

INFORMATION MANAGEMENT: STRATEGY, SYSTEMS, AND TECHNOLOGIES

# E-DISASTER RECOVERY MEANS CONTINUOUS AVAILABILITY

Janet Butler

INSIDE

From Recovery to Planning; Sources of Downtime; Planning Ahead; Infrastructure Management; Getting Started

## INTRODUCTION

For the past two decades or so, disaster recovery has meant recovery of a company's information technology (IT) function after a natural disaster, such as a hurricane, flood, or earthquake. About ten years ago, companies began to realize that traditional resurrection of the data center was only one part of a company's recovery, because it made little sense to have a fully recovered data center if customers could not contact the company or employees could not get into the facility to do their work. Thus, terms such as impact analysis, business recovery, and business continuity gained popularity.

The emergence of the Internet and E-commerce has also sparked comprehensive changes in the disaster recovery/business continuity (DR/BC) industry. In fact, the new business venue has changed the very definition of disaster from that of a natural emergency to one that includes man-made happenings and, indeed, all forms of disruption to business — even events. And such business disruption can be very costly. Lloyd's of London, for example, reports that E-commerce companies lost more than \$20 billion worldwide in 1999 due to computer outages, downtime, and hackers.

### PAYOFF IDEA

Companies that depend on their World Wide Web site for revenue are realizing that uptime and speed of delivery of content are essential. Therefore, continuous availability has become the key phrase for large and small E-businesses alike. Industry analysts are advising these organizations to build continuous availability into their Web architectures, using E-business continuity plans, redundancy, off-site data protection, geographic load balancing, and a consistent backup and restoration strategy. After-the-fact retrofitting of Web site availability is far less feasible.

---

The need for high availability has also become far more prevalent with E-commerce. While organizations have been creating high-availability architectures for back-end systems for years, only recently have they begun to architect high-availability solutions for front-end World Wide Web environments. In the past, only a few applications, such as customer support, enterprise resource planning (ERP) systems, or select financial applications, required high levels of availability. Now, however, because business-to-consumer (B2C) and business-to-business (B2B) systems require tremendous traversing and integration of different applications, a far greater proportion of applications must be available. Nor can widespread, transaction-oriented customer relationship management (CRM) applications tolerate downtime. As a result, the requirement for continuous availability, which was unusual only two or three years ago, has become typical.

#### **FROM RECOVERY TO PLANNING**

E-commerce has also changed disaster recovery from a reactive, separate exercise, to a proactive part of the application development process. Ten years ago, companies considered it creditable to recover from a disaster in 24 to 48 hours; they later squeezed the recovery window to hours and minutes. Now, however, the term “recovery window” has become obsolete; most Internet requirements allow no downtime at all.

With after-the-fact recovery no longer an option for many companies, E-businesses must build continuous availability into their application architecture. Because this is a highly complex undertaking, companies must invest in and think about DR/BC planning up front during the design phase, rather than during the production rollout stage of a new application or business process, as was traditionally the case.

Backup, an important part of the traditional disaster recovery process, has also changed with E-commerce. For starters, few dot.coms are doing a rigorous job of backing up information. While they might be doing internal failover, their backup servers might well be located in the same facility as the primary servers. Although that guards against hardware failure, the site is still out if the facility is not available.

Furthermore, E-businesses must be concerned not only with the database, but also with content. Traditional IT shops back up their database in real-time, but they are rarely required to back up application code. However, the requirements are different for E-commerce application code, which changes rapidly.

For example, if Amazon.com lost transaction data, it would be painful, so the company must keep track of every one of its transactions. However, it is the HTML or application code that contains the descriptions of books and pricing information. If Amazon lost this content, it would be fatal; it would not have a business.

---

## **BUSINESS CONTINUITY GAINS VISIBILITY**

In the past, apathy made disaster recovery a hard sell because few companies would make the large investment required by solutions until they experienced a disaster. Today, E-commerce has made availability, outages, downtime, and thus business continuity highly visible and therefore worrisome to dot.com executives. However, while they are receptive to learning about their vulnerabilities, their main concern is gaining revenue, so getting their time and attention remains difficult.

