Chapter 5

Identify Data Storage and Recovery Sites

Chapter Objectives

After reading this chapter and completing the exercises, you will be able to do the following:

- Determine the best way or ways to back up your data so that it can be recovered later.
- Evaluate your offsite storage options.
- Acknowledge information as well as hardware and software as an asset.
- Determine recovery site options.
- Examine recovery site types.
- Develop recovery site selection criteria.
- Outline a recovery solution.

Introduction

Once you have determined what you need to restore, you need to have an idea of where you will be able to restore it. Depending on your particular hardware configuration, this decision may be more difficult or easier than you anticipate. This chapter will help to determine the alternatives and the wisest choice for your situation.

Data Backup

Backing up the data—how, when, and how often—is one of the first considerations that come into play in a disaster recovery plan. Many decisions need to be taken into account when you are looking at a backup strategy, more than can be covered here. However, we can look at the main points to take into consideration.

It is important when doing your disaster recovery testing that you test with a normal backup. Don't schedule a special backup just for the purpose of using it for DR drills because you believe that you can restore from it. It is important to use ordinary backups for testing.

How to Back Up Your Data

There are many ways to back up data. You can opt to invest in enough storage to maintain a backup online, on a hard disk. You can look at burning data to a DVD or CD, depending on the amount of data to be backed up and the recovery method that you end up choosing (although that choice is often made based on the type of backup rather than vice versa). Or you can follow the time-honored tradition of backing up to tape. Let's look at these different options in the light of recovery and recovery ability.

If you back up to disk, and the restoration is to the same server from which the backup was taken, recovery is simple, elegant, and quick. Disk-to-disk recovery is the quickest method because your access times do not include the time that it takes to transfer data from another medium. Looked at through the lens of disaster recovery, however, if you store your backups at the same location as the data that was backed up, and a disaster makes this site inaccessible, then your backups are inaccessible as well. Further, if there is a hardware failure or a catastrophe that destroys the hardware on which the backups are stored, then there is no way to recover.

If you back up to tape, you have a backup methodology that has been proven over time. The restore time is longer than for disk-stored recovery because the tape drive needs to access the media, and you have to unspin the tape to retrieve the data from it. Restoring an entire tape, as would likely be the case in a disaster situation, is less labor and time intensive than recovering a single deleted file. This calls into play the definition of disaster that you are using in each case. If the loss of a partial file system constitutes a disaster, but not a disaster that requires restoration in an alternative location, then there are different concepts to take into account.

Please note that the types of backup methods as well as how they are deployed are covered later in this chapter.

IN PRACTICE: Storing the Backups

Another consideration to understand with the use of tape backups is tied, to a great extent, to the location where the backups are taken and stored.

A company in Texas found itself unable to guarantee that it could restore from tape backup because the company that it contracted with to transport its tapes from its data center to the secure offsite location where the backups were stored hired new people and found that one of their carriers, rather than delivering tapes to the secure, climate-controlled location, was carrying the tapes in the trunk of his car in July and August.

A company in New Orleans chose to back up to tape and to store the tapes offsite with a company that chose to store the media in the basement. Drying the tapes out after that company experienced a flood caused by a broken water pipe was a time-intensive job, and one that no one was sure was not going to destroy the tape's ability to be read and restored from.

Another option is to make backups to CD or DVD. This media has faster seek times than tape, historically is less volatile and prone to degradation than tape, and is more portable than a SAN or a drive array. The disadvantage, however, is the amount of data that can be stored to the media. The DVDs we are referring to have much higher capacity than those that you write to from your home computer (9.4 GB compared to the 4.7 GB that you likely find with DVDs created for home use), but when backing up terabytes, petabytes, or exebytes of data on large systems it will still take a considerable number of DVDs (but then, it would take a large amount of anything to hold that much data).

DVDs are good for backing up the data for home offices or small businesses. It is important, however, to note that they are not nearly large enough to back up large amounts of data. Tape backups are still the most common method of backing up large amounts of data. Network-attached storage is also a method that is growing in popularity. Although network-attached storage doesn't necessarily meet the needs of an offsite recovery effort, it is still the backup method of choice for many large organizations.

Backups that are taken to disaster recovery testing should be the same backups that are done for the purpose of production recovery, not special backups taken for the purpose of recovery testing.

When to Back Up Your Data

When you schedule your backups (either manually or automatically) is a matter of complying with company requirements and with what you are doing and how. If you need to shut down all the systems on your server in order to take an adequate backup, then scheduling the backups for a Tuesday afternoon would

probably not be a good idea. However, if you can leave all systems up, active, and available while you are taking your backups, this may be a viable time to perform the backups. It would be advisable, however, to schedule them when the system is less active and has fewer processes running and fewer users accessing the data.

Although it is always a best practice to perform backups when servers are least active (typically after hours), in a truly global organization there is often not enough slow time for the server to allow for this. Research on server load for most servers in the organization is advised so that backups can take place during the lowest average load.

Keeping all this in mind, it is also important to remember that there are some systems that either cannot be backed up when the files are open and being accessed by the applications or that must have special processing in order to have the backups be consistent and in a manner that allows for a successful restoration of the system.

The most obvious example of this is when working with databases. Many database management systems are structured in such a way that the database has to be shut down in order for backup and subsequent recovery to be successful. Either the database has to be shut down or special settings have to be set in the database to allow for accurate backups. If these procedures are not followed, the database likely will not recover to a point where it can be opened.

Various techniques can be used to back up data that is in use, or to minimize the downtime to seconds or minutes to facilitate the backup, including snapshots or third-party utilities.

How Often to Back Up Your Data

How often you end up taking your backups ends up being a combination of math and company policy. Many companies have a records retention policy that dictates how long you must retain records and be able to recreate those records. Many governmental regulations come into play with these policies. To take into account all factors when making these determinations, you should consider the following parameters:

- The ***recovery point objective (RPO)*** is the point in time to which systems and data must be recovered after an outage as determined by the business unit. Your backup methodology must take into account the need to lose no more than an hour's worth of data, or exactly no data loss in case of disaster or the ability of the organization to recreate a week's worth of data after the systems are fully functional.

- The ***recovery time objective (RTO)*** is the period of time within which your systems, applications, or functions must be recovered (either because of contractual obligations or because of organizational requirements to withstand the loss of productivity) after an outage (e.g., one business day, 72 hours, or 2 hours). RTOs are often used in combination with service level agreements (SLAs) as the basis for the development of recovery strategies, and as a means of determining whether to implement

the recovery strategies during a disaster situation. This parameter ties in closely with SLAs and with the expectations of the organization. If you must have your business fully functional within 72 hours, then you can't rely on a backup methodology that takes 60 hours to restore because that leaves you only 12 hours to get both personnel and media to the recovery location and for troubleshooting and problem resolution.

- The ***maximum allowable downtime*** is the absolute maximum time that the system can be unavailable without ramifications to the organization, either directly or indirectly.

Where to Store Backups

This is one of the most important decisions surrounding backups that you will make. Considerations need to include the accessibility to the backups when needed; the time it takes to retrieve the backups; and the climate, geography, and topography of both the primary business and backup storage locations (whether or not these are the same location).

On Site Storing your backups on site makes for easy access to the organization. This has advantages and disadvantages. Having the backups handy means that you can restore a single file whenever someone needs it to be restored. It means that, in the case of failure, you don't have to factor in the retrieval time for the media from an alternative storage location. You can simply retrieve the files whenever necessary.

However, there are two disadvantages to this scenario. First, storing the media on site means that you have added security precautions to take into consideration, as you need to store the backups in a location that is inaccessible to most people in the organization. Retrieval of the media for any reason other than to rotate or destroy the backups at the end of the retention time needs to go through the process of getting permission from the disaster recovery team captain or whoever is responsible in your organization for such a sign-off. In this way, you can ensure that no one retrieves the backups for nefarious purposes. Second, storing the media on site means that whatever disaster may befall the organization is likely to befall the backups as well. Although this may not be true for a Denial of Service (DoS) attack, virus, malicious data tampering, or hardware failure (by far the most common "disasters" that you will have to deal with), in the event of a fire, flood, or other physically disruptive disaster you will likely be no more able to access the backups than you are to access the primary system.

Often, organizations use climate-controlled fireproof, waterproof, and even tornado-proof safes or shelter structures to store tapes on site. This lessens the risk of the tapes being damaged in a disaster and still allows for rapid access by the organization when necessary. It also allows the organization to guarantee the chain of custody and minimize the points of access to the data that is contained on those tapes. This can be of particular interest for legally sensitive data.

Offsite Offsite storage, on the other hand, is more expensive because you have to pay rental and service fees for the offsite location, as well as any services associated with that storage. There are many levels of offsite storage to consider. Depending on the size, budget, and complexity of your organization, one of these solutions should be optimal for you.

Owned offsite storage, depending on the complexity of the organization and its requirements, could be as little as storing the backups at the home of a responsible individual or at a rented storage locker. Each option brings with it advantages as well as disadvantages. Storing the backups for an organization at the home of a responsible employee has the advantage of very limited cost, but it also means that this person has to be available should a disaster occur, and also that the backups are stored in such a way that a disaster at that individual's house cannot destroy the organization's tapes.

If the organization has more than one physical site, even separate physical sites within the same town, the tapes can be stored in a vault at the secondary site. In some cases a site may have another building at the same complex, one that is an acceptable distance from the primary site or main server building, that could serve as an alternate site. The organization needs to determine the acceptable distance between server locations and the offsite location. It is important to note that this can also be tied to the RTOs and the SLAs previously mentioned.

