The Democratization of IT Disaster Recovery

September 2014

FREEFORM DYNAMICS

Executive Briefing Guide

In association with

commvault

Introduction

For some, the term 'Disaster Recovery' (DR) conjures up visions of racks of equipment and rows of empty desks with phones, screens and keyboards sitting there doing nothing, waiting for the day when a fire, flood or natural disaster takes out your offices and systems. When disaster strikes, you ship the business wholesale to the DR location, and after a period of recovery you eventually get everything back up and running again.

This kind of arrangement and process defined the nature of DR for many years, and that legacy remains with us as people today frequently associate DR with words like 'complex', 'expensive', 'wasteful' and 'luxury'.

Yet with modern businesses being so reliant on IT systems and electronic data, the need for effective DR is arguably more important now than ever before. In the remainder of this paper we will therefore be looking at IT-related DR through a more up-to-date lens. We'll explore how advances in communications, virtualization, cloud computing and management tools have radically changed the game. The notion of DR has essentially been democratized through a dramatic reduction in cost and complexity. This means that protection can now be cost-effectively applied to a much broader range of your systems and data – conceivably even all of your IT.

With modern businesses being so reliant on IT systems and electronic data, the need for effective DR is arguably more important now than ever before.

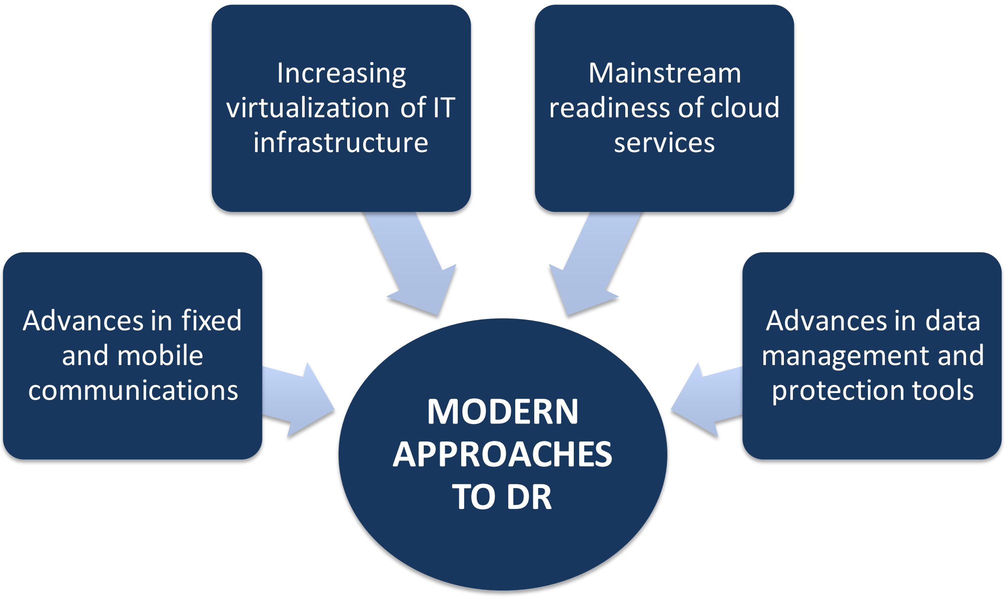
So, if it has been a while since you thought about your DR measures, or a review has been prompted by a risk assessment, compliance audit, actual disaster or some other scare, it's worth taking some time to understand what can be achieved in light of important changes that have taken place over the past few years.

Important trends and developments

Beyond an increasing reliance on IT systems, and the relentless growth in data volumes with which you will undoubtedly be familiar, a number of more specific trends and developments have come together to open up a range of new, more modern approaches to DR (Figure 1).

Figure 1

Trends and developments impacting DR needs and practicalities


Increasing virtualization of IT infrastructure
Mainstream readiness of cloud services
Advances in fixed and mobile communications
Advances in data management and protection tools
MODERN APPROACHES TO DR

Each of these trends and developments has impacted the DR domain individually, but they also build on each other to generate a much greater overall impact. Let's walk through how this works.

Advances in communications

Pervasive connectivity means employees in many businesses have been liberated from their previous dependency on fixed work locations.

