



2

IN A SERIES
OF BUSINESS
CONTINUITY
WHITE
PAPERS

“Pushbutton Recovery”

By Jon William Toigo
CEO, Toigo Partners International
Chairman, The Data Management Institute

 CA XOsoft

ABSTRACT

The holy grail of business continuity is the ability to failover—at the press of a button—the complete IT infrastructure required to support a mission critical business process. With the increasing commoditization of infrastructure components and standardization of communications protocols, this dream seems to be within reach. However, planners are discovering that complexities at the application layer and at the storage layer continue to obfuscate their efforts.

At the application layer, there have long existed impediments to efficient failover in the form of hardware-centric middleware approaches selected by application developers without concern for their impact on recoverability. Client-server application designs that use hardware-centric message controls linked to MAC addresses, machine IDs or IP addresses to communicate between application components often force planners to replicate infrastructure on a costly one-for-one basis at the recovery facility—a costly approach that is prone to failure with each change to hardware in the production environment.

Physical clustering configurations and virtual server environments, while touted as more resilient, also introduce complexity and risk into the disaster recovery strategy. They limit the manner in which recovery can be accomplished over distance and also the options available to the planner for designing a recovery environment. For whatever these application architectures may contribute to operational efficiency in the production shop, they can become a nightmare in a recovery situation.

Storage itself resists commoditization. Vendors are keen to embed their own proprietary replication approaches onto hardware to lock in the consumer, lock out the competition and justify huge mark-ups on what are essentially boxes of commodity disk drives. This situation also adds complexity to continuity plans, often requiring the management of multiple processes for failing over application servers and software on the one hand, and storage components on the other.

These constraints are driving interest in “wrapping” continuity approaches, using third party software wrappers to provide a unified approach for managing the multiple processes involved in continuity failover and to deliver a holistic pushbutton recovery solution on a process by process basis. CA XOsoft is a leader in this space.



jtoigo@toigopartners.com

INTRODUCTION

Increasingly, companies are seeking to ensure that their always-on business requirements are supported by a robust IT failover strategy that is closer in concept to high availability than traditional disaster recovery methodology. The idea is simple: for each key business process, there are supporting applications and infrastructure. When a catastrophic interruption occurs in production environment, these applications and their associated infrastructure need to failover automatically, or at the push of a button, to an alternative location. That way the critical business process can continue without protracted interruption.

It's a great idea, but one that has been challenging—technologically and financially—to implement. Fortunately, affordable tools are beginning to enter the market and other support elements, including access to less expensive broadband networks, are propelling the strategy to center stage.

Of course, distance failover isn't required (or even appropriate) for every business process: only for those that are critical. This paper examines the requirements for distance failover in business continuity planning as suggests parameters for building such a capability within business IT infrastructure.

TABLE OF CONTENTS

Introduction	3
Distance Failover: HA meets DR	4
About High Availability	4
Defense in Depth	5
Angels in the Architecture, Devil in the Details	5
CA XOsoft as a Wrapper	6

DISTANCE FAILOVER: HA MEETS DR

Given the risk of regional disasters, companies seeking always-on business continuity need to place their recovery environment at considerable distance from their primary or production environment. There is no commonly agreed upon "minimum safe distance," but the rule of thumb in continuity planning is 50 miles or 80.45 Kilometers. Separating primary and recovery facilities by this distance usually ensures access to alternative utility services and wide area networks that will not be impacted by the same disaster that may affect the production environment.

"Distance failover," as the strategy is called, requires careful planning and provisioning. At the outset, the following requirements must be satisfied:

1. The planner must know what will be failed over: he/she must know what business processes are critical and which data supports each process. The planner must also know how the application is hosted in the primary environment and how it will be hosted in the recovery environment. And he/she also needs to know where the data used and produced by the application is stored in production and how it will be stored for access by applications in the recovery setting.
2. The planner must know how a failover will be triggered: what is the logic that will guide failover? Will all business critical infrastructure fail over under all circumstances, or will selected components fail over in the face of differing interruption scenarios? A more flexible strategy may take some of the cost out of recovery by limiting the number of personnel who must migrate to the recovery site.
3. The planner must know where processes and support infrastructure will failover to: a branch office, an ISP, a hot site, a company-owned recovery facility? For distance failover to work, planners must know in advance their failover target environment so that it can be maintained in a constant state of readiness.
4. The planner must know how the environments will remain synchronized in terms of hardware infrastructure capability (bandwidth, capacity and processing capability), software (version and patch level), and data (state and platform). This means monitoring change in the production environment and ensuring that necessary changes are also made in the shadow infrastructure. It also means establishing and monitoring ongoing asynchronous data mirroring across the WAN.
5. The planner needs to know how the strategy can be tested without disrupting normal operations.
6. Finally, planners need to know how fail back will be accomplished. Once disaster conditions have subsided, there needs to be a way to transfer operations back to

the production environment that will not create a second disaster in the process.