Depending on the criticality of the data and the need for assurance of both safety and accessibility, greater distance may be required (1 mile, 5 miles, 500 yards, half a state away). Only the organization can make that kind of decision, and then only after careful consideration. These decisions need to be made on a case-by-case basis.

IN PRACTICE: It Is All a Matter of Location

Depending on the organization's location or the resources available to it, it is possible that there may be opportunities to own the offsite storage location for the backups. In many locations there are old, abandoned mines that are sufficiently climate-controlled. These include limestone, coal, iron ore, and salt mines.

Underground storage initially became attractive as secure storage of documents and people during the Cold War when people were more interested in the end results of nuclear attack. The fear of bombs has lessened, but the massive underground structures live on. They have since been transformed into locations for tape and other records storage.

CONTINUED ON NEXT PAGE

CONTINUED

These facilities, 200 or more feet underground, are highly secure and access is easily controlled, with only one way in and one way out, and they are virtually immune from nearly any natural disaster (fire, flood, hurricane, or civil disorder). This makes them very attractive as storage locations (either owned or contracted), and organizations choosing these locations can rest assured that their backups will be safe from nearly any eventuality.

FYI *Contracted*

One thing to remain aware of with the hiring of an offsite vendor is that the vendor must be able to assure you that it will be around when the time comes for you to declare an emergency, and that it can deliver what is promised in the timely fashion that you need it.

Don't just rely on vendors telling you that they can. See if you can get references from other people who have used them in the past. See if there is proof, one way or another, of their ability to deliver. And above all, get it in writing that is legally binding.

The contracted organization needs to be in the business of data archival and retrieval or the contract will likely not work in the end. SLAs need to be established with the vendor to ensure the organization's RTOs can be met successfully and in a timely manner.

Information as an Asset

This is all well and good, when you are looking at hardware as an asset, but it is important to remember that people and information are also assets of a company, and that companies often overlook these critical assets in DR planning. They consider, at great length, what it means to have duplicates of their data, but it is not always the hardware that fails.

Often it is software that fails, and software failures often go undetected for extended periods of time. Where hardware failure is very apparent and gets noticed by virtually the entire company, bugs in software can cause insidious corruption that may go unnoticed for days, weeks, or even longer.

IN PRACTICE: Homegrown Applications

Software often is not designed by an external company (such as IBM, Oracle, or Microsoft) but is designed and implemented internally. Sometimes, although precautions are taken and testing is done, the inevitable happens, and a set of circumstances come together that causes a calamity to the business.

For example, a very large organization had a homegrown purchasing and accounting system. This system processed purchase orders, receipts, and invoices daily. The system, with few changes, had been in place for over a decade. Receipts were processed every night, via files handed off from each division to headquarters. These files went through several COBOL programs before they were added to the database and before the costs for the receipts of purchases were attributed to the different plants.

One of the plants was in Northern Minnesota, and due to connectivity issues several days passed during which no files were handed off. When they finally were, an extremely large file was in the batch to be processed. Mainframe processing often bases the size of output files on the average file size for a given period of time. When the allocated file size grows larger than the allocated file, the program running abends (abnormally ends), and the operators restart processing after adding additional space to the output file. Ordinarily processes are put into place to account for the restarting of programs and the condition would be handled efficiently.

This time, however, the file was big enough that the program had to be restarted several times, and there were no precautions in the program to skip the already processed receipts and move ahead to where the program abended.

The run started, ran for several thousand receipts, and stopped. Operations added space to the file and restarted the program. The first several thousand receipts processed again and then several more thousand receipts processed before the program broke again. Operations added more space and restarted the program. The first several thousand receipts processed a third time, the second several thousand processed for a second time, and the batch finally ran to completion.

A sequence number that was generated whenever the program ran is the only thing that made these receipts unique.

No one noticed that the database and resulting files were incorrect for over a week, because the accounting departments at the

CONTINUED

plants were understandably concerned when their anticipated and realized costs were radically different.

Recovery in this case couldn't be to restore the system and reprocess—too much time had passed and too much processing had been involved. New programs had to be written to determine how to back out the duplicate and triplicate receipts and not affect any other part of the system.

Not only was this a lesson in recovery and the differences in recovery, but the differences in the types of disaster from which you have to recover (in this case, internally programmed software failure) and the ways in which you can recover. Further, it was a lesson in the fact that, no matter the size of the company and the amount of effort put into disaster recovery and the effectiveness of backups, there is always something that you can learn about your system.

It is important to remember that part of the information that many organizations forget to bring to the disaster recovery site is the information that will be necessary for restoring systems and data. It is easy to remember that you need the backups of the data, but you may need information on the licenses that you have purchased for software or software-hardware combinations. This is particularly true for systems like IBM mainframes, where the license is tied almost directly to the hardware and alteration of where the operating system or other software is installed will mean that the software cannot be used. These vendors fully understand the need for disaster recovery and disaster recovery testing and will work with you to provide temporary licenses for new hardware configurations, but having your existing license key available at the DR site will mean that you can more effectively and efficiently work with your vendors to expedite the process. Time is critical in a disaster situation, and the time that you save could mean the difference between meeting and missing SLAs, and the cost can be extensive.

Metadata is just as important as licensing information. The company must determine what data is important and how important each different piece of information is. Wasting time early in the recovery process recovering information that isn't as critical as other information is a less than effective use of resources. Apply your efforts where they are going to make the most difference, especially early in the recovery process (whether it is a real disaster or a drill). Part of the planning process should include attention to which information has to be available and in what order to have the data ready for the users.

Critical data is data that has to be retained and recovered for legal reasons and for the restoration of minimum work levels. Within this definition of data, it is important to note that the critical data necessary to restore minimum work levels is the data that should be concentrated upon. You have to be able to access

the data that needs to be retained for legal reasons, but that data does not necessarily have to be among the first recovered. The definition of minimum work levels is situation specific, but among this information would likely be accounting information and access to purchasing, receiving, invoice processing, and payment processing functionality.

Vital data is information that has to be retained and recovered to maintain normal business activities. This information and data represents a substantial investment by the company in time and effort, and the recreation of the data may be difficult or impossible. This data may or may not be necessary in a disaster recovery situation. This is determined situation by situation (the data in question, the criticality of the restoration, and the duration of the disaster all play a part).

Sensitive data is the data or documentation that is necessary for normal daily operations of the organization but for which there are alternative sources of the same data or data that can be easily reconstructed from other data that is readily available (critical or vital data may be sources for this data).

Noncritical data is data that can easily be reconstructed at minimal cost or that has its source in critical, vital, or sensitive data but that has less stringent security requirements.

Recovery Site Alternatives

Every organization is faced with many alternatives, not only from a backup and recovery storage perspective, but also when dealing with the determination of recovery site alternatives.

Function

Disaster recovery sites function as a scaled back office for your key functions so that your company can continue its business until either its primary site can be restored to operational functionality or until your company can come up with a viable alternative for long-term functionality. A disaster recovery site is a stopgap location so that you can keep from having to close the business while the long-term solution can be accomplished.

That means that when you are looking at the different options and the different locations, you need to keep your business and its functions and needs in mind. Make sure that what you are looking at will meet the needs of the company for at least a couple of months. If you are back up and running as usual in a shorter time, that is wonderful, but you should look at everything with realistic expectations.

Hot Backup Sites A hot backup site is a location, other than your typical business site, where a duplicate of your data is stored, ready for access in a moment's notice. Real-time data transfer occurs between the primary and the

hot backup site, which means that the site has to be contracted continuously or that the company has to own both the primary and the hot backup site.

The advantage of a hot backup site is that the company can guarantee that there will be at most a certain period of data loss (for example five minutes) that can account for the lag between the two sites. This can be a core competency or a selling point for an organization to its clients.

The disadvantage, of course, is cost. Not only do you have the cost of either contracting or acquiring the backup location, but also the cost of the network traffic that is required to keep the two environments continuously in sync.

Warm Backup Sites A warm backup site is a facility that is already stocked with all the hardware that it takes to create a reasonable facsimile of what you have in your primary data center.

In order to restore the organization's service, the latest backups from your offsite storage facility must be retrieved and then delivered to the backup site. Once this is accomplished a bare metal restoration of the underlying operating system and network must be completed before recovery work can be done.

The advantage to the warm backup site is that it can be gotten to and a restoration accomplished in a reasonable amount of time. The disadvantage is that there is still a continued cost associated with the warm backup site because you have to make sure that you maintain a contract with the facility to keep hardware up-to-date with that which is found in the organization's data center. This is the only way to ensure that the backups can be read when the time comes and the only way to be sure that the restorations can be accomplished successfully.

The warm backup site is the compromise between the hot backup site and the cold backup site.

Cold Backup Sites A cold backup site is a building space that can be leased or located (if you are very lucky) when a disaster occurs. Everything that will be required to restore your organization to productivity will have to be procured and delivered to this site when the disaster has been declared. Only after the physical resources have been acquired can the process of actual recovery begin.

The advantage of a cold backup site is that the cost is minimal. Even if the organization chooses to contract the floor space long term, the cost is far less than that associated with the others.

The disadvantage of a cold backup site is that the organization can find itself unable to recover operations for an extended period of time, particularly if the organization relies on resources that are difficult to obtain or that require extended lead times to be configured. If an organization chooses to rely on a cold backup site, and it is not reliant on commodity hardware, that organization may fail to recover in a timely enough fashion to remain a profitable business and may fail.

