

WHITE PAPER

The 5 Critical Steps for an Effective Disaster Recovery Plan

Introduction

In today's climate, most enterprises maintain some form of business continuity plan. Business continuity plans are designed to provide a way for an enterprise to continue business operations in the event of a catastrophic disaster, such as a flood, an earthquake, or a widespread power outage, that shuts down business operations at one or more primary locations. Business continuity plans cover information technology (IT) infrastructure recovery, people issues that must be dealt with when business operations must be restarted at a remote location, and physical infrastructure issues, such as re-establishing telecommunications, ensuring physical security, and providing appropriate work areas at remote locations.

IT infrastructure recovery, sometimes referred to as disaster recovery (DR), addresses the issues involved with recovering computing equipment (servers, storage, etc.), data, and application services. DR provides a necessary foundation for business continuity plans, but is not a substitute for them. This white paper focuses on the key elements of creating an effective DR plan, and is appropriate for IT management as well as technical staff.

Setting The Stage

Business information is the crown jewel for most enterprises, regardless of size. For today's highly computerized businesses, business information is maintained as data, and it is a rare enterprise that is not taking daily steps to ensure the recoverability of both new and archived data. Historically, local data protection was addressed through some form of tape-based backup. Copies of backup tapes were retained locally to meet daily recovery requirements for lost files, database tables, etc. and copies of some of those tapes were periodically shipped to remote locations where they were often stored for years to ensure data recovery in the event of catastrophic disasters which may shut down primary sites. Long term, off-site storage of tapes was the conventional "DR" plan of record. If business operations needed to be re-started in some location other than the primary one, data on tapes at these remote locations would be shipped to the new "primary" location, application environments would be manually rebuilt, the data would be loaded onto the new servers, and business operations were transacted from this new location until such time as the primary location could be brought back on-line.

Inherent in the description above are a set of "DR" requirements which in the past were not often consciously considered. The single most important point in this white paper is that, if you are going to implement a DR plan, you need to start your planning by understanding your requirements, and the implications of those requirements, rather than starting with one or more technologies that are often associated with the term "DR." Ultimately, what you want to do is to understand your requirements, and then implement a plan which reliably meets those at the lowest cost. To do so, you'll need to transcend the idea that "off-site backup tapes = DR" and move towards matching your specific requirements with the right technologies that can cost-effectively meet them.

Step 1: Understand Your Business Priorities

While enterprises have a number of business processes, certain of those business processes are more critical than others. Generally, any business processes that are directly related to revenue generation or customer support are deemed critical, but this will vary by business. It is incumbent upon the CIO (or ranking IT management) to work with executive management across the company to determine critical business process areas.

Understanding the time-sensitivity of recovery and how it relates to business priorities will help you to focus on those areas for which a recovery plan truly must exist. Determine not only which business processes have an effect when they fail, but also the scope of that effect. Does it impact revenue directly? Does it have a workaround? How onerous is the workaround? The loss of some business processes will begin to immediately affect the business in appreciable ways, such as through lost revenue or an inability to support customers, while others may not have an appreciable impact if they are lost for days, weeks, or even months. If you identify processes that do not start to impact the business until they have been gone for several months, and their impact is small, you may choose not to implement any recovery plan at all for them.

Create a prioritized list that includes all major business process areas, and then map those business processes to the relevant supporting IT infrastructure. What you want to end up with is a list of the applications, servers, and storage that must be available to support each business process. Executive management participation in the process of defining critical business priorities will be very helpful later in securing the budget necessary to put the right recovery plans in place for your company, based on consensus agreement of what the critical business priorities and their associated processes are, and what the impact to the business is of those processes failing.

Step 2: Assess Your Recovery Requirements

Once major business process areas have been prioritized, you will clearly know which ones need to be focused on first. Since we are focusing on DR here, you automatically know that any candidate application environment must be valuable enough that you will allocate budget to maintaining updated copies of its data at one or more remote locations. In business terms, recovery solutions should minimize data loss on recovery, be able to meet very short recovery times, and provide simple, reliable recovery. The key recovery metrics that support these requirements and need to be evaluated for every critical application environment are recovery point objective (RPO) and recovery time objective (RTO). RPO is a measure of how much data you are willing to lose on recovery. If your target RPO is 1 hour, that means that you target to lose no more than the most recent hour of data. RTO is a measure of the longest recovery time you are willing to accept. If your target RTO is 4 hours, that means that your target is to have that particular application environment up and running again in no longer than 4 hours.

In the event of a catastrophic disaster, you will likely need to recover both applications and data. Many enterprises implement a DR plan for just data, assuming that servers and application environments will be manually rebuilt and recovered if they need to be. DR plans that provide for automated application recovery will be able to meet much shorter RTOs than those which just recover data and then depend on administrators to manually recover applications. Those plans will also be more reliable, and perform more predictably, because they will not be as dependent upon the skill of the administrators that are actually performing the recovery (some of whom may not be available when a real disaster hits).

