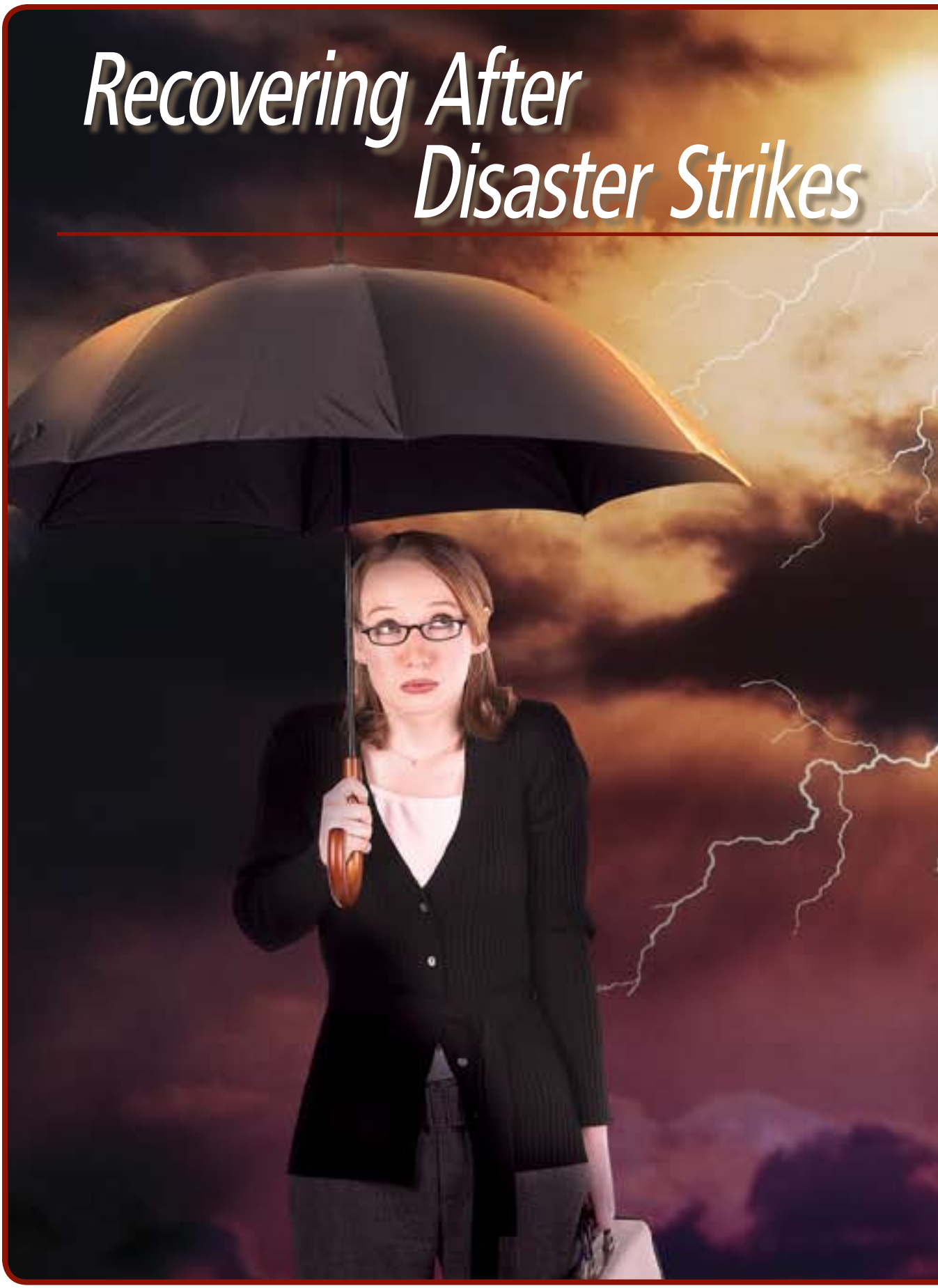
Recovering After Disaster Strikes



Developing a storage strategy that protects and recovers data requires a solid understanding of what you need to protect and where you need to protect it.

You can't prevent all disasters. Hurricanes, floods, earthquakes: these are only a few likely culprits that could wipe out your network. According to IDC analyst Rhoda Phillips, it's not just about planning for the disaster brought on by a storm or fire, but the human element as well. "Someone could trip over a cord and disconnect the server causing data loss," she says. Schools and government offices alike need an appropriate backup and storage strategy that can offset both natural disasters and human error.

"Disaster recovery is protection from the 'big one,' and is often related to a site failure," says Rob Emsley, director of product marketing for EMC. "It involves restarting from a different location. Operational recovery from local errors such as hardware or software or human error is about recovering operations onsite."

