in the farm and the user will be directed to a new shared desktop and new applications, with all of their personalization intact. As with a physical desktop failure, the user will lose any open and unsaved edits, but recovery to a new virtual environment will be much faster.

- As can be seen, the act of transforming the desktop into a virtual desktop overcomes productivity loss associated with desktop device failure.

Operating System

The second major reason why users of physical desktops have a loss of productivity is due to the corruption of the operating system, which results in failure to boot, blue screens or system crashes. Using XenApp to deliver hosted shared desktops, administrators can deliver a controlled, locked down desktop environment based on Windows Server 2008 R2. The combination of a more robust server class operating system and allowing Users less access to the operating system itself minimizes the possibility of an operating system corruption. In the event of a problem with a XenApp server, users can connect to another server in the farm with minimal disruption. For users who require more control over their operating environment, or for situations where a full Windows 7 desktop is desired, XenDesktop can be used to deliver a proven, optimized image to the virtual desktop through the FlexCast for Hosted VDI Desktop models. This paper focuses on delivery of applications and hosted shared desktops using XenApp. For more information on Citrix FlexCast models, visit flexcast.citrix.com.
Applications

Applications can be delivered into the virtual desktop in three different ways. Users can be protected from failures within the application due to corruption, deletion, or other means via the processes used to deliver those applications, particularly when dealing with virtual desktops, or dedicated desktops with streamed or hosted applications.

- **Installed**: Installed applications are part of the base desktop image. In a hosted shared desktop model, the image is delivered as read-only and user changes are not allowed. When the shared desktop is restarted, the changes are lost and the user receives a new desktop environment. This keeps the operating system and the installed applications in a pristine format, free from any corruptions or misconfigurations by the user. Physical desktops or dedicated virtual desktops manage installed applications through traditional systems management model.

- **Streamed**: Streamed applications are delivered to the XenApp server or the users physical or virtual desktop over the network as requested. The applications are stored on a file server (Application Hub) as an application profile. When users launch the application, portions are delivered to the desktop or XenApp server. The streaming process verifies the files exist in the correct state during launch. If not, the correct files are streamed from the Application Hub automatically and seamlessly.

- **Hosted**: Hosted applications are virtualized and executed on a XenApp server; therefore the application does not impact the physical or virtual desktop. Hosted applications can either be installed or streamed. If the applications are streamed to the XenApp servers, they will self-heal due to the file checks built into the application streaming process.

**High Availability**

With a traditional XenApp design, users have a base level of fault tolerance through the delivery of applications and hosted shared desktops from multiple XenApp servers. However, to provide a greater level of availability for the delivery of hosted applications and desktops, the environment can be augmented based on the following areas of focus:

- XenApp server virtualization
- XenApp delivery
As can be seen in the following figure and subsections, these areas of focus provide a highly-available virtual desktop solution.

![Diagram](image)

**Figure 1: XenApp High-Availability**

**XenApp Server Virtualization**

A virtual XenApp server delivery platform can be designed to be able to accommodate planned and unplanned outages. The system must be able to recover seamlessly. Using XenServer as the underlying infrastructure platform allows for the following levels of fault tolerance:

- **Hypervisor Failure:** Hypervisor failure needs to be considered in order to avoid situations where the loss of a single physical server can affect users running on the virtual XenApp servers in the configuration. When configured for high availability, the loss of any one server within the virtualization pool will not impact the other servers within the pool. While a XenServer pool has a dedicated master server responsible for the proper functioning of the pool, any other server in the pool can become the master so the loss of any server does not create a single point of failure.

- **Imminent Failures:** Imminent failures are issues with the physical hardware that slowly degrades service. These types of failures can be overcome without any user interruption with XenServer’s XenMotion technology, along with proper storage design. XenMotion allows any running virtual server to be migrated to another XenServer within the same resource pool without disrupting the user.

- **Critical Failures:** A critical failure is an unforeseen failure that causes the physical server to crash. These types of failures typically have no warnings, which results in locally running sessions being lost. XenServer incorporates high-availability features to allow for the
auto-restart of different systems based on priority levels. This type of functionality has benefit for XenApp servers, as infrastructure components can be started before XenApp application servers, and critical application servers can be given priority to make the most used applications available soonest. It is advisable to set priority levels for other critical infrastructure components like: XenApp XML Controllers, Web Interface servers, etc. Best Practice recommendations can be found in the [XenApp Planning Guide – Virtualization Best Practices (CTX129761)](#).