Reciprocal Agreements Reciprocal agreements are often a viable and useful alternative to a warm recovery site, but you have to make sure that the

company you are working for and the company that you are anticipating entering into such an agreement with are not only aware of the advantages and disadvantages of such an agreement, but that the company that you work for is a good fit for the agreement long term. A company that is not a good fit long term may still use a reciprocal agreement as a short-term solution to the disaster recovery scenario, but all parties have to acknowledge that the solution is, indeed, short term.

A reciprocal agreement is the agreement of two (or more, but this is less frequent) companies to join together and act as partners in disaster recovery. One kind of business that can make use of this kind of agreement particularly well is the home office or startup business. These businesses often have shoestring budgets and can little afford to be without access to their information for extended periods of time, and can afford even less the outlay of money for the purpose of making sure that there is a place where they can recover their information should a disaster or unplanned event prevent them from using their own resources. Two or more of these companies can enter into a simple agreement so that they can share the cost of resources over the duration of the agreement.

For example, Larry is a freelance author. He is working nearly full time at writing while also holding down another full-time job to pay the bills and provide insurance and support to his family until he feels able to get into writing as a single full-time profession. Adam is a tax accountant who is just starting to venture out on his own. He, too, is working at another job until he can build his client list to such a degree that he can see his way clear to make his own business his only profession. Amandya runs a home crafting business making custom scrapbooks for customers' weddings, their children, or other significant life events. This is her only official job. Finally, Joy has an established, though limited, enterprise that provides housekeeping services to a given set of clients. All these people have their own set of information that is relevant to their particular businesses and their own needs from a hardware and software perspective. However, many of the needs overlap or are exceeded by the needs of another (one application has greater hardware requirements that others may be able to take advantage of even if the need is not present) and can therefore be easily met. Some, such as the tax accountant, have specialized software requirements that the others don't have. These four organizations can enter into a reciprocal agreement that can include the cooperative effort to purchase a backup system that is stored in a central location (another person's home, or in a storage locker somewhere) that can be used by one or more of the organizations (at one of the four offices in question, or at another agreed upon location) for the purpose of disaster recovery. For this to work well, the parties involved cannot reside in very close proximity to each other (for example, the same street or the same housing area, and likely not even the same town) so that there is a good chance of recoverability should a hurricane affect Amandya's location or a forest fire affect Larry's access to his home office.

Such an agreement is typically made up of two parts, agreed upon by all parties. The first is a letter of understanding that lays out, ahead of time, any costs associated with the agreement and a general agreement of what responses

each company expects to offer and receive during an outage. This letter of understanding should also provide information on proposed recovery sites for each party in the event of a disaster. For example, should Joy need to recover on the backup system, she would have a certain amount of time to either inform or request from any other party (Larry, Adam, or Amandya) that she is in need of temporary recovery facilities, and office space could be made available to her at one of the alternative locations. The second part of the agreement is a completed configuration questionnaire that covers current and anticipated hardware and software requirements, and an anticipated schedule for updating the software and hardware and making the associated changes to ensure that all parties' needs will be met and that all parties will be able to use the systems in question in the event of a disaster.

At some point, the parties involved may find this to not be an adequate solution or they may find that their needs and the needs of the others begin to diverge to an extent that the agreement is no longer anticipated to be beneficial. In this case, it is important that the agreement be equitably dissolved and the parties that choose not to remain in the agreement find other recovery solutions.

Another interesting situation where reciprocal agreements are profitable is when key personnel in the organization have their own resources at home from which they can perform many of their ordinary work duties temporarily. The use of these resources as auxiliary disaster recovery resources can be contracted as a reciprocal agreement with the reimbursement for the personal computer and supplies needed to be part of the disaster team. The benefit to the employee would be financial; the benefit to the company would be the readily available resources should they prove to be in short supply at some point.

Two Data Centers This alternative can be seen as a logical extension of either hot or warm backup solutions; however, it can be defined differently when the organization in question isn't sufficiently large to contract either for a hot or warm backup site or to own a hot backup site in another location. If this is the case some organizations, typically smaller ones, can equip a limited second data center.

For smaller organizations second data centers can mean (if the company is small enough) a couple of laptops equipped with sufficient hardware resources and the proper software necessary to conduct business. It can mean a duplicate set of hardware (if a couple of laptops are not sufficient for the amount of resources or if laptops are not robust enough to provide sufficient power for the organization's need) set up with software stored in an offsite location, such as a storage locker or a rented space in a warehouse.

This solution is limited in cost, but can be sufficient. If space is required temporarily to recover the business, office space is easily contracted on a limited time basis either from official locations or, in a pinch, from a hotel in a safer location.

Consortium Arrangement If we consider a consortium to be a group of organizations that have formed some kind of association or formal combination with the purpose of engaging in a mutually beneficial venture, then we can extend that to a disaster recovery consortium arrangement and see that the venture would be to provide disaster recovery facilities to the other members of the consortium. This isn't totally unlike a reciprocal agreement, but as the group is often formally joined together, the venture ends up being more formalized and longer lasting than the reciprocal agreement.

Since it is unlikely that any of the members of the consortium are in the group for the sole purpose of providing the other members with backup facilities, it is more likely that each member would provide a given amount of space on existing hardware to other members for the purpose of all or part of their temporary recovery efforts.

Because few organizations, especially in today's business environment, can afford to dedicate significant chunks of their resources to remaining available for use in a disaster situation for one of the other members of the consortium, this type of agreement is not typically used.

Vendor-Supplied Equipment Agreements are often made with vendors to be able to provide emergency equipment to organizations in the event of a declared disaster. Space and the utilities to make a functional representation of the organization are often easier to come by in alternative locations, and often are contracted for extended periods of time, but the hardware on which the systems run is contracted only to be made available in the event of a disaster.

This is much easier to do if you are requiring the use of commodity type equipment rather than custom equipment. For example, PCs, even several dozen or several hundred, are far easier to come by than two or three PDP-11s. With the increasing popularity of Linux as a reliable operating system for industry, this commodity kind of hardware is becoming more prevalent, but the reliance on mainframes has not diminished and it may be the fact that, depending on the extent and severity of the disaster, some equipment may not be as readily available, even to fulfill vender contracts, as it could be.

Caution should be taken when entering into this type of agreement to make sure that the vendors in question will be able to fulfill a need such as this in a timely enough fashion for the organization.

Combinations *In disaster recovery, as in many areas of business, exactly what happens and how is not always cut and dried, not always black and white.* In a lot of cases, one solution does not fit every aspect of a company's disaster recovery needs. In these cases combinations of solutions are created to custom fit the needs of the organization.

IN PRACTICE: Custom Solutions for Solutions Reference

Solutions Reference is a research company that assists companies with marketing research and with research and development based on that marketing research. It has determined that, in different situations, it needs to employ different recovery methodologies.

In the event of a natural disaster that affects the locations where it has its primary computers, the company will take its daily backups to an offsite recovery location in a different state so it can recover while its primary location is being brought back online.

The company has determined, however, that (because it is in a hurricane prone location) it can opt to head off the effects of a disaster by temporarily relocating the servers for a period of time to the servers in the winter home of one of the company managers in California. The manager simply needs to fly himself (and potentially his family so he has that piece of mind) to California, attach his system via the extended network to the network at work, and make a copy of the system. The switch over can occur at that time, meaning that there is literally no downtime for the servers and the clients are unaware of any potential for disruption.

Written Agreements

Written agreements for the specific recovery alternatives selected should be prepared, and should include the following special considerations.

Contract Duration The duration of the contract needs to be set and agreed upon by all parties. Close attention should be paid to when the end dates are approaching so that renegotiation of the agreement can occur and the parties in question are not left stranded without an agreement and a disaster pending.

This is one reason why a single owner of the disaster recovery plan and all associated materials is usually critical to an organization. One person (with a backup, of course) owning all disaster recovery material means that that person is responsible for keeping up with agreements, licenses, and contracts and that the organization can rest assured that it will not be left without an agreement should it need to rely upon it.

It is particularly convenient if agreements and contracts can be written so that they all expire at nearly the same time, so that all negotiations can occur around the same time. This will make remembering easier for the responsible parties.

Termination Conditions All contracts and agreements should have termination conditions that allow either party to terminate the agreement, although

there are usually financial repercussions for whoever decides to terminate before the expiration date of the contract.

Conditions for termination include the insolubility of one of the parties or the occurrence of too many declared disasters without justified cause in too short a period of time. Although some disasters are unavoidable, such as floods and tornadoes, others are often preventable in part or completely by taking reasonable precautions.

One example is the case of Crystal Dogs. The Crystal Dogs company was a dot-com startup that sold crystal dogs online. It was a home business retailing crystal dog figurines and Christmas ornaments, an Internet-only business whose founders ran the Web site and all associated systems. These owners contracted with another service provider (DR_Cheap.com) to be their backup for a nominal fee so that, in the event of an emergency, the business could be up and running again without having to resort to spending a lot of money on a backup site.

Crystal Dogs, however, failed to take the minimal necessary precautions to secure its network and install firewalls and filters on its systems, and it fell victim several times within a 6-month period to DoS attacks and viruses and had to rely on the services of DR_Cheap.com an inordinate amount of time.