These are the six essential ingredients of planning for a distance failover strategy. They are commonsensical, as is most continuity planning, but they are also more challenging than they seem, especially the first step. Defining business process criticality may be simplified by having senior management indicate which processes require, in their view, always-on operation. However, considerable additional research and investigation may be required to break down the processes into tasks and workflows in order to identify interdependencies with other processes that confer upon ancillary processes critical status as well. Investigation is also required to identify applications and their data assets and infrastructure support so that the right protection targets are identified and the right resources are provided at the recovery site.

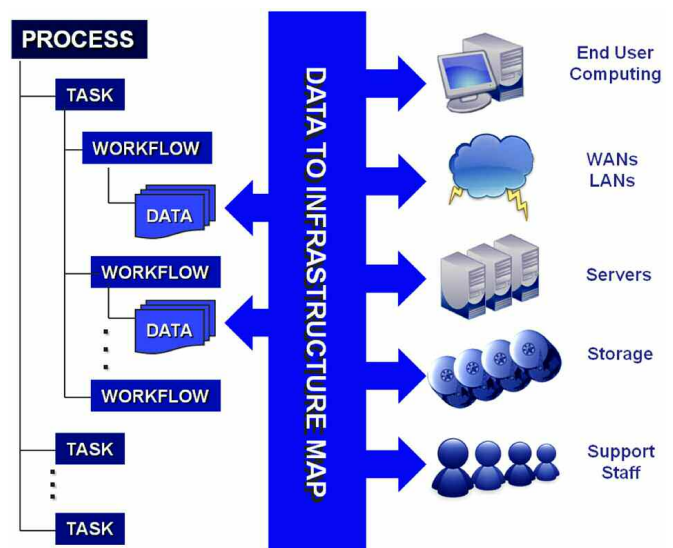


Figure 1

ABOUT HIGH AVAILABILITY

Traditional high availability (HA) concepts apply in architecting all distance failover scenarios. Specifically, there is a need for:

- Simple logic to guide the failover decision
- Instrumentation of infrastructure with "heartbeats" to help trigger failover scenarios
- Suitable hosting platforms (both hardware and operating system environments) at both locations whose consistency is maintained over time

- Appropriate application software at each location, properly patched and updated over time
- Consistent data sets at both locations with replication and validation processes operated on an ongoing basis
- Network switchover capabilities, including voice and data networks used by customers, workers or others to access application interfaces to conduct business
- Holistic visibility into all of the above for management and monitoring of the architecture over time

Many vendors talk about the high availability characteristics of their products. However, these discussions tend to focus on local protection scenarios—failovers that occur within the same sub-network, physical plant or inside a piece of hardware. In short, what are commonly represented by vendors as “high availability features” of their hardware or software products are usually not appropriate for distance failover scenarios.

Localized high availability capabilities have their place, of course. They can help to ensure that a comparatively minor interruption event such as a failed disk drive, server, array or switch or an application or operating system glitch does not cascade into a larger problem.

DEFENSE IN DEPTH

Localized HA may well be part of a “defense in depth” strategy for business continuity that provides both the capability to cope with a localized failure, and also recovery from a facility-wide or region-wide disaster. One way to conceptualize defense in depth is as a set of flexible responses to different types of interruption events ranging from data corruption or disk failure (which may be addressed with RAID and internal hardware redundancies), to array failures (which may be addressed by array mirroring and tape backup), to access interruptions and facility outages (addressed through resource replication and self-healing networks), and finally to milieus or regional disasters (requiring distance failover).

