

A QuorumSoft Perspective:  
Disaster Recovery in a Virtualized World, What Every SMB Should Know



QuorumSoft





## Executive Summary

While the first wave of virtualization has consolidated server hardware into virtual hosts and recognized significant savings in physical machines, small and mid-sized organizations must now consider the impact of this consolidation on their backup and disaster recovery (DR) plans. It is possible to continue using agent-based backup and other traditional tools in virtual environments, but these methods do not allow for much of the value that a virtualized environment can offer.

Virtualization decouples the traditional link between software and hardware, allowing for recovery and automation techniques that previously were impractical, cost prohibitive or impossible. By encapsulating workloads that can be easily shifted between physical hosts and even 'clouds', new options for performance, redundancy and disaster recovery exists for companies that have implemented virtualization. In the past, these techniques have been available only to very large organizations with sufficient budgets, staff and equipment and have been out of reach for the smaller operations. That has changed. There are now strategies, tools and business processes available to small and mid-sized organizations that make it possible to cost-effectively implement a comprehensive disaster recovery plan in their virtualized environment. As a leading provider of backup and disaster recovery solutions for small and mid-sized organizations, QuorumSoft has developed an in-depth understanding of the DR best practices available to organizations with limited information technology (IT) resources. With small and mid-sized organizations in mind, this document explains recovery scenarios, discusses how to set realistic objectives, and recommends the tools to use. This guide is intended as an overview of disaster recovery considerations for small and mid-sized organizations, and not a means to achieve conformance to any specific regulatory environment.



## Disaster Recovery Planning: An Overview

A formal disaster recovery plan can enable an organization to survive a serious event by eliminating confusion during a disaster, and setting clear expectations for what systems and services will be brought online when, and how. Many factors go into the creation and maintenance of a DR plan, including the systems, network, budget, staff and resources and, of course, business requirements. The process of developing a DR plan is complex. Of the many books and professionals dedicated to the topic, most recommend starting with a simple audit of your environment. Itemizing and documenting your services, systems and their dependencies can be a big help in understanding your environment, and paves the way for the rest of the work. Identifying the 'fragile artifacts' within an organization allows you to plan for their protection.

### Setting Objectives: Beyond RTO and RPO

When most industry experts talk about recovery metrics, they discuss Recovery Time Objective (RTO) and Recovery Point Objective (RPO). RTO describes how long it takes to restore a service after a disruption, and RPO describes the maximum amount of data loss you are willing to tolerate.

For example, if your organization requires an RPO of 24 hours and an RTO of six hours for a particular system, you would need a minimum of one backup per day, and you would have six hours to bring it back online. In this case a tape restore may be perfectly feasible. However, if the same system has an RTO of one hour, tape may no longer be an option, and a more complex and expensive solution may be required. Typically the lower your tolerance for data loss and time, the higher the cost to accommodate the requirements.

The problem with these two metrics is most businesses would answer "as quickly as possible" to RTO and "as little as possible" for RPO. You will need to explore and articulate other objectives and tradeoffs to correctly determine what is realistic, how much it will cost, and when it can be rolled out.

### Cost vs. TCO?

For organizations on a tight budget, price is a crucial factor when introducing new backup products and solutions. IT organizations with a broader mandate and bigger budget to pursue business continuity objectives should consider the total cost of the solution over time, in addition to the up-front sticker price. Additional questions to consider in determining the total cost of ownership (TCO) for a solution are: How many IT staff members are required to maintain a system? Does the system require specialized knowledge? Is the system scalable, and what are the ramifications if I expand? All things being equal, a simpler solution will cost less in the long run.



## Understanding Recovery Scenarios

Not all disasters and accidents are created alike. Small and midsized organizations should plan for three basic uses of their backups, and distinguish their costs, effort, and priority.

