



Optimizing Citrix XenApp high availability – A new approach

Using tiered availability to achieve your SLAs with fewer resources and lower costs

Executive summary

Many businesses today are delivering applications to users as an on-demand service. This reduces the cost of application management and can improve application and data security. Citrix® XenApp™ empowers companies to achieve these results as well as other benefits. Challenges do however remain around maintaining application availability and protecting the integrity of your data. This white paper explores what availability really means, and how availability correlates to RTO and RPO baselines. The concept of levels of availability is introduced, which is an industry analyst driven definition for multiple levels of availability used to protect different applications from downtime to fulfill SLA commitments. XenApp and XenApp infrastructure components are analyzed to determine how they can be protected so you can achieve your SLA commitments. Marathon Technologies everRun® software is introduced with XenApp deployment samples depicting how everRun can be used to protect XenApp from downtime and data loss.

“everRun VM is a game-changing availability offering.”

Simon Crosby, CTO
Virtualization & Management Division
Citrix

After reading this white paper you will understand:

- Availability terminology and industry standard definitions for availability
- XenApp availability challenges and how to identify the level of availability you need to achieve your SLA commitments
- Marathon Technologies solutions and how they apply to protecting XenApp as well as virtualizing your XenApp farm to save money on hardware costs
- Deployment samples for protecting XenApp with everRun solutions

Business challenge

Business applications drive processes and enable people to be productive. More and more companies are delivering these critical applications to users remotely. The benefit to delivering applications remotely is that it allows a company to be more agile and efficient in rolling out applications. XenApp is a primary enabler for this business need. XenApp farms are set up to house business applications that get streamed to users as they require them instead of requiring applications to be installed and managed locally for each user. As the requirements grow to add more applications and users to the XenApp farms, you will need to add one additional server for every 30-50 users. As the datacenter environment grows, IT needs to be able to roll out and maintain XenApp servers more efficiently and cost effectively. Many companies are using XenApp for high value applications such as a CRM application for sales force automation, back office operation applications for order processing, billing, and shipping, and customer support applications. In all cases, as XenApp becomes a more strategic part of your business, the infrastructure that hosts your applications must be protected. In addition, the application silos and all the data must also be protected.

Defining availability objectives

Creating availability objectives is an important step in formulating your protection strategy. This is typically done by establishing Recovery Time Objectives (RTO) and Recovery Point Objectives (RPO) baselines. RTO and RPO are defined as:

- RTO (Recovery Time Objective) is the time it takes for an application to be running again (potential downtime)
- RPO (Recovery Point Objective) is the point in time at which you can recover data (potential data loss) in case of a failure

RTO and RPO baselines establish the SLAs you commit to for the overall company, business units or specific internal groups. You may have different application SLAs for different users within your company.

Understanding the terminology

When it comes to availability, simply understanding your options can be challenging. Vendors often use terminology loosely and inconsistently. Clarifying terminology is an essential first step.

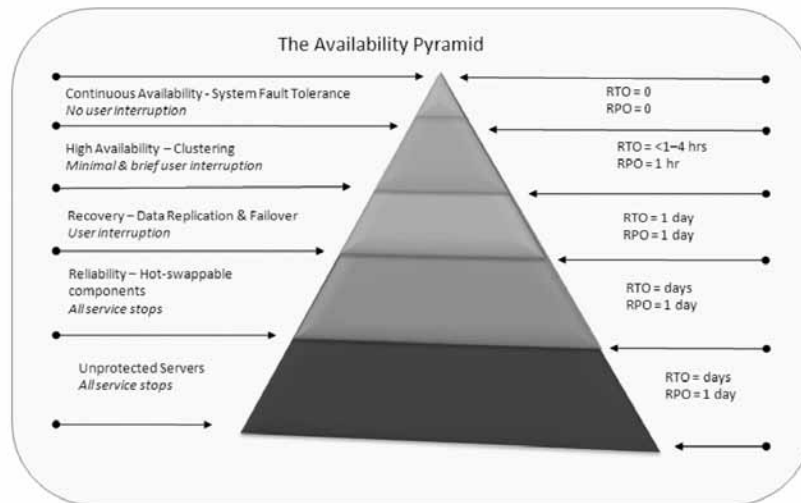
- **Continuous availability** – Continuous availability is the highest level of availability, offering non-stop operations without interruption during failures and outages. A continuously available system eliminates downtime, data loss and transaction loss.
- **Fault tolerance** – Fault tolerance is the ability of a system or a component to fail without interruption to the application or users. Fault tolerance is a means to achieving continuous availability.
- **High availability** – High availability (HA) configurations maintain a high degree of application uptime, minimizing downtime but not necessarily eliminating it completely for all types of failures. Applications should recover quickly from localized failures, with minimal or no interruption.
- **Disaster recovery** – Disaster recovery (DR) is the ability to recover systems and data after a major disaster (i.e., hurricane, tornado, flood or regional power outage). Recovery may take hours, days or even weeks, depending on the type of outage and how you have implemented DR.