Of course, it is also recommended that high availability of the network and storage components be considered when designing and building a hosting platform for virtual desktops and applications. While beyond the scope of this whitepaper, Citrix has developed [best practices for planning and designing storage for XenDesktop](#), which includes storage recommendations for XenApp and Provisioning Services. Network infrastructure should always be deployed with redundant configurations, from the server side (redundant NICs) through to the end user infrastructure to avoid single points of failure.
XenApp Server Delivery

Delivering the resources to the user, whether those resources are desktops or applications, is the responsibility of Web Interface and the XenApp controllers. As discussed in the following sections, providing high-availability to these components is possible with the use of smart monitors.

**Web Interface**

In a full-scale XenApp infrastructure, the Web Interface servers are responsible for delivering applications and hosted shared desktops to the users. Initially, a user makes a connection to Web Interface. Whether used through a browser or by a desktop appliance/thin client, Web Interface is a critical component for users. From the interface, users enter their credentials and select their desktops or applications. Based on the user interactions, Web Interface communicates with the XML Service for the XenApp sites to fulfill the user request. In situations where the sever hosting Web Interface fails, or the IIS service fails or Web Interface encounters issues, users would be unable to connect to the environment.

NetScaler provides an intelligent monitor for Web Interface. By launching a connection to the Citrix Web Interface, the monitor determines if the server is available, if the web service is running and if the Web Interface site is functioning and responding. If disruptions in the service are identified, NetScaler generates an alert.

The alert is then used as part of the NetScaler load balancing algorithm. If a Web Interface server is not responding correctly, the server is removed from the load balancing pool until the problem is corrected. New user requests are routed only to the available Web Interface servers.

Providing high-availability for Web Interface goes beyond providing an available server. Configuring high-availability for Web interface allows an administrator to determine the best way of balancing users across servers. For example, one Web Interface server could be the least loaded server but is currently busy with another process. NetScaler could direct new user requests to another Web Interface server that responds first instead of directing based on user connections. Intelligent monitoring incorporated into load balancing allows for users to get the fastest application delivery experience.

**Application Delivery**

Applications, which are not installed within the base operating system image, are delivered into a physical or virtual desktop via the Application Delivery (XenApp) controllers. Without fault tolerance at the application delivery layer, users could end up with a desktop and no applications. To mitigate this risk, XenApp application delivery has built-in redundancy based on the following:
Hosted Applications: For hosted applications, the application logic and processes occurs on the XenApp server. As a recommended best practice, XenApp servers should be configured in an N+1 configuration where there are enough XenApp servers to accommodate the loss of one. When a XenApp server fails or the application crashes, the user re-launches their application. The application is launched and hosted on a different, available XenApp server.

Streamed Applications: For streamed applications, the application logic and processing occurs on the desktop, although it is not fully installed there. When users request the application, portions of the application profile are delivered from the Application Hub via the network to the desktop. If the Application Hub fails, users do not receive the application stream. Implementing Windows DFS, which is a component of the Windows Server 2008 R2 operating system, provides a clustered solution so the loss of one server does not impact application delivery to the users.

**XML Service**

A critical component of any XenApp environment is the XML Service. The broker service is responsible for user authentication, resource enumeration and resource launching processes. A failure in the broker will result in users being unable to start their virtual desktop. The following diagram shows how critical the XML Service is to users.

![XML Broker Process Flow Diagram](image)

The XML service is the link between the users and the XenApp infrastructure, which makes it critical. Monitoring the broker service is not a trivial task as the monitoring must go beyond simply identifying if the service is running to identifying if the service is responding appropriately. If the broker responds incorrectly, the Web Interface server could get stuck in a request/response loop resulting in users not gaining access to their resource.
NetScaler provides intelligent monitoring of the XML Service through the use of pre-configured monitor templates for the XenApp. The monitoring determines if the XML Service is running and if the requested information from the broker responds in a timely manner and with expected information. If the monitor determines an unexpected result or complete failure to respond, NetScaler creates an alert, which is used in the load balancing algorithm. NetScaler dynamically adjusts the environment to bypass the failed XenApp server. If the XML Service functionality is restored, NetScaler automatically detects and incorporates the XenApp server back into the environment.