DR_Cheap.com felt that they were losing money on the proposition (kind of like an insurance company when you have an unusually large amount of accidents in a short amount of time) and dropped the contract with Crystal Dogs because it felt that Crystal Dogs wasn't taking the necessary precautions to prevent declarable disasters.

Testing Any agreement should have, as a part of the services provided, the facility to allow your organization the ability to test your disaster recovery plan. It is important to note that, unless you have fully tested the plan and are confident in your ability to recover from a disaster with the given plan, you really can't say that you have a disaster recovery plan. You can say that one is in process, but without knowing that you can rely on the plan in case of a declared disaster, you can't say that you can actually recover.

Costs Another critical feature of the written agreement needs to be the provision for all costs associated with the agreement. The contract needs to detail anything that isn't covered by the agreement, and you need to make sure that it includes anything that you think ought to be covered. Usually hardware is covered, but are there any software costs that you think ought to be covered that aren't apparent in the agreement?

Although this may not seem to be overly important for many software products (such as Microsoft Office), there are times when not having the correct license key can mean the difference between being able to recover during a test *or declared disaster and not being able to.*

For example, there are keys associated with VMS and ZOS operating systems without which you cannot install the operating system. These keys are costly

and often associate a given version of the operating system with a given configuration of hardware. Deviation at any level from the original configuration can invalidate the keys. Special keys may need to be acquired in the event of a test or a declared emergency, and it is best to know upfront if you are to be responsible for acquiring these or if there will be additional costs associated with the acquisition of them should the contracted company be required to acquire them.

Special Security Procedures In many cases, the agreement will have a clause or clauses concerning security measures that are required or being provided. It is important to make note of these measures and procedures as they may or may not be sufficient in your organization's given situation, and special provisions may need to be made to ensure the safety of all parties and the security of property and information in the event of a test or of a declared disaster.

Notification of Systems Changes There should be a process set forth in the agreement documents providing a process or procedure by which you can change the configurations of your systems covered by the agreement. There will likely be charges associated with these changes; however, as a means to protect your organization's interests in being able to upgrade and stay current on systems, it is worth the extra money to have written into the agreement the clauses that will allow you to make such changes.

Hours of Operation What happens if you have to declare a disaster at 3 A.M. on a Sunday morning? Will you be able to reach anyone at the DR company to let you in and help you get set up? Will you be able to find someone to call? Will you even be able to get into the site before 9 A.M. on Monday morning? What will it do to your company if you are unable to start your disaster recovery efforts until 30 hours after the disaster happens? Will you be able to access your backup tapes in that time frame, or will you have to wait until Monday morning for that, too? The hours of availability for all agreements should be spelled out so that you can be sure you have access to the resources required when they are needed, not necessarily just when it is convenient.

Specific Hardware and Other Equipment Required for Processing Does your organization have special hardware needs? Are you running on a PDP-11 (a DEC computer from the early 1970s and still in use in many companies today) and you need to have that available to meet your disaster recovery efforts? If you have such a need, it should be spelled out in detail in any agreement. This extends to any odd tape devices required for recovery, printers that are programmed for in your application (do you only print your reports on certain-sized paper, do you need a special printer to handle your check or invoice printing requirements, or is your application set up to send commands to a DEC Laser printer or an HP printer and the company with whom you have a contract has standardized on something else?), or any other quirky hardware that may be difficult to obtain on short notice.

Personnel Requirements Many companies will contract to provide hardware and even software to another company in the event of a declared disaster or for the purposes of testing a disaster recovery plan, but what happens when you need people to do the job? First, is your disaster recovery plan sufficient that it would allow someone walking in off the street to do the job for you? Then, what are the security requirements for the people performing the recovery if they are not employed by your organization? Finally, where are you going to get these people, how much are you going to have to pay for them if the need arises (the Year 2000 "scare" taught many people, COBOL programmers in particular, the real facts about supply and demand, and looking for people who can fill roles with no warning in a disaster may mean paying an extreme premium for the service), and how much notice is going to be needed to get qualified people to perform the recovery or the tests? These questions should be spelled out in any agreements that you have that could be impacted by the answers to the questions.

Circumstances Constituting an Emergency *Arguably one of the most important parts of the agreement is the definition of situations that may arise that constitute an emergency.* One company's total fiasco might only be a bump in the road to success for another. Planning far enough in advance for what your company might see as a sufficient disaster or emergency not only allows you to plan but also those organizations with whom you sign agreements to plan for such a contingency.

It is important to outline what constitutes an emergency. Although you don't actually have to detail everything that might happen, you should at least make sure that you find a way to quantitatively detail what the impact to your organization needs to be in order for you to declare a disaster and therefore make use of the agreement.

Process to Negotiate Extension of Service Because any agreement is limited in duration, the process for renegotiation of services needs to be outlined in the agreement or in auxiliary documents to the agreement. This is important so that both sides of the negotiation know what to anticipate and when to anticipate it.

Nonmainframe- or Nonserver-related Resource Requirements Although many people, and indeed many companies, think of mainframes and midrange computers or Linux and Windows servers when they think about disaster recovery (some even take the printers into account), many don't think about the smaller and less glamorous resources such as desktop computers, laptop computers, phone systems, and other devices that allow people to actually do their jobs.

Because many disaster recovery service providers don't provide for the computers that the users will need in order to get back to work so that the company can continue with business as usual, those facilities need to be planned for as well.

For example, IBM doesn't necessarily have the facilities at their Boulder location to house a hundred or more people sitting at desktop computers without special arrangements being made to make the facilities available. However,

it likely would be possible to find, in the same general location, office space or warehouse space that could be made ready fairly quickly so that users could be brought to the area of the disaster recovery site for the duration of the recovery efforts and until the main office site could be made safe for business again.

Priorities Recall that in Chapter 1 we looked at breaking necessary functions into groups of prioritized jobs for the purpose of making the recovery and the recovery planning more efficient and organized. These priorities can be translated into the agreement. The resources necessary for a tier 3 recovery need not be available at the same time as those that are necessary for a tier 1 recovery. In the best-case scenario, those resources may never be required because the organization's main office may be available by that time.

If the resources required can be put off, then cost can be spread out over the duration of the recovery effort and concentration can be paid to the job at hand rather than worrying about the extra hardware that may be around the area.

If there is hardware around, someone in IT will often try to recover everything available rather than concentrating on the job at hand. If there are spare cycles, a technologist tends to try to apply those cycles to any job at hand. It is better, in this case, to pay close attention to the job at hand.

Other Contractual Issues There are likely other issues that need to be put in writing. Legal assistance should be sought with this as with any legal agreement. If you are with a large organization, there will likely be legal council already associated with the organization. If you are with a smaller company or you own the company, you may need to find legal assistance for the purpose of these agreements.

Alternative Site Selection Criteria

Whenever you are selecting possible sites for your disaster recovery site, you will be faced with many alternatives. It would be helpful to have a list of prioritized features that you are looking for in a recovery site. This section provides you with some of the criteria that are commonly used in the determination of provider and site selection criteria.

Number of Sites Available

Although you will likely want to recover to the closest location available to limit the cost of travel and the time needed to get to the recovery site, a single disaster may affect more than one location, and different disasters may happen in different locations at the same time. The availability of more than one location can mean that you have choices in the event of a declared disaster. You may have to agree on a primary recovery site so that the provider can plan ahead for the eventuality of your needing to recover at this location.

Many companies can't offer a wide selection of recovery sites, but they will be able to give you at least two or three different locations to which you can relocate your functions.

Distance from Site

Keeping in mind what your company will need in order to recover and get back up and running in a timely manner, and the time frame that you have committed to be able to do just that, you will need to determine the best places to find a recovery site. If you are in Minnesota and you have 72 hours to have your recovery done, it isn't likely that you will want to use a recovery site in Australia. Colorado may be a viable alternative, or even Atlanta.

Another reason to look at distance from the site is that if the site is too close to your business, you may find the recovery site affected by the same emergency that is affecting your business's primary location. If a hurricane were to affect Tampa, choosing a disaster recovery location in Tallahassee may not be an optimal solution. Houston or even New Orleans may not even be good solutions. Pittsburgh, however, may be a solution for you or New York or Delaware.

Balancing the need to stay far enough away to allow you to escape the disaster and being close enough to make getting to the site in a timely manner practical can take on the air of an art rather than a science.

For example, if you are recovering a business located in Amarillo, Texas, or Hibbing, Minnesota, you may need to take into account that there are a limited number of locations to which you can fly directly, even though there are international airports available. You may have to plan in time to get from your location to another location, stopping at interim airports with an indeterminate layover at any of them.

In narrowing your choices, look at current flight schedules and notice not the shortest time from point A to point B but the longest time it could take you to get there given the schedules. Then double that time to account for possible delays or cancellations. It is better to have a realistic idea than to suggest the best possible picture to the team and to upper management and be proven wrong at the expense of the company. The tighter the timeline required and the tighter the budget for the recovery effort, the more important it is to err on the side of pessimism.

Facilities

Look at the complete picture of the recovery site and the surrounding location. How hard is it going to be to get what you need while you are there? Don't just consider what will be needed immediately for the recovery proper. What will be needed for the period of time after the recovery for employees to get back to work and be productive? Computers are a given, but what about notebooks, pens, and envelopes? And people can't work 24 hours a day. What is the local environment like? Let's look at some of these considerations.