Pervasive connectivity means employees in many businesses have been liberated from their previous dependency on fixed work locations. Wide area networks, broadband access and various forms of wireless technology have made it much easier to work effectively from remote sites, home offices, hotels, trains and so on. In the context of DR, this means you don't have to worry so much about providing an alternative physical place of work should an office or building be rendered inaccessible. Of course there are exceptions, such as hospitals, manufacturing plants, and distribution facilities, but a site-wide disaster with these will often generate a range of non-IT related headaches that are outside the scope of this paper.

Beyond workforce relocation flexibility, wide area networks, and even the internet (with appropriate controls), may be used to exchange data between facilities. This opens the door to data protection strategies that back up or replicate data across the network, whether site-to-site, between home offices and the data center, or between company and third party facilities. At a more micro-level, the degree to which end user devices are now connected also means you can better protect against and recover from disasters with individual employee desktops, laptops tablets and smartphones. This may sound trivial in comparison to protecting a core business system, but with so many employees now reliant on such devices to do their jobs, this can be a significant consideration.

Infrastructure virtualization

A major trend in IT over the past few years has been the virtualization of infrastructure components, particularly servers. The motivations for this have ranged from cost savings resulting from server consolidation, to improved flexibility and responsiveness to changing business requirements.

When a server and the software stack installed on it is virtualized, it is essentially transformed from a physical asset to a data object.

An often overlooked spin-off benefit of server virtualization, however, is in relation to DR. When a server and the software stack installed on it is virtualized, it is essentially transformed from a physical asset to a data object. For many purposes, virtual machines (VMs) can simply be regarded as another type of data file or package which can be moved around just like any other type of data. In a DR context, this means servers can be easily replicated to any location and spun up on alternative hardware when you need to recover.

There's a bit more to it than that, and we'll be looking at the practicalities and dependencies shortly, but the main thing is that DR is no longer necessarily dependent on having identically configured physical servers waiting around 'just in case', or recovery being dependent on acquiring and building server stacks to get the business back up and running. A VM can be recovered on any hardware that presents a compatible hypervisor, internally or in 'the cloud'.

Readiness of cloud services

Building on the idea of virtualization easing the job of backing up and recovering servers, the maturing of cloud services from a range of mainstream and local providers takes us to the next level of DR flexibility. One of the big selling points of cloud is rapid acquisition of resources. Depending on the arrangement with your provider, the compute and storage capacity needed to recover the servers and data associated with a business application can be available within minutes or hours. Indeed, it's perfectly possible to have dormant servers sitting in the cloud on stand-by, waiting to be spun up when necessary. Either way, this is a great improvement over the service levels typically written into traditional DR contracts. Cloud-based DR, of course, also has the added advantage that you often only start paying for resources when the servers are activated.

Apart from speed and flexibility, cloud-based DR has the added advantage that you often only start paying for resources when the servers are activated.

But there are some things to think about, such as how to keep cloud based servers and data in sync, and/or how to get them to the target cloud environment in the event of an emergency. Any use of hosted cloud services must then take account of the security and compliance implications, with a keen eye also being kept on variable costs to do with storage and transport of data.

Such considerations bring us onto the topic of the management software necessary to implement modern distributed DR of the kinds we have been discussing.

Advances in management software

At the highest level, when it comes to software, we are in the area of backup and restore, and/or solutions to replicate or synchronize data between environments. Such operations are pretty straightforward on slow moving systems, but the chances are that your business is dependent on one or more applications that would be classified as high throughput. While historically more of a challenge to deal with, this is an area in which management software today is much more capable. Modern tools are able to cope with data stores that are changing rapidly, allowing a continuous approach to data protection, which minimizes the loss of transaction data in the event of failure and avoids imposing onerous levels of intrusion and overhead on live systems.

Modern tools are able to cope with data stores that are changing rapidly, allowing a continuous approach to data protection.

In order to exploit the DR advantages offered by virtualization, the latest tools are also able to deal with system level dependencies. If you are going to recover a web-facing application in a DR location, for example, this is likely to involve more than one server. More capable solutions are therefore now able to capture and act upon dependencies at an application landscape level without the need for extensive manual scripting. As a simple example, a rules or pattern based approach can be used to ensure that the various VMs (and application data) used in a service all reflect the same recovery point state. From an activation perspective, other rules are then used to specify the order in which VMs need to be spun up (e.g. database server, then web server, then application server, etc).