### Disaster Recovery

Disaster Recovery encompasses a broad set of practices within the broader discipline of business continuity. From an IT perspective, Disaster Recovery concerns itself with the processes needed to provision for and recover from incapacitation of IT capabilities at a primary site. Incapacitation could result from a natural disaster, terrorism or intentional sabotage, massive hardware failure, and any other adverse event.

### Partial Outage

Partial outages are a common occurrence in IT organizations, and backups can help serve the critical function of restoring lost capabilities when redundancy and fault tolerance have failed. A partial outage covers a range of scenarios from a single system or service failure to more sweeping outages across the network. Partial outages can occur from hardware errors, failed upgrades and patches, human error, viruses, or deliberate attack, to name a few. Organizations can mitigate outage through redundancy strategies, but backups are a necessary complement to hardware and service redundancy.

### Routine Recovery

Routine recovery includes restore operations necessitated by a localized data loss or mishap, rather than a true disaster. Situations such as recovering a single file overwritten accidentally or restoring a deleted email can be vital to preventing lost productivity and time, but rarely factor into a full DR plan. Good recovery procedures in this area can save your business money and time, as numerous small restore requests can divert valuable IT resources from more vital tasks. For example, offering self-serve access to previous versions of a file via Microsoft's Volume Shadow Copy Service on a file share, or careful configuration of Exchange's deleted item recovery setting, can save a tape restore for a file or email. When considering routine recovery, remember that efficiency is critical. Time saved by avoiding unnecessary restoration from backups can add up quickly.



## Not all systems are equal

Consider the recovery needs of two different systems: a web load balancing server and a batch order entry system. Both may be critical components of a business' infrastructure, but they may have vastly different requirements.

The load balancer, for example, front ends a real-time web service, but it retains little to no historical data of its own. Because its configuration and data rarely change, its RPO may be weeks or more. However, it is a vital component to the website, so its RTO is extremely low because when it is down, the whole website is inaccessible to its users.

Contrast this with the needs of a hypothetical batch-order entry system. This system processes orders taken by salesmen on their mobile devices whenever they choose to sync their data. Once uploaded to the entry system, the data is removed from the mobile devices to save space. In this case, the RPO is dramatically more important because lost data may mean lost orders. However, the RTO is less critical since the orders can remain queued on the mobile devices until the system comes back online.

## Recover to or from a Secondary Location?

A distinguishing feature of most DR plans is the attention placed on recovering critical services in the event of a site disaster. To accomplish this goal, a secondary recovery site is typically established to be used during a disaster. The role of this secondary site can range from simple off-site storage of backup tapes, to data vaulting, to a full infrastructure build-out and hosting of a replicated environment that mirrors production.

In the past, a secondary DR required a significant investment in data center space, server capacity and networking and IT staff. Often similar or identical server hardware had to be carefully provisioned and maintained in the off-site location to match the production configuration. This infrastructure had to be managed and maintained over time, just like any other. Thankfully, with the increasing popularity of server and desktop virtualization, the cost of entry and difficulty of implementation have dropped dramatically. By understanding the requirements of virtual machines (VMs) and their workload, it is simple to plan for capacity of your host servers in the secondary site. Furthermore, many Managed Service Providers (MSPs) and private cloud hosts can eliminate the cost and complexity of maintaining a dedicated secondary recovery site entirely. Leveraging virtualization and low-cost vaulting and replication techniques, can lower the cost of entry to within reach of many smaller organizations.



## Common off-site techniques

### Data Vaulting

In data vaulting, backup data (usually from disk-based backups) is sent over public or private networks to an offsite location. This allows for a high level of recoverability in the event the main site is lost or destroyed, and systems must be rebuilt. This is typically not the fastest way to recover services, but can provide a very cost effective means to provide protection for the systems and data being vaulted. Vaulting solutions come in both hardware and software forms, and generally incorporate compression, encryption and deduplication.

