Disaster Recovery for Virtual Environments:
One Simple Solution for Five Common SAN Replication Challenges

Hani El-Qasem
System Engineer, Veeam Software
Contents
Introduction
Background
Scope and Definitions
SAN Replication Overview

Challenge 1:
Hardware Dependency

Solution 1:
Software Replication to Dissimilar Hardware.

Challenge 2:
Achieving Application Consistency

Solution 2:
Granular and Agentless Application Consistency

Challenge 3:
Bandwidth Constraints

Solution 3:
Bandwidth Optimization

Challenge 4:
Failover and Failback Management

Solution 4:
Simple Management for Failover and Failback

Challenge 5:
Replicating Corruption

Solution 5:
Multiple Restore Points

Conclusions

About the Author

About Veeam Software
Introduction

It would be no overstatement of fact to say that in the last five years virtualization has radically changed the landscape of IT infrastructure for the better. Workloads encapsulated into standardized virtual machines have significantly increased our ability to optimize and use physical resources in a way that saves much time and money. In addition to these economic benefits, new avenues have opened up to tackle data protection and disaster recovery, allowing us to increase service uptime while also reducing business risk. This white paper focuses on some of the common challenges experienced while implementing and using SAN-based replication for disaster recovery and it examines an alternative approach to disaster recovery to help resolve these issues.

Background

Pre-virtualization disaster recovery plans were underpinned by application-level features hooking directly into specific hardware to achieve the required business recovery goals. To ensure that disaster recovery could be achieved, network infrastructure, hardware, software and application data were replicated to an off-site location, commonly referred to as the Disaster Recovery (DR) site. Depending on an application’s required Recovery Point Objective (RPO) and Recovery Time Objective (RTO), costs could spiral upwards to achieve small improvements in both RPOs and RTOs. When you increase the amount of application uptime provided from 99.99% to 99.999%, the cost increase is not linear, it’s exponential.

With the advent of virtualization, the infrastructure stack gained a new layer, enabling the movement of workloads between geographically dispersed locations. Importantly, this is achieved without requiring application-specific engineering, because workloads are compartmentalized and encapsulated into virtual machines. In a virtual machine, everything needed to support that workload can be encapsulated into a set of files in a folder and moved as a single contiguous entity.

Scope and Definitions

The virtualization and storage layers are examined below; i.e., virtual machines (VMs), hypervisors and storage area networks (SANs). Application-level replication is beyond the scope of this document.

There are many potentially overlapping terms, which people often interpret differently. For the purposes of this paper, I will use “Continuous Data Protection (CDP),” “synchronous,” “fault tolerant,” “asynchronous” and “high availability.”

CDP consists of synchronous replication, which in turn involves double-writing to two different devices where an application (or hypervisor) only receives a confirmation of a successful write to storage when both devices have acknowledged completion of the operation. CDP can help you achieve a zero RPO and RTO but requires strict hardware compatibility at both source and destination sites. This allows you to deploy VMs in a cross-site, fault-tolerant configuration, so if you have an infrastructure problem, you can failover to the DR site without any downtime.
Synchronous solutions are expensive and require a lot of network bandwidth but are appropriate for some mission-critical applications where no downtime or data loss can be tolerated. One issue with synchronous replication is that data is transferred to the DR site in real time. This means that if the disaster is driven by some kind of data corruption, malware or virus, then the problem that brings down the production site simultaneously does the same to the DR site. This is why synchronous implementations should always be combined with an asynchronous capability.

This paper primarily is concerned with asynchronous replication of virtual infrastructures for disaster recovery purposes. An asynchronous strategy takes a point-in-time copy of a portion of the production environment and transfers it to the DR site in a time frame that matches the required RPO/RTO goals, and this may be “near real-time/near CDP” or scheduled (hourly, daily, etc.). This is more akin to high availability than fault tolerance. High availability in virtual environments refers primarily to having cold standby copies of VMs that can be powered on and booted in the event that the live production VM is lost. This approach underpins most currently implemented DR strategies.

The next section examines how SAN technologies approach asynchronous replication and the differences between SAN-level and VM-level strategies to achieve DR objectives.

**SAN Replication Overview**

SAN devices are typically engineered to aggregate disk resources to deal with large amounts of data. In recent years, additional processing power has been built into the devices to offload processing tasks from hosts serving up resources to the virtual environment. The basic unit of management for a SAN device is a Logical Unit Number (LUN). A LUN is a unit of storage, which may consist of several physical hard disks or a portion of a single disk. There are several considerations to balance when specifying a LUN configuration. One LUN intended to support VMs running Tier-1 applications may be backed by high-performance SSD disks, whereas another LUN may be backed by large, inexpensive disks and used primarily for test VMs. Once created, LUNs are made available to hypervisors, which in turn format them to create volumes; e.g., Virtual Machine File System (VMFS) on VMware vSphere and Cluster Shared Volume (CSV) on Microsoft Hyper-V. From this point on, I will use the terms “LUN” and “volume” interchangeably. A LUN can contain one or more VMs.
For SAN devices, the basic mechanism for creating a point-in-time copy of VM disk data is the LUN snapshot. SANs are able to create LUN-level snapshots of the data they are hosting. A LUN snapshot freezes the entire volume at the point it is taken, while read-write operations continue without halting to another area of the array.

Once a snapshot is taken, its data blocks are initially replicated to a sister SAN device in a DR site, thus having a standby copy of the virtual machine disk files in a ready-to-boot state. Subsequent replications of the LUN should only take changes between the last snapshot and current snapshot.

During a recovery operation, the replicated snapshot needs to be mounted correctly at the DR site and made available to the appropriate hosts. Once made visible to the hosts, the VMs hosted on the datastore can be booted up.

Sounds simple? Read on.