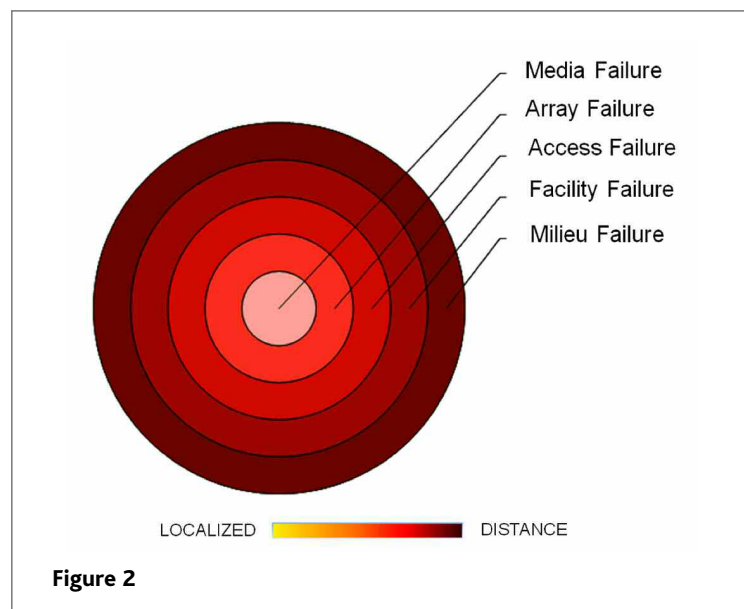


Figure 2

Distance failover is the counterpart to localized high availability but, as the name implies, over distance. Scenarios in which distance failover may be required run the gamut from building fires, power outages and distributed denial of service attacks to civil disturbances, terrorist attacks, severe weather events (hurricanes, ice storms, regional fires, etc.) and chemical or biological events (an overturned tanker truck carrying chlorine, for example). In the face of such events, the company must be prepared to continue operations from a site that is outside the “event horizon” of the disaster.

ANGELS IN THE ARCHITECTURE, DEVILS IN THE DETAILS

Distance failover can be complicated by many factors. First, applications themselves may not have been designed to accommodate such a strategy readily. Complex *n*-tier client/server applications, whose components are stitched together with middleware that establishes message passing conduits by hardcoded server addresses (either machine ID- or IP address-specific remote procedure calls), can be difficult to configure for failover. In such cases, equipment configurations must be replicated on a one-for-one basis at the recovery site, and maintained in a parallel state on an ongoing basis. Just maintaining two sites in hardware congruency can be a Herculean task.

Custom *n*-tier client/server environments are commonly found in companies that have built their own applications over time and under the auspices of several designers, each with his or her own preferences in architecture and middleware. However, even in environments where commercial software products have been used, problems may exist. This reflects the competitive nature of the Enterprise Resource Planning (ERP), Manufacturing Resource Planning (MRP), and Customer Relationship Management (CRM) software market. When one vendor introduces a new feature, the competitors must quickly add comparable functionality to their wares (usually within one quarter). The path typically taken by vendors is to purchase the necessary technology from another independent software vendor and bolt it to their core product using whatever type of middleware that works. The result is much the same as home grown *n*-tier client server applications: a system held

together by a combination of hard coded and message oriented middleware that is difficult to provision for HA failover.

A second obstacle to distance failover may be HA features already deployed in the local environment. Many vendors, including Microsoft, have developed failover clustering architectures for their applications to protect against localized interruption events. If server A fails, workload is shifted to server B, which is part of the same cluster. The devil in the details has been that the two servers share the same storage, creating a single point of failure in the solution. When applied to failover over distance, the method requires a second process for data replication that may not be synchronized to provide a realistic recovery.

Server virtualization vendors, such as VMware, have also begun to gain mindshare in the market because of their products' purported benefits in reducing server sprawl and simplifying resource management. However, these products often come up short as guarantors of application high availability and provide an incomplete solution, again related to shared storage, for distance failover.

Both physical clusters and virtualized operating environments can add complexity to distance failover strategy when they are not replicated in the recovery environment exactly. Many consumers have clustered email servers, for example, in their primary environment but have a virtual environment in their recovery center that will serve as their failover target. They are concerned about the impact of having to reconfigure application hosting platforms on the fly if distance failover is required.

The bottom line is that HA features built into applications today can create obstacles to the cost-efficient and operationally-effective continuance of always-on business processes. Something more is needed; specifically, a way to

"wrapperize" the entire infrastructure supporting the business process, with all of its component (storage, server and network) replication and redirection processes included.

CA XOSOFT AS A WRAPPER

CA XOssoft is a leader in what we term distance failover. The product family provides what amounts to an "infrastructure wrapper" around a set of application, data and hardware to enable its failover between local and remote environments.

CA XOssoft supports defense in depth strategies in several ways. Using the product, planners can:

- Build a flexible set of scenarios for failing applications and their infrastructure between known points that enables partial and complete failover response in the face of different disaster events.
- Consolidate the monitoring and management of the many replication processes (including backup, mirroring and continuous data protection) that exist in the environment, via either through direct application programming interfaces (API) with other CA and third party applications or using CA XOssoft's script language.
- Test and refine failover strategies non-disruptively and across non-like infrastructure (physical to virtual, for example) under certain circumstances.

With CA XOssoft's native data replication functions, planners can bring sanity and order to the myriad processes that need to be coordinated to accomplish distance failover within the timeframes sought by corporate management. Separate scenarios can be developed on a process by process basis and managed and tested on an ongoing basis to move the organization closer to the realization of "pushbutton recovery" for always-on business requirements.



Jon William Toigo
CEO, Toigo Partners International
Founder, The Data Management Institute
1538 Patricia Avenue
Dunedin, Florida 34698
jtoigo@toigopartners.com