### Warm Standby/Server replication

Having duplicates of servers at the ready in a secondary site can provide an extremely rapid form of service recovery. There are many ways to achieve this, but generally the cost and difficulty of a warm standby solution increase as your RPO goes down and your data change rate goes up. Given the proper workloads, a scheduled replication of virtual machines can form the backbone of many DR plans.

### Realtime Replication

For systems requiring a minimal RPO and RTO, real time replication may be the only viable solution. These systems will replicate (journal or log ship) changes in real time, as they are made to the production systems, sending them to the secondary site. In the event of a disaster, the secondary site can assume hosting production services very rapidly, if not immediately, with minimal data loss. These systems are among the most costly, and require constant care to ensure they are functioning properly. They can also fall prey to vulnerabilities that affect many replication techniques, where data corruption or loss in production can replicate to the secondary system. For this reason, it is usually wise to support any replication strategy with a solid backup plan as well.

An often overlooked side to planning a secondary site is mapping out how to recover back to the primary site. You will eventually need a mechanism to move services back from the secondary to the primary site. Planning for this eventuality will also assist you in capacity planning for your secondary site. If you intend to run on your secondary system for two months, you may need more substantial systems than if you only intend to stay there for two days.



## Differentiating Among Backup Sources

How a backup is conducted should be dictated by the data you are trying to protect, as well as your other business requirements. Three common sources of backup data are considered below.

### Operating System Images and VM Snapshots

In a DR situation, it is most common to recover entire systems, not just files or applications. Traditionally, backup solutions address this issue by providing system recovery products that allow administrators to restore to “bare metal.” This allows an administrator to restore an entire server, operating system (OS) and all data. Bare-metal recovery is challenging, because operating system configuration and drivers must match hardware configuration, and in many instances, the hardware cannot be matched exactly. Fortunately, virtualization provides tremendous flexibility in this area, allowing virtual machines to be recovered freely between physical hosts, and in some cases virtual operating systems, commonly called hypervisors. These ‘image level’ restores are invaluable in a disaster recovery situation.

Virtual backup adds another distinct advantage over physical backup. Its flexibility in deployment and re-deployment makes it easy to address many contingencies. If an OS upgrade or patch goes wrong and a partial outage occurs, virtual backup makes it easy to recover the entire VM. Since upgrades touch hundreds or thousands of system files, snapshotting a point-in-time image of the entire VM is the most convenient way to protect data. Virtualization also makes it possible to clone a VM and test changes on the copy, or migrate the VM to an isolated environment before making alterations.

### Transactional Applications

Backup approaches have differing abilities to capture the state of running applications. If you need to back up live applications and services, be careful to distinguish between “crash-consistent” data and “application consistent” data.

Crash-consistent data refers to the state of the disk at the time the backup was conducted, and is comparable to unplugging the computer the instant after the backup was conducted, since that is how the application in question will respond after you restore it. For some applications, this is fine. But for others, it is not good enough.



Capturing a transactional system's disk in the midst of its operation and attempting to recover it at this point in time may create inconsistent data, or worse, data corruption. In relational databases, this is usually overcome by scheduling exports of data using an agent. In the case of an outage or failure, the most recent export is imported into the database. This export may be supplemented with transaction logs in order to replay recent activity that occurred after the export. Database backup is a complex topic and requires specific expertise and planning. Make sure to consult and involve the appropriate technical subject matter expert.

Microsoft offers a uniform approach to application consistent backup called Volume Shadow Service (VSS). VSS can capture the state of a file, even if it is open for writing. Furthermore, applications can notify VSS when they are ready to be captured, making it possible to use a single backup mechanism across diverse transactional applications. VSS is leveraged by many backup mechanisms, including backup agents and whole VM backup. When used in conjunction with VM backups, VSS helps provide a consistent view of an entire virtual machine, known as a quiesced snapshot. This "one-size-fits-all" strategy to Microsoft server backup has merits, particularly when capturing simple transactional systems. But VSS can also add complexity and risk to your backup strategy, may not be available for all your applications, and may be unwieldy for larger or busier systems.