Office Supplies One of the most overlooked things in a disaster recovery plan and in the selection of the disaster recovery site location selection is the ability to acquire office supplies. Although in a perfect world we could exist in a paperless society, in reality, this really isn't much of an option. People still prefer to take notes on paper with a pen. Even a stylus and a PDA don't usually offer people much compensation for being able to write with a paper and pencil.

Printer paper and toner may or may not be included in the cost of the recovery site. You should know upfront what is and isn't provided and you should take into account where you can purchase, locally or quickly, the products that you will need.

Meals You have people; they have to eat. If you are looking at being at this location for an extended period of time (a month or more), having everyone subsist on vending machine food isn't practical or healthy. Paying attention to the restaurants and supermarkets in the proposed disaster recovery site area means that you can prepare for the fact that you will have many people who will need to be fed and in turn can prepare those going to the recovery site for what they will find there.

One company I know of created a disaster recovery journal. It is a three-ring binder with menus from local restaurants, maps to the restaurants from the DR site, and maps to local shopping establishments (malls, Wal-Mart, grocery stores) so that anyone who ends up at the disaster recovery site (for testing or for a declared disaster) can go to a central location and find information on what they need to know. Every time there is a DR test, one person on the testing team is designated to go to all the local establishments and get a new menu for the journal. This may not be a part of the true disaster recovery, but it is something that will make those involved more at ease.

Living Quarters Look at the proximity of hotels, motels, and short-term housing to the disaster recovery site. A site is less practical if it has only long-term housing (a one or more year lease) close to it.

Again, making an up-to-date list every time there is a DR plan or having someone maintain a list of these local residences will allow those traveling to have a starting point both for the testing and for a declared disaster.

Postal Services If your organization is going to have to send and receive mail (bills, invoices, payments), having ready access to postal services will be a must. If you are going to be at the backup location for more than a few days, you will likely find yourself having to rely on the postal service in the recovery location. Knowing where it is and how to access it and creating a working relationship with the people there will go a long way toward easing pains should disaster declaration become necessary. Knowing where to access and how to locate these facilities should be part of the criteria what you base your decision for the choice of location.

Recreational Facilities Okay, so maybe it isn't one of the things that people think of first when looking at disaster recovery sites, but when you are dealing

airport in Lubbock, plus the wait time for the flight because you never know how close to the flight time the disaster will be declared, and you could be chewing up a significant number of man-hours, not to mention recovery hours, in the process.

These are all hours that you can't spend recovering your data or your systems. But they are hours that count against those hours in your SLAs.

FYI *Air Travel*

Keep in mind that, in the event of a declared disaster, you won't be able to rely on getting the best prices from somewhere like Orbitz or Priceline. You will likely have to walk up to the ticket counter at the local airport and pay whatever the asking price is for the next flight out. And you may not be lucky enough for that flight to be leaving right when you want it to. You may find that you will have to wait extended periods of time for the flight that you need and pay a premium for that wait. All these costs are costs associated with the choice of recovery site.

Cost of Temporary Living No employee, no matter how good-hearted, will foot the bill for the cost of temporary living expenses incurred while they are either at a recovery test drill or at an actual recovery. The company will have to compensate the employee for the temporary living costs. The more people involved, the more that expense can add up. This is yet another expense to be taken into account. The cost of living in Clarendon, Texas, is likely to be somewhat less than the cost of living in Boulder, Colorado, and the cost associated with living in Boulder during peak tourist season for extended, though temporary, periods of time can get to be excessive.

This is not usually a cost that is considered in association with the location of the disaster recovery site, but it ought to be taken into consideration.

Contract

Naturally, the contract should be considered when looking at the recovery site. Terms of the contract, what is included in each contract that the company is evaluating, the duration of the contract, and any additional costs for goods or services that may not be included in any given contract should be considered.

Designing Recovery Solutions

Without the ability to recover, there is no purpose for any backup other than to waste media. Going hand in hand with whatever backup solution is in place, or that you are putting into place, is the knowledge that at some point you will likely have to recover using these backups.

We have to establish a recovery site, select backup and recovery strategies, identify the tools necessary to meet our storage needs for the backups we have taken, and identify those places where creative solutions need to come into play.

Establishing a Disaster Recovery Site

One of the first decisions that has to be made in the disaster recovery planning process is how you intend to structure your backups and your planned recovery. Will there be automatic failover and no intervention will be necessary? Or will you have to retrieve tapes from offsite storage and have them transported to an alternative recovery site? Each alternative comes with its own advantages and disadvantages. Each also comes with trade-offs and costs. Some costs are direct, as in the physical outlay of capital, and some are indirect, as in the cost of bodies and time to recover the organization.

Choosing a Site: Hot, Warm, or Cold Standby The choice of sites often comes down to a matter of money and practicality. Many companies would rather have a *hot backup site*, where they can fail their entire business over within a matter of minutes to hours. This kind of assurance, however, costs money—typically a lot of money. It means providing an entire office complex in another location (almost always another city, and usually another state). If the company is a manufacturing company, this could also mean the contracting of another facility or subcontracting work out in case the primary facility becomes unavailable.

This choice may be influenced by clients, stakeholders, and shareholders, but ultimately it is a matter of measuring the return on investment and determining how long your organization can withstand the risk of being without its ability to process work.

There are cases where a hot backup is necessary. Recall that we undertook a risk analysis in Chapter 3. We can begin to think in terms of identification of that risk as it pertains to the identification of recovery solutions. The acceptance, or the non-acceptance, of risk will drive many of the decisions that the organization makes surrounding the backup and recovery solution decisions. Recall that risk is the chance that an event will occur. This chance of an event occurring, coupled with the loss of revenue that will occur if that event occurs and the cost that is associated with it, are all factors that have to play into the organization's determination of whether the cost of the recovery solution is justified.

If you determine that there is little chance that a disaster or even an emergency event will occur, then you have minimized the risk from the beginning, and the choices that you make concerning recovery solutions is very much simplified. If, however, the organization determines that there is a significant enough risk surrounding the organization, there will be additional amounts of investment that the organization will have to be willing to make.

IN PRACTICE: Small Business

Risk, and how an organization chooses to deal with mitigating that risk, is relative. Big businesses tend to be able to throw more money at a solution than smaller businesses, but the size and complexity of solutions is relative. They may not be throwing any larger a percentage of their money at the solution than a small business or a home-based business might be in similar circumstances.

Small businesses are sometimes less adaptable to disaster than large companies. Someone who has turned a room in their house into a small-scale manufacturing line may not be any more or less able to withstand the occurrence of a disaster than Ford, Dodge, or Westinghouse. If all the capital and thought is tied up in creating the product, and little is tied up in learning what they would have to do if a disaster levels their house, they are in little better position than a large organization in a similar situation.

However, an accounting company that is based out of the home office of the accountant may be totally recoverable with minimal downtime if that company invests in the time and material necessary to make a periodic backup of the software and records on its system and stores those backups in a waterproof, fireproof file cabinet in a closet and makes two duplicate copies (DVDs are even getting cheaper today) and sends one to a relative in Pennsylvania and puts the other in a safety deposit box in a local bank.

Simply replacing damaged or destroyed hardware from a local commodity hardware store (electronics store, office supply store, or even from the mall) and recovering the backups would be all that it would take to recover the business in a storefront three miles from the primary location or in an apartment or hotel room nearby.

It is up to every individual organization to assess the risk and determine what it will mean to that organization and how best to recover should the worst happen.

Build vs. Rent or Share The decision over whether to build a recovery site (whether dedicated to nothing but being a recovery site or as a line of business of its own) or to rent or share a recovery site with other companies is often a difficult one. The level of trust that has to be involved between two organizations in order to share a recovery site implies a special kind of relationship. Although this kind of relationship may exist fairly easily between two home-based businesses where there may already exist a trust relationship, the same level of trust may not exist between two larger organizations. Not only is there a level of trust that has to exist, but there has to be a small chance that both organizations will need to

use the recovery site at the same time and that both are going to be able to uphold their end of the agreement over time. On the other hand, it is typically more affordable to establish a share relationship than to shoulder the entire cost of the recovery site. In the end, the decision to rent or share comes down to total cost.

Selecting Backup and Restoration Strategies

The choice of backup strategies in order to make the optimal use of both time for backup and time and effort for recovery is often a decision that will be based as much on guess work and art as it is on facts and figures. Many times the backup strategy that is chosen by a company is simply the favorite of one of the administrators or is just something that everyone seems comfortable with. These, however, are not valid and sound justifications for a backup strategy.

Full backups are complete backups of a system. They are a snapshot of how a system looked at a given point in time. Full backups (often referred to as cold backups) are invaluable to an organization for many reasons, and should always be a part of any backup and recovery solution. They may not, however, be the all-inclusive solution to a company's every need. In order to take a full backup, all the applications and databases on a system need to be shut down. This means that time needs to be taken to complete the shutdown and the restart after the backup is taken, and a backup taken of the full system. The disadvantages of full backups are that they take time and that you can only recover to that specific point in time, if that is the only backup on which the company is relying.

Incremental backups are those backups that gather together, in a perfect scenario, only those changes that have occurred since the last full or incremental backup. This kind of backup is not always possible without the assistance of third-party backup tools to assist with the backing up of the data and surrounding application. Incremental backups provide the fastest backup scenario, because you are only backing up the data that has changed in the period of time since the last incremental or full backup. What's more, because you are backing up the minimal amount of data that is practical, the storage space for these backups is the smallest. The trade-off here is that your recovery scenario can be longer because you have to incrementally restore each of the files from the original full backup to the point in time to where you need to recover. If you don't take full backups very often, this can be extensive.

