



WWW.NEVERFAILGROUP.COM

PREDICT · PROTECT · PERFORM

Continuous Availability

High Availability
Disaster Recovery
Data Protection



Business Continuity:
Choosing the Right Technology Solution

CONTENTS

CONTENTS	2
INTRODUCTION.....	3
WHAT ARE THE OPTIONS?	3
HOW TO ASSESS SOLUTIONS.....	5
WHAT TO LOOK FOR IN A SOLUTION.....	7
FINAL THOUGHTS	8
ABOUT NEVERFAIL	8
APPENDIX I.....	9

Introduction

The relentless expansion of the Internet has resulted in 24x7 demands on business globally. Developments such as web 2.0, mobile computing, and wireless hotspots mean that application and system availability requirements become more and more critical. In turn, the processes and tools required to protect those applications have evolved as well.

Today there are a myriad of technologies offering different approaches to data protection, application availability, high availability and disaster recovery. These technologies typically have at least one thing in common: they are IT-based solutions that are built to protect IT assets. When it comes to business continuity, it is imperative that choosing the right solution is a business decision based on the level of risk and disruption that can be tolerated by the different parts of the business.

For example, email is ubiquitous and preserving access to email through any type of disruption should be a priority, with 100% uptime the goal. Database applications such as sales order processing or on-line collaboration and content management may also require 100% uptime as the impact of downtime will be too much of a risk to the business. Other applications, such as purchase order processing, may demand no data loss, but a recovery time in the region of one hour may be acceptable. There may also be applications that are non-critical, where data can be recreated from original sources, or that are low risk and downtime measured in hours or even days is acceptable.

Business continuity requirements will vary according to business type and function. There is unlikely to be a "one size fits all" solution for all applications used in business.

Ultimately the risk to the business will be the driving factor. Assessing business need requires taking into account multiple factors. Data protection with extended recovery times may be acceptable for some functions, immediate data access for others. Protection through planned maintenance

may be vital in some instances, 100% availability through disasters for others. Technology selection must address gaps between business expectations and existing IT capability. Closing the Business Continuity gap ensures IT delivers what business expects.

This paper explores some of the factors which will govern the selection of the right solutions to deliver an appropriate solution for business continuity.

What Are the Options?

There are two approaches to business continuity: recovery centric or availability centric. Quite different technology is used to deliver the two approaches.

Today there are two classes of technology which can be adopted in a recovery centric strategy: backup or replication. Both are typically focused on data protection.

Ranging from legacy tape technology to continuous data protection, there are a complete set of backup technologies that will protect data. Whether held in tape format or on disk, recovering from a backup will require rebuilding databases and file systems then reconnecting with applications, which themselves may need rebuilding. Although backup technology can approach a Recovery Point Objective (RPO) of zero data loss, a Recovery Time Objective (RTO) measured in seconds will not be achievable. This is because of the focus on data protection and the separation (or lack of) application protection. Of course, backup provides great flexibility for Disaster Recovery as tapes can easily be protected off site, and shipped to alternative sites on demand, but recovery of the business service will likely take days.

Today replication is rapidly becoming an alternative approach for availability. Host or storage-based replication allows exact copies of operational data to be taken. Synchronous replication provides for no data loss, but considerations such as

performance, cost and bandwidth requirements for off-site protection must be taken into account. More widely spread is asynchronous replication, which has much lower operational implications and provides near zero data loss. The only loss would occur from potential transactions in flight at the time a failure occurred.

The big attraction of replication is that data recovery is not required. The online copy of data can be used immediately for failover. This is likely to require manual intervention, or significant scripting, and may require applications to be rebuilt. There is also a risk that application datasets may be missing from the replica copy if administrative processes have broken down and application upgrades have failed to be identified to administrators.

Protecting data off-site for Disaster Recovery also requires closer consideration. There will be bandwidth considerations, and remote systems must be available to

hold an operational copy of the data.