## **User Data Backup**

User data is typically the intellectual property at the core of your business. For most organizations, these documents are in office file formats. Other businesses heavily use graphics or media files, and still others depend on specialized or proprietary formats. When backing up user files, first survey the types of files users currently store—on their laptops, desktops, and file servers—and plan accordingly. You may wish to determine whether you require search features or retention policy managers to fit your regulatory environment. While image-level restores are far more effective during a disaster recovery scenario, file-level backup and recovery can be considered part of your routine recovery system.



## Understanding Backup Targets

In order to meet your various recovery scenarios, backup objectives and backup targets, you will probably end up with more than one backup approach. The options include tapes, disks, and network backup. The role of tapes, long the mainstay of backup strategies, has been in decline but still represents a good option in some situations. Quoted in *Network World*, Dave Russell of research firm Gartner, Inc., noted that the percentage of enterprises relying exclusively on tape backups has declined from 63% to 13% in the past six years. He also notes that today 65% of enterprises have adopted a disk-to-disk-to-tape approach, which involves backing up to disk first and then writing the backup to tape.<sup>1</sup>

*Key backup targets and techniques, and their strengths and weaknesses, include:*

### **Tape Backup**

Tapes are the classic backup target, and persist today because of their durability, economy, portability, wide support, and decades of precedent. The downside of tape is the complexity of managing tape rotation, slow backup and restore times, and fickleness of the media. As a means of archival, tapes are hard to beat, but as a form of short- and mid-term backup, they are being eclipsed by many of the newer options and technologies available.

### **Disk to Disk Backup**

Disk-to-disk backup uses live, spinning disks to house backup data. Disk backup can offer much faster backup and restore times than tape, though disks are not often considered a form of long-term storage. The cost of disk-based backup can be considerably more than tape, although techniques such as data deduplication can help minimize the cost while providing the performance advantage and convenience of disk. Whenever possible, isolate disk backups from production disks.

### **Portable Disk Backup**

Portable disk backup, usually with USB drives, is effectively a hybrid of tape and disk backups. It provides an inexpensive means to store backup data on a portable medium, which can be rotated and transported offsite. This technique can be convenient and effective, if planned and reviewed carefully. The risks of this system

<sup>1</sup> Russell, Dave, research vice president, Gartner, Inc. quoted in "Disk or Tape? How about both", by Ann Bednarz, *Network World*, February 21, 2011.



arise from the fact that hard drives (excluding solid state) contain many moving parts that can be easily damaged by constant movement and rotation. Because of this, there is a very high failure rate with commodity USB drives, so they can not be considered very durable or reliable.

## Network Backup

Network backup uses local area networks (LANs), wide area networks (WANs), or public networks to send backup data from a primary site to one or more secondary locations. Network backup is generally considered a form of disk based backup that provides geographic and hardware diversity. Network backup encompasses both hardware and software solutions, such as replication, data vaulting, and cloud backup.

## Cloud Backup

Cloud backup has become a popular topic in the industry lately, and is usually a form of hosted network backups, where the off-site storage is provided by a service provider. In deciding whether or not to back up data to a third-party vendor, you will need to consider your regulatory framework as well as other business factors. A key consideration for cloud-based backup is privacy and security for your data. Encryption of the backup data is generally a requirement, especially when the service housing the data is a shared or multi-tenancy system. Performance concerns are very common among cloud-based offerings, especially for the 'initial load' issues. Many vendors will offer a means to ship a tape or portable drive containing 'seed data' to help offset the otherwise large uploads over internet connections. For many organizations, using public cloud storage to meet at least part of their back up requirements has provided business benefit. According to Dick Csaplár of Aberdeen Group, organizations that moved at least part of their back up to the cloud recovered from downtime events on average four times faster than companies with no formal cloud storage program.<sup>2</sup>

## Data Deduplication

Data deduplication is a storage technique that identifies redundant data and consolidates it. Because backups inherently contain a lot of redundant data, especially over time, data deduplication is an excellent fit for archiving. Since rotating tapes or drives will impact data deduplication, it is best to combine data deduplication with disk backup.

