


Realtime  
publishers

# *The Shortcut Guide<sup>™</sup> To*



Untangling the  
Differences Between  
High Availability and  
Disaster Recovery

*sponsored by*

**MARATHON**  
Run to Infinity

*Richard Siddaway*

Chapter 2: Disaster Recovery Is Not High Availability.....	18
It Can't Happen Here.....	18
Disasters Are for Everyone.....	18
Types of Disaster.....	19
Damage and Risk Assessments.....	20
Don't Assume.....	20
What Is Disaster Recovery?.....	21
Disaster Recovery.....	21
Disaster Recovery Compared with High Availability.....	21
Business Continuity.....	23
Do You Still Need to Back Up?.....	24
What Do You Need to Recover?.....	26
Communications.....	27
Data.....	28
Processes.....	29
Involving the Business.....	29
Planning.....	29
Testing.....	31
Time Is Money.....	32
Disaster Recovery Configurations.....	32
Cold.....	33
Warm.....	33
Hot.....	34
Mobile Data Center.....	34
Convergence with High Availability.....	34
When Should High Availability Be Used?.....	35
When Should Disaster Recovery Be Used?.....	35

Combining High Availability and Disaster Recovery .....	35
Geographic Clusters.....	36
Replication.....	36
Mirroring.....	36
Virtualization.....	37
Summary: Stop the Disaster Being a Disaster.....	37

## **Copyright Statement**

© 2010 Realtime Publishers. All rights reserved. This site contains materials that have been created, developed, or commissioned by, and published with the permission of, Realtime Publishers (the “Materials”) and this site and any such Materials are protected by international copyright and trademark laws.

THE MATERIALS ARE PROVIDED “AS IS” WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, TITLE AND NON-INFRINGEMENT. The Materials are subject to change without notice and do not represent a commitment on the part of Realtime Publishers its web site sponsors. In no event shall Realtime Publishers or its web site sponsors be held liable for technical or editorial errors or omissions contained in the Materials, including without limitation, for any direct, indirect, incidental, special, exemplary or consequential damages whatsoever resulting from the use of any information contained in the Materials.

The Materials (including but not limited to the text, images, audio, and/or video) may not be copied, reproduced, republished, uploaded, posted, transmitted, or distributed in any way, in whole or in part, except that one copy may be downloaded for your personal, non-commercial use on a single computer. In connection with such use, you may not modify or obscure any copyright or other proprietary notice.

The Materials may contain trademarks, services marks and logos that are the property of third parties. You are not permitted to use these trademarks, services marks or logos without prior written consent of such third parties.

Realtime Publishers and the Realtime Publishers logo are registered in the US Patent & Trademark Office. All other product or service names are the property of their respective owners.

If you have any questions about these terms, or if you would like information about licensing materials from Realtime Publishers, please contact us via e-mail at [info@realtimepublishers.com](mailto:info@realtimepublishers.com).

**[Editor's Note:** This book was downloaded from Realtime Nexus—The Digital Library for IT Professionals. All leading technology books from Realtime Publishers can be found at <http://nexus.realtimepublishers.com>.]

## Chapter 2: Disaster Recovery Is Not High Availability

---

Chapter 1 looked at various aspects of high availability—what it is and the implications to your organization of not implementing it. In this chapter, the focus turns to the other side of the coin—namely, disaster recovery.

The first chapter clarified that disaster recovery is not the same as high availability. It is now time to consider disaster recovery in its own right and discover what is meant by that term. In order to cover the topic properly, let's consider what is meant by *disaster*.

### It Can't Happen Here

Many organizations do not have a disaster recovery plan. If you take an honest look at the organizations that you know, how many of them actually have a disaster recovery plan that is tested and would actually work if it was invoked? My expectation is that the percentage of organizations with a workable disaster recovery plan is relatively small. The majority of IT administrators will go through their careers without being involved in a major disaster recovery incident. They may have to recover a server occasionally, but that is about the limit. If most organizations don't have a disaster recovery plan and most IT admins never get involved in a full disaster recovery operation, you don't have to worry because it's not going to happen to you. Right?

Wrong.

If you play the odds, eventually they will catch up with you and disaster will strike. "It can't happen here" is not a defense. If your organization is caught in a disaster without a recovery plan, the chances are that your organization will cease operations. Industry statistics suggest that one out of every five organizations that suffers a major disaster never actually recovers and goes out of business.

### Disasters Are for Everyone

Disasters don't respect geographic boundaries, public holidays, or sick leave. They can occur at any time and in any location. Some disasters are extremely unlikely (for example, a tsunami in the middle of a continental land mass) but similar damage could occur if a dam were to burst up river of your location. You need to consider what sort of disasters can occur, in a generic sense, and then assign probabilities regarding risk and damage. You can't protect explicitly against every possible occurrence, but you can ensure that you have a plan. It may not matter too much why you've lost the data center as long as you can keep the business working.

## Types of Disaster

Disasters can be grouped into two major categories: natural disasters and man-made disasters. Natural disasters come in many forms. Your geographic location plays a large part in determining the types of natural disaster with which you may need to contend:

- Earthquake
- Flood
- Storms
- Lightning strike
- Fire

Can you protect yourself against these events? The answer is, as so many times, an emphatic maybe. An obvious answer is avoidance, for instance don't build in a known earthquake zone and check that your building is not on a flood plain; however, the natural event may be outside of the normal range or even completely unexpected.

