

Breaking through the Noise around Application Availability

Availability. Everyone is talking about it as apparently everyone can provide it, one way or another. Unfortunately, once you dig just slightly below the surface, it becomes apparent that there are nearly as many definitions for availability as there are vendors touting it. Some consider availability of the data while others availability of the server or storage subsystem. However, they all use the same buzz words like “disaster recovery,” “real-time,” “business continuity,” and “continuous” or “high” availability, which has caused a lot of confusion and misconception within the market on what all these things really mean. Many will even try to convince businesses that a few hours of downtime are acceptable. All the noise and hype have made life very difficult for those businesses looking to better protect themselves and their applications from downtime due to failures and outages.

It wasn't long ago that there were only a few essential systems that required a high degree of uptime, and those were generally only within very large organizations. Today, however, with the extremely competitive marketplace and the ever-growing compliance regulations, businesses of all sizes must maintain a high level of availability for their systems and applications.

When researching the options, it's important to cut through some of the noise and false claims to really understand their value and, more importantly, what they will actually do for your business.

At its core, availability is defined as “present and ready for use; at hand; accessible.” The gray area that is exploited by solution vendors is the degree, or timeframe in which systems should be “ready for use.” Within most any organization the necessary level of availability varies between different servers and applications. Some applications may require continuous availability for zero downtime while the majority can likely afford a brief interruption. Availability certainly is not a “one size fits all” solution. To understand their real needs, businesses should classify their systems by performing a business impact assessment and risk analysis to help define appropriate service levels. Once the businesses needs for availability are understood, appropriate solutions can be researched and identified.

Breaking through the Noise around Application Availability

There are literally dozens of products on the market that provide some level of application protection, and if you read the marketing literature, most of them sound about the same. For ease of understanding the myriad of products in this space, they can generally be categorized into three groups as follows:

- Fault-tolerant servers
- Shared-disk clusters
- Replication & Failover

FAULT-TOLERANT SERVERS

Fault-tolerant servers provide continuous availability through hardware failures by housing redundant components in a single chassis. These solutions are generally built on proprietary hardware and in some cases proprietary operating systems. This architecture requires additional training and maintenance procedures as well as additional hardware components to manage. It also comes at a cost that is often much higher than other solutions.

One of the major limiting factors of fault-tolerant servers is that they can not protect against site failures as all components are in a single chassis and can not be separated into different locations. Should the datacenter or building suffer an outage, the fault-tolerant server will also become unavailable, resulting in downtime for your applications.

Generally only 15–20% of businesses applications and data require complete fault tolerance and zero downtime offered by fault-tolerant servers. Companies would then be required to implement other solutions for those applications that don't fall into this category.

SHARED-DISK CLUSTERS

Shared-disk clusters are a common solution for providing application availability. There are only a few true shared-disk clustering solutions for Windows environments, as many that claim to be "cluster-like" utilize replication and failover architectures. The typical cluster model can consist of multiple servers (although are often configured with just two) connected to a shared storage device. Each server is fully configured, including OS and applications; one is "active" and the other is the "standby." When a failure occurs in the active server, the clustering software will start the standby application to take over processing. While they can't provide continuous availability, clusters can provide some level of availability by restarting applications on a standby server. Whether this meets the defined service levels depends on the particular application and its role within the company.

Breaking through the Noise around Application Availability

Clustering solutions have been around for a number of years however have not been widely adopted due to some fairly substantial downsides.

For one, the shared-disk model does not provide data redundancy or any level of protection of the actual data, imposing a single point of failure in the solution. In order to protect the data, you must implement a RAID system or complex and expensive SAN.

Secondly, clusters don't support just any application. Applications must be "cluster aware" to work in the clustered environment and to fail over and restart when necessary. This greatly limits the market for these environments as many applications simply can't support running in a cluster. Of course applications can be made "cluster aware;" however this is very costly as it can take weeks of development and implementation as well as ongoing testing and maintenance. This avenue is not generally an option for most businesses.