Levels of availability

As you establish your availability objectives for all of your different groups of XenApp users, you more than likely will have varying requirements for your users and the overall XenApp farm. This emphasizes the need to have multiple levels of availability to assign to the components of the XenApp farm and XenApp user groups and, additionally, you need the flexibility to change those levels as your business changes. The concept of multiple levels of availability for different applications is portrayed in the pyramid below with the addition of recovery objectives to help determine where your specific needs fall.

“Teaming with Marathon enabled us to provide enterprises with a range of availability solutions for the Citrix XenServer platform delivering the best performance, protection and value.”

Lou Shipley,
Vice President and General Manager,
Xen Products Group
Citrix



The bottom two tiers of the pyramid describe an application that does not need availability planning beyond hot-swappable hardware components. While you may need to recover it at some point after a major failure, the business can continue operating with one of these servers missing. There is no hard RTO requirement.

Next, in the level labeled Recovery, are those applications for which recovery time (RTO) within a day is often acceptable. If one should fail, you want it to recover automatically if possible. But some downtime is acceptable, and even significant downtime won't have a detrimental effect on the business.

The next level, labeled High Availability, is the home of the majority of applications that run the business, such as e-mail, CRM, financial systems and databases. These are systems with high downtime costs and, therefore, short RTO requirements. Data loss is expensive—rolling back an hour can be costly. You need to minimize interruptions and be confident that recovery will be consistent. This level requires redundant systems and components.

The highest (and smallest) slice of the pyramid belongs to those applications with the most stringent RTO requirements, where even brief moments of downtime are extremely detrimental to the client or business. These may include process control, security systems, and trading and banking applications. Some businesses may not have any applications in this category.

Assigning levels of availability to your XenApp environment

You may want to consider assigning levels of availability to your XenApp environment and application silos. There are many components in a XenApp infrastructure, including but not limited to the Receiver Update Server, SmartAuditor and Data Store. All of the components are important, but some are more critical than others as they relate to XenApp availability, and would therefore require a higher level of protection. Based on a typical XenApp deployment, the following XenApp components may be flagged as particularly critical and potentially needing High Availability protection.

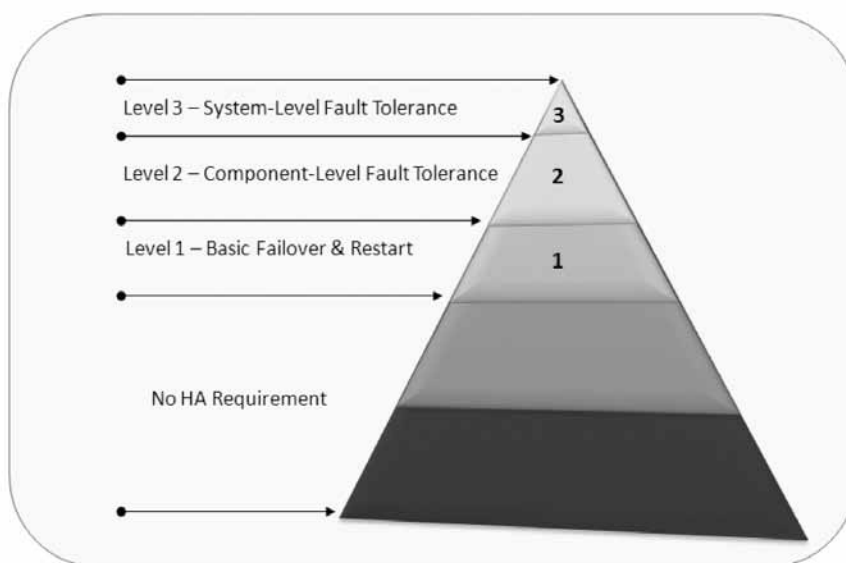
Citrix Web Interface is used as the entry point for users to launch applications. It is critical that the Web Interface always be accessible to users, or they will not have access to their applications and they will not be productive. The Web Interface is often deployed on two Web servers with a network load balancer, in order to create a redundant environment.