Man-made disasters can be divided into two main categories: deliberate and accidental. Deliberate is anything of a criminal or terrorist nature. Physical security can go a long way to mitigate these threats. If the bad guys can't get into the building, they can't cause damage.

Can high availability help to protect you against these threats? The answer is "Yes to a certain degree." It all depends on the technologies that are used. Standard Windows clustering won't help because the second node is on the same subnet, which usually means the same building and often the same server room or even the same rack! Geographic, or stretch, clustering may help to a certain degree, but how is the data to be replicated? Does that make the solution too expensive?

Applications that can replicate, or mirror, data such as SQL Server or Exchange can provide a high availability and disaster recovery capability, but it means using the latest versions and, in SQL Server's case, ensuring that the application portfolio used within the organization can take advantage of these capabilities.

What do you need to consider when thinking about potential disasters?

## Damage and Risk Assessments

There are a few things to consider when attempting to assess the risks of a disaster. Geography can play a major part in reducing or increasing the possibility of certain disasters. However, don't get fixated on the detail—just because your building is a few hundred feet above the nearest watercourse doesn't mean you can't be hit by a flash flood because the drainage system is blocked. A more likely possibility is that a fault in the water system in your building causes water to leak into the server room. Arranging umbrellas over servers because it's a critical time of year seems funny now but it wasn't at the time.

This boils down to a reminder to start the thinking and planning from a generic viewpoint: Consider water damage rather than flooding—it doesn't matter where the water comes from; and think about fire suppression rather than arson but make sure that fire suppression doesn't lead to water damage. The next step is to consider the details. How could that water damage happen? One last point in all of this you have to remember is that “The ‘storm of the century’ has to happen sometime.”

## Don't Assume

One of the biggest problems with disaster recovery planning is the assumptions that have to be made. Unfortunately, you don't have all the information when trying to plan for these events, so you have to make assumptions. This can lead to problems, as the following examples will show.

I was once involved in setting up a mini-data center for a subsidiary of a financial services organization. It was a relatively small operation with only a handful of servers. However, this part of the business totally depended on IT and therefore on those servers.

Some disaster recovery planning had been performed before I joined the project, and it was a little while before I had a chance to review it. It was a pretty good plan apart from the fact that it had been assumed that in the event of the building becoming unusable, the organization could use the next building. At their closest, these buildings were about two car lengths apart. It took a long time to explain that in the event of the building being unusable, there was a good chance next door would be as well.

The second assumption to be aware of is that you will be able to access your building. An organization I worked for was based on a business park at the edge of town that had two entrances. At that time, the biggest identified threat was a military jet aircraft crashing on the park. The assumption was that if this happened, it would only affect one entrance to the business park and we would still be able to carry on working. The threat of unexploded munitions, the fires produced by the crash, and the area being closed for crash investigation and recovery hadn't been considered.

It doesn't matter if your servers are still running. As Chapter 1 pointed out, if users can't access their applications and data, the system is down.

## What Is Disaster Recovery?

Having thought about the types of disasters that could affect your organization and some of the assumptions you need to overcome, it is time to turn attention to disaster recovery itself. A lot of people talk about it and some organizations even plan for it, but what exactly is disaster recovery?

### Disaster Recovery

A quick search on the Internet will reveal a number of definitions for disaster recovery. One that I like is that it is the processes used to return a business to a normal operating state after the occurrence of a catastrophic event. However, you need to think about the difference between disaster recovery and business continuity (we'll visit this topic shortly).

Disaster recovery, as a term, is normally applied to the recovery of an IT system, or infrastructure, after a disastrous event has struck. As with any IT activity, there is the option to do nothing. In this case, you are trusting luck and the skills of your technical people to restore operations. The worst possible time to be working out how to restore a system is when it is unavailable.

Remember the statistic about organizations that go out of business after a disaster—that could be the price of doing nothing. The price could become due earlier as many business sectors are regulated and are required to have disaster recovery plans in place. Regulators don't believe in luck; they will need to see your plans.

### Disaster Recovery Compared with High Availability

The differences between disaster recovery and high availability have been mentioned but have not been fully defined. The aim of both techniques is to keep the business working, but they approach this goal in different ways. In real terms, the difference comes down to a simple difference of approach:

- High-availability techniques are about *preventing* the everyday failures that cause downtime (network card failure, disk crash, and so on)
- Disaster recovery techniques are designed to help you recover after true disasters (floods, hurricanes, fires)

Can you just use one or the other? Do you always need both?

You could just use disaster recovery techniques. In the event of a system failure, you might, depending on the techniques adopted, need a significant period of time to recover the system. Conversely, if you just use high availability techniques, you could keep the system running through “normal” events but you could have a long period of downtime while you rebuild the system in the event of a disaster.

Active Directory (AD) provides a good example of a layered approach. Assume that you have two sites that are of a sufficient distance apart that a disaster would only affect one site. It can also be assumed that you have sufficient network bandwidth between the sites for your purposes.



