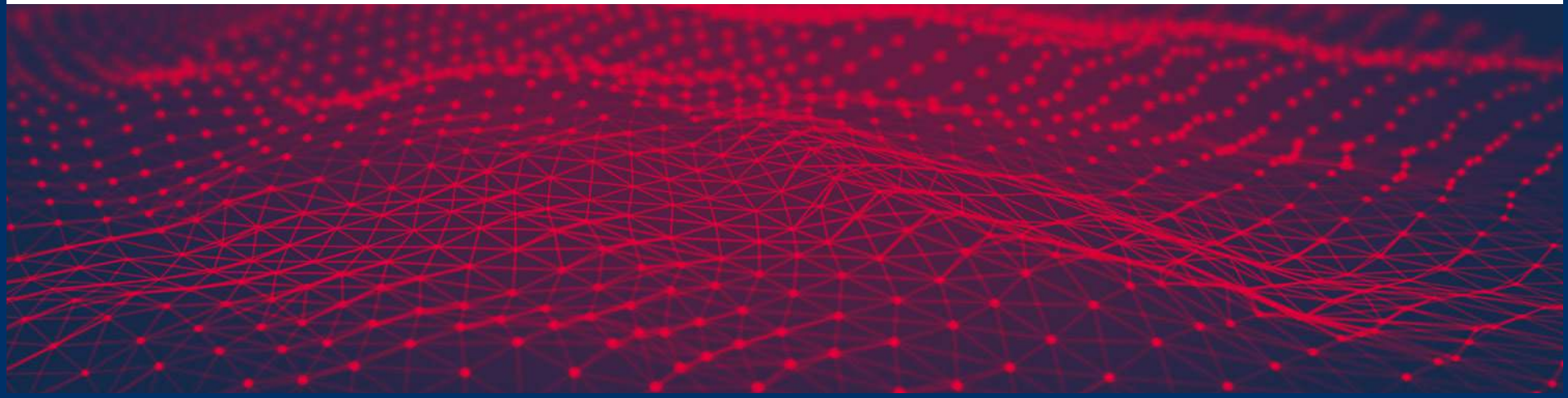
The Essential Guide to

Data Protection & Disaster Recovery

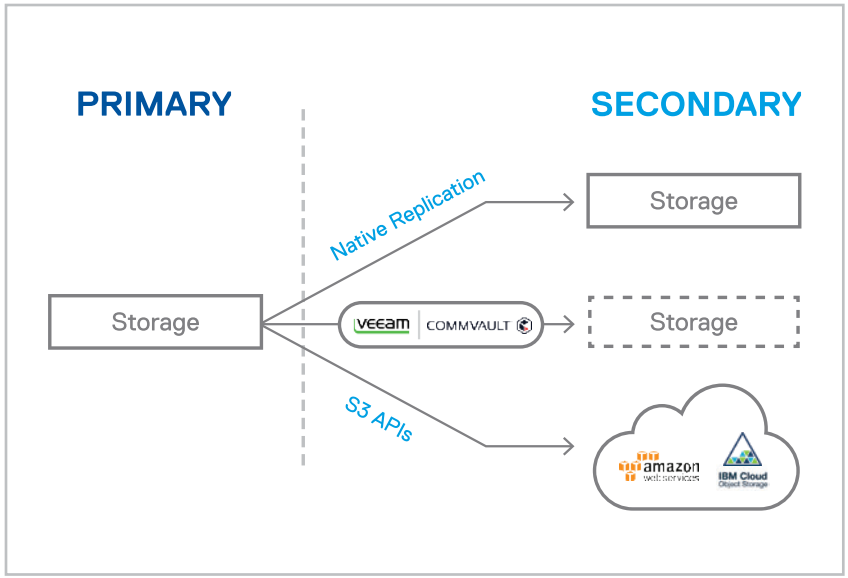


CONTENTS

In Brief

The available technologies for Data Protection and Disaster Recovery are evolving rapidly. It's important to understand your storage options, so you can build a strategy that maximizes both control and flexibility.

Methods of data protection and disaster recovery (DP / DR) for the enterprise change constantly in response to growing data volumes, increasing business demands, and faster recovery requirements. IT teams must make smart choices to leverage the latest innovations and ensure their businesses stay online under all circumstances.