Citrix Data Store holds information critical for XenApp to function, and the XenApp farm cannot be managed without it. It is considered the brain of the XenApp farm. Some of the information that it holds is user load, printer driver information and session counts. The Data Store is often deployed on a SQL database with SQL clustering to protect it.

Citrix License Server is used to manage and distribute all licenses for all Citrix products to the entire farm. Users worldwide may access this server to obtain the licenses they require for the applications they need to be productive. This server is considered critical and should be treated as a major server. If this server should go down, users can only connect for a grace period of 96 hours or until the XenApp servers are restarted, which would be a considerable disruption.

Protecting application silos is a critical aspect to XenApp availability. Consideration should be given to the SLAs you have in place for specific business units. If you have a CRM program accessed by users with an SLA of no downtime, then Continuous Availability protection is recommended. If you have a CRM application that can have some downtime, then High Availability protection may be enough. In some cases, Basic Failover may be enough protection for an application. It is all dependent on your SLA commitments.

everRun – Protecting all applications at any level



Level 1	Level 1 protection offers basic failover and recovery, with restart capabilities.
Level 2	At the next level, everRun offers component-level fault tolerance, eliminating downtime due to I/O component failures and guaranteeing recovery from system failures.
Level 3	At the highest level, Level 3 Lockstep protection provides system-level fault tolerance for continuous availability of the most mission-critical

Level 1 – Basic failover

The first level of availability, basic failover and recovery, is ideal for applications that require high availability but where immediate recovery is not absolutely critical, and manual intervention, while not desirable, is acceptable.

Your environment is monitored and, if a host failure is detected, Level 1 protection attempts to restart the application on another host within that pool. Manual intervention may be required if, due to failed or unavailable resources on the recovery host, the application cannot be restarted.

Level 1 protection provides:

- Basic failover to another host within the same pool, with resource calculation to determine whether adequate resources are available within the pool to handle a defined number of simultaneous host failures. (Level 1 protection does not check the health of available devices, such as network and storage.)
- Monitoring of health of the hosts within a pool. No storage or data protection; using this level requires a shared-storage configuration.

Level 2 – Component-level fault tolerance

For applications with business-critical roles, everRun provides component-level fault tolerance: the ability to withstand the loss of an individual network or storage component without interruption or downtime.

everRun monitors all components within the system, using Marathon’s proprietary ComputeThru® technology. everRun creates and synchronously maintains a cloned application environment on a secondary host within the pool for full redundancy. everRun can detect and respond to a wide range of I/O and component failures, and identifies faulty I/O pathways before they are needed to be placed into active mode. everRun performs continuous Active Validation of all components on both the primary and secondary hosts to ensure proper recovery should a failure occur on one of the primary host’s components.

The attributes of Level 2 availability include:

- Automated fault management policies handle system, network and disk I/O failures without IT intervention
- Assured recovery of the application
- Zero downtime due to I/O failures
- Synchronous data mirroring between hosts; no need for shared storage
- Zero data loss
- Continuous active validation of all components on production and standby system to ensure complete redundancy at all times for recovery in the event of a failure
- Comprehensive protection for system, network and data availability, all in one integrated solution

everRun is appropriate for many business-critical applications, for which data loss and downtime are extremely costly. Customer Relationship Management, ERP, back-end databases and financial software are all candidates for Level 2 using everRun.

Level 3: System-level fault tolerance

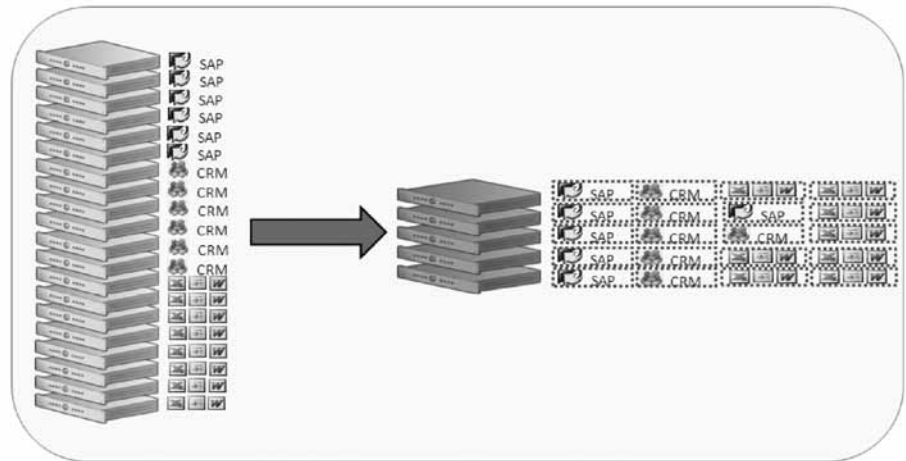
For the most mission-critical systems, Level 3 protection provides system-level fault tolerance, with continuous availability in the face of component or system-wide failures. Level 3 offers protection for systems that cannot experience any downtime and must maintain transaction state at all costs. (RTO=0, RPO=0)