You start off with a good design and, by following best practice, you have two domain Controllers at each site. For this discussion, let's ignore global catalogs and any other complication. The following events have been identified as occurrences you need to guard against:

- Loss of a single domain controller
- Loss of a site
- Loss of a single item of AD data
- Loss of a significant amount of data (for example, an organizational unit—OU—containing 100 users)
- Complete loss of AD through corruption of the database

Your recovery plan (in outline) looks something like this:

- For a single domain controller, we will replace the machine, join to the domain, promote to a domain controller and let replication populate the AD database. We have another domain controller on site (high availability) so we can carry on working until the second domain controller is back online.
- For the loss of a site, we have the other site available so that people can work there or over the Web. Authentication can occur because we still have domain controllers (high availability). *We are assuming deliberately that all other applications and data are available.*
- If a single item of AD data is lost, we can perform one of the following:
  - a. Recreate
  - b. Perform authoritative restore
  - c. Recover from Recycle Bin if using Windows Server 2008 R2
  - d. Recover using third-party tools
- In the event of losing a significant amount of data, we would usually perform b, c, or d from the previous options depending on the technologies available. (Now we are getting into disaster recovery.)
- The complete loss of AD is the big money problem. We have lost all domain controllers and no one in the organization can access email, applications, or data. We can officially label this event as a disaster. This is where the recovery plan is desperately needed.

The first four scenarios you can cope with on an *ad hoc* basis if necessary. It is better to have plans and processes in place that are documented and tested, as the recovery will be far quicker. If the last scenario strikes, you really will need to have the recovery plans in place. The information required to create these plans is available from the Microsoft Web site but needs to be tailored to your environment. At this point, the whole business is affected so you are dipping into the area of business continuity.

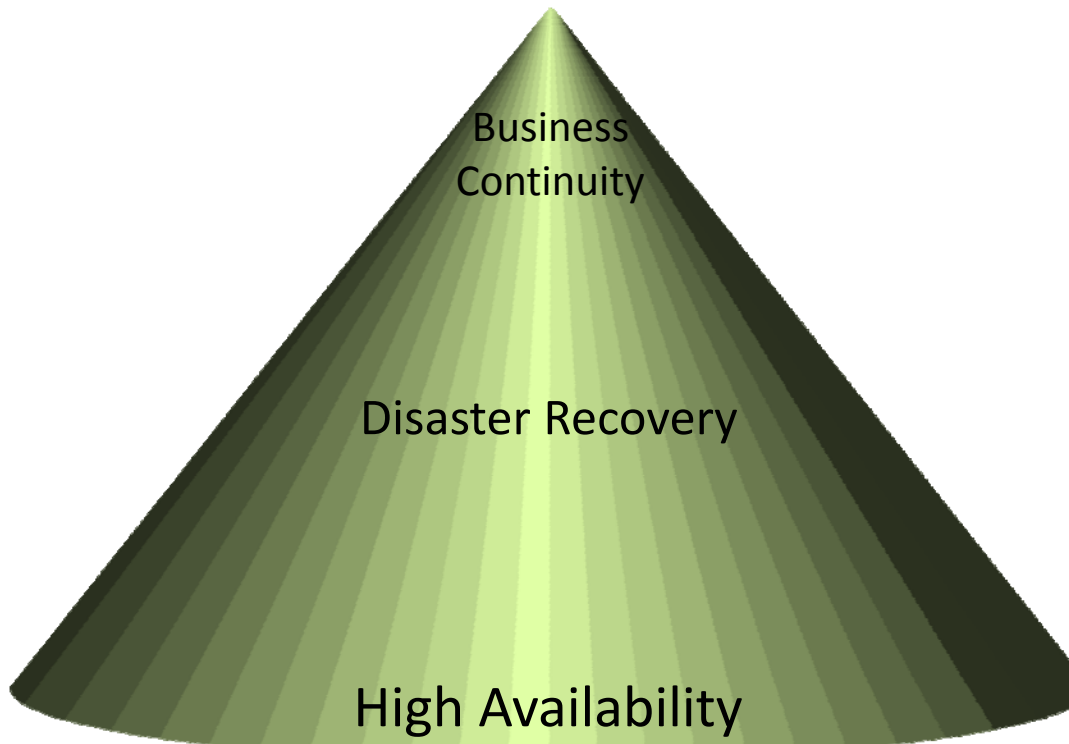
## **Business Continuity**

Business continuity and disaster recovery are confused for each other even more than disaster recovery and high availability. In simple terms, the Business Continuity Plan (BCP) is the processes and procedures that will be followed to enable the business to function in a reduced, or preferably normal, capacity. It is involved with a large number of areas including:

- Office space and furniture
- Supplies
- Communications
- Business process systems
- Transport
- Procurement of replacement equipment
- Possibly even the hiring of new staff

There will be an IT input into many, if not all, of these areas. Disaster recovery is best regarded as a component of the BCP that is concerned with restoring access to the IT systems the business requires to continue to function. This may be all systems or it may be a critical subset of the application suite.

High availability can be viewed as a way to prevent the disaster from happening, or preventing some of the smaller disasters. Thus, from one viewpoint, it is a subset of disaster recovery (see Figure 2.1).



**Figure 2.1: The relationship between business continuity, disaster recovery, and high availability.**

The diagram should not be taken to mean that disaster recovery or high availability is less or more important than BCP. There is a spectrum of activity that spans these areas. The important point is that your organization is protected, to the correct level, from both large-scale and small-scale events that could prevent business processes occurring. High availability is the most significant part of the plan that protects against the “normal” everyday type of interruptions that will happen with the greatest frequency. Disaster recovery and business continuity build on this concept to protect against the bigger, less frequent episodes that could potentially close down the organization on a permanent basis.