Whether you're looking to protect against operational or disaster recovery, there is a solution. According to EMC's Emsley, in operational recovery, many customers are making use of backup to tape, which provides copies of the production environment that can be recovered. Others have adopted a disk-based technology for backup and recovery processes, which are kept online for some time.

"Once that time period is exceeded, they copy those backup copies from disk to tape media for a slightly longer period — sometimes up to 30 days or longer," he says. "What we often find is that some customers have actually made use of virtual tape libraries," he says. EMC's CLARiiON Disk Library integrates ATA drives and tape emulation software that enables disk-based backup and restore without changing current operations or backup infrastructure. "The reason some customers use them is because they don't need to make any changes to their backup processes." he says. "The virtual tape libraries appear to the backup applications as an additional tape library that's been brought into the network and made available for use."

Another approach is to use snapshots and clone technology to make point-in-time copies within the array. "You can implement a schedule where on a periodic basis the array replication software creates a snapshot or a clone of the production environment," says Emsley. "A lot of customers create a snapshot or a clone on a daily basis, but they are increasing to more frequent snapshots or clones."

One solid operational recovery solution is a failover disk drive. With failover, if a device gets unplugged or a server fails, the hardware and software that allow data access to continue switches the workload to an alternate path.

Various solutions offer different recovery-level capabilities.

"Customers often want to use different levels of recovery for different types of data," says Emsley. "We're making sure that if they choose to do that, they don't have to manage each individual recovery service independently." EMC's Replication Manager lets users manage different types of replication technology. It supports snapshots and clones, and continuous data protection (CDP) devices.

Backing Up Remote Sites

Most government agencies and educational institutions have the added challenge of backing up and protecting data at remote sites. With more branch offices and remote facilities to support, the public sector sees more decentralization, making solutions for disaster recovery challenging.

A good example is the Federal Motor Carrier Safety Administration (FMCSA) within the DOT, which is responsible for implementing critical safety programs. The FMCSA struggled to find a better way to protect critical data collected at 90 remote offices that had limited IT staff and an insufficient backup infrastructure. "Typically, [agencies] provide individual remote protection solutions to each site," says Emsley. But then there's the challenge of managing those solutions at each site where trained IT staff is often in short supply.

> "The strategy you use for disaster recovery depends on the nature of the information you have to store and the kind of applications you're running. If the data is relatively stagnant and doesn't change much throughout the day, then once-a-day backup is probably enough. But those days are going away."
>
> — Bob Farkaly, Director of Product Management, Overland Storage

To resolve this and bring in a lower-cost alternative, the DOT used EMC's RepliStor for real-time data replication of those 90 remote offices to their centralized data center in Washington, D.C. Then all 90 replicas were backed up to EMC's NetWorker using its tape management features. "NetWorker provided them with the ability to maintain copies locally within the tape libraries for operational recovery purposes, and they also created copies of those tapes and moved them offsite to provide a disaster-recovery solution," says Emsley.

According to Bob Farkaly, director of product management at Overland Storage, school districts have to grapple with the same issue of backing up remote sites as many government agencies do. "We deal with school districts that use their central administrative site as the repository for data and put devices into the schools themselves using existing communications facilities," he says. And while the financial loss of data may not be the same as in business, the loss of student data or government information could cause lost time in trying to recover that data as well as the loss of valuable information that's needed to function properly.

At the San Dieguito Union High School District in Southern California, the average amount of data storage per student rose from 5MB to more than 200MB in the last three years. While each of the district's 12 sites had its own tape backup, the system had become time-consuming to use and offsite data retention requirements were not being met consistently. Using Overland's REO 9000 at the main site and REO 1000 at each of the 12 remote sites, backups are automated in the field, eliminating the need for management onsite and freeing three to four hours per week in backup and storage time. In addition, restores now take 10 minutes or less, rather than hours or days.

Emerging Technologies

One growing area is continuous data protection. "CDP basically keeps a time stamp of all data changes and can provide granular recovery," says Heidi Biggar, analyst for Enterprise Strategy Group (ESG). "End users don't run into problems if files are deleted or a virus hits like they could with traditional asynchronous and synchronous mirroring techniques, because with CDP you simply roll back to the last known 'good state,' and off you go." This TiVo-like effect has many manufacturers adopting the technology. Emsley sees CDP as an operational recovery solution. "It's a little like snapshots and clones on steroids," he says. "Customers can configure a protection window so they can roll back the production environment to the time before issues occurred," he says.