<sup>2</sup> Csaplár, Dick, "Small and Mid-Sized Organizations Gain Disaster Recovery Advantages Using Cloud Storage," Aberdeen Group, December 31, 2010. <http://www.aberdeen.com/aberdeen-library/6827/RA-disaster-recovery-cloud.aspx>



## Backup Tiering

Using tiered backups, organizations can gain the advantages of more than one backup target while increasing redundancy. The most common tiered strategy is to back up first to disk, then archive the backup set to tape (i.e. “disk to disk to tape”). The tape tier can be used for disaster recovery purposes, while the disk tier supports routine recovery and partial outage.

Tiering creates complexity, which may conceal gaps in your backup coverage and create a false sense of security. For example, it is tempting to use extra space on production storage area networks (SANs) for backups, which is then backed up to tape. But using spare production disk as a backup target is a risk because if that disk becomes unavailable, you have compounded the problem of a production outage with a new problem of missing backups. If other systems fail while you are recovering from tape, a partial outage has grown into a catastrophe.

Furthermore, tiering has implications on RPO. If you back up files to a SAN on Sundays, and back up the SAN to tape on Saturdays, the tape tier will contain backups that are nearly two weeks old. Unless you carefully synchronize the two backup schedules, and make sure they remain in sync over time, you can seriously hamper your RPO. When tiering backup products or backup procedures, it is key to calculate RPO end to end, not just the RPO at each tier. Finally, tiering multiplies the potential for failure, as a failure anywhere in the backup tiers will result in an inability to recover at bottom tier. All backups fail occasionally for a host of reasons, such as running out of space, networking issues, or media problems. You must periodically test the entire chain to verify restorability.

Despite these problems, most organizations should consider some form of tiering, especially in virtualized environments. Tapes aren’t going away quite yet, and marrying them with the speed and convenience of disk is practical and expedient approach if pursued thoughtfully.

## Test your DR plans

The longer a Disaster Recovery plan has not been tested the more likely it is to fail. Systems and services change constantly over time, and so too must your plan. A routine review and update of your DR plan is vital to ensuring its viability. Scheduled tests and dry runs enacting your plans will iron out any wrinkles or shortcomings before the plans are put into action. This will also help eliminate any confusion or uncertainty during an emergency. Ideally, your DR systems will be auditable, allowing you to test your backups, replicated or failover systems without interruption to your production services.



## Conclusion

It is now possible for small and mid-sized organizations to access and use many of the same powerful IT capabilities once available only to larger organizations with more resources. Implementing virtualized IT environments is one area. As smaller organizations leverage the potential of virtualization, it is critical they consider the implications on their disaster recovery planning. By gaining an in-depth understanding of specific strategies for disaster recovery, small and mid-sized organizations can leverage the potential of virtualization without creating a hidden, potentially costly, risk to their systems and data if a major disruption to their IT systems occur.

## About Quorum Software, Inc.

Quorum Software provides deduplicating backup and Disaster Recovery products specifically tailored for organizations with a virtualized environment. Its popular Alike™ product line is the first deduplicating archival solution for Citrix XenServer. The solution is available in Free, Standard, and the Alike DR Disaster Recovery editions. Alike DR is ideal for small and mid-sized organizations that need a cost-effective disaster recovery solution for their virtualized environment. Alike DR can vault multiple point-in-time backups offsite and efficiently replicate entire virtual machines to a secondary XenServer pool. Alike DR also includes an Enhanced Replicate feature that snapshots and transfers virtual disks from one XenServer host to another directly over a compressed data stream.

To get more information about Quorum Software and the Alike product family,  
visit [www.quorumsoft.com](http://www.quorumsoft.com) or call **800.688.8769**.