As we will see later, you can’t always do everything you would want to do and some hard choices need to be made regarding what is protected and how it’s protected. The final line of defense from disaster is your backup. We asked in the previous chapter whether backup was still necessary. The answer is yes when you are thinking about high availability, but is it the same when you consider disaster recovery?

### **Do You Still Need to Back Up?**

If backup was your last best hope for high availability, it is required even more so in a disaster recovery scenario. There are a number of possible ways that a disaster recovery site can be configured, as we will see later. Some of them involve the use of a restore from backup to bring the services online. Will your backup support this scenario?

One topic that is often overlooked is disaster recovery for the backup system. Imagine a simple environment in which you have a number of servers all backed up to a central tape system. What happens if that tape system breaks down? How will you back up and more importantly how will you restore?

Now let's expand this thought to losing the data center. You can't access your tape system. It may be destroyed or damaged or just be inaccessible. How do you restore your tapes that you have in offsite storage? You do have your tapes offsite, don't you?

If any part of the recovery plan involves restoring from tape, the most critical item is to ensure that a tape system is available that can read your backup tapes. I would say that if you don't have a known, tested, and proven restoration system, your backups are a waste of time, effort, and money.

Your backups should be stored offsite. The ideal location is at your disaster recovery site. The idea of storing tapes on site can seem appealing to the organization because it is cheaper and to the technical staff because everything is readily at hand and they don't have to worry about the tape retrieval process. The statement will be made that the tapes are protected because they are in a fireproof safe. The safe itself may be fireproof but what about the contents. Papers will spontaneously combust at 451 Fahrenheit (186 Celsius). If the paper starts to burn, what happens to the tapes? Even without a fire in the safe, will the tapes survive the elevated temperatures in a usable state? How will you know that the tape is usable? Can you trust the data from a suspect tape?

One point that is often missed regarding backup tapes is their security. If a tape is stolen, it could be read and any confidential information used against the organization. Do you encrypt your backups? Who has access to them? In the UK, organizations are liable for a fine of up to £0.5 million if a data breach occurs. It's not going to cost anywhere near that to ensure your data to a safe location.

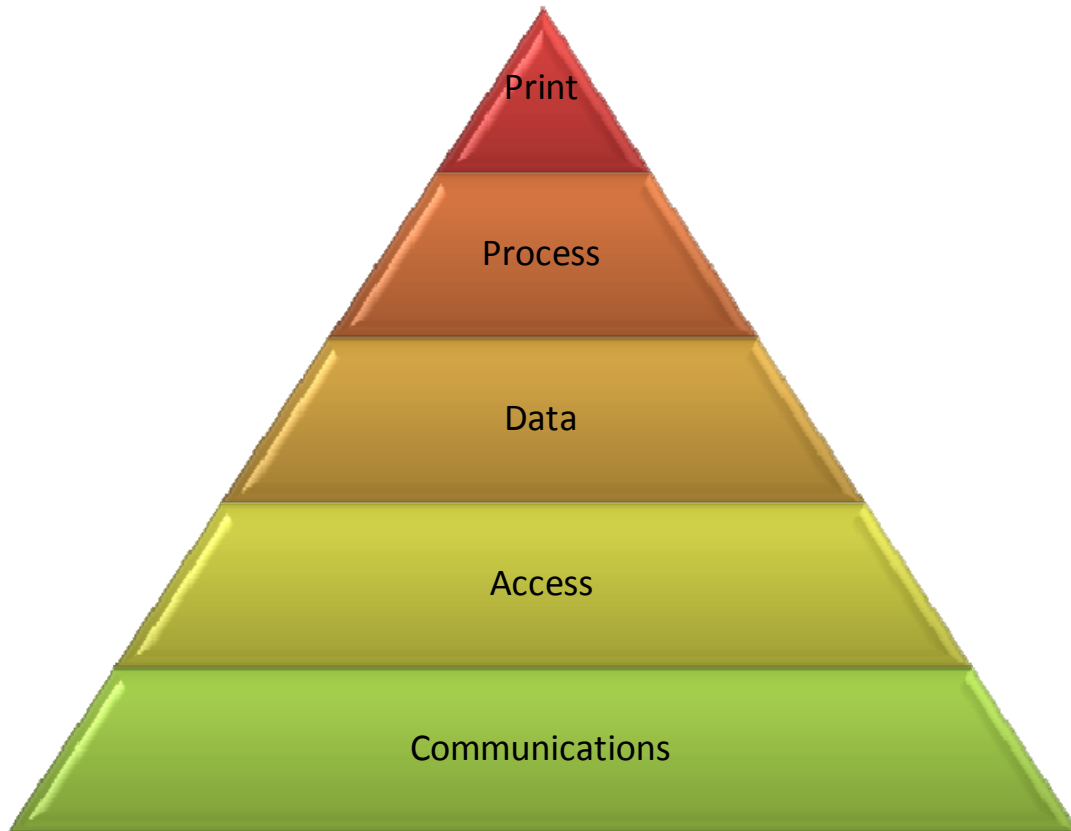
When did you test your backups? The question keeps coming up in any discussion of high availability or disaster recovery. If your ability to restore is not tested on a regular basis, how do you know it will work? The backup system may be able to verify that the tape is readable and that data on tape matches what was on disk, but that doesn't make it usable.