Storage systems, which once played a more-or-less passive role in DP / DR, are now becoming key players. If you're in the market for new storage, paying attention to a few key factors will help you address your DP / DR challenges.

Any storage you choose must include snapshots, replication, and cloning. Inline deduplication and compression are also essential in helping you conserve both storage capacity and network bandwidth. These technologies must operate at the right granularity to control DP / DR cost and increase efficiency.

Primary storage must give you the tools you need to control and monitor DP / DR processes and the flexibility to utilize a variety of secondary storage targets. You may need to be able to replicate and recover to and from a public cloud or a cloud service provider, storage from the same vendor, or secondary storage such as a backup appliance (physical or virtual). Your storage must integrate with the DP / DR tools and processes you currently use or plan to use in the future.

This guide will help you understand the storage features that are essential to modern DP / DR practices and provides specific guidelines on what to look for.

What Are Data Protection and Disaster Recovery?

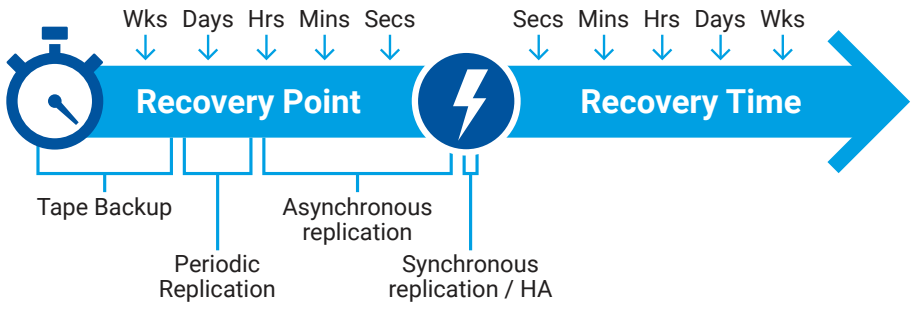
DATA PROTECTION

Data protection (aka backup and restore) is the process of protecting data against:

Corruption	**Accidental Deletion**	**Loss**

Examples where data protection comes into play include coding errors that cause a database table to become unreadable, a user who accidentally deletes an important file or directory, or hardware failures that result in minor data losses (as opposed to full-on disasters).

Traditional methods of data protection include tape backup and asynchronous replication.



DISASTER RECOVERY

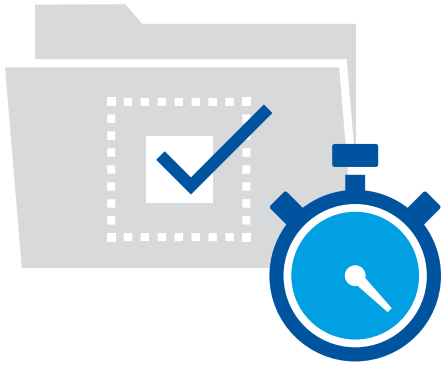
Disaster recovery is the process of restoring a failed application, service, or the operations of an entire data center to full operational status after a human-caused or natural disaster occurs. Having a disaster recovery plan is critical for ensuring business continuity.

There have been a number of high-profile outages recently that are making enterprises rethink DR:

- A data center outage at Delta Airlines in August 2016 grounded over 2000 flights over three days and cost $150 million.
- A storm-caused power outage at AWS Sydney knocked out a number of large websites and online services for up to 10 hours.

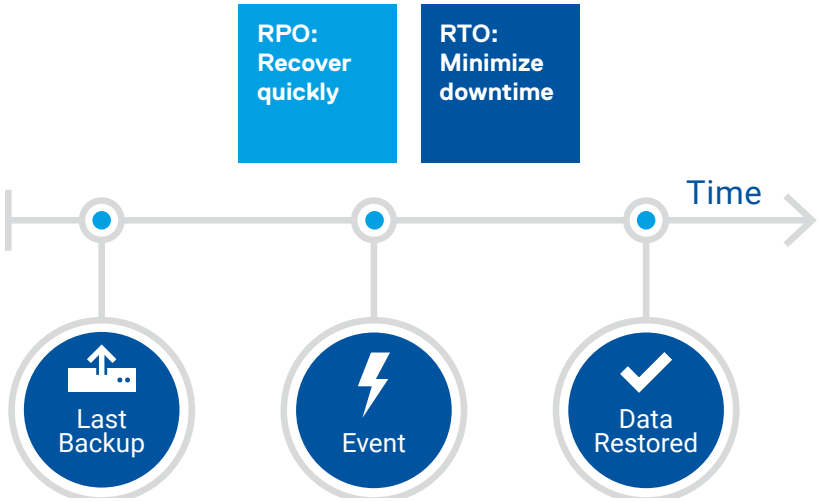
Methods of disaster recovery include recovery from offsite tape (slow) and replication (asynchronous or synchronous) to a secondary site. Organizations choose synchronous replication to protect their most mission-critical applications.