Lastly is the very complex nature of clusters. They require very specialized expertise to be implemented and properly managed. If this level of expertise does not exist within your organization, it is not recommended to implement a cluster solution.

REPLICATION & FAILOVER

This is the category that has caused the most confusion within the market, yet provides the lowest level of availability with a fairly heavy implementation process. Replication products are primarily designed to move data from one server to another server using an asynchronous model to allow for unlimited distances between the two servers. For business needs that require disk-to-disk backup over long-distance WAN, replication is certainly a viable choice. But since we are talking about achieving availability and not simply performing backup, using one of these replication products for availability will not likely satisfy the requirement.

As with clusters, replication products require two fully configured servers and simply restart the application on the standby server should the active server have failure. Unlike clusters, applications do not have to be cluster aware, although they do require extensive scripting to allow the failover process to happen and for applications to restart on the standby server. Due to their implementation of the failover mechanism, restart times can take upwards of 90 minutes. A 90 minute interruption is probably not what most businesses would consider "high availability."

Claiming "real-time" and "2 minute failovers" sounds reasonable, but it's important to understand the realities of their architectures to be aware of how they will actually perform in your environment.

Asynchronous replication solutions inherently lose data upon failure and failover. Even if few minutes of data loss may be acceptable to the business, it could result in corruption of the underlying database due to missing transactions. This corruption can take hours to manually rebuild from, during which time the application is unavailable to your users and customers. While a few minutes of lost data may sound acceptable

MARATHON

Run to Infinity

Breaking through the Noise around Application Availability

WORLDWIDE HEADQUARTERS

Marathon Technologies Corporation
295 Foster Street, Littleton, MA 01460
Tel 1.800.884.6425 / 1.978.489.1100
Fax 1.978.489.1101
Email: info@marathontechnologies.com
web: www.marathontechnologies.com

EMEA HEADQUARTERS

Marathon Technologies UK Ltd
Regus House, Trinity Court
Wokingham Road, Bracknell
Berkshire, RG42 1PL
Tel +44 (0) 1344.706.241
Fax +44 (0) 1344.706.242
Email: emea@marathontechnologies.com
web: www.marathontechnologies.com

at first glance, looking a bit deeper unveils the potential for excessive downtime and manual data recovery.

Disk-to-disk replication has its place, but it's not the best option to provide availability for your business applications.

A NEW STANDARD

With a very stagnant availability market over the past ten years or so, businesses have had few viable options for enabling the varying degrees of availability they require. Due to the increasing demands for data and application availability, along with the confusion within the market, an unfortunate trend of indecision has occurred; indecision on what to implement, resulting in insufficient protection for the company.

Marathon everRun™ defines a new standard in application availability, breaking away from the hype and hyperbole to bring real solutions and real value to a starving market. Using patented virtualization technology, everRun provides a single Windows application environment with complete redundancy of all underlying hardware and data, all while utilizing standard x86 servers. Marathon offers businesses Infinite Availability™ for all of their Windows applications without the need for customization or scripting, allowing businesses of any size and from all industries to realize the benefits of simple yet industrial-strength availability. everRun delivers:

- Software solutions that utilizes standard x86 servers
- Support for any Windows application without requiring modifications
- Simple deployment with minimal ongoing management
- Synchronous data redundancy for data and transactional integrity
- Long-distance separation of servers for protection against site failures
- A single Windows environment that requires only a single application instance
- Hands-off operation with very low maintenance
- Flexible solutions for high availability and complete fault tolerance

Marathon everRun cuts through the noise and hype and actually delivers as advertised; simple solutions for SMB to enterprise businesses that make Infinite Availability affordable, practical, and real. everRun provides the solutions to ensure your applications are always "present and ready for use," maintaining accessibility for your users and your customers.

Sound too good to be true? Contact us for more information or to take test drive.

We think you'll agree; everRun is as good as it sounds.

MARATHON
Run to Infinity