Have you tried to restore using the equipment at the disaster recovery site? Have you tried to restore and determined that users can access the data. The availability issues still apply in a disaster recovery situation. If the users can't access the data and applications, the system is not available or in this case recovered.

Having thought about the backup and restore process in some detail, you need to consider just what you do need to recover in the event of a disaster.

## What Do You Need to Recover?

When asked what you need to recover, you will receive exactly the same answer as you did for high availability. Everything! In this case, it is correct. You need to restore everything to have a fully functional organization. But, and it makes a huge difference to how you go about disaster recovery planning, does your organization need everything recovered at once or is there a hierarchy of needs (see Figure 2.2)?



**Figure 2.2: Hierarchy of restoration.**

This figure assumes that the BCP has identified a suitable location and it is outfitted to meet your needs. It may be that you don't need a single location and that many of the users can work from home as long as they can access their applications and data. This example is purely hypothetical—the needs of your organization will have to be thought of in this way to produce a plan that is suitable for your organization.

At the bottom of the pyramid is communications. This can encompass voice, email, instant messaging, Web access, and data access. Without the right level of networking and communication in place, nothing else will work.

The second need is access. This is your authentication and authorization layer. Your plan does include the recovery of AD, right? If you can't log on to the systems, you can't do anything.

Data is a fundamental need. As the majority of business processes require access to data in one form or another, it has to be available before you can recover the process. This includes access to the process. The plan must include the provisioning of client machines so that the users can access applications.

I have put print at the top of the pyramid because it isn't needed until you have something to print (this, until you can run the process).

This is only an overview of the recovery hierarchy. It could be a useful exercise to take one of your critical business processes and produce a pyramid like this for recovery. I guarantee there will be more on it than you expected by the time it's finished. There are some critical points that need to be raised regarding the main parts of this hierarchy.

### Communications

Communication is critical to the success of any organization. The supporting layer of your whole endeavor is the network. If you don't have connectivity, everything else is a waste of effort. There can be a significant lead time (weeks or months) to get new external links into a location. Don't wait until a disaster occurs before investigating the requirements.

If the disaster recovery site belongs to the organization and data is being replicated to the site, then the links obviously exist. If it is a specialized disaster recovery site, the links should exist, but check with the vendor. When the site belongs to your organization, adequate network links must be part of the disaster recovery plan.

Email is the medium of choice for business-orientated person-to-person communication. Many organizations rely on email for communicating with suppliers and customers. A significant amount of business can be lost if email is unavailable.

One question to ask when planning your disaster recovery strategy is just what is required from the email system. If the goal is to rapidly restore the ability to send and receive emails, it may be possible for the users to be given a "dial tone" mailbox. This is an empty mail database and mailboxes that the users can access to send and receive mail. It takes the identity of the user's mailbox. The original mail database is then restored to a Recovery Storage Group. The recovered data can be merged into the new mailbox to restore full functionality.

This is a prime example of the thinking, and mindset, required for disaster recovery planning. What can I do to restore as much functionality as quickly as possible? What is the minimum functionality that is required? What do I need to do to restore full functionality?

The last point on email is to consider who needs restoring first? The obvious answer is that the organizational structure should be followed, but this isn't always the best strategy, especially if it impacts revenue generation.

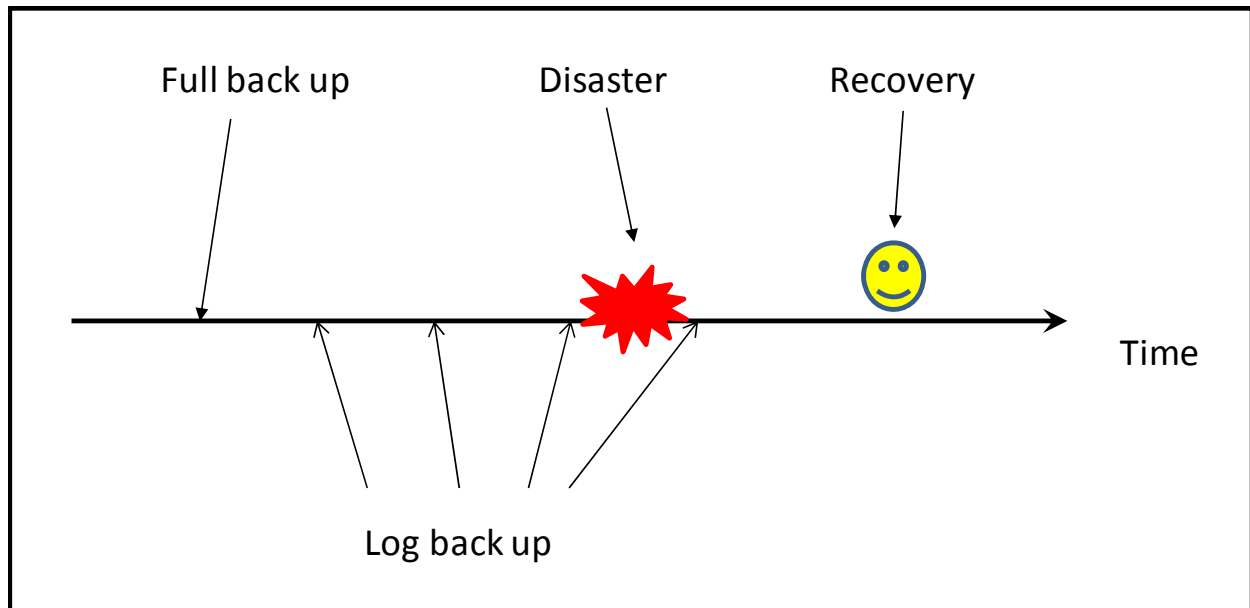
A full voice implementation with voicemail, IVR, and all the other bells and whistles may not be available early in the recovery operation. What is needed? Would a simple system based on mobile phones be sufficient? Is there a need for multiple locations?