What Are RTO and RPO?



RECOVERY TIME OBJECTIVE (RTO)

Recovery time objective or RTO is a service level that specifies the time needed to restore an application or service to operation after a failure takes it offline. For example, an RTO of one hour is a commitment to have an application or service back online within one hour of failure.



RECOVERY POINT OBJECTIVE (RPO)

Recovery point objective or RPO is a service level that specifies the target point in time to which an application or service can be recovered. An RPO of one hour means that you will be able to restore all data for an application or service to a point in time no more than one hour before the outage occurred. An alternative way of thinking about RPO is that it is a measure of the amount of data you are willing to lose, i.e., an application or service with an RPO of one hour will lose up to an hour of data.

As a general rule, the shorter the objective, the harder it is to meet and the more expensive it is to achieve. For example, for applications that have a zero RPO and near-zero RTO, synchronous replication is the only option.

The Current DP / DR Landscape The landscape and the players in the DP / DR field are changing rapidly. Keep your options open so you don't get left behind.

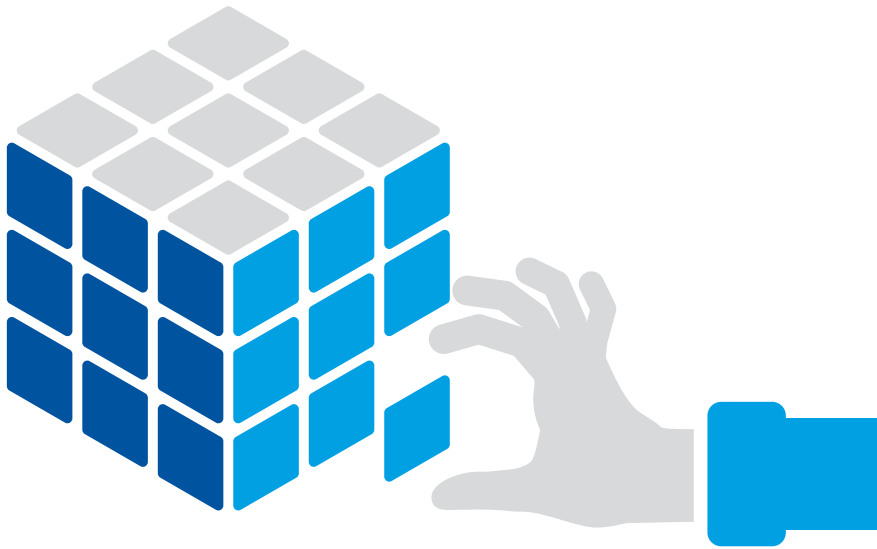
Once upon a time, DP / DR processes relied almost exclusively on tape media. But managing tapes was complicated and as the total amount of data grew, restore times became extended. Applications could end up being down for days if not weeks. To shrink backup windows, accelerate restores, and improve service levels, IT teams turned to disk-based solutions for protecting their most important data. Over time, special backup appliances that use deduplication and compression to increase efficiency and lower storage costs have become extremely popular as backup targets.

Today, it's common to have multiple copies of data on different media: disk-based copies for fast recoveries, tape and/or cloud copies for longer term retention, DR copies in remote locations for disaster, and so on. Several other important trends are affecting the DP / DR landscape as well:

- The use of snapshots as a first line of defense is increasingly accepted—and even expected.
- The enabling technologies that led to the backup appliance—compression and deduplication—are now widely available in modern storage systems and elsewhere, potentially making specialized appliances less valuable.
- IT teams are being asked to somehow shrink RPO and RTO further at a reasonable cost.

In order to navigate this changing landscape, you need to choose storage for your data center that will keep your options open. As you evaluate new storage options, you should ask the following questions:

- How easily will this storage fit with my current DP / DR approach?
- Does it support snapshots as a first line of defense?
- Will the storage allow replication to a variety of secondary storage targets?
- Does the storage integrate with leading DP / DR vendors?
- Does the storage integrate with the public cloud?