Evaluate the types of “disasters” that you are most likely to encounter given where your facilities are located. If you are in an area where you may be forced to deal with natural disasters that can affect large areas, such as floods, earthquakes, or widespread power outages, you may want to follow the DR best practice guideline of locating your remote recovery site at least 200 miles away from your primary site. If in fact this is your requirement, this will affect any decision you make to implement replication technologies to help address your DR requirements. Replication technology comes in two flavors: synchronous and asynchronous. Synchronous replication keeps a source and target synchronized in terms of data states, but because of latency issues can inflict an onerous performance impact on production applications if the source and target are more than about 30 miles apart. Because of this limitation, asynchronous replication is much more widely used to meet long distance DR requirements. Asynchronous replication keeps a source and target in sync over literally any distance, but the target may lag the source by up to several minutes (depending on the write volumes and network latencies). Still, asynchronous replication provides the kind of RPO performance necessary to meet 99.9% of all DR requirements, and it does so in such a way that does not impact the performance of production applications in any way.

Recovery tiering is an approach that is often used when evaluating the recovery requirements associated with various business processes. Instead of evaluating and setting recovery requirements individually for all major business process areas, a small number of recovery tiers is defined. Each tier has a set of recovery performance metrics that are associated with all application environments within that tier. For example, IT management may define three tiers as follows:

TIER 1	RPO 5 minutes, RTO 1 hour	Application Environments A, B
TIER 2	RPO 6 hours, RTO 8 hours	Application Environments C, D, E
TIER 3	RPO 1 day or greater, RTO 1 day or greater	All other application environments

These numbers are not suggestions for your business since your recovery tiers will vary based upon your business and regulatory mandates. But the general idea applies: there will be a small number of critical application environments that require very low RPO and very short RTO; then there will be another set of very important application environments that require stringent RPO/RTO, but not as stringent as Tier 1; then there will likely be all the rest of the application environments which are not critical and may only need to be recovered within one or two days or more. You may have only two tiers, or you may have more than three tiers, depending on your requirements, but keep in mind that what you don't want is one tier: clearly all your application environments do not merit the highest priority in terms of recovery. You don't want to pay the price premium to meet your most stringent recovery requirements for application environments that don't need it. By the same token, you don't want your critical application environments supported by the same multi-day RPOs or RTOs that you use for relatively unimportant applications.

Asking your end users what their recovery requirements are is an important data point, but not the only data point. Generally, meeting more stringent recovery requirements requires more expensive solutions. When not thinking about costs, most end users will respond that they want very rapid recovery, when in fact they may easily be able to deal with less stringent recovery capabilities. But a trade-off needs to be made between meeting recovery requirements that are truly required by the business and the costs associated with that. That's why you need an accurate understanding of the business priorities mentioned in step 1.

Step 3: Match The Right Solutions To Your Recovery Requirements

Once you've determined the key recovery metrics of RPO, RTO, and recovery reliability, you'll need to consider just what type of IT infrastructure you need to meet it. The first consideration is likely to revolve around the remote location: do you already have a corporate location that can be used as the DR site for one or more of your primary locations? For companies that do not already have such a site, you'll need to consider how to address this issue. You may be able to rent access to a facility from a legacy DR services provider which can be used in the event of a disaster but be cautious: if you must declare a disaster to have access to a site which the legacy DR services provider shares across multiple customers, and the disaster is widespread enough that multiple companies in the same city as your primary site will be affected, you may find yourself without a chair when the music stops. Other options include newer managed service providers that may rent compute or storage resources for DR purposes from a large, shared infrastructure that supports multi-tenancy. Your resources may be dedicated, but you will likely have less flexibility to actually run business operations from this location if the outage to the primary site spans more than a few days. Still, for some smaller companies that just do not have access to a remote site location, these types of managed service providers may offer a good option.

If you do have a location that can be used as a remote recovery site, then you will likely need to evaluate some additional options. The use of tape vs disk as a recovery medium will likely be another critical early decision, so let's take a closer look at that. Tape-based DR meets a lax set of recovery requirements. If you're making extra copies of tapes once or twice a week that you ship via ground transportation to a remote location for long term storage, then the best RPO you will likely be able to achieve is going to be no more recent than several days to a week. The best RTO you can hope to hit will likely be several days to a week as well. If this meets your recovery requirements, and you are not concerned about your ability to recover data from tapes, then you may have what you need. If this type of recovery performance does not meet your needs, you'll need to think about another approach.

To meet more stringent recovery requirements, you'll want to consider the use of disk in your data protection and disaster recovery plans. Disk is much better suited for backup and recovery tasks because it can accept data at varying speeds, supports random access, and is a much more reliable medium than tape. When disk is used, all major recovery performance metrics (backup window, RPO, RTO, recovery reliability) improve over that achievable with tape, and disk provides access to other technologies, most particularly for DR asynchronous replication and the automation of certain recovery processes, that support the implementation of DR plans that can meet RPOs of

several minutes and RTOs within the same range.