Access to Web traffic requires a two-pronged approach. Incoming traffic needs to be accommodated in the case of a Web-based trading system. The longer that system is down, the more revenue is lost, including future revenue if the organization gets a reputation for an unreliable Web site. If the organization's Web site is purely for informational purposes, a simple site with a message to ring a specific number may suffice until it can be fully restored.

### Data

In many cases, the data has to be recovered before an application can be made available to the users. This raises a huge question: How much data can the organization afford to lose? If all transactions are performed via the system, there may not be a paper trail to fall back on. Does your back up system allow you to recover to a known point in time? How frequently do you back up the database and log files?

The recovery point is the point in time to which the database can be recovered. The business must understand the implications of any decisions that are taken regarding potential data loss. This is illustrated in Figure 2.3.



**Figure 2.3: Recovery point and potential data loss.**

In a traditional database backup scenario, you would take a full back up and a series of log backups. Your disaster will occur at sometime between backups. Assuming that you have the backups available, you can restore to the point of the last backup. If the database logs are available, you may be able to recover information from them, but their availability is not guaranteed, in which case you have a potential for data loss.

Can the business afford that loss and are there recovery mechanisms? Even if the answer is yes, there is a time and effort expenditure involved that may not be affordable during the recovery from a disaster. This is especially true if the recovery prevents new business from being processed.

The protection afforded the database needs to be re-engineered so that data loss is minimized. Replication or mirroring techniques can be used, as we will see later.

Does the organization require all the data restored before processing can commence? Is the database organized so that a phased recovery can be implemented? If this is possible, it reduces the time between disaster and business resumption. There needs to be careful communication with the business to ensure they understand the implication of a phased recovery. Once you have recovered the data, you need to think about the business processes.

### **Processes**

Ideally, you will have sufficiently skilled technical personnel that can be restoring the business process systems at the same time as they are restoring the data. But do you have the restoration capacity from your backup system to accomplish this? Speed of backup is important but restoration speed, which is always slower, is even more important.

Do all the processes need to be recovered before work can recommence? Can some of the work be done manually until the full system is available? These are more questions to take to the business. The questions raised in this section should highlight that disaster recovery is not an IT issue, it is a business issue. IT implements the recovery, but there are a lot of business decisions required.

### **Involving the Business**

We have seen the questions that need to be answered regarding disaster recovery, but how do you bring the business into the process? The most important point is that this process requires sponsorship from the highest levels of the organization. It cannot be done by IT and then presented to the business. It must be a business-driven process.

As with any IT activity, you need to start with the planning process. This involves supplying answers that are specific to your situation to the questions posed in the following sections.

### **Planning**

Let's assume that you have decided which applications and business processes you need to recover to enable the business to function. If this step hasn't been completed, the guidelines in earlier sections need to be applied before proceeding further with disaster recovery planning. The assumption will also be made that the relevant decisions have been made regarding locations to be used for recovery.



The invocation of the disaster recovery process is a major step. Guidelines need to be established as to when this step is taken. Possibilities include:

- Loss of access to a building either through damage to the building or the blocking of access routes.
- Damage to the data center rendering all or part of it unusable. If some of the equipment is undamaged, it may be possible to relocate it to another part of the building. Are you prepared to relax your data center standards during a recovery?
- Loss of network connectivity to the data center. This is more common than might be imagined especially if the routing of underground cables has not been recorded properly. Lightning strikes can destroy even armored cables. Do you have redundant network links to the data center?
- Major data corruption affecting a critical business process.

Is your disaster recovery plan flexible enough to have a partial invocation? Who has the authority, and responsibility, to invoke the disaster recovery plan? If it's a single individual, what happens if he is unavailable? High availability is needed regarding decision making as well as systems! How will the invocation of disaster recovery procedures be communicated? By whom and to whom?

The recovery plan dictates what is to be recovered and in what order. The people available to perform the recovery need to be confirmed and their skill sets matched against the recovery plan. Are there any gaps that need filling? How can they be filled? Can your IT vendors or partners help?

When planning the order of recovery, don't forget the dependencies. I was involved in a disaster recovery planning session some years ago where the initial proposal was to restore Exchange on day 1 and AD on day 3. It took a while to convince the planners that setup wouldn't work.