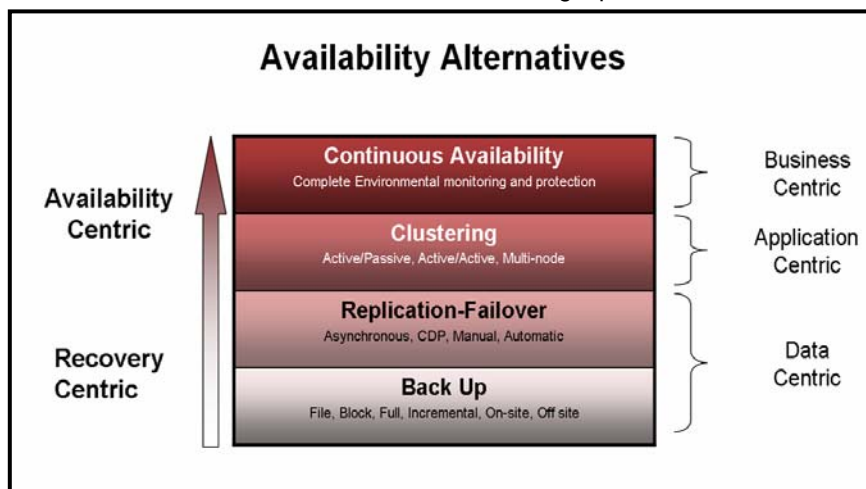
A recovery centric strategy will, by definition, be disruptive to the business. Recovery centric approaches are applicable to less important applications as business services will stop while recovery takes place. Although the level of disruption will be reduced with a replication/failover solution, it will still not be suitable for delivering an acceptable level of availability for mission critical applications. For such applications, an application or user centric approach is required.

Historically such approaches have depended on clustering technology. Clustering allows several

machines to run the same copy of the application which is accessing its data on shared storage. Clusters may consist of multiple physical and/or virtual machines and provide a platform that protects against physical or virtual machine failure. In some situations, it may also address availability for planned operations where individual machines in the cluster may be disconnected, allowing maintenance to take place.

The limitations of cluster centric approaches relate to application and processor failure. Failure situations that address the whole site, such as natural disasters, power outages and facility upgrades are not covered. Because clusters rely on shared storage and shared facilities, it is important to guard against failures at that level. In turn, this means protecting the storage from being a single point of failure. This can be costly,

requiring storage virtualization and/or replication to be implemented concurrently. Additionally, virtual clusters may suffer from corruption of shared application images.



Provisioning applications across machines from the same virtual image will not guard against application corruption, and not allow application maintenance, thus limiting the level of high availability that can be delivered.

As mentioned in the introduction, there is an increasing realization that there is a disconnect between the reliance of the businesses on business critical applications and the IT approach to business continuity. The business continuity gap exists because the solutions discussed above ignore the needs of the end-user- uninterrupted

access to applications regardless of the cause of failure.

Results of a recent survey indicate that in regards to email, over half of organizations depend on the users to notify IT of an issue. By this time, email access has been interrupted. Addressing the needs of the user has resulted in a new discipline of Continuous Availability.

Continuous Availability solutions typically use redundancy of data and hardware, combined with data replication, in a “shared nothing” approach. While replication solutions share this approach, the difference comes when looking at the impact on the user, and hence the business. Continuous Availability solutions will provide pro-active application awareness.

Application availability will be monitored through embedded best practice facilities with a degree of self-healing provided, changes in application configuration and data dependencies will be catered for, and automation will be an option to avoid the need for manual intervention. The level of protection will embrace the end-to-end service, not just an individual software component such as Exchange.

The choice of availability strategy will depend on many factors. Taking into account complexity in operation, total cost of ownership, skills available and the risk to the business of failure may mean combinations of the above technology are required to address business risk.

How to Assess Solutions

When looking at mechanisms to protect applications, any IT decisions need to be based on a firm foundation of business risk. It helps to look at application availability solutions in the context of four pillars of risk.

Recovery Profile

How much business disruption is acceptable? Will a backup/recovery based approach deliver against the Recovery Point and Recovery Time

Objectives? The definition of Recovery Point is data based; how much data loss is acceptable?

A daily backup may lose 24 hours worth of data while a snapshot approach may lose only 15 minutes of data. Replication technology will deliver no data loss, if synchronous, or limit data loss to in-flight transactions, if asynchronous. But recovery is not limited to data. How long will it take to get the business up and running again?

Operating systems and applications will need to be rebuilt. Recovery Time Objectives should focus on minimizing or eliminating business disruption and should address data and application availability requirements.

Scope of Protection

The scope of protection directly affects the level of business disruption that can be tolerated. Limiting the scope to data backup accepts that recovery will be required, there will be data loss and there will be significant disruption to affected business services as data and applications are rebuilt. Implementing replication based solutions will eliminate disruption from loss of data, but applications will still need rebuilding and users will require reconnection. Manual intervention will be required, but the business downtime will be reduced.