Providing high-availability to the XML Service is more than load balancing in the event of a failure. NetScaler also load balances connections across multiple XML Services to help spread the load and to provide a better and faster application initialization experience. In many organizations, a major shift change starts at 8 or 9 AM. This results in a huge load on the XenApp servers providing the XML service. Integrating NetScaler into the environment distributes the load across multiple Brokers while also monitoring and allocating requests based on availability.

**SQL Server**

The SQL database hosts the XenApp data store, which provides the foundation for all configuration information in the XenApp farm. All configuration information about published applications, users, printers and servers is stored in the database. If the data store fails, no configuration changes to the farm are possible. It is recommended that the SQL database be made highly available through SQL Mirroring and/or SQL Clustering technologies. For more information on SQL mirroring and clustering see the Microsoft whitepaper on [High Availability with SQL Server 2008 R2](#).

**Active Directory Infrastructure**

AD and DNS services are required to resolve FQDNs for web interface and XML communications, and authenticate user access to the resources delivered by XenApp. As AD and DNS services are critical components of an enterprise IT architecture, they are usually designed with high availability in mind. Active Directory has built in availability features such as multi-master replication and Active Directory integrated DNS. Generally, using these features will address availability needs for AD and DNS.
Business Continuity/Disaster Recovery

As users continue to be spread across different regions, there is a requirement for many XenApp environments to provide business continuity and disaster recovery. Oftentimes, this results in one of the following actions by the administrators:

- Users are told to use a different address in the event of a failure at the main data center
- A manual change is made to the DNS table, which will direct user requests to the backup data center

Of course these items are not automatic based on monitors and they also idle roughly 50% of the hardware stored at the backup data center because these business continuity solutions are configured in an active-passive mode. In the active-passive mode, only one site is active at a time. When a failure occurs, the passive site becomes active, but until this occurs, the backup site is in standby mode. However, in an active-active model where there are multiple data centers, users require a solution that not only provides them access, but also provides them access to their applications, with considerations for optimization based on where the user data is located. If the user cannot access the data required to perform their job roles, then failover makes little sense, whether automated or manual. Failover can be automated through a combination of these functions, depending upon the specific requirement:

- NetScaler Global Server Load Balancing (GSLB)
- Roaming User Support
- Disaster Recovery

Global Server Load Balancing

The first step towards providing a business continuity and disaster recovery solution for XenApp is to utilize the global server load balancing functionality of NetScaler.

![Figure 3: High-Availability Architecture](image-url)
External users make a request using a fully qualified domain name to the organization’s Access Gateway login page, which is delivered from NetScaler. If the user is internal, the user would make a fully qualified domain name request to the Citrix Web Interface, which is load balanced by NetScaler. Providing the best site is the responsibility of the NetScaler. Each site communicates with the other sites using the Metric Exchange Protocol (MEP). This lets each NetScaler know the status of the other sites so each NetScaler can provide an accurate response to incoming user requests. The responses must take into account the following:

- **Site Availability:** Does the data center have all critical components available (Web Interface, Access Gateway, XML Broker, etc.)? NetScaler uses monitors to check the availability of the Access Gateway, Web Interface and XenApp XML Service components within the sites in the GSLB configuration.

- **Best Available Site:** Based on the sites with critical component availability, which one is the best for this particular user? The best site is based on the current situation and configuration. With NetScaler, the best site could be the site which responds the fastest, the site that is closest in geographical proximity to the user, or it could be in a round robin flow. The choice should be based on the organizational architecture and needs.

Once these items are taken into account, the user is presented with an Access Gateway logon page at the best site.
Roaming User Support

When it comes to roaming user support, there are multiple considerations that need to be made to design and deliver an optimal solution for the users. The speed of the network and the location of user data relative to the published application must be considered. The global server load balancing functionality gets the user to the best responding site, but sometimes this is not the best site from the user perspective. The closest site from a proximity perspective is measured from the DNS server location, not the user location. For example, if a user’s data is located in a data center in North America and they are currently travelling in Asia, global server load balancing would direct the user to the Asia data center. This may require the application running on a XenApp server in the Asia data center to communicate with the North America data center to retrieve the user’s data. If this network link is slow, the result would be an underperforming application from the user perspective.