Given the people available and what needs to be recovered, is there any option for parallel working or is everything performed sequentially? Are there any alternative ways of working? Can people work from home or Internet hotspots? Can some work be done manually? And what level of recovery is required? Do you need to re-create the full high availability environment or can you accept the increased risk of running without it while you restore the rest of the systems?

We have already discussed that communications are an essential, and priority step, on the way to recovery. Using those communications to explain what is happening can help the recovery process. There are a number of parties interested in the progress of the recovery operation:

- Staff not involved in the immediate recovery activities need to be informed of progress. They also need to be told when and where they will be required to work as the recovery progresses.
- Customers and suppliers need to be kept informed to ensure business confidence isn't lost.
- Partner organizations need to be informed.
- In some cases, the press may need to be informed to prevent the spread of rumors affecting your business.

In all cases, there needs to be a communication plan stating who can communicate on behalf of the organization during the recovery period. And finally, when planning your recovery, don't forget to include the celebration at the end when it's all worked out and the organization is back on its feet!

Having created your plan, you need to think about how you can test it.

### Testing

Disaster recovery testing is a necessity. If your plans aren't tested, how do you know they will work? The people involved in performing the recovery operations need to practice their actions so that they can be performed quickly and efficiently when necessary. Don't use the same people for each test, as they become a single point of failure in your recovery.

Disaster recovery tests need to be planned. Do you perform a full test involving users accessing the systems or is it just enough to rebuild the system and test the restore process? Is one big test involving the whole environment better than multiple little tests each of which only recovers a portion of the organization's functionality? This is another case of there not being a single correct answer—choose what works best for your environment.

Whatever type of test is chosen, it needs to be planned:

- Test scope agreed and communicated
- Roles and responsibilities assigned
- Record keepers assigned—keeping records of what happens is a good way to discover flaws in the recovery plan that need to be corrected
- Equipment and locations agreed and communicated
- Timings communicated
- Lessons learned recorded, communicated, and used to modify the recovery plan as required

The frequency of tests will generate a lot of debate. Too frequent, and they become an expensive overhead on the operations staff who will also get bored by the whole thing. If the tests are too infrequent, the skills and knowledge may become rusty.

One idea is to spring the odd surprise test. Walk into operations and announce that you are invoking a disaster recovery scenario for the email system and watch the reactions! Again, not something to be done too often.

### Time Is Money

The key thing for any organization to remember is that downtime costs money:

- Lost orders
- Lost revenues
- Unproductive staff time
- Loss of business confidence.

These may not always be amenable to exact quantification but indicative values should be calculable.

There will always be pressure to reduce recovery times to the minimum that is possible. Planning and testing will help with this, but there is a finite speed at which a tape drive can restore data. If the recovery times need reducing, further spending is required to provide the equipment or to utilize hot standby configurations.

### Disaster Recovery Configurations

There is a spectrum of disaster recovery configurations you can adopt (see Figure 2.4). A cold disaster recovery site just supplies space. You have to build all of your systems and restore the data. At the other extreme, a hot disaster recovery site is ready for you to move in and start work immediately. The speed of recovery is proportional to the cost. A hot standby site will have a much higher cost than a cold equivalent.



Figure 2.4: Relative costs of cold and hot recovery sites.

## Cold

When using a cold recovery site, you have to start from the beginning:

- Arrange networking
- Procure the hardware
- Build the systems
- Restore the data
- Configure client access
- Etc

This all takes time. It may be possible to reduce some of the time by having contracts in place to supply networking and hardware, but there is still the setup and restore time. It is the cheapest option.

## Warm

A warm standby site costs more but recovery time is shorter. You can perform some replication to ensure that data is available. Possibilities include:

- A domain controller to ensure AD is available
- Email database replicas
- Database replication through log shipping or mirroring
- File server data via replication

It would be possible to perform the replication via a backup/restore process but that adds manual effort to the process and additional cost. It also raises the possibility of an error being made.

If data is being replicated, is it secure on that site? If not, why are you using that site!

Third-party sites need auditing to ensure that your servers and data cannot be accessed by other people. Physical security at the building, server room, and rack level is required with auditing of who requests access and the keys to the room or rack.

If a warm standby facility is being considered, what is going to be protected? Remember the dependencies discussed earlier. There is no point in replicating data if you can't access it because AD is unavailable.

## Hot

A hot standby disaster recovery site could be considered to supply high availability at the data center level. Full duplication of all data and systems is performed with an automated failover to the disaster recovery site in the event of the primary site being unreachable.

A variation on this is to utilize both data centers and to run 50% of the processes from each. In this way, the failover time is reduced as only part of the application suite has to fail over. A further advantage is that some processing is still possible if there is a problem with failover.

## Mobile Data Center

One final possibility to consider is the mobile data center. In this case, a trailer, equipped as a data center, is delivered to a site ready for use. The applications and data still have to be restored to the supplied machines.

If you are contemplating this approach, there are a number of questions to consider:

- Will your organization fit in a single trailer?
- What are the essential applications that you need to recover?
- Is there a suitable space on site and will it be available after a disaster?

If not adopted as a solution in its own right, it may augment other disaster recovery provisions.

## Convergence with High Availability

This chapter has looked at disaster recovery as a separate activity, but taking it together with the first chapter, there is an obvious convergence between disaster recovery and high availability. They can be thought of as related, as Figure 2.1 showed. This raises the question: When is it best to use high availability or disaster recovery?