Nevertheless, there is a current trend for E-businesses to build disaster recovery and business continuity into their application architecture — and with good reason. Jupiter Communications has found that when Web users have technical problems on a Web site, 24 percent switch to a different, competing service, and 9 percent permanently leave the problem site. As a result, leading E-commerce corporations today are targeting 99.999 percent availability, roughly five minutes of downtime a year, according to Forrester Research.

## **AVAILABILITY MANAGEMENT IS LAGGING**

There is no doubt that E-business is big business. Forrester Research predicts that industries will see 17 percent of their B2B revenues coming from E-commerce by 2004. And the Gartner Group forecasts that Net marketplaces will handle \$2.71 trillion in E-commerce transactions by 2004, or 37 percent of the overall B2B market.

Furthermore, on average, most E-businesses spend approximately \$6.9 million to build an Internet-based nonstop infrastructure and \$2.9 million a year to run it. Many are outsourcing their Web site function; Forrester predicts that by 2003 the Web hosting industry will be a \$15 billion market.

Despite all the money being spent on developing Web sites, however, availability management is lagging. Thus, to date, only 14 percent of companies have an effective plan in place to ensure availability for their Internet business applications, according to Comdisco's Vulnerability Index. In the retail industry, Keynote Systems reports that during the 1999 holiday season peak, the best of the largest retail E-commerce sites averaged 99.4 percent availability, about half an hour of downtime. Although that might be acceptable, the worst of the retail E-commerce sites, with 86.7 percent availability, were unavailable for nearly 12 hours.

Companies are finding that downtime is expensive. The average 87 hours of downtime for a 99 percent available site transacting \$5 million a day in E-commerce represents \$3.7 million in lost revenue per year, according to Forrester Research, even if only 20 percent of the transactions are actually lost. In addition, E-businesses are finding that brand erosion and customer dissatisfaction can add millions more to the cost.

---

## SOURCES OF DOWNTIME

Therefore, organizations must evaluate the risks associated with E-commerce; specifically, the business impact of downtime. For example, the cost of downtime per hour by application has been estimated at \$227k for eBay, \$125k for AOL, \$113k for home shopping, \$90k for catalog sales, \$89.5k for airline reservations, and \$28k for package shipping.

Furthermore, bad publicity hurts an E-business's reputation with customers, suppliers, and investors. For example, E\*Trade experienced a series of outages in 1999 over a month-long period, whereupon its stock price dipped 22 percent. At auction site eBay, an operating system failure caused a 22-hour outage in June 1999, costing the company \$3 to \$5 million in lost revenue, after which the stock price fell 26 percent.

Three major sources of downtime in E-businesses are:

1. *Unplanned disruptions*: including utility company failures and natural disasters, as well as equipment failure, human error, and fire and other infrastructure failures.
2. *Planned disruptions*: scheduled downtime for hardware, software or facility maintenance; adds, moves, and changes; or for data backup or restoration.
3. *Peak demand*: unexpected, high-peak traffic spikes due to sudden interest in a site, stemming from special events or other occurrences. When a company does not build scalability into the architecture, its infrastructure is inadequate to process such high demand. Although the site may be up, customers are unable to complete transactions on a high-volume day, so in their experience, the site is down.

## MITIGATING RISKS

To protect against downtime, E-businesses must architect a high availability Web infrastructure. In this way, they will be able to transparently recover from system and site failures, ensuring that their sites have a continuous Web presence. From an architectural perspective, the solution should provide redundant components, off-site data protection, geographic load balancing, and a consistent backup and restoration strategy.