"CDP keeps an ongoing record of your data," says Farkaly. "As you change data, a copy is retained somewhere else. It's not backup or replication but a new technique, an optimal solution that lets you dial back in time."

CDP is gaining some ground in the storage arena, but it's a slow process. According to a recent survey by IDC, while 53 percent of respondents had heard about CDP, they were not clear about the meaning of the term. Only 25 percent thought they had a good understanding of CDP, and 20 percent of those surveyed said their organizations currently use CDP.

Asynchronous vs. Synchronous

For many institutions, the disaster recovery plan involves restoring data from a set of tape media that they have created and moved offsite. Some keep copies local and also clone copies and send them to a second site. But many more are starting to make use of replication technologies that fall into two categories: synchronous replication and asynchronous replication.

Synchronous replication ensures that when you have an application that writes a piece of data in one location, that write is not acknowledged until the replication software has written that same piece of data into a second location. The application then gets confirmation that your write has been successfully completed, and you can do another write.

"Synchronous replication has some limitations because of the time it takes for data to be transferred over a long-distance link," says Emsley. "You have to factor in the latency that supports a synchronous application."

With asynchronous replication, the application is told that the write has taken place as soon as the write to production is complete. It doesn't wait for a full confirmation that the write has been completed. Therefore, you don't have the same guarantee as with synchronous replication, but you have the ability to move data over longer distances more quickly.

"Many are using disk-based remote mirroring technology for synchronous and asynchronous applications as part of their disaster-recovery planning," says Emsley. "They look

Make a Plan

According to experts, there are at least five things to consider when developing a disaster prevention and recovery plan.

1. What do you need to recover?

Recovery point objective (RPO) is the point in time from which you want to recover your data. For example, do you always want to be able to recover your data from five minutes ago (for which you would use a snapshot), or is five hours OK? Or is it 24 hours? This is something that all organizations need to determine based on how frequently their data changes and how critical those changes are. Once that's determined, they can better select a disaster-recovery solution.

2. How fast do you need to recover your data?

Recovery time objective (RTO) is how long it actually will take to recover your data and return to normal operations. If your data is on disk, you can recover it in minutes. If it's on tape in a vault, it could take hours or days to recover. Keeping your most critical data accessible is recommended.

3. What applications need to be brought online immediately?

Applications are the lifeblood of any organization. You need to look at the whole scope of your applications to determine your disaster-recovery strategy. For example, if you have a database with transactions that change constantly, you'll protect that application differently than an application with static information.

4. How synchronized does your backup need to be?

Synchronous recovery is used when you have to keep the remote site in absolute lockstep with the primary site. Every time data is written at the primary site, an exact duplicate write occurs at the remote site. Asynchronous replication means that the write operations at the primary site do not have to wait for the data to get to the remote site. Changes to data are batched and sent periodically. Only the most critical data is synchronously replicated.

5. Is the backup site hot, warm or cold?

A hot disaster-recovery site means everything is running — servers, connectivity — and the site can automatically take over operations if the primary site goes down. A cold disaster-recovery site is one where someone needs to physically bring up everything, initiate disaster-recovery activities, and get the operations up and running. Warm is in between, where some key systems are running but others need to be brought online.

upon disaster recovery as an insurance policy. It's designed for the plan being called into action in the event of a disaster that could wipe out the entire network."

What You Need to Consider

"The strategy you use for disaster recovery depends on the nature of the information you have to store and the kind of applications you're running," says Overland's Farkaly. "If the data is relatively stagnant and doesn't change much throughout the day, then once-a-day backup is probably enough. But those days are going away." Losing 24 hours of work can be devastating for any organization. "I can't imagine too many scenarios in which taking one snapshot a day would suffice," says ESG's Biggar.

"Many organizations are seeing the need for more aggressive backup strategies," says Chris Caprio, manager of technical services for Imation. "Many have been saying that tape is dead and that it's been supplanted by disk drives, but truthfully, it has its place in the hierarchy of storage solutions."

Many users who are looking for the best way to get the most out of multiple-blended storage solutions most likely will consider all options. "You need to keep in mind what your recovery objectives are and structure a plan based on data criticality and recovery time," he says. Emsley agrees that it is important to decide what critical applications need to be brought online in the fastest possible way should a disaster occur.

ESG recommends that organizations first evaluate their data volumes and define their recovery point objective and recovery time objective for their different classes of data. "Once they've done this, they can begin to match technologies to these data classes, keeping in mind trends, existing technology investments, compliance requirements and budgetary concerns," says ESG's Biggar. "Ultimately, it boils down to a feature, cost, performance and potential disruption analysis."

CDW·G®

Did you know that CDW•G offers configuration, product support and customized professional services? Call your account manager for details.