Differential backups are those backups that contain all changes since the last full backup. The advantage of a differential backup is that you only have to restore the cold backup and one additional backup in order to restore the system to any given point in time. The disadvantage is that you have to back up and store the backups of redundant data. For example, if you take full backups every Saturday night, and differential backups on Sunday, Monday, Tuesday, Wednesday, Thursday, and Friday, by Friday you have duplicated the data that you backed up on Sunday five times.

IN PRACTICE: Small Business

Risk, and how an organization chooses to deal with mitigating that risk, is relative. Big businesses tend to be able to throw more money at a solution than smaller businesses, but the size and complexity of solutions is relative. They may not be throwing any larger a percentage of their money at the solution than a small business or a home-based business might be in similar circumstances.

Small businesses are sometimes less adaptable to disaster than large companies. Someone who has turned a room in their house into a small-scale manufacturing line may not be any more or less able to withstand the occurrence of a disaster than Ford, Dodge, or Westinghouse. If all the capital and thought is tied up in creating the product, and little is tied up in learning what they would have to do if a disaster levels their house, they are in little better position than a large organization in a similar situation.

However, an accounting company that is based out of the home office of the accountant may be totally recoverable with minimal downtime if that company invests in the time and material necessary to make a periodic backup of the software and records on its system and stores those backups in a waterproof, fireproof file cabinet in a closet and makes two duplicate copies (DVDs are even getting cheaper today) and sends one to a relative in Pennsylvania and puts the other in a safety deposit box in a local bank.

Simply replacing damaged or destroyed hardware from a local commodity hardware store (electronics store, office supply store, or even from the mall) and recovering the backups would be all that it would take to recover the business in a storefront three miles from the primary location or in an apartment or hotel room nearby.

It is up to every individual organization to assess the risk and determine what it will mean to that organization and how best to recover should the worst happen.

Build vs. Rent or Share The decision over whether to build a recovery site (whether dedicated to nothing but being a recovery site or as a line of business of its own) or to rent or share a recovery site with other companies is often a difficult one. The level of trust that has to be involved between two organizations in order to share a recovery site implies a special kind of relationship. Although this kind of relationship may exist fairly easily between two home-based businesses where there may already exist a trust relationship, the same level of trust may not exist between two larger organizations. Not only is there a level of trust that has to exist, but there has to be a small chance that both organizations will need to

use the recovery site at the same time and that both are going to be able to uphold their end of the agreement over time. On the other hand, it is typically more affordable to establish a share relationship than to shoulder the entire cost of the recovery site. In the end, the decision to rent or share comes down to total cost.

Selecting Backup and Restoration Strategies

The choice of backup strategies in order to make the optimal use of both time for backup and time and effort for recovery is often a decision that will be based as much on guess work and art as it is on facts and figures. Many times the backup strategy that is chosen by a company is simply the favorite of one of the administrators or is just something that everyone seems comfortable with. These, however, are not valid and sound justifications for a backup strategy.

Full backups are complete backups of a system. They are a snapshot of how a system looked at a given point in time. Full backups (often referred to as cold backups) are invaluable to an organization for many reasons, and should always be a part of any backup and recovery solution. They may not, however, be the all-inclusive solution to a company's every need. In order to take a full backup, all the applications and databases on a system need to be shut down. This means that time needs to be taken to complete the shutdown and the restart after the backup is taken, and a backup taken of the full system. The disadvantages of full backups are that they take time and that you can only recover to that specific point in time, if that is the only backup on which the company is relying.

Incremental backups are those backups that gather together, in a perfect scenario, only those changes that have occurred since the last full or incremental backup. This kind of backup is not always possible without the assistance of third-party backup tools to assist with the backing up of the data and surrounding application. Incremental backups provide the fastest backup scenario, because you are only backing up the data that has changed in the period of time since the last incremental or full backup. What's more, because you are backing up the minimal amount of data that is practical, the storage space for these backups is the smallest. The trade-off here is that your recovery scenario can be longer because you have to incrementally restore each of the files from the original full backup to the point in time to where you need to recover. If you don't take full backups very often, this can be extensive.

Differential backups are those backups that contain all changes since the last full backup. The advantage of a differential backup is that you only have to restore the cold backup and one additional backup in order to restore the system to any given point in time. The disadvantage is that you have to back up and store the backups of redundant data. For example, if you take full backups every Saturday night, and differential backups on Sunday, Monday, Tuesday, Wednesday, Thursday, and Friday, by Friday you have duplicated the data that you backed up on Sunday five times.

Mirror backups are those backups that directly copy all selected files, directories, mount points, and file systems from one set of disks to another. A mirror backup is conceptually identical to a full backup except that the data in question cannot be compressed in any manner and cannot be password protected. Mirror backups are often made when using RAID by splitting the mirrored device and maintaining one of the mirrors (or backing up one of the mirrors) as the backup. The space required for a mirror backup is equal to that of the data that it is backing up, but typically (as long as the backup stays on disk) the recovery from failure is faster. This solution, however, is not typically practical in a disaster situation because a disk array is difficult to store and transport and a disaster at the primary site will likely wipe out the backup as well. Rather, a mirror backup is normally used as the source for an incremental, differential, or full backup so that the backups can be carried out without impact to the source system. The requirement for a duplicate set of hardware (or triplicate, depending on the RAID configuration and redundancy requirements of the organization) remains, but the recoverability of the system is greater when the mirror is used as a backup source rather than the backup itself.

Most organizations will come up with their own combination of these backup strategies. They may delineate their backup strategy based on the amount of activity on a system or on the features available to a given system by the underlying architecture. It may be easy to back up the entire system weekly with a full backup and then forget about most of the executable programs because they don't change much. The data in a transactional database (a database that has users inserting data into and altering data already resident in the database) will likely change minute to minute. That data will need to be backed up frequently, either incrementally or differentially (the choice of how has to be based on your organization's recovery requirements), in order to be sure that you can recover the data as quickly as possible to any given point in time. The data in a data warehouse, because it is far less volatile, may change only once or twice a week, depending on your load strategy, thus a single differential or incremental backup after those loads would be sufficient and similar to each other in recoverability and storage.

Hot backup is another alternative that is becoming more popular with larger organizations. With a hot backup, the entire decision (backup site location selection, backup and recovery strategy, vendor, etc.) is a package deal because the recovery location is in constant communication with the primary systems. This means that, in the event of an emergency or disaster, connectivity to the systems automatically fails over to the backup site and there is effectively no downtime noticed. This is an optimal solution for many organizations that have multiple data centers or multiple locations that can be used as data centers.

Although any backup solution has to effectively fit within the window available to you for the backup, the primary consideration in these different options needs to be the recovery needs of the organization. The most elegant backup strategy is worthless if you cannot effectively recover using the backup

in a timely manner, and in a manner that meets all the needs of the organization. It is not practical to believe that, if it takes 10 hours to take a full backup of a system and each incremental backup takes 5 hours, and you need to restore a full backup and five incremental backups, that your recovery window will be any less than 35 hours, and this does not take into account the fact that tapes will likely have to be changed out at the end of each step or that restoration from tape usually takes longer than backing up to tapes. If you have a contractual agreement with your customers that your data and applications will be available for them within 72 hours, your recovery strategy, barring any difficulties along the line, takes at least half of that time. In a case like this, differential backups may be a solution to your problem, at least for some of your systems.

The combination of the consideration of backup strategies to fit within the confines of the organization's requirements and the ability of those backups to meet the requirements of the organization is considered to be matching the strategy to the operational constraints. In all aspects of backup and recovery, disaster recovery, or business continuity planning, it is critical to remember the operational constraints of the organization. If any part of the plan doesn't fit within those constraints it may be functional, but it will not likely meet the needs of the organization and will therefore not be a valid solution or consideration.

Although the backups we have talked about so far can provide a solution to the recoverability of the electronic files stored by a company, the DR team must be aware that not all the company's records are stored on computer systems. There will be a certain number of records that are stored on other media, such as paper or microfiche. Planning for the storage, backup, and recoverability of these records is as important as for the computerized records.

Backing up paper documents is as simple and as labor and space intensive as making copies and shipping those copies offsite to be stored in a climate- and access-controlled location. Alternatively, an organization can choose to digitize these paper documents, having those documents scanned and stored electronically on tape, disk, CD, or DVD. Similarly, with microfiche and other non-electronic documents, duplicate copies, in one form or another, should be created and stored in a secure location.

These non-electronic documents are often overlooked in the disaster recovery effort, and are often critical pieces of information that get the organization into trouble. Older data, data from before much of the information was stored digitally, is required by governmental regulations to be available for inspection for up to 20 years or more. If an organization finds itself in need of this information and finds that it is later unavailable, it can be catastrophic.

Storage area networks (SANs) and network-attached storage (NAS) have had an impact on the backup and recovery decisions that an organization makes. Because the amount of data that can be stored on a SAN or a NAS can be significantly greater than that stored simply on internal or regular RAID array, the time required to back up these devices can be significantly greater than on common storage devices. Further, because on any given network storage device

an organization can store data from multiple heterogeneous operating systems, the backup and recovery solution needs to take these differences into account.

For more information, see the following article: **http://www.ameinfo.com/39672.html**.

Storage Backup and Recovery Tools

There are many backup and recovery tools on the market today. It is important to not tie yourself to a tool that is proprietary enough that you can't find a recovery site easily that will be able to handle your recovery needs or that you couldn't change recovery sites after a period of time if an earlier site no longer meets your needs.