If you are moving to the use of replication technologies, consider the potential impacts on network bandwidth. How much network bandwidth will be required to meet your RPO requirements? There are a number of WAN optimization technologies available which reduce the overall amount of data that has to be sent to remote locations to support recovery at those sites, and they include technologies such as TCP optimization, compression and other storage capacity optimization technologies, and bandwidth shaping and other quality of service tools. If you know exactly what the I/O rates of the application environments are that you will be protecting, it makes it much easier to determine if you can meet your RPO requirements with your existing networks or if you will have to purchase additional bandwidth when adding replication. If you don't know these rates, investigate how you can obtain that data prior to solution deployment.

Other considerations to take into account include what operating systems, server, and storage hardware need to be supported. Will you have to have equivalent hardware at both the local and remote site locations, or will you be looking at technologies that can support heterogeneous environments? Given that most IT shops today have a range of heterogeneous equipment, the use of DR solutions that support heterogeneous environments tends to result in more cost-effective implementations. DR solutions that support heterogeneity not only help to preserve existing investments, but they can also help to maintain maximum freedom of choice in new server and storage purchases as you move forward.

Step 4: Test Your DR Plans

There is a big difference between theory and reality. We've probably all heard the story about the bumblebee. Scientists evaluating the aerodynamics of the bumblebee, given what we know about aeronautics, would have to conclude that it could not fly. And yet it does.

Even if you have meticulously planned your DR implementation, going over it again and again in theory, to be sure it will work you have to test it. And testing it means more than just testing it once. To ensure that you will obtain predictable recovery performance from your DR solution – in other words, that it will perform as you expect, with no surprises, in meeting the RPO/RTO requirements you designed it to – you need to test it regularly. Due to changes that inevitably occur in systems, storage, and software, deployed DR configurations can potentially “degrade” over time in their ability to perform predictably. A small change on a production system at the primary site may result in an inability to recover data and/or applications at the remote site. The worst time to find this out is when you are actually in the middle of a real recovery. Very strict change management may be able to address this “degradation” issue, but this is a real risk. Replicated configurations can be complex, involving hardware and software from different vendors that must all work perfectly together to meet your recovery objectives. Although they may not admit it to their management, most DR administrators who use manually intensive recovery processes have very little faith in their ability to perform a disaster recovery without taking into account a huge dose of failure identification/fault isolation and troubleshooting/iteration in getting things up and running at a remote site.

There is a simple reason why many companies do not test DR plans. Testing them is disruptive to production application environments, and can be expensive – particularly if you are using some form of outsourced DR service where you are charged extra anytime you fail over to or recover from data located at their facilities. If you have deployed a DR solution using your own infrastructure, newer technologies such as server and storage virtualization, continuous data protection (CDP), and asynchronous replication, when combined with DR test automation, can help to address both concerns. Virtualization technologies can help minimize the hardware requirements at the remote site, lowering the overall cost of DR deployments, while technologies like CDP will allow you to perform DR tests without impacting production operations in any way. Automation minimizes the human risk element in performing recoveries, making the success of those recoveries much less dependent upon the sophistication of the administrator. Regular testing also helps you fine tune and improve your recovery capabilities, evolving them over time as your own recovery requirements evolve.

Step 5: Create A Disaster Recovery Run Book

A “run book” includes workflows that support system and network operational processes. Run books are applicable across all IT management disciplines and supply a way to deliver and prove higher IT operations efficiencies with respect to issues like provisioning IT resources and mean time to repair (MTTR), among others. Creating a DR run book is the first step in creating a set of repeatable processes that result in predictable recovery outcomes, and can be created in either electronic or physical book form. It will contain the procedures to perform IT infrastructure recovery operations, as well as descriptions for special request and contingency handling. Once you have defined your recovery plans, document them, and keep several updated copies – at least one at both the local and remote sites – available. This does several things:

- It ensures that the same recovery processes (hopefully the ones that you have thought a lot about and are regularly testing) are followed no matter who is doing the recovery
- It provides a baseline against which incremental improvements can be made over time

Conclusion

If you approach it with a good understanding of your recovery requirements, putting an effective DR plan into place is straightforward. The 5 general steps that should be followed to put a DR plan in place are:

- Understand your business priorities
- Assess your recovery requirements
- Match the right solutions to your recovery requirements
- Test your DR plans
- Create a DR run book

About InMage Systems

InMage is the leading independent software vendor in developing and delivering comprehensive, disk-based and scalable business application recovery solutions that allow companies to meet stringent disaster recovery requirements, eliminate the impacts of local backups, and manage application uptime to meet high availability needs. Targeting enterprises, InMage solutions provide solid data protection for high growth data environments while eliminating backups, minimizing data loss on recovery, shortening recovery times, and increasing recovery reliability in heterogeneous environments. InMage uniquely offers a low impact hybrid recovery technology and the use of AppShots to further support its customers’ remote disaster recovery, local backup, and high application availability requirements.



Headquarters

3255-1 Scott Blvd, #104, Santa Clara, CA 95054
Phone: 1.800.646.3617 | Local Phone: 408.200.3840 | Fax: 408.588.1590
Email: info@inmage.com | Web: www.inmage.com