When Storage Selection Goes Wrong… The wrong storage fails to deliver performance when tasked with DP / DR and production work together.

Traditional Storage Fails to Meet Cybersecurity Lab Backup Needs

A government cybersecurity laboratory tests capabilities and trains personnel to prevent and defend against network intrusions.

Traditional storage using LUNs and volumes was not meeting either their provisioning or backup performance needs. The laboratory needed to be able to provision hundreds or thousands of applications to support training events and safely and efficiently back up all event data during exercises.

It was taking a full day or longer to back up the entire site on the existing storage platform, and they had to stop all other functions—including rollouts of other events—because it would effectively lock up the entire storage system and use all available resources for the backup.

PROBLEMS:

Existing storage lacked the performance to handle backup and production workloads simultaneously.

Backup and DR Interfere with Virtual Desktops at International Law Firm

A large international law firm was early to adopt a VDI-only environment with over 500 virtual desktops.

As VDI grew, legacy storage failed to keep up with the load. Users were constantly complaining about application lags.

Complicating matters, the IT team could never get vendor replication tools to work properly, in part due to performance. Whenever the team tried to do replication, it used up so much bandwidth and resources that it further degraded VDI performance. As a result, data protection suffered.

PROBLEMS:

Vendor DR tools were complicated and too resource intensive to run in conjunction with VDI on existing storage.

DP / DR too Complex with Traditional Storage at Public University

A leading public research university was struggling with all aspects of its legacy storage environment including overall performance, storage analytics, and the complexity of the backup environment.

To satisfy data protection needs for a variety of virtual machines, the IT team had to create and manage extra datastores in order to address a variety of recovery point objectives and performance requirements.

The storage monitoring software was unable to tell them which applications were causing problems, making it impossible to meet service levels.

Because of the complexity, all admins had to be sent to training. There were too many knobs to turn and a lot of pitfalls.

PROBLEMS:

Traditional storage made data protection too complex, lacked useful analytics, and required too much training.

What matters? The primary storage you deploy in your corporate data centers and other locations can make a huge difference to the success of your DP / DR efforts. Look for storage with these attributes:

Core Building Blocks

Storage should include snapshots, replication, and cloning, which create the foundation of your DP / DR strategy.

Data Efficiency

Compression, deduplication, and other efficiency features are essential technologies to save storage capacity and WAN bandwidth.

Flexibility

With the DP / DR landscape in flux, it's important to choose storage that leaves your future options open.

Automation

All DP / DR functions must be easily automated to reduce errors, save time, and enable private cloud and DevOps.

Control

Visibility into all aspects of DP / DR and other storage functions can eliminate surprises.

Ecosystem Integration

Your storage should integrate with all DP / DR elements including other hardware, data protection software, hypervisors, and more.

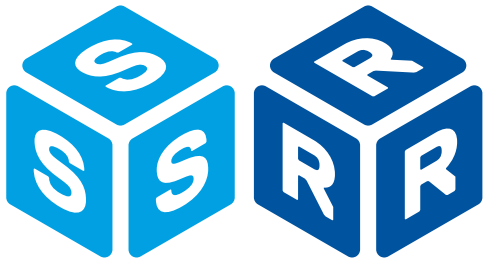
Each of these topics is explored in more detail on the pages that follow

Building Blocks Several storage capabilities are essential in enterprise data centers. You should not consider storage that lacks these features.

Snapshot

A fast and space-efficient snapshot is your first line of defense in data protection. In many storage implementations, snapshots are the foundation of most data protection and data management functions. Only recently has this technology been taken seriously as a first line of defense ahead of other data protection solutions.

It's worth noting that there are significant differences in the implementation of snapshots between vendors. Some implementations may still rely on time-consuming data copies. This is not what you want. New implementations track changed blocks. Some may use copy-on-write (COW) algorithms that will decrease storage performance after a snapshot has been created. The most efficient implementations use a redirect-on-write (ROW) algorithm.