Other requirements dealt with by modern management tools include the optimization of storage and network traffic through the use of compression and de-duplication techniques, along with highly selective approaches to replication. When changes occur, only the differences are sent across the network to maintain the remote backup or standby environment, rather than whole objects or data sets. Such optimizations make a huge difference to both performance and cost when coordinating activity across a wide area network or using cloud-based DR.

More capable solutions are increasingly hypervisor-agnostic.

And related to the use of cloud-based services from a flexibility perspective, more capable solutions are increasingly hypervisor-agnostic, allowing restore of data across major hypervisors or cloud platforms without complex manual processes.

DR and cloud in more detail

Cloud is meeting DR in a number of different ways.

Cloud services are becoming an integral part of the IT delivery activities of an increasing number of mainstream organizations. Some are exploiting options at an infrastructure or platform level, using virtual cloud resources to augment internal servers and storage devices. Others are using Software as a Service (SaaS) to consume business application functionality. Along the way, organizations are mixing and matching according to their needs, with most ending up with a mix of on-premise systems and multiple cloud services from multiple providers.

Against the backdrop of a distributed hybrid world, cloud is meeting DR in a number of different ways, and some of the most common are summarized in Table 1.

Table 1
Common cloud-related DR scenarios

Scenario	*Description*
On-demand cloud-based recovery	When recovery of an on-premise system is necessary, virtual servers and storage are created/allocated in the cloud, and backups are used to recover applications and data temporarily in the hosted environment. Systems and data may then be brought back on-site once the local infrastructure is up and running, at which point cloud resources can be released.
Cloud stand-by environment	Virtual servers and/or storage are set up in the cloud, and backups/replicas are kept up to date in that environment on a periodic or continuous basis. Then, in the event of a failure, these cloud servers are available for immediate failover from primary systems. Unlike on-demand recovery, the cloud DR environment here is persistent, so represents an ongoing cost.
Cloud-to-cloud recovery	Cloud services can provide resilience, but they are still subject to failures so may need additional protection. The same approaches as described above (on-demand cloud-based recovery and cloud stand-by) can be used with hosted virtual applications. The only difference is that the primary systems being protected are cloud-based rather than on-premise.
Reverse cloud DR	In some circumstances it can make sense to use your on-premise infrastructure as a recovery target in the event of cloud system failure. This may be appropriate if you have a private cloud running on your internal infrastructure that allows rapid, flexible resource re-allocation. Using available capacity in your data center for recovery can avoid paying for a temporary setup on an alternative cloud service.

It is possible to keep some of your options open.

The examples we have given here are not exhaustive; they are intended to provide an illustration of what's possible. The approaches listed are also not mutually exclusive, and it is possible to keep some of your options open, e.g. by deferring the decision on recovery target until the time a disaster occurs. The right option may then be selected based on available in-house private cloud capacity and/or the most favorable rates offered by alternative cloud providers.

Practical considerations

You need to avoid nasty surprises.

It is important to be aware of a number of practical considerations, both positive and negative, when using or considering cloud services as part of your DR plans. We have hinted at some of these already, but let's take a closer look. When going through the following sections, please note that not all of the areas listed will be relevant to you. Even where they are, the level of requirement or challenge may not be that great, depending on your circumstances, but you need to avoid nasty surprises.

Scope of protection and recovery time objectives (RTOs)

Cloud-based DR potentially enables RTOs to be optimized across the board.

Physical systems running Tier 1 (e.g. business critical) applications are likely to need fast recovery times in the event of any disaster, while other systems may have less rigorous requirements. Depending on criticality, you may therefore elect to use 'hot', 'warm' or 'cold' standby techniques. By enabling provisioning, spin-up and/or scale-up on demand, however, cloud-based DR potentially enables RTOs to be optimized across the board, allowing warm stand-by or rapid recovery in particular to be implemented cost-effectively for a broader range of applications. It is therefore worth working through your application and service portfolio and setting new objectives based on shifting cost and complexity lines.

Recovery point objective (RPO) related practicalities

Tier 1 applications may demand continuous protection so that no data is lost in the event of a failure. Other applications may be able to tolerate a certain amount of data loss following recovery (e.g. relating to the last few minutes or hours before the failure), permitting less expensive snapshot type approaches. While the use of cloud services per se changes little in terms of RPO principles, it is necessary to work through the mechanics of how data is moved into and out of cloud environments. Appropriate processes and tools are required to bulk transfer large VMs and data sets (including via physical media), then keep them in sync with incremental changes while avoiding runaway transport and storage costs and/or high admin overheads. Willingness and ability to support required schemes practically and cost-effectively need to be understood when selecting cloud service providers for DR purposes.