Implementing cluster technology provides maximum protection against business downtime caused by server hardware failures, but site outages, data failures, application corruption and user errors will all cause significant business downtime. Outages come from network failures, processor load, data loss, application issues, human error and any number of other reasons. It's also worth remembering that protecting email is not just about protecting Microsoft[®] Exchange or Lotus[®] Domino[®]. Email, as a business service, needs to embrace anti-virus and anti-spam tools[®] and mobile platforms such as RIM Blackberry[®].

Understanding the risk means understanding availability of these various components, and the

gaps in protection that will bring business downtime.

Operational Capability

Not every organization has the expertise available around the clock to deal with outages at multiple levels. They may not even have the expertise and processes to ensure data is protected in the first place. Application administrators may introduce new databases or files to be protected, but unless these administrators are also responsible for high availability, will they remember to request that data protection be added? Will administration of the backup or replication regimes be updated accordingly? When failures occur and the pressure is on, are experienced personnel available to be relied upon to take the right action.

Furthermore by adding a second (failover) server into any environment, IT staff must also consider the procedural changes necessary to support the new server. Even the smallest, seemingly harmless configuration change to one server may affect the reliability of failover operations. Furthermore, changes elsewhere in the IT environment (for instance to network routing tables or IP subnetting) may also have an impact on operations. Unless the availability solution is designed to account for such changes automatically, you may in fact be implementing nothing more than a false sense of security.

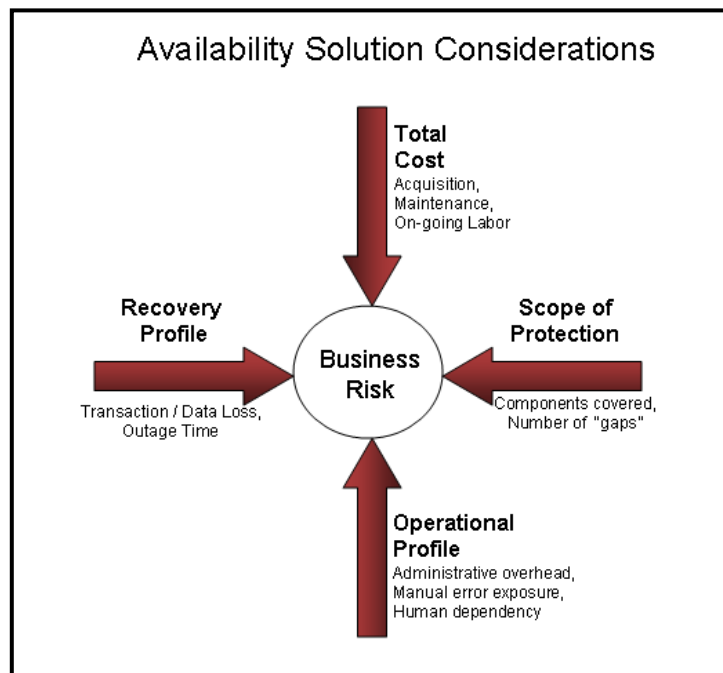
If user reconfiguration is required, how skilled are the users themselves? The operational capability may demand complete end-to-end automation. Only a minority of organizations do full scale

disaster recovery testing because of the complexity and risk involved. Of those that are tested, it is not uncommon that the test fails, again due to complexity.

Total Cost of Ownership

The true cost of availability comes at many levels. Upfront costs are important. If existing hardware can be re-used, the ongoing cost will be reduced significantly. Solutions which replicate data asynchronously and which offer advanced data compression can also keep bandwidth costs at a minimum, dramatically impacting recurring monthly costs. The implementation costs are equally

important. If significant effort is required pre-install and configure software on failover systems, make changes to DNS or Active Directory topologies, develop custom failover scripts and further customize the implementation to account for installed auxiliary applications, these will require upfront investment and on-going maintenance costs which should be factored in.



Once the solution is in place, if application configuration changes on one server need to be duplicated manually on the failover system this will incur additional personnel costs (especially if the failover system is located remotely). Where the chosen solution requires manual operation, what are the incremental costs of training and maintaining personnel on site, or at least with remote access? Ultimately, what is the true cost of downtime to the business, where cost is not just associated with lost productivity, but also with lost opportunity and reputation?