Replication

Storage with built-in replication allows you to make copies of important VMs, applications, and data regularly, providing multiple points of recovery and protecting against disaster. There are two types of replication: asynchronous and synchronous:

Asynchronous replication is usually based on periodic snapshots. Once an initial data copy is created, only the changes that have occurred since the last snapshot are replicated, saving time and bandwidth. This significantly lowers the burden on production storage systems (and networks) versus full data copies. Asynchronous replication may be able to support a variety of bidirectional topologies including one-to-one, one-to-many, and many-to-one, giving you more flexibility in how and where you want to protect your data.

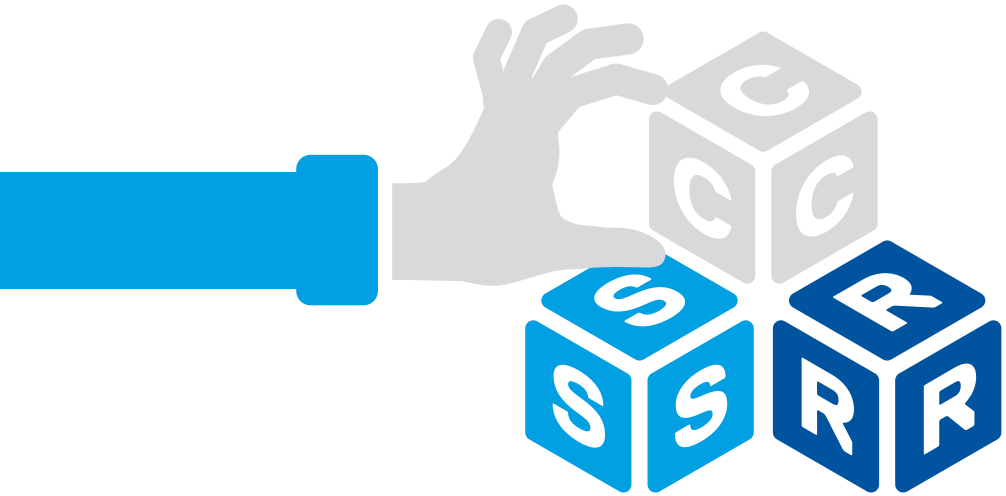
Synchronous replication guarantees that all data is written synchronously to two separate locations. Even if one replica is destroyed in a disaster, no data is lost. Synchronous replication may affect the rate at which data can be written.

Either type of replication can be performed within a data center, across a campus, within a metropolitan area, or across greater distances, depending on need. The greater the separation, the greater the latency challenge and as distances increase, bandwidth costs increase. Synchronous replication typically has a distance limit to minimize the effects of latency on write performance.

Building Blocks

Cloning

Cloning gives you the ability to create a writable “virtual” copy of a storage object such as a LUN or VM without requiring a full copy. Cloning is extremely useful for data protection, disaster recovery, and other data management functions. A failed VM can be quickly restored, in either a primary or secondary location, by cloning a backup snapshot. Use of cloning enables you to do DR testing without disrupting ongoing replication. Cloning also simplifies many other functions including a variety of development and DevOps tasks such as creating multiple copies of test data sets, creating identical work environments for developers, or making a test environment identical to your production environment.



Choosing Storage with the Right Building Blocks for your Needs

As you consider these features, it’s worth paying close attention to the differences between implementations. The most important factors include:

- **Time to completion.** How quickly does the operation execute? If it fails, do you get notified?
- **Performance impact.** Does use of the feature have a big impact on storage system performance either during or after the operation? (Especially important for production storage.)
- **Space efficiency.** How efficiently does the operation use storage space? Greater efficiency means less cost.
- **Granularity.** At what granularity can the operation be performed? LUN, volume, file, VM, vDisk? If your operations are largely virtualized, you’ll likely be happier with storage that can operate directly at the VM level.

DP / DR in a Modern Data Center

In a modern data center, the IT team typically takes snapshots of important data and applications on primary storage as a first line of defense every hour or two. This provides many potential recovery points versus daily backups. For some data, snapshots are enough. More important data or applications are replicated to one or more secondary locations (DR data center, cloud, service provider) at regular intervals. These secondary copies serve both as backups and provide DR.

Should a failure occur, and depending on the severity of the outage, the IT team can:

- Restore necessary data or applications locally from a snapshot.
- Restore necessary data or applications from a secondary copy.
- Restart the affected application(s) or service(s) at the secondary location.

Clones of applications at the secondary location facilitate testing of DR processes.

Data Efficiency Choosing the most efficient storage for your needs can significantly decrease overall costs—not just in the amount of storage capacity purchased, but also in WAN bandwidth expenses.