### Redundant Components

Organizations should use multiple Internet service providers (ISPs) and routers, as well as redundant switches and firewalls; they should also have dual and backup power in place. However, because having totally redundant components at different sites is very costly, there are options. For example, if organizations are using a managed hosting service, the provider could provide a standby server at the time of failure, or during the planned downtime for a primary server. Those companies using dual production sites could use a standby local database server, whereby the

---

two sites back each other up. In this case, however, the applications must be “connection aware,” so that if the primary server fails, the applications will automatically connect to the standby server. Another strategy is to use a server that performs staging and testing as a backup server for the primary production site.

Companies that use ISPs or application service providers (ASPs) to host their E-business environment should focus on three areas of redundancy: the network infrastructure, application servers, and data content. Because recovery aims at identifying and restoring applications and data that are most critical to the business, these hosting providers should use techniques that couple “hot” standby application systems with mirrored real-time data servers. They should also inventory customers’ technical environments, providing a contingency plan that illustrates how backup sites will operate.

### **Off-site Data Protection**

This means moving database images off-site. Here, an organization might use disk mirroring or electronic vaulting, depending on how quickly it must recover and whether any loss of data is acceptable. Disk mirroring is a technique in which data is written to two duplicate disks simultaneously. Then if one disk drive fails, the system can instantly switch to the other disk without data loss. Electronic vaulting is the process of maintaining duplicate data and systems at a recovery site. This is done by remote storage shadowing or mirroring, which replicate information as it is created, transaction by transaction. The information is simultaneously transmitted via high-speed fiber optic circuits to a remote site, so the information is stored at two locations. Then the information is immediately available in the event of a processing disruption.

### **Geographic Load Balancing**

In geographic load balancing, which provides overall continuity of operations, dual production sites are in operation. This allows organizations to geographically distribute the workload across centers, thus helping them transparently recover from a site failure. At issue is how to replicate the databases between the production sites so they are synchronized. To do this, organizations might use multi-mastered databases that replicate the data in real-time or use a single database master that both sites can tap into.

### **A Consistent Backup and Restoration Strategy**

While many organizations perform backups, they might not know if they could restore their systems in the event of a disaster or disruption. Therefore, they should consistently practice restoring their servers from daily, weekly, and monthly backups.

---

To answer these needs, traditional disaster recovery vendor Comdisco has a premier offering that features load balancing between geographically distributed sites, database synchronization, high-speed networking, and at-time-of-peak (ATOP) services. For less time-sensitive applications, Comdisco uses redundant technology, standby operations, and alternate sites to restore a company's Web presence within a predefined time period.

For its part, SunGard also offers high-availability design solutions, as well as a new eSourcing unit that provides Web hosting and monitoring, geographic load balancing, and rapid recovery. IBM Business Continuity and Recovery Services provides consulting expertise to design continuous availability and recovery into application architectures. And premier "complex Web hoster" Exodus introduced Managed Web Hosting and Web Application Management services in October 2000 to help answer the increasing customer demand for high-level service.

### **PLANNING AHEAD**

In addition to architectural concerns, organizations should have an E-business continuity plan so that customers never see interruptions in service and processes. Unfortunately, however, only about 29 percent of companies surveyed have a plan in place to recover their Internet/intranet applications (according to an industry survey on business continuity, conducted in June 1999 by Ernst & Young and Contingency Planning Management). In contrast, many of these companies have plans for network data recovery components (73 percent) and internal data center recovery (65 percent).

In planning ahead, organizations should conduct an E-business impact analysis, determining points of exposure and the consequences of unplanned downtime — both financial and intangible. Quantifying the effects of outages helps generate buy-in for the E-business continuity program from executive management.

Impact analysis also helps organizations select the best recovery options. For example, an organization that depends on its E-commerce site for most of its revenue might well require continuous availability, justifying a "ghosted" copy of the entire site, which might be co-located or off-site. If data is critical, the organization should seriously consider real-time data mirroring. Alternatively, the site might not merit real-time solutions. However, organizations should monitor their revenue flow and technical developments to determine if and when that situation changes.