Managing Stored Data and Applications Different vendors bring a variety of products to market. A simple search for data backup solutions yields links to many products that could provide you with the solutions that you are looking for. There are many others, and a fairly exhaustive search should be carried out to determine the optimal fit not only for an organization's situation but also for the environment it has as its infrastructure. Many solutions are targeted to Windows and Unix, whereas other vendors target primarily IBM-based products.

The following is a list of storage solutions that are applicable for many businesses.

- Veritas (**www.veritas.com**) provides storage protection, automation, and performance solutions for a wide variety of operating systems and databases.
- EMC (**www.emc.com**) has a full line of data management and backup and recovery solutions.
- EMC Legato (**www.legato.com**) provides data protection and availability products for an assortment of operating systems and platforms.
- Network Appliance (**www.netapp.com**) is a high-performance file server, filer, and caching system for Windows and Unix systems that provides the ability to back up and recover data in an extremely rapid manner. It also provides access to analyst reports, customer stories, data sheets, technical reports, videos online, and a substantial glossary of terms.
- Storage Tek (**www.storagetek.com**) brings with it data protection, archiving, storage productivity, and industry solutions. It also has a substantial library of case studies and technical and white papers.

There are, of course, other alternatives, and there is always the home-grown method that relies on scripts and utilities that are resident in either the operating system or that can be built and stored in the organization's software and script library.

FYI *Backing Up Scripts for Recovery*

If you are maintaining your own backup and recovery scripts, you should take care to have an external source for those scripts—one that is not reliant on the recovery of the data to the server as a means to access the scripts. It would be very difficult to recover the scripts if the scripts are needed to recover the data. A minimal backup would include a script library that is maintained independently and retrievable with just the operating system utilities such as copy (regardless of the OS there is a copy command).

The Impact of Storage Area Networks (SANs) on Recovery It is often difficult to find an elegant and cost-effective method of disaster recovery solution for an environment that uses a SAN infrastructure. SANs are being used more and more as a solution to an organization's data storage challenges.

A SAN is a network of storage disks. A SAN typically connects multiple servers, usually with heterogeneous operating systems, to a centralized pool of storage. The idea is that rather than having to manage hundreds or thousands of servers, each with their own set of storage disks, by using SAN technology you can improve overall system administration. Organizations choose this type of architecture so that they can centrally manage the storage resources with a single backup and recovery solution, a single disk maintenance solution, and a single schedule over the entire organization. This architecture can create issues, however, when it comes time to perform disaster recovery because of the heterogeneity of the operating systems and the pooled storage model.

As a pooled storage solution, the SAN doesn't care what is put on it, what the format is, what the operating system is, or whether the data is ASCII or EBCDIC. But you will have to take care that your backup and recovery solutions take into account that there may be data included from multiple operating systems and for multiple purposes.

Further, because the storage solution is likely to be used by nearly every system in the organization, additional care needs to be taken when considering the recovery timeline, the order in which the data needs to be recovered, and the criticality of the data. It may be that none of your data will be available until all of your data has been recovered, and this may not be acceptable to the business as an entity.

Determining what a cost-effective DR solution is when you have standardized to any extent on a SAN storage solution is based on your organization's needs. Some solutions are based on host-based data replication over a network (LAN or WAN) to a remote and often dissimilar SAN technology. Alternatively, storage virtualization is seeing growing acceptance and can allow for the replication of the data across the SAN or to dissimilar storage arrays.

FYI *Storage Virtualization*

Storage virtualization sounds like a complex concept, but it is simply the transparent amalgamation of multiple storage devices into what appears to the end user and the other hardware to be a single storage unit. A high-speed network is typically key to allowing for the elegant use by multiple heterogeneous operating systems and hardware solutions to access this virtual pool rapidly. Storage virtualization is typically handled by software masking the underlying hardware details.

Current Trends in Recovery In recent years, more and more organizations have realized that any event that is disruptive of their business can be a crisis. Acknowledgment of the impact of the Sarbanes-Oxley Act, the Gramm-Leach-Bliley Act, and HIPAA (the Health Insurance Portability and Accountability Act) as realities of everyday life and of nearly every business's daily life has driven many of these realizations. These kinds of regulatory issues have been the impetus for many organizations to launch their own disaster recovery planning.

FYI *Governmental Regulations*

The Gramm-Leach-Bliley Act, also known as the Financial Modernization Act of 1999, includes provisions to protect consumers' personal financial information that is held by financial institutions. (**www.ftc.gov/privacy/privacyinitiatives/glbact.html**)

The Sarbanes-Oxley Act was designed to radically change the way that publicly traded companies recorded and reported inancial information. (**www.ey.com/global/download.nsf/Russia_E/EY_Sarbanes_9_12_02e/$file/EY_Sarbanes_9_12_02e.pdf**)

HIPAA addresses health and insurance information privacy protection for individuals. (**www.hipaa.org/**)

More governmental regulations will likely come in the near future, and with these regulations will come evolving requirements for disaster recovery and implications for additional organizations.

Terrorism is becoming more widespread and is impacting more and more lives daily, but until recent events (9/11/2001 in the United States, *London's transportation system bombings in 2005,* and the continuing global strife over the recent years) it was not always looked on as an event that

would touch many organizations. Organizations are starting to take this as a wakeup call. Even if they don't believe themselves to be at risk of being affected by a terrorist attack, they are starting to realize that they are not beyond the effects of disasters. They are also realizing that they are relying on organizations that might be affected by terrorist attacks (utility companies or transportation companies, for example) and are taking steps to mitigate these risks.

Data and information are becoming more of a critical concept for more organizations, and the demand for realistic 24/7 access to data is becoming the norm rather than the outlier in the statistical analysis of organizations. This demand for continuous access to data and information is impacting more disaster recovery plans, and service level agreements are equally impacted by the tightening timelines. This means that backing up data has to evolve to meet the need, as do recovery techniques and attempts to meet the tighter requirements.

Because email is becoming more critical as a means of business communication, and because it is becoming more relied upon as evidence in litigation, lost emails, or emails that are not retained in a sufficiently secure manner, can be a disaster recovery event. With more organizations based entirely on the Internet and with email being the primary means of communication with even a greater number organizations, the necessity of maintaining accessibility to the email system and to the historic information gathered from this email information is critical. This is true not only from a disaster recoverability perspective, but from the perspective of security as a whole. We not only have to be able to accurately recover and provide access to email information, but also be able to accurately report on and restrict users who have had access to that email and provide its relevant change history.

Restoring Communications and Recovering Users

Recovery of the business is nearly pointless unless you can restore communication with and recover all processes to the ultimate end users. Not only will this allow your business to continue its line of business and to hopefully grow and thrive, it will foster the goodwill that is necessary in doing business with both your customer and every stakeholder in the organization.

Determining Vital Users with BIA Recall that you previously have done a business impact analysis (BIA). When it comes to determining which users to restore connectivity to and in what order, this will become, yet again, an invaluable tool. One of the areas of analysis for the BIA would have included those users, both known and implied, that would be impacted by an event. Assigning weights to these users allows you to determine in what order to recover not only the software and hardware resources, but in which order to restore connectivity and accessibility to given users or given user types. It is important to restore

user access in a logical and correct order not only to critical systems and functionality but also to those areas that will be most directly impacted by service level agreements.

Rerouting Voice, Mail, and Goods Delivery So you have relocated your systems and you have relocated many of your users and business units if necessary, and you need to get back to business as usual. But to really be back to normal, communication (other than strictly electronic digital communication) also needs to be re-established. Mail (payments and invoices, naturally, but other communication as well), voice communication, and delivery of goods to alternate locations need to be taken into account.

Voice communication is a vital link in nearly any organization's business. What's more, there is no acceptable alternative to voice communication in many situations. It is the manual alternative in most cases for any other alternatives, and it remains the communication solution of choice for most people. The ability to know from the tone of voice that both parties understand what you are discussing is important. The sound of another human voice on the other end of the line to provide information or to hear out a situation is often necessary to allow all parties to maintain a comfortable relationship.

There are different techniques that allow you to minimize the chances that voice communication will be affected, including redundancy and diversity of implementation. But these don't address what to do if there is a catastrophic disruption in service. In this case, you will need to work with the telephone company to determine the best way to handle the situation. Discuss location with them *before* there is a disaster situation. If you wait until there is a disaster, you may not be the only one making similar requests, and this could cause delays. An ongoing relationship with the telephone company or companies involved will prove useful when the disaster is declared.

Rerouting voice communication can be a challenge, depending on your location and your method of communication. The simplest solution is to standardize your organization on the technology of Voice Over IP (VOIP). VOIP is a method of taking audio and analog signals, converting them to digital signals, and transmitting them, at least part of the way, over the Internet. In this way voice communication can be routed to anywhere, the IP address implies routing. Redefinition of the IP address in a routing table can mean that the voice communication can be routed to an alternative location.

Mail is much easier to reroute. Simply submitting the proper forms will allow mail to find its proper locations, and this solution is customizable *enough to route mail directly to where the* receiving location needs to be, even when multiple locations are involved. Again, however, a working knowledge of where to get the proper forms, or maintaining a stock of the proper forms with the disaster recovery materials, can minimize downtime.