What to Look For in a Solution

Selecting an appropriate solution means considering multiple options. Ultimately, different solutions may be required for different business functions. It's important to have a clear understanding of the user and application mode of operation and the relevance in the context of availability.

The chart below is intended to be a quick reference for the key areas to consider before, during and after a business and technology review.

This chart is not meant to be a recipe card for choosing one set of technology addressing Business Continuity. It may be that a backup based solution is suitable for some less critical applications whereas mission critical solutions need continuous availability. In between, there may be semi-critical applications that need a replication/failover solution. In the end, the business and operational needs will drive the decision. For more information and more detail, Appendix I contains a series of bullet points which will be useful when addressing requirements and evaluating solutions.

		Backup	Replication/ Failover	Clustering	Continuous Availability
Operational Profile	No Business Disruption			●	●
	Users not impacted: No Service Interruption				
	No Reconfiguration			●	●
	Manual changes not required: Standby system ready-to-go				
	Continuous Connectivity			●	●
	No client reboot: Users remain connected. No Application Restart				
	Automated Operation			●	●
	No manual scripting; Unattended failover; Automated Discover				
Scope of Protection	Continuous Operation			○	●
	Health check; Resource monitoring; Fault correction: Planned Maintenance				
	Recovery Not Required		○		●
	Immediate failover: Seamless switchback: Synchronization				
	Configuration			○	●
	Validation and Monitoring				
	Server		○	●	●
	Hardware and OS monitoring and protection				
	Data Protection	●	●		●
	Replication. Corruption Recovery				
	Application			○	●
	Monitoring. Healing. Configuration. End-to-End				
Recovery Profile	Network				●
	Access monitoring, protection, optimization				
	Performance			○	●
	Monitoring. Correction				
	Disaster	●	●	○	●
	WAN aware. Secondary Site support				
Total Cost	Recovery Point	Hrs:Days	Sec:Min	Sec:Min	Sec:Min
	No data loss. Application and Data Protection				
	Recovery Time	Hrs:Days	Min:Hrs	Sec:Min	Sec:Min
	Planned. Unplanned. Disaster. Application and Data. Business downtime.				
	Total Cost of Ownership	\$\$\$	\$\$\$	\$\$\$\$\$\$\$	\$\$\$\$\$
	Software. Hardware. Implementation. Management. Business Impact.				

○ Partial Solution
● Full Solution

Final Thoughts

Data Protection, High Availability and Disaster Recovery are all important constituents of Business Continuity. Combining the best attributes of these disciplines will make the difference between a full Business Continuity solution addressing the range of applications in use, and one with gaps in expectation and delivery.

Critical applications, ranging from email and websites to databases and mobile information platforms, are in continuous use and need to be continuously available. Continuous availability demands that these applications are highly available, their data is continuously protected and that in the event of planned or unplanned IT outages (including disaster scenarios) they continue to operate without user disruption. Other applications may require lower levels of protection based around backup and/or failure.

One thing is clear: there is a mission critical class of application for which legacy discussions about Recovery Point and Recovery Time Objectives alone are inappropriate. Legacy approaches to availability that rely on clustering and data recovery strategies are no longer acceptable for mission critical applications.

About Neverfail

Neverfail provides affordable software that delivers Continuous Availability and Data Protection for critical applications. Neverfail's predictive approach protects businesses during planned maintenance and IT outages. Regardless of the nature of the problem, from a single system component failure to full site disaster, critical business applications will continue to run without disruption.

Continuous Availability brings together High Availability, Disaster Recovery and Data Protection software with best practice monitoring and automation facilities to avoid disruption caused by failures affecting critical applications such as email, databases, BlackBerry® and websites.

Appendix I

Points to Consider when Evaluating Business Continuity Technology.