You ideally want a combination of DR automation with recovery flexibility.

Requirements for automation and flexibility

Building on this last point, you ideally want a combination of DR automation with recovery flexibility. Even though these two requirements may seem at odds with each other, modern tools go a long way towards providing the best of both worlds. More comprehensive solutions allow dependencies between system components, along with 'run books' and workflows for executing recovery processes at a detailed level, to be defined in an abstracted manner. This enables automated recovery regardless of the underlying nature of the target environment, cloud or local. As an example, it should be possible to recover a primary system landscape running on Hyper-V to a cloud service that presents virtual servers based on a KVM, Xen, ESX or any other mainstream virtualization layer, or vice versa. Look out for availability of embedded templates and best practices relating to common scenarios when evaluating tools, and make sure you understand relevant limitations and constraints.

You need to think through the security implications.

Assuring security of recovered applications and data

If you are going to use a totally independent third party environment to recover systems in the event of a disaster, you need to think through the security implications. Assuming you are working with reputable cloud providers, the concern here is less about the inherent physical security of your data as the providers will probably be doing at least as good a job here as you do (often better, given the nature of their core business). The challenges are more around making sure that existing security policies and access controls continue to be applied following recovery. In practical terms, this means including security and access components in your recovery landscape, and/or ensuring that existing security infrastructure (if it's still running) is available from the DR environment. Again, modern tools can ease the pain of doing this through capturing relevant dependencies and workflows.

Ensuring data integrity at a business level

Effective protection of business systems is dependent on making sure that backup or standby data sets aren't littered with incomplete transactions or other partially committed data. An important DR solution capability to look out for is therefore 'application awareness', which is the key to ensuring data integrity at a business level. The same mechanism makes it possible to define restore points then recover to the latest 'good state', e.g. in the event of corruption.

Those outside of North America may object to the use of US data centers, and others may be concerned about data being held in countries that are politically unstable or in which a high degree of corruption exists.

Dealing with compliance and reputational risks

One of the hot topics in relation to cloud services at the time of writing is data location. This sometimes matters because of regulatory requirements which might dictate where, geographically, certain types of data may or may not be stored. You also need to consider the reputational side of things, e.g. by asking yourself what your customers would think if they found out where you were storing sensitive information. Following the Snowden revelations, those outside of North America may object to the use of US data centers, and others may be concerned about data being held in countries that are politically unstable or in which a high degree of corruption exists. Fortunately, there are regional providers and larger players will often allow you to specify the region in which your services and data will reside, but if data location is an issue for you, it is wise to do your homework and identify options before the disaster occurs. Coming back to compliance, it's also necessary to ensure that providers have the certifications in place that are relevant to your business.

Reinstatement of primary systems

While it's natural to focus on the emergency recovery part of the process, it's important not to forget the practicalities of reinstating primary systems once the disaster is over. In the event of corruption or temporary outage, for example, it makes more sense to recover on cloud infrastructure, then stream back deduplicated data to repair the problem at the primary site instead of trying to perform a full recovery from the cloud. Apart from being quicker, this is more economical. Moving data into the cloud is typically free, but transferring it out again in bulk is expensive. Of course if the primary DC is a 'smoking hole in the ground' the transport costs are probably palatable, but for every such major disaster, organizations generally experience many relatively minor ones.

It's important not to forget the practicalities of reinstating primary systems once the disaster is over.

Other considerations

In addition to the above, other practicalities to consider include:

- The use of data encryption throughout the protection and recovery cycle, and how encryption keys are managed along the way.
- Network speed and latency and its impact on both the recovery process and subsequent operation of recovered systems.

- Network addressing and directory access to ensure consistency and continuity of access internally, over the WAN, and via the Web or mobile devices.
- Management and operations visibility in relation to the cloud environment, particularly the ease of administering resources and supporting/troubleshooting recovered systems.

There is a clear need to test your DR arrangements to make sure you can actually recover your systems and data in the event of a real disaster.

Of course, overarching all of the above, there is a clear need to test your DR arrangements to make sure you can actually recover your systems and data in the event of a real disaster.