Rerouting deliveries of goods and services may not be as simple as rerouting mail delivery. Special arrangements will need to be made for delivery, either in postponing the delivery or in rerouting deliveries during the time that you are in a temporary location. This should include not only the time you are in the new location but also the period shortly after, to take into account orders that are in process and not yet delivered at any point in time.

There may be orders that are in process at the time of disaster and already out for delivery. These orders may need to be rerouted and there may be additional cost associated with either the return of the orders or with rerouting the deliveries to an alternate location. This will be a sunk cost that won't likely be recoverable, but most suppliers will work with you.

Eliminating Network Single Points of Failure This is by far the most important thing to take into consideration when looking at backup and recovery. For every single point of failure that you can eliminate, that is one place where your system may be robust enough to make disaster recovery, regardless of the level of loss, irrelevant.

A single point of failure is any point in your system where a failure will result in loss of the system, loss of the network, or loss of ability to access information. This single point can be hardware or software, and is often not even thought of as a point where something might fail. Some single points are difficult to eliminate. Even if they can't be eliminated completely, recognizing them as points where failure can occur and taking steps to monitor and maintain these points is a way to keep them healthy and to recognize that if any issues appear to be presenting themselves they can be taken care of as quickly as possible.

Elimination of these single points of failure doesn't come without cost. You will be duplicating hardware and, often, software systems. The cost-benefit trade-off in each case needs to be considered at each point that you uncover.

But wait. Not every single point of failure is directly connected to your computer. Often the single point of failure is a user who is overly tired, under trained or who thinks that he or she knows more than they actually do. Several years ago, the I LOVE YOU virus wreaked havoc on many organizations. An accounting user at one company knew just enough to understand the language in which the virus was written but didn't realize exactly what the ramifications might be if the virus was opened. He associated the virus with Adobe Acrobat, thankfully, and called the local help desk to help him to unassociate the virus with the software so he could find another way to open it and read the code.

Just as often, a user whose IT department is less diligent in keeping computers updated with the current security patches and virus definitions can be the

cause of failure in a system or in an organization. Remember, a disaster is not necessarily the catastrophic loss that accompanies a hurricane or tornado, it can be simply the extended inability of an organization to continue with the work that needs to get done.

One way to eliminate this kind of single point of failure situation is open communication and training. Another is diligence in the maintenance of the systems connected to the system. Your system is only as robust as its least secure single point of failure, regardless of whether that point is hardware, software, or humanware.

This is just as important when recovering the organization as it is when examining the organization's home network. When you are looking at where to recover, examine where there might be single points of failure. When constructing the contracts, make sure that you read carefully and negotiate appropriately so that you eliminate as many points of failure as possible. Once you are at the disaster recovery site, there are few alternatives if there are further disasters.

Connecting End Users One of the things that is often overlooked when considering disaster recovery is where your end users will be located, and what kind of connectivity they will be able to rely upon. Although you will likely have control over the definition of IP addresses assigned to the new hardware, you may have to take into account firewall ports and VPN or dial-in access to the network.

Further, you may need to find a location to house your users. Many organizations rely on temporary trailers or set up temporary office space in conference rooms in locations not affected by the event. It is important to think through where these users will need to be, and plan for the rapid implementation of establishing these people into their temporary office space.

Not only is it important to make sure that these end users can connect as rapidly as possible so that the company can return to business as usual as rapidly as possible, it will allow your users to regain a sense of control over their lives and over their surroundings. At a time when they are likely feeling that things are out of their control, the security that you provide to them will allow them to get back to normal as rapidly as possible.

Summary

You should now understand the critical steps in assessing a system to help you best address any vulnerability in the system. It is also imperative that the person conducting the security audit document the specific steps taken as well as any flaws found, and what corrective actions were taken.

Test Your Skills

MULTIPLE CHOICE QUESTIONS

1. One of the first considerations that comes into play in a disaster recovery plan is _____.
   A. data recovery
   B. data backup
   C. data recording
   D. business process

2. The quickest recovery methodology is _____, because your access times do not include the time that it takes to transfer data from another medium.
   A. tape backup
   B. floppy disk
   C. disk-to-disk
   D. portable drive

3. When determining a backup schedule, you need to take which of the following factors into account?
   A. Recovery point objective
   B. Recovery time objective
   C. Maximum allowable downtime
   D. All of the above

4. Which of the following types of failure often goes undetected for extended periods of time?
   A. Software
   B. Hardware
   C. Technical
   D. File recovery

5. Locations, other than your typical business site, where a duplicate of your data is stored and ready for access in a moment's notice are known as _____ sites.
   A. backup
   B. hot
   C. warm
   D. cold

6. A facility that is already stocked with all the hardware that it takes to create a reasonable facsimile of what you have in your primary data center is known as a _____ site.
   A. backup
   B. hot
   C. warm
   D. cold

7. The agreement of two or more companies to join together and act as partners in disaster recovery is known as a _____ agreement.
   A. reciprocal
   B. mutual
   C. coordinated
   D. group

8. Written agreements for the specific recovery alternatives should not include _____.
   A. contract duration
   B. termination conditions
   C. testing
   D. specific personnel names

9. Which of the following is most overlooked in a disaster recovery plan and in the selection of the disaster recovery site location?
   A. Location
   B. Office supplies
   C. Shelter
   D. Record storage

10. The overall cost of the site should include _____ cost.
    A. site and travel
    B. site and supply
    C. travel and supply
    D. None of the above

11. The choice of a backup site often comes down to a matter of _____
    A. money and location
    B. money and practicality
    C. practicality and location
    D. location and site design

12. Complete backups of systems are known as _____ backups.
    A. mirror
    B. full
    C. disaster
    D. precautionary

13. Backups that directly copy all selected files, directories, mount points, and file systems from one set of *disks to* another are known as _____ backups.
    A. mirror
    B. full
    C. disaster
    D. precautionary

14. A network of storage disks is known as a(n) _____.
    A. ASCII array
    B. RAID file
    C. NET
    D. SAN

15. The act that includes provisions to protect consumers' personal financial information that is being held by financial institutions is the _____ Act.
    A. HIPAA
    B. Gramm-Leach-Bliley
    C. Sarbanes-Oxley
    D. Privacy Modernization

16. Imagine what can be overlooked *when considering* disaster recovery is _____.
    A. where your end users are going to be located
    B. what kind of connectivity the end users are going to rely upon in connecting
    C. finding a location to house your users
    D. All of the above

EXERCISES

Exercise 5.1: Pick Backup Software

1. Imagine a data center that has about 300 computers, roughly divided as 75% Windows, 20% UNIX variants, and 5% running other operating systems (Novell, etc.).

2. Thinking back to the backup considerations, make a list of possible backup software to use in your data center.
3. Prioritize your list of backup software, using restore performance and options as your primary consideration.
4. Assuming a different primary consideration, such as price or some other feature of your choice, decide if your priorities would be different.

Exercise 5.2: Pick Backup Media

1. Using the search engine of your choice, make a list of the various tape technologies available, and pick the best tape technology for the data center in Exercise 1.
2. Compare the tape drives available for the media you have chosen and pick a favorite drive.
3. Look at the tape robot systems available from the major vendors, and see if your media and drives will fit into that system.

Exercise 5.3: Offsite versus On Site

1. Your organization has approximately 10TB of data, and you need to decide if your organization should have on-site or offsite tape storage.
2. Your organization must be able to easily recover data no older than one month, as an operational requirement.
3. Your organization's further requirement is that recovery operations must resume at minimal levels for all systems within two weeks of a total catastrophe at the data center.
4. Decide how your organization should house its backups.

Exercise 5.4: Asset Classification

1. Imagine a small business with no more than 50 employees.
2. Develop a list of the hardware assets for the company you have imagined.
3. Develop a list of the other assets that the company has, from the information assets to the human resources.

Exercise 5.5: Picking an Alternative Site

1. Your company has decided to contract with an outsource center for a warm backup site.
2. Assume that your company needs 20 critical employees to oversee the technical recovery at the warm backup site should a disaster occur.
3. Develop a budget for a recovery using the warm backup site and the ongoing operational costs for the site. Does using a warm backup site make the recovery costs lower?

PROJECTS

Project 5.1: Develop a Complete Backup Solution

1. Assume you have a large data center consisting of over 300 systems with multiple operating systems.
2. The entire data center produces approximately 10TB of data today, grows approximately 1TB per year, and only changes about 10% per week.
3. Develop a complete backup system for the data center that allows the organization to quickly recover any system to data no older than one week with tapes in the facility.

Project 5.2: Bench Testing the Window

1. Consider a data center housing 150 computers with a normal performance distribution, that is, some systems run very slowly and a few others very quickly, but a graph of all the systems' performance is a smooth curve.
2. Using available performance data for common tape and optical storage technologies, attempt to determine how long the backup window needs to be to completely back up the entire data center.
3. Each system must be fully backed up each week, and you cannot need more than two sets of media to restore a system to within 1 day of a complete failure.
4. Divide the data for these systems in the following volumes: 5TB, 100TB, and 500TB.

Case Study

A large multinational company has three regional data centers to support the operations of the three major geographic regions where the company operates. A fourth data center serves as the central data warehouse for the company and resides in one of the three geographic regions, but government regulations prevent the two data centers from being located within 250 miles of each other. Consider the discussion of hot, warm, and cold sites in this chapter and write a brief position paper on how this company should approach the data recovery of their data centers with respect to an alternate site strategy. All four centers have equal importance in the processing of information for the company, and a disaster at any one of them impacts all four.