

July, 2007
A White Paper by Stratus Technologies

All Data is Not Equal



All Data Is Not Equal

Popular opinion maintains that backup data doesn't have the same importance as the original information. But is that a true assumption? Perhaps the answer is "it depends on the business situation." That is, sometimes backup data is indeed just another form of business information – and sometimes it is mission-critical.

Classifying the value of information to a business, understanding information service objectives, and then steering information to resources to achieve those objectives is what Information Lifecycle Management (ILM) is all about.

No where is ILM more relevant than availability management and disaster preparedness. The difference between these two concepts is small but significant.

Availability management focuses on all of the IT-related elements of keeping a business workflow going. A corrupted file (e.g. because of a logic error in a program) or lost data (e.g. hardware error or, more likely, an operational error) constitute issues of operational availability. The line between a "simple" availability problem and a disaster is usually the magnitude of the problem and the expected duration.

Disaster preparedness identifies the large threats and vulnerabilities that affect a data center and ameliorates their effects through countermeasures. As planners, our objective is to develop countermeasures to as many vulnerabilities as possible ... for as many workloads as possible ... based on probability of occurrence and its effect on business.

ILM and its emphasis on service level objectives enables planners to establish priorities and anticipate resource requirements during contingencies. Without a thorough understanding of the application and the information used by the application throughout its lifecycle, it is virtually impossible to set recovery points, recovery time objectives and meet availability requirements.

Time is the Enemy

Installations are frequently tempted to divide work into "mission-critical" and "other" categories. No one wants to allow their application to occupy the "other" category because that implies lesser importance with respect to business criticality. However that also ignores the fact that, as time passes and the duration of the outage lengthens, the pressure to restore normal service escalates to the point where even test and development work will become "mission-critical."

The polar opposite of the two category philosophy is to treat all information the same. The problem here is that the "80/20 rule" really does apply to most installations. Only 20% (or less) of the information is really likely to be time-critical in most cases. Treating everything the same way is expensive and wasteful of valuable time.

Time truly is the enemy of availability. In cases where access to data is especially time-critical, business continuity measures may even be established to sustain the processing environment throughout the disaster. It is no accident that the metrics associated with availability are all time-oriented.

Given all of the incidents that have occurred in recent years – tsunamis, hurricanes, terrorist bombings and monsoons – it hardly seems necessary to justify spending on disaster preparedness. However, studies continue to show that data are not routinely backed up, stored offsite and located well away from a potential disaster area. At least one offsite storage facility was submerged by the Katrina disaster, and it is all too common for earthquakes to affect data centers, offsite storage locations and the transportation resources needed to get media to a recovery location.

Assessing Threats and Vulnerabilities

Disasters take many different forms. Oftentimes the threat is the easy part to assess. The vulnerability is generally much more difficult to evaluate. For example most data centers have some warning of an impending hurricane. And several clients of the submerged storage site wisely noted that recovery media in the path of the storm was no better than being on-site and therefore pre-positioned their media at distant recovery sites, avoiding the “compound/cascading disaster.” Recognizing vulnerabilities (e.g. the collateral damage associated with the storage location) is frequently difficult to imagine.

Even those data centers that survived Katrina relatively unscathed suffered when roads were impassable for days – preventing the delivery of diesel fuel for generators – and when telecommunications networks were down. Probably one of the least recognized vulnerabilities to large threats like Katrina is the loss of IT personnel. In this case it was because they were needed to take care of their families. However, in other situations, such as a terrorist bombing, the IT personnel may be permanently incapacitated.

Testing is a Pre-requisite

When there is an on-line recovery facility, the ultimate form of testing is to swap workloads or migrate the production workload to the recovery facility on a routine basis. Many of the large companies in the U.S. and Europe do scheduled failovers to handle maintenance and essential activities at the primary data center. The more aggressive companies do unscheduled failovers and lockdown subject matter experts so that the situation is as realistic as possible. The lessons learned from these types of exercises go a long way toward ensuring that a company meets its recovery point and recovery time objectives.