Level 3 protection offers all of the benefits of Level 2 protection together with:

- Zero downtime even for complete host failures
- Application state maintained during failures
- Memory state maintained during failures

everRun protecting your XenApp environment

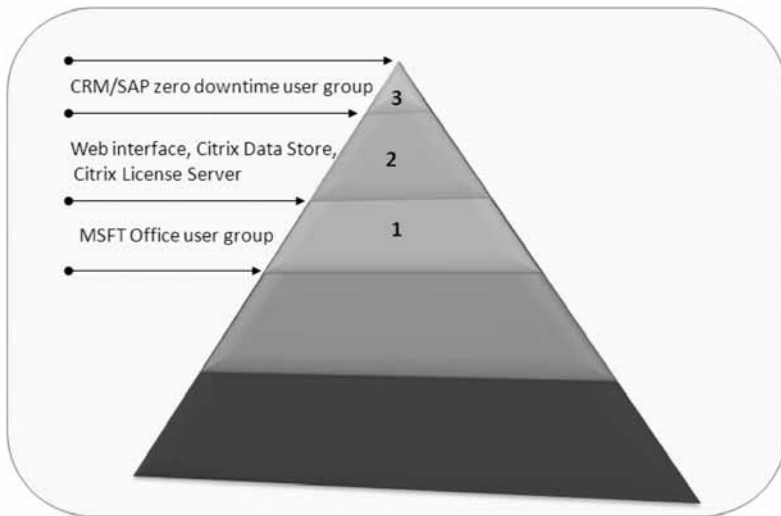
everRun has the capability to protect your entire XenApp environment and silos at the availability level necessary to achieve your SLA commitments. In addition, everRun can also resolve a challenge stated earlier of maintaining and managing the growth of XenApp farms. everRun runs on Citrix® XenServer™, which enables virtualization and hardware consolidation. For example, a company may have a silo with six SAP servers, six CRM servers and eight Microsoft Office® servers, supporting 500 users spread across 20 physical servers. XenServer can be used to consolidate the servers down to just five or six instead of 20 to support the same 500 users.



This significantly reduces the cost of hardware and enables efficient growth with less hardware required to add users.

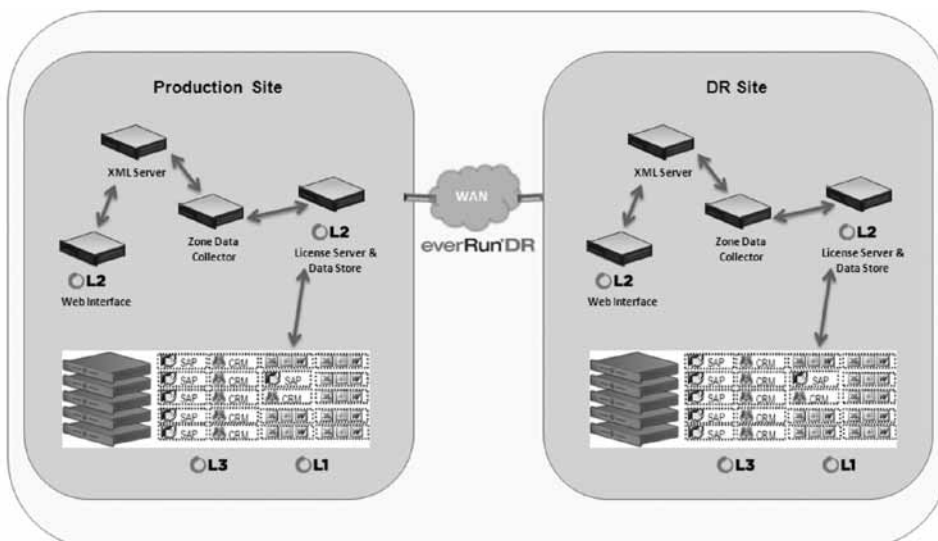
everRun can protect the entire XenApp farm from downtime by combining the resources of two physical servers to provide a completely redundant compute environment. Each component of the XenApp infrastructure protected by everRun can be configured as an independent virtual machine running on two servers. All components including the storage and storage devices, networks and even the servers are transparently paired and components automatically activated should a failure occur. For the most critical applications sensitive to even a moment of downtime, everRun can deliver zero-downtime for zero RTO requirements. everRun also provides an option to geographically separate the two physical hosts the virtual machines are running on. everRun SplitSite® enables this capability and although the hosts may be in different cities, any network or storage failure is automatically recovered with no downtime or data loss. This protects your XenApp farm from a regional disaster, fire or power outage at your primary site.