1. What impact does the solution have on the end user? This could also be considered under the Recovery Time section below, but it is important to understand that system outages that impact end users, even for a short period of time, have operational Impact on IT. Help desk calls are made. Staffs are disrupted and possibly even pulled from their homes in the middle of the night to "fix the problem"
2. When outages occur IT is under a lot of stress to bring systems back up quickly. Some solution may require manual configuration change before recovery can occur. The skills and knowledge required need to be understood.
3. Do "restarts" have to occur in the environment to complete a recovery? This might need to occur on a Server as well as a client.
4. Requiring a server to restart before a recovery can occur introduces another level of risk into a solution. What happens if the "reboot" fails? Client restart might mean loss of data on the individual's machine and will be disruptive to the user
5. Does the solution require manual scripting? This is particularly prevalent in replication-failover solutions. Manual start/stop scripts which must be maintained can cause an operational burden over time as environments change. A script which is out of date can cause a system to simply not work. And this might not be discovered until an actual recovery is necessary.
6. Any change in a steady state of a system means an increase in risk. Change = risk. With this understanding it is probably better to avoid having to incur a failure than recovering from one. Fault avoidance is important.
7. Solutions with some form of pro-active intervention to detect and correct problems before a recovery action is required will reduce the overall risk to a system and reduce operational overhead.
8. Recovery is one aspect of availability, but often overlooked is the overhead and impact to an organization after an outage and what actions must be taken to restore the system back to its pre-failure state. In some instances the risk and overhead of configuring an environment back to its "normal" production state can be significant as operating systems and applications must also be restored, and the backup cycles for these may not be synchronized with data backups.
9. System Configuration is one aspect of a system which can often cause availability problems and is many times over looked is evaluating availability solutions. A misconfiguration of the environment can actually cause an outage. Many times these types of issue can be proactively identified before they cause an issue. Configuration validation and monitoring is a key element of availability.
10. Monitoring and protection of the Server OS and hardware is fairly standard on most availability solutions. Although even in here the level of sophistication on the solution can vary widely. For instance some solutions might perform a basic ping against a server to ensure that server is available and consider this "server protection". In fact a server might be having all sorts of issues that are causing an outage for the user and still be able to respond to a ping.
11. Data protection is the core function required for availability, but data protection is not just about recovering from system failures, data corruption can be an issue. Copying bad data can be as bad as not copying any data at all. So solutions should have some form of rollback facility which enables recovery from corrupt data.
12. Application monitoring and performance is critical if complete high availability is required. Application slow downs or an application process hanging is the same as a server crash from the

user perspective. Clustering solutions will deliver on this within the limits of their monitoring abilities, Continuous Availability solutions certainly deliver in this area.

13. Networking is another key component in the scope of Protection needed for availability. Servers can be running fine, but if users cannot access the server that has the same effect as a hard server crash to the organization. In many instances, network monitoring is a completely separate discipline and the user is responsible for integration and leverage of the different technologies for availability in their environment. A solution focused on Continuous Availability should not leave this action to the user but should monitor network access and take corrective actions if there is a problem.

14. Performance is a key area for availability. Many time slow downs and stoppages in performance can be occurring within an environment and the first time an IT professional knows about it is when an irate user calls. Understanding what is happening within a system can be as important as knowing if the system is up. Clustering solutions can deliver some visibility into the performance of the cluster. To deliver Continuous Availability a solution must have the ability to monitor and correct performance problems as well. For complete protection monitoring system resources is not sufficient, the ability to validate the end user experience through test transactions should be considered

15. Obviously one of the significant components of availability is the protection from a disaster. Disaster recovery protection is important enough that it has spawned its own industry, but essentially site protection is just another component which must be considered when evaluating availability alternatives. Most data protection solutions can deliver site protection by duplicating data to a secondary site either on a streaming basis or simply shipping back-ups off site. Clustering solutions have been traditionally difficult to use for site protection as “stretch clustering” is a very expensive and a complex proposition. A Continuous Availability solution must

address site protection as well. The impact of data replication available bandwidth must form part of the assessment.

16. The desired recovery point will vary across solutions. Backup typically has the greatest lag. Clustering may have the shortest unless there is physical data loss and/or corruption in which case replication and/or backups may be required. Some replication-failover solutions deliver periodic snapshots so a recovery point might be several minutes. Generally speaking solutions which deliver on-going asynchronous replication will have recovery points measured ranging from a few seconds to minutes.

17. Recovery time is how long it takes to bring a system back to full operational state. The greatest range in recovery time falls within backup and replication-failover solutions. Depending on the replication-failover solution and the amount of manual intervention required to execute a failover, recovery times could take a few hours or just a few minute. The solutions focused on Availability rather than data protection will have the shortest recovery times.

18. The final consideration when evaluating availability alternatives is total cost. Solutions can range from a few hundred dollars to hundreds of thousand of dollars. In some instances the on-going administrative overhead for a solution might be considerably greater than the initial acquisition cost so a thorough understanding of the on-going support and maintenance of a product is essential.