If this is the case, the ideal solution would be to have the user connect to the Asia site, based on GSLB, and then use worker group preference to route the user to the North America data center for the application, thus keeping the user’s data local to the North America data center. The user would interact with the virtual desktop via the Citrix HDX, which is optimized for WAN environments.
This approach provides global redundancy and can provide the user with the best experience as the only thing crossing the WAN link is HDX instead of application data. This can be accomplished by

1. Creating Active Directory groups corresponding to each geographic site
2. Creating XenApp worker groups corresponding to application servers within a given site.
3. Adding the appropriate application servers into the defined worker groups. These groups are used to define the user’s “Home” site.
4. Configuring load balancing policies in XenApp to assign user groups to preferential worker groups, and backup worker groups as required.

When a user connects to the GSLB configuration, the NetScaler selects the best site based on the load balancing schemes configured. The user authenticates with the Access Gateway, and is connected to a load balanced Web Interface server. Based on the user’s group membership, the Web Interface directs the user to the correct application servers based on the load balancing policy. The user’s HDX traffic will traverse the corporate WAN structure behind the Access Gateway; no HDX traffic transits across the AG.

Depending upon the particular architecture in use, other options may be required to provide roaming user support for the XenApp solution. If multiple farms are used in the organization’s environment, it may be preferable to use the web interface to define preferred farms based on Active Directory group membership, as discussed in the High Availability for XenDesktop whitepaper. Careful consideration of the particular architecture and data access requirements are needed to determine the best approach to roaming user support.
Disaster Recovery

If users are directed to their home site with component-level availability, there is still the potential of not enough capacity to support the user’s access requirements. This could be due to a large percentage of servers being offline or communication failing between components. Regardless of the reason, if the Web Interface is not able to provide the user with an available application or desktop, it can be configured to automatically fall back to a defined recovery site. The recovery site architecture could look like the following:

![Figure 6: Fallback Data Center](image)

The recovery site, defined by the administrator, allows the user to work with virtualized applications or desktops, but if the user’s data is not synchronized between the home and recovery sites, the user will experience a degraded experience. It is advisable for the administrator to configure the recovery site with a message informing the user of the potential degradation. By using automated recovery sites, the user will receive access to their virtualized resources, even in the event of a catastrophic failure. Use of automated recovery should be carefully considered within the context of an organization’s IT operational model and change management process. It needs to be determined whether it is optimal for an organization to have an automated failover and deal with potential issues with performance, especially if the failover is not required (i.e. no actual disaster) or if a semi-automated process should be used by manually reconfiguring site preferences.
Conclusion

As the desktop operating environment is moved from the endpoint into the data center, fault tolerance, high availability and disaster recovery are even more critical. If the infrastructure suffers a failure, a large percentage of the user population is impacted. Being able to provide fault tolerance at each level mitigates the risk of centralized computing. When providing high-availability to a desktop virtualization solution, the architecture must include availability at each of the three layers:

- Desktop availability
- High Availability
- Business Continuity/Disaster Recovery
Revision History

<table>
<thead>
<tr>
<th>Revision</th>
<th>Change Description</th>
<th>Updated By</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>Document released</td>
<td>Daniel Feller – Lead Architect</td>
<td>October 28, 2009</td>
</tr>
</tbody>
</table>

About Citrix

Citrix Systems, Inc. (NASDAQ:CTXS) is the leading provider of virtualization, networking and software as a service technologies for more than 230,000 organizations worldwide. It’s Citrix Delivery Center, Citrix Cloud Center (C3) and Citrix Online Services product families radically simplify computing for millions of users, delivering applications as an on-demand service to any user, in any location on any device. Citrix customers include the world’s largest Internet companies, 99 percent of Fortune Global 500 enterprises, and hundreds of thousands of small businesses and prosusers worldwide. Citrix partners with over 10,000 companies worldwide in more than 100 countries. Founded in 1989, annual revenue in 2010 was $1.9 billion.

©2011 Citrix Systems, Inc. All rights reserved. Citrix®, Access Gateway™, Branch Repeater™, Citrix Repeater™, HDX™, XenServer™, XenApp™, XenDesktop™ and Citrix Delivery Center™ are trademarks of Citrix Systems, Inc. and/or one or more of its subsidiaries, and may be registered in the United States Patent and Trademark Office and in other countries. All other trademarks and registered trademarks are property of their respective owners.