As was discussed earlier, there are many critical elements of a XenApp deployment that may warrant an extra level of availability protection. XenApp users with zero downtime SLAs for an application can be grouped together and protected at Level 3 for Continuous Availability to an application silo. Other users with less stringent SLA requirements can be protected at Level 2 or Level 1. Important XenApp infrastructure components can be protected at Level 2 for high availability protection for the XenApp farm. Protection levels in a sample deployment may be as follows:



everRun provides the flexibility to choose a level of availability for each component or application that needs protection. The XenApp farm can be protected against component and system failures with automatic detection and automatic responses based on the level of protection assigned. Additionally, everRun is flexible enough to adapt to your business changes. It is a quick, simple procedure to change protection levels up or down on a virtual machine if the need arises.

As a final safeguard, protecting your farm from a wide-spread disaster may also be a requirement. This may be done for user productivity purposes, as well as legal, regulatory and business requirement reasons. Disaster recovery ensures that your application silos and data are available at a remote location hundreds or even thousands of miles away. With everRun, you can implement a complete DR solution with remote replication and restart capabilities at the DR site. In the following diagram, all components and applications at both the production site and the DR site are protected at the same level to ensure a highly available environment should a restart be necessary at the DR site.



Summary

Given the increasing importance of Citrix XenApp to many organizations, achieving your XenApp SLA commitments has become essential. This white paper has provided the information you need to achieve your SLA commitments:

- Understand availability options and their objectives
- Establish RTO and RPO baselines for SLA commitments
- Determine the level of availability protection required to achieve your SLA commitments
- Protect XenApp, all components of XenApp and XenApp users at these levels

everRun provides a way to meet your availability and recovery objectives by offering multiple levels of availability in a single software solution. You can protect XenApp to meet your different SLA commitments even if they change over time.



Worldwide Headquarters

Citrix Systems, Inc.
851 West Cypress Creek Road
Fort Lauderdale, FL 33309, USA
T +1 800 393 1888
T +1 954 267 3000

www.citrix.com

Americas

Citrix Silicon Valley
4988 Great America Parkway
Santa Clara, CA 95054, USA
T +1 408 790 8000

Europe

Citrix Systems International GmbH
Rheinweg 9
8200 Schaffhausen, Switzerland
T +41 52 635 7700

Asia Pacific

Citrix Systems Hong Kong Ltd.
Suite 6301-10, 63rd Floor
One Island East
18 Westland Road
Island East, Hong Kong, China
T +852 2100 5000

Citrix Online Division

6500 Hollister Avenue
Goleta, CA 93117, USA
T +1 805 690 6400

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. Its 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 prosumers worldwide. Citrix partners with over 10,000 companies worldwide in more than 100 countries. Founded in 1989, annual revenue in 2008 was \$1.6 billion.

MARATHON

Run to Infinity

About Marathon

With more than 2,500 global customers, Marathon is the award-winning provider of automated, fault-tolerant high availability and disaster recovery software for physical and virtual servers. Marathon everRun® software prevents outages and data loss—without IT intervention or specialized IT skills. Organizations using everRun achieve continuous availability, 100 percent data protection, and rapid disaster recovery – all through automated click-to-protect operation. In the past two years, the company has been recognized with more than a dozen major industry awards including “Best of VMworld 2007 – New Technology”, “Search-ServerVirtualization 2008 Products of the Year,” CIO.com “10 Virtualization Vendors to Watch in 2009.” The company was also recognized as a 2008 Computerworld Honors Program Laureate, a winner in eWEEK's Eighth Annual Excellence Awards Program and finalist for 2009 Citrix Ready Solution of the Year. For more information visit www.marathontechnologies.com

©2009 Citrix Systems, Inc. All rights reserved. Citrix®, XenServer™ and XenApp™ 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. The Marathon logo, SplitSite® and everRun® are trademarks or registered trademarks of Marathon Technologies Corporation. All other trademarks and registered trademarks are property of their respective owners.