This is another area in which templates, workflows and best practices embedded in the latest generation of information management solutions come into their own. Testing DR mechanisms that have been designed from first principles specifically for your requirements can be extremely onerous, especially if there is a significant reliance on scripting and manual activity. If most of the processes, integration and automation functions have been constructed to operate in a robust manner by the tools vendor, then testing and remediation overhead can be dramatically reduced – it's mostly a case of verifying that solutions perform as specified.

Final thoughts

DR solutions have advanced quickly in recent years. The expansion of virtualization platforms running mainstream business applications and the rapid development of cloud offerings provide organizations with a growing list of options for IT DR. The potential reduction in cost offered by cloud based and off-site managed DR solutions is clearly attractive.

More importantly these developments provide organizations with the potential to expand DR capabilities to systems which could not previously justify the costs and complexity associated with traditional approaches. Modern data protection systems, network connectivity and cloud solutions combine to provide options able to meet many, if not all, DR requirements. This is certain to interest line of business managers charged with keeping their departments functioning at all times, who may well find themselves under pressure from external regulators to be able to demonstrate adequate business continuity and DR arrangements.

Success depends on using the right technology, possibly in conjunction with cloud services, and implementing robust operational processes with as much automation as possible.

But while things have become a lot easier and cheaper, implementing DR effectively is still not a trivial matter. Success depends on using the right technology, possibly in conjunction with cloud services, and implementing robust operational processes with as much automation as possible.

This will translate into obtaining suitable tools that are functionally able, but which also have best practices embedded in them in the form of templates, policies and so on. Such solutions should also be capable of automating both data protection workflows and, more importantly, disaster recovery processes. In order to help you assess your existing capability, or to formulate selection criteria when reviewing potential new solutions, we have summarized some of the more important requirements in Appendix A for your convenience.

In the meantime, we hope our discussion in this paper has provided some useful input as you look to take advantage of the latest developments in this highly important area.

References and further reading

The following research reports are available for free download from the Freeform Dynamics website (www.freeformdynamics.com).

1. **Enabling Rapid and Effective IT Recovery**
   DR insights and tips for small and mid-sized businesses

2. **Disaster Recovery in European SMBs**
   Insights for vendors and the channel

Appendix A

Software tool requirements

Success with cloud-based DR is dependent on making the right management tool decisions. When selecting solutions in this area, here are some of the key capabilities to look out for:

- Ability to protect data in a variety of methods from snapshot type replication to near continuous data replication.
- Capability to save data to a variety of target environments including on-site, cloud shared resources, cloud dedicated resources, managed service providers, co-located storage facilities etc.
- Ability to store virtual machine systems in a format that permits VMs to be recovered to a platform based on a different hypervisor.
- Capability to de-duplicate certain data sets before moving data to the target location for saving.
- Software designed to optimize recovery times for both uncondensed data and de-duplicated data sets.
- Capability to recover systems with detailed granularity. i.e. recover entire system(s), single data sets, single file, single VM or a linked series of data sets that form an operational system.
- Ability to automate data backup / replication processes.
- Ability to create workflow processes for automatic recovery operations.
- Ability to test data protection and recovery capabilities to a variety of target environments.
- Ability to recover systems / data / VMs to a range of target environments including on-site, cloud shared resources, cloud dedicated resources, managed service providers, co-located storage facilities etc.
- Broad community of suppliers of services based on selected technologies.

It is also worth noting that many suppliers are happy to provide advice and guidance as part of the pre-sales process on a no obligation basis, and this is certainly something you should look to take advantage of where possible.

About Freeform Dynamics

Freeform Dynamics is an IT industry analyst firm. Through our research and insights, we aim to help busy IT and business professionals get up to speed on the latest technology developments, and make better informed investment decisions.

For more information, and access to our library of free research, please visit www.freeformdynamics.com.

About CommVault

CommVault provides companies with a better way to protect, manage, and gain business value from their data. Today, with more than 17,000 customers and counting, CommVault is liberating companies worldwide from chaos, excessive costs and complexity.

CommVault is a publicly traded data and information management software company headquartered in Oceanport, New Jersey. It made its mark with the industry's leading backup product, Simpana software. Customers choose CommVault because of its Solving Forward® philosophy and ability to deliver complete solutions with infinite scalability and unprecedented control over data and costs.

Leading technology companies worldwide have formed strategic partnerships with CommVault, including Dell, Hitachi Data Systems, Microsoft, NetApp, VMware, Novell, HP, Oracle and Bull.

To learn more about CommVault, please visit our website at www.commvault.com.

Terms of Use