### **When Should High Availability Be Used?**

High availability is used to protect a single system. You may have multiple components in the system, including Internet links, a Web farm, firewalls, network links, and a database. Each component requires the application of a high availability technique. The techniques will be different for each component (for example, clustering for the database and NLB for the Web farm). High availability is usually applied within a single location.

### **When Should Disaster Recovery Be Used?**

Disaster recovery is used to protect all the systems in the organization, or at least the subset of systems that are critical to the organization's continued ability to function. The planning for disaster recovery has to assume that other locations are required and that systems, applications, and data must be made available in those locations.

### **Combining High Availability and Disaster Recovery**

High availability and disaster recovery are both concerned with protecting the organization and reducing, if not eliminating, downtime. Both techniques involve cost to the organization—in some cases, significant cost. The ideal solution would be to combine your high availability and disaster recovery requirements and arrive at a single solution that satisfies both sets of requirements.

Consider the simple case of a database that is at the core of an organization's principle trading system. All revenue earned by the organization is dependent on this database. The traditional approach to this scenario is to treat high availability and disaster recovery as two separate disciplines. High availability would be supplied by using clustering. This raises the following issues:

- Cost of creating a clustered solution: hardware and software
- Storage requirements
- Skills to create and maintain the solution.

Disaster recovery would be supplied by one of the following:

- Backup and restore
- Replication
- Database mirroring

This involves additional cost.

Why not combine the two approaches and use database mirroring to provide high availability and disaster recovery. The mirror can be created at another location and, depending on the application, be set to provide automatic failover. There are reduced costs as the cluster isn't required—nor is the SAN-based storage to support it.

There are a number of technologies you could utilize to supply a combined high availability and disaster recovery capability.

## Geographic Clusters

Native Windows clustering has, until recently, been confined to a single location. Windows Server 2008 changed this with the introduction of geographic, or stretch, clusters. These clusters work across multiple locations because the nodes can be on different subnets and the heartbeat between nodes can be adjusted to accommodate latency (within limits) on the network links between locations.

The application can failover between the nodes of the cluster, but how is the data to be handled? Is it possible to replicate the data, at the storage level, so that it is available in the secondary data center?

There are some points at which this approach may fail:

- The location of the cluster witness may prevent failover
- Replication may not happen
- Cost may be prohibitive, especially for the storage
- The appropriate skill sets may not be available within the organization
- Network failure or problems may interfere with failover

Other techniques working at the data level may solve the problem.

## Replication

Unless you are going to rely solely on backup and restore, replication will feature in your disaster recovery strategy. The question that then arises is what else can you do with it? If you can replicate the data, can you get applications to access it? Can the failover be made automatic? How much manual intervention is required?

Data replication can be managed at the storage level or at the application level. If you use your storage systems to manage replication, what does that do to your cost models? Allowing the application to manage replication may mean that you can use it for high availability as well as disaster recovery.

Another possibility is to replicate the data to the disaster recovery site and perform the backups on that copy of the data. If replication is continuous, or timed to happen before the backup, then you have an up-to-date copy of the data. Performing the backup at the disaster recovery site means the tapes are immediately offsite from the primary site's viewpoint and the effect of a backup window can be ignored as you won't be impacting the primary site.

## Mirroring

Mirroring is enhancement of replication in that you have duplicate systems that are kept synchronized. This is employed at the individual system level, but if combined with replication for other parts of the environment such as AD and file stores, you then have high availability and disaster recovery.

The advantage is that either end of the mirroring can supply access to the users. They neither know nor care what is happening—they just keep accessing the data. In the event of a failure, the other end of the mirror supports the full load.

Downtime is still possible with this approach, but its probability is much reduced. One potential hazard is if the mirroring is interrupted by a network problem. This could mean the whole system has to be reset, which could involve significant work and imply a period of increased risk as the high availability has been degraded.

Another technique for added protection includes geographically separated mirrored systems. These types of systems in some ways combine the benefits of both high availability and disaster recovery. Locating servers in different buildings or data centers protects your business from localized power failures and building-wide problems. Separating them geographically protects your business from major downtime due to major problems as well, such as hurricanes, flooding, and earthquakes. Using synchronization links routed over a switched WAN, you can split sites between data centers in different locations, at varying degrees of geographic separation, depending on bandwidth and latency. If a problem strikes one data center, applications and data are available, up-to-date, and fully operational at the other location via the mirrored system.

### **Virtualization**

Virtualization is the “big thing” in IT at the moment. By itself it does not supply any high availability or disaster recovery capability. However, virtualization hosts can be grouped into “clusters,” and the virtual guests can be moved between the hosts without downtime.

This can be extended to movement between data centers, but you still need to replicate the data. There may be a small amount of downtime due to a need to reconfigure the virtual machines to utilize the storage in the alternative data center.

### **Summary: Stop the Disaster Being a Disaster**

Disaster recovery is an exercise in planning for possible problems that would impact the organization’s ability to function as well as supplying solutions to overcome those problems. As you have seen, there are a number of approaches with the correct one for your organization only being determined by performing the planning exercise. If your disaster recovery plans aren’t established, your organization’s future depends on you:

- Preparing a disaster recovery plan
- Testing that plan
- Reviewing the plan on a periodic basis to ensure it is still relevant.

You should be starting now.