Conserving Storage Capacity

Modern storage systems, whether for primary or secondary storage use, should include a range of storage efficiency technologies to reduce the overall capacity you'll need:

- Thin provisioning reduces the amount of storage you need to allocate to applications and services upfront, increasing storage utilization.
- Deduplication can reduce the capacity needed for many common data sets. If your environment is virtualized, depending on how it's configured it may contain a large amount of duplication in the form of identical operating system and application files across many virtual machines. Backup data often has a very high degree of duplication, though this can depend on the backup methods you're using.
- Compression also does a good job of reducing capacity requirements for most data types. (The exception is data types such as audio and video that are frequently already compressed.) Compression can reduce requirements for database storage where levels of data duplication are typically low and deduplication may not be much help. Since the effectiveness of compression is highly dependent on the data type, look for algorithms optimized for colder backup data.

Conserving WAN Bandwidth

Any storage you choose should include efficient replication at the right granularity. If you are moving or plan on moving to a container model, choose storage that can operate at that granularity. There are two ways replication can conserve WAN bandwidth and reduce costs:

- By ensuring that data is always sent compressed and deduplicated over the WAN.
- By reducing the amount of data that must be sent. As an example, you may want to replicate a single application or set of applications, but LUN-based storage requires you to replicate an entire LUN with many applications—or re-architect your storage layout.

Efficient DRaaS Offerings for Service Providers

If you're a cloud service provider, controlling costs and maximizing resource utilization is critical. To offer DR as a Service (DRaaS), however, you may have to over-provision storage to satisfy customer needs in case of a failure. And you may have to provision storage systems for each customer or offer different tiers of service.

Modern storage systems with per-application quality of service (QoS) can solve this problem. Per-application QoS allows you to establish tiers of service on a single array; setting minimum or maximum QoS thresholds for groups of (or even individual) applications.

Now you can consolidate data from multiple customers onto a single array and guarantee that each customer will receive the level of service they are contracted for.

DRaaS is a fast-growing and increasingly important use case for cloud service providers, but you need storage that can provide the SLAs and policy management necessary to manage these advanced services. Companies are attracted to DRaaS offerings because they are able to eliminate the expense and complexity of maintaining a data center for DR and shift CapEx to OpEx.

Flexibility Given the evolution occurring in DP / DR, it's essential to select storage that gives you the flexibility to update your strategy to meet business needs and minimize costs.

The key factor in ensuring flexibility is to keep your options open in terms of where you will save DP / DR copies of your critical data. There are three main choices:

- **Secondary storage from the same vendor.** Using primary and secondary storage from the same vendor has several potential advantages. Since you don't have to manage different storage in the secondary role, it can simplify management. A vendor's native replication may deliver maximum efficiency and could eliminate the need for third-party software.
- **A backup appliance or other secondary storage device.** If you already have such a solution in place, you'll certainly want your new storage to be able to work with it—and you want to be sure your storage integrates without making the environment too heterogeneous. Sourcing a secondary storage solution from a different vendor will likely require third-party software.
- **Cloud storage either from a public provider or other service provider.** Businesses today are looking at ways to utilize the cloud to reduce backup and disaster recovery costs.

Choosing Storage for Maximum Flexibility

To be sure that the storage you choose doesn't hem you in, consider the following:

- **Broad support.** Ideally, make sure the vendor you choose has broad support for all three options described above. See the later section, "Ecosystem Integration" for related information.
- **Vendor limitations.** Make sure to understand any limitations in a vendor's DP / DR tools. For example, how much flexibility do you have in choosing a secondary device? You may choose all-flash storage for your primary storage needs, but it will likely be more cost-effective to replicate to disk-based or hybrid storage as the secondary target.
- **Synchronous replication.** Even though you may not currently be planning to use synchronous replication, it may be wise to choose a vendor that offers it as an option.
- **Cloud.** Be sure to find out which public cloud vendors the storage can integrate with and find out how extensively each storage choice is used by cloud service providers. It can be advantageous to choose a service provider that uses the same storage you'll deploy in your data center.

Automation To meet your DP / DR requirements now and in the future, you'll want storage that is easily automated. This will facilitate DP / DR processes and testing, offload IT staff, and enable you to better support private cloud and DevOps needs.