If the recovery environment is more constrained than being able to swap sites at a moment’s notice, then try to maintain some semblance of realism. Is it worthwhile staging a test with “Application A” in Q1, “Application B” in Q2, etc? What is missing is the chaos of trying to resume “X” 1st-tier applications in “Y” amount of resources. This is especially true when the recovery site is viewed by management as available “spare” capacity.

Examples of Some Business Continuity Scenarios

Case Study 1: Stock Brokerage

A large, U.S. based stock brokerage firm needed to ensure business continuity 24x7 to support trading operations. There was a large web presence as well as voice response and live support. The company built two geographically separated data centers in a geologically stable portion of the U.S. with neighboring Internet backbone access, as well as being serviced by multiple electrical grids, and telecommunications.

- The data centers were served by two dark fiber routes, each capable of handling 100% of the data and voice traffic.
- All the data from each site were synchronously mirrored to the other site so that if a failure occurred, the mirror site (not really a secondary site) would take over operations. Achieving this level of failover capability comes at no small price; specialized monitoring systems kept watch on hardware and software systems. Automation systems were capable of performing site failover with or without operator intervention.
- Finally, all data including tape-based data were mirrored between the two sites. This required the use of virtual tape libraries, with integrated disk and robotics.
- Customer goals were no loss of data – except in-flight transactions – with a recovery time objective of 15 minutes or less.

Case Study 2: Retailer

This large international retailer operates primary, secondary and tertiary data centers. Unlike the stock brokerage firm, the retailer's business continuity environment is very selective with effect to the business as the discriminator for applications chosen for business continuance.

- Warehousing, inventory control and distribution systems receive first tier support. Second tier support systems include replenishment system ties to suppliers. Third tier systems are accounting and labor reporting.
- Asynchronous mirroring is used because of the distance between data centers. Dense wave division multiplexers (DWDM) increase the effective bandwidth of the dark fiber between data centers. Unlike the brokerage firm, systems are predominantly handling transactions between internal applications with little or no end-user interaction. The recovery point objective is almost as critical as the brokerage firm; applications are expected to resume processing at the last committed transaction and distribution centers need to be up and operating within four hours of an outage to get trucks back out and rolling again.
- For extended outage control, backups from one data center are stored at another data center in a "round-robin" fashion.
- The failover process is not nearly as seamless as in the brokerage company (where it was automated). The operations staff is expected to handle the recovery process when the occasion arises.

Case Study #3 – Electric Utility

Business continuance at the electric utility is critical. This utility operates nuclear power generating stations which are regulated by several different Federal, State and Local government entities. Furthermore, because electric power is critical for health care facilities and the safety of the public, there is a great deal of scrutiny over this (and most other) electric utility companies.

- To meet the stringent availability requirements of regulators, and to ensure that regional storms don't effect data processing capabilities, the electric utility uses a "multi-hop" configuration for availability. Multi-hop means that data are synchronously mirrored from one location to another (generally at or near the effective limits of synchronous mirroring technology (i.e., about 100KM) and then asynchronously mirrored from that location to a remote location hundreds or even thousands of kilometers away from the primary site.
- Applications protected by the business continuance plan are carefully analyzed to ensure they meet strictly defined criteria for necessity during emergency situations. Furthermore, the upstream feeds to these applications are also included in regular reviews to ensure that critical applications do not "starve" for data.
- Because of the electric utility's location it gets to exercise its disaster preparedness plans many times every year. However, corporate policy requires that simulations and tests be executed for a wide variety of scenarios no less frequently than monthly to ensure staff readiness.

Summary

There are many different methods of preparing for disasters. The effort and expense is directly related to the risk of loss incurred by an outage. As seen by the three real customer scenarios described above, an assessment of the applications and their effect on the overall business almost has to be taken into consideration. It is the vulnerabilities and risks that need to be weighed to ensure that the solution that gets adopted is appropriate to the magnitude of the threats. An "information lifecycle management" approach to addressing these issues helps to identify the resources, performance requirements and service level objectives that are factored into the disaster preparedness equation.

Specifications and descriptions are summary in nature and subject to change without notice.

The Stratus Technologies logo is a trademark of Stratus Technologies Bermuda Ltd. All other trademarks and registered trademarks are the property of their respective holders.

© 2007 Stratus Technologies Bermuda Ltd. All rights reserved.

X938