The business continuity plan should include incident detection, notification, and response mechanisms to inform the correct personnel of site problems so that predetermined steps can be quickly taken to resolve the issues. In addition, there should be provision for reassuring customers and investors that the problem is under control.

The plan should take into account how the site is affected by legacy systems and specify the contingency plan for those systems. In the case

---

of a supply chain, organizations must determine the impact of an outage up and down the chain. In fact, some companies are withholding a portion of payment if an alternate manual mode of order processing is not provided in the event of a Web site failure. For B2C, advanced data protection and restoration processes can minimize the loss of customer data.

The E-commerce continuity plan should consider such critical business continuity issues as backup strategies, predicted increases in traffic, whether the communications infrastructure can handle the spikes, and what to do in the case of a catastrophic event. Organizations should consider using outside expertise for disaster recovery equipment, data services, and plan creation. Finally, all plans should be regularly updated and tested.

### **INFRASTRUCTURE MANAGEMENT**

As many E-businesses have found, Web site traffic can be unpredictably high, due to such seemingly innocuous factors as targeted marketing or good press coverage. To handle traffic spikes and growth, organizations need a *scalable* infrastructure. Because the server is the most likely source of problems, organizations should balance site traffic using such techniques as server switches, load balancing, and Web caching.

Organizations also need a *reliable* infrastructure, in order to run 24×7 operations and handle the increased load on current resources stemming from E-business. The systems hardware points that are most vulnerable from a reliability standpoint are mainframes, servers, switches, and routers. In addition, organizations should make sure their ISP can provide the required reliability and bandwidth.

E-commerce sites feature such *security* measures as firewalls, authentication devices, public key infrastructure components, virtual intrusion detection systems, and virus protection. Thus, organizations' E-business continuity plans must be able to replicate these safeguards if they are forced to move off-site.

Finally, users judge Web sites negatively by downtime and positively by the speed of the site: how fast the site appears and the speed of the interaction. Therefore, organizations must ensure *scalability* and *availability*. The latter requires the tight coordination of all hardware, software, high-speed network connections, and external vendors, as well as a comprehensive recovery plan.

### **HOW TO START**

Those organizations that cannot afford to develop a high-availability infrastructure in-house might consider a managed hosting provider. The Web hoster's environment offers multiple ISPs, redundant switches, firewalls, and gateway signal routers. Hosters also provide bandwidth options that allow companies to meet their peak bandwidth demands, while subscribing to their normal capacity requirements.

---

If full-scale geographic load balancing is not an option, E-businesses might consider load balancing their Web server configuration by purchasing a local load balancer. As a result, Web servers will be more maintainable because one can be taken down for maintenance, staging, or testing.

Organizations should also protect their database data in real-time off-site. This not only ensures data currency, but provides a starting point for an eventual move to a secondary, geographically load-balanced production site.

## **CONCLUSION**

Customer relationships are critical in E-commerce, so E-businesses are being challenged to establish and develop customer trust in technology-based transactions. For example, E-tailers can differentiate their product offerings via a stable and reliable site. However, this is double-edged sword. For established brick-and-mortar companies, as well as dot.coms, the E-commerce site advertises a brand. If the site fails, the brand may well be permanently damaged.

For their part, investors are beginning to require documented business continuity plans when an initial public offering (IPO) is issued. They have recognized the fact that high revenue returns for E-businesses are directly linked to high system availability.

Continuous availability, backed by a stable and robust technical infrastructure, is a primary illustrator of both customer commitment and investor risk mitigation in the Internet environment. Therefore, E-businesses should architect their sites early on for availability, putting the requisite pieces in place. Then, as they scale up, they can more easily move to higher or even continuous availability. By contrast, companies that already have high-volume, lucrative sites will have trouble retrofitting their Web infrastructure for high availability.

---

Janet Butler is a freelance writer and consulting editor for Auerbach Publications. She is based in Ranchos de Taos, New Mexico, and can be contacted at [jbutler@laplaza.org](mailto:jbutler@laplaza.org).