IT teams recognize automation as a way to streamline tasks and reduce or eliminate the chance of operator error. This is certainly the case with DP / DR operations, most of which are highly repeatable. You want to be able to automate regularly scheduled backup and replication activities, as well as ad hoc functions. For instance, it's common to create a snapshot of a application before installing a patch or upgrading software. Then if something goes wrong, you can immediately roll back. You will also want to be able to automate DR testing processes as much as possible.

IT teams are being asked to provide private cloud services through self-service portals with infrastructure resources such as applications and storage capacity. Storage should provide easy automation to make it possible to include DP / DR choices through the self-service portal. For example, a user might want to provision a new application and have snapshots of that application created every two hours with nightly replication.

Many enterprise IT teams are moving to a DevOps model that bridges the gap between IT teams and internal software developers to deliver software features more quickly and with higher quality. If your company is moving towards DevOps, you'll need to be able to incorporate snapshots, replication, and clones as part of automated workflows and testing.

What to Look for in Storage Automation

- **REST APIs.** The key to storage automation in the cloud era is REST APIs. Look for storage that offers a full set of REST APIs to allow you to automate all storage functions including snapshots, clones, replication, and monitoring.
- **Deeper Integration.** Some vendors may offer specific integrations on top of REST APIs to make it simpler to automate storage using popular tools such as PowerShell, Python, VMware vRealize Orchestrator, and various Microsoft tools.
- **Easy orchestration.** Storage infrastructure that's easy to orchestrate can be a big advantage. Once again, it comes down to a question of operating at the right granularity. In a virtualized environment, application-level operations will simplify the automation process. The same is true for container-level operations, if your company is moving in that direction.

Testing Your DR Plan

Most organizations find it extremely difficult to effectively test a DR plan. Bringing up a DR site to make sure that everything is working properly inevitably interrupts ongoing replication processes.

Cloning solves this problem. You can clone everything you need for your DR testing to create writable point-in-time copies. Replication processes can continue to write to the active volumes while you carry out your DR testing using the clones.

And since additional storage space is only consumed as changes are made, the whole process is extremely storage-efficient.

IT teams need a testing ground with the infrastructure resources to simulate failure scenarios without impacting production.

Control Storage should give you the tools you need to stay on top of your DP / DR environment.

In a dynamic IT environment, things are changing all the time. The right storage tools will give you the control you need to make sure your DP / DR processes are working as expected. Here are a few things you might need to do:

- Monitor all primary and secondary storage to ensure it has the headroom (capacity and performance) to keep up with production workloads and/or DP / DR processes.
- Monitor all DP / DR processes to make sure they are completing as expected.
- Identify data that is not being protected.
- Troubleshoot any problems that arise.

The right tools can greatly enhance your ability to implement a successful DP / DR strategy and save your team significant time and trouble.

Choosing Monitoring and Troubleshooting Features

Look for storage that provides the following features:

- **Basic monitoring.** At a minimum, the storage you choose should provide monitoring and alerting for all its core capabilities including backup and DR, so you can easily verify that everything is operating correctly.
- **Advanced monitoring.** The ability to monitor both primary and secondary storage and understand the relationships with networks, servers, applications, and applications can help you make sure that everything is receiving the correct level of protection. For example, it is often difficult to identify all the applications and vDisks associated with a particular application.
- **Predictive analytics.** The ability to forecast your future capacity and performance needs so you know exactly when to add resources can be extremely helpful.
- **Integration.** Storage should integrate with other monitoring tools in your environment. The ability to access information through API calls is particularly helpful for managing heterogeneous environments.
- **Application-level or container-level monitoring.** Most storage monitoring tools operate at the LUN or volume level. Visibility at the application-level and container level can simplify interpretation and speed up troubleshooting.

Ecosystem Integration Because storage is such a critical part of DP / DR success, it needs to work with all the other elements of your environment. Choose a solution with wide integration from a company that understands data protection and can support your efforts.

Make sure the storage you choose integrates with the other elements of your DP / DR environment and your data center:

- **DP / DR software.**
  - Look for vendors that work with leaders including Veeam and CommVault.
  - If you're a VMware Site Recovery Manager (SRM) user, check for integration.
  - Ideally, your chosen storage vendor should have a good relationship with other vendors whose software you intend to use.
- **Hypervisors.** Your storage needs to integrate with the hypervisor(s) you run:
  - VMware vSphere, including VAAI and VASA support
  - Microsoft Hyper-V, including ODX support
- **Automation and orchestration.** You'll want to be able to automate DP / DR functions as described earlier.

Some storage vendors offer plugins for greater levels of integration with the products above and direct access to DP / DR features.

When Storage Selection Goes Right… The right storage can help your DP / DR succeed. By making informed storage choices, these organizations overcame their storage problems.

Cybersecurity Laboratory Solves Data Protection Woes

When its existing storage failed to meet performance and backup needs, this government facility found storage that could keep up.

Deploying the hundreds of applications needed for a customer event can now be completed quickly.

Replication between production and DR storage takes just minutes, compared to a full day previously and backups can now be done behind the scenes without affecting production.

Additional benefits include QoS to guarantee performance at the level of individual applications—a 50% footprint reduction—reduced power and cooling, and reduced training needs.

SUCCESS FACTORS:

Guaranteed performance, ability to run DP / DR and production workloads, simplified management.

Modern Storage Accelerates VDI, Enables DP / DR at Law Firm

After viewing online demos, talking to peers at other companies, and reviewing industry blogs and reports, the law firm chose a new storage platform. Migration went quickly, and the IT team soon decided to upgrade all its existing storage based on phenomenal results.

With built-in replication and tight integration with Veeam, the team simplified its backup and DR strategy and reduced the complexity of recovery.

Given the lower costs, they were able to expand DR to all virtual desktops and reduce RTOs for many more systems. They also have more time to focus on strategic initiatives like security and service innovation instead of infrastructure.

SUCCESS FACTORS:

Vastly better DP / DR with no impact on VDI performance. Lower costs, better security.

Public University Slashes Backup Complexity

For this busy research facility, conventional storage was too slow and data protection was cumbersome.

Because the old storage only understood LUNs, the team had to maintain many different LUNs with different data protection and performance profiles. New applications were assigned to the LUN that was the closest fit.

With application-level storage, they can now enable snapshots, cloning, and replication on a per-application basis. The DP / DR needs of every application can be met exactly without the complexity. Per-application quality of service (QoS) addresses application performance requirements. Analytics provide visibility at the application-level, simplifying troubleshooting.

SUCCESS FACTORS:

Application-level storage simplifies DP / DR, removes storage complexity, and eliminates need for storage training.

Create a Successful DP / DR Strategy The advice in this book should help you narrow the field of potential storage vendors. As you compare the options in terms of functionality and cost, keep the following guidelines in mind:

Core Building Blocks

Space-efficient snapshots, replication, and cloning are must-have features for any new storage deployment. Be sure they operate at the same level of granularity as your operations.

Data Efficiency

Thin provisioning, deduplication, and compression can greatly reduce both your primary and secondary storage capacity requirements and reduce WAN bandwidth needs. That all adds up to significant cost savings.

Automation

You'll want the ability to automate DP / DR functions as much as possible including automating DR testing processes. You might need to incorporate automated DP / DR functions in a private cloud offering or as part of your development or DevOps practices.

Flexibility

The storage you choose should give you as much flexibility as possible to evolve your DP / DR strategy in the future. Look for storage that gives you a variety of secondary storage options including storage from the same vendor, backup appliances, and cloud storage.

Control

Storage should provide tools that simplify monitoring and management of all DP / DR processes. Ability to map from storage to the VM to the application can be extremely helpful in making sure you are getting data protection right and will allow you to quickly fix any problems. Predictive analytics simplify capacity planning.

Ecosystem Integration and Support

Your storage must be able to integrate with other parts of your DP / DR environment including backup and recovery software, existing secondary storage, and so on. You'll also want storage that integrates explicitly with other elements of your environment such as hypervisors and orchestration platforms.

Thanks for reading!

We hope the Essential Guide to Data Protection and Disaster Recovery got you thinking. Now it's time to make those wheels turn even faster—get hands-on with storage built specifically for cloud environments. We've created a mock-up of the Tintri UI, so you can see how easy it is to guarantee performance, scale-out, replicate and more. Just visit:

Explore.tintri.com

Tintri maximizes performance for your applications and the people who manage them. With all-flash storage and software for virtualized workloads, Tintri automatically manages each application, so you don't have to. That means you're free from decades-old storage constraints, so you can spend your time on high-impact projects.

Tintri

For more information, visit www.tintri.com and follow us on Twitter @Tintri