**Challenge 1: Hardware Dependency**

In order to replicate a SAN snapshot between two sites, there must be a second SAN at the DR site. Not only must this be a SAN from the same vendor, it often has to be the same model with the same firmware. Setting aside concerns around vendor lock-in, this is an expensive approach as reasonably specified SANs come at a premium as a result of their pivotal role in the modern datacenter. In addition to the duplicate SAN, the physical infrastructure is mostly duplicated too, with the same specification hardware being required for hosts and network infrastructure. One core objective of virtualization is to abstract workloads away from the hardware required to run them, thus creating hardware independence. Reintroducing hardware dependencies into virtual environments is not something most organizations desire and, where possible, it should be avoided.

So far, we have assumed one production and one DR site; additional costs should be expected for any additional, remote or satellite sites. The purchase of an additional SAN will be required at each geographical location. Where a site is too small to justify the cost of a SAN, there can be no DR for that site.

**Solution 1: Software Replication to Dissimilar Hardware**

Using a VM replication approach gives scope for removing hardware dependencies. By stepping up the replication operation to the hypervisor level rather than working with just the raw bits in storage, it’s possible to leverage the inherent capabilities of the relevant hypervisor, not only to snapshot individual VMs, but also to recreate entirely new copies (i.e., replica VMs) at the DR site. This means you can be much more granular in your approach and more selective about which VMs you replicate. Not only does this mean you can use a different specification of hardware at the DR site, but you can also introduce vendor independence. This is a great use case for repurposing old hardware during an upgrade.
As this approach centers around the use of built-in hypervisor features, wherever there is a hypervisor there is the ability to replicate. Therefore, bringing additional, remote and satellite sites under DR protection is a very simple, easy and cost-effective step.

**Challenge 2: Achieving Application Consistency**

Three levels of consistency can be achieved in any asynchronous replication operation:

- **Crash consistent** – This is analogous to yanking the cord out of the machine supporting the target application
- **File consistent** – The files undergoing file operations are completed and paused before a replication operation is initiated
- **Application consistent** – The files are put into a consistent state, and anything currently occurring in memory is also committed to disk.

Although not always possible, it makes sense to implement application consistency as this will result in the least possible data loss and the shortest possible restore time in the event of a disaster. File system consistency is the second most desirable technique, with crash consistency being used as a last resort.

SAN vendors are not inherently from a background of application awareness, so they tend to start at the crash consistency end of this spectrum of techniques. In order to achieve application consistency, SAN device vendors often insert agent software into each VM’s guest operating system, alongside some basic interaction with the hypervisor. While this approach may achieve the required level of consistency, it is complex, difficult to implement and requires careful orchestration of the pausing process among the application, the hypervisor and the SAN. In addition, there is overhead – resources (CPU, RAM), performance and management – in any approach that requires agents to be installed inside VMs.

Perhaps the greatest challenge is the lack of granularity with which SAN snapshotting occurs. As snapshots are processed at the LUN level, for example, if there are 30 VMs sitting on the same LUN, you have to put all 30 VMs into a consistent state “at the same time.” If you can imagine 30 VMs dumping whatever they have in memory to disk at the same time, you may not be surprised that this causes a massive spike in resource utilization, in some cases grinding production hardware to a standstill.
In order to avoid such issues and spread out the load, recommendations usually include the separation of VMs onto different LUNs based on the level of data consistency they require. While this may resolve the issue, what if you have already placed VMs on specific LUNs for performance reasons? Should you then subdivide LUNs into multiple LUNS to ensure that both application consistency and performance requirements are achieved? Increasing complexity can only lead to further challenges and problems. It seems like you are being forced to change your virtual environment in order to meet the limitations of the storage devices. Not a good place to be.

**Solution 2: Granular and Agentless Application Consistency**

By addressing this issue at the hypervisor level, you can be more granular in how and when these tasks are processed. You can invoke application consistency and hypervisor-level snapshots on a per-VM basis rather than in large, LUN-level groups. This means you can spread the activity over longer periods, avoiding flash points and optimizing the process to match the capabilities of your host and storage hardware. Some software solutions can do this without agents, therefore removing any inherent management and resource issues that may be encountered.

Such an approach would match itself to the virtual environment that is presented, rather than requiring reconfiguration of the environment to support the limitations of the SAN.

**Challenge 3: Bandwidth Constraints**

Most SANs are perfectly capable of processing large quantities of on-device data; however, moving that data off-device to a DR site is another matter. The fact that SAN technologies are typically not optimized for network transmission and that the smallest unit of management is a LUN snapshot means that it isn’t easy to create an effective process to move the data.

Initial replication of the source LUN snapshots can be painful, often requiring the first pass replication to be done while the two SANs are in the same physical location. The sister SAN is then physically transported to the DR site. Beyond the first run of the replication process, most SANs are able to process incremental
changes, so only the changes since the last replication cycle are sent across. Unfortunately, because of the propensity to work with large blocks of data, changes are often tracked at a large block size, sometimes as large as 16 MB. This means that if a small change of 10 KB occurs on one block and a change of 20 KB occurs on another block on the same LUN, a chunky 32 MB of data is replicated across the WAN for what are relatively small changes.

As bandwidth constraints are identified, most organizations will seek recommendations and workarounds. These come in the form of attempts to reduce the data on each LUN by identifying and removing data that doesn’t need to be replicated. Examples of this may be swap files or in-guest page files. If we split these out of the core VM folder onto different LUNs, they will not be replicated to the target site. Unfortunately, while achieving a reduction in data, this also breaks one of the core tenets of virtualization – encapsulation. By splitting down VMs and distributing VM files to separate pieces of storage, we lose portability and manageability. Reports of lost data and of administrators losing track of where their VM files are stored are not uncommon. For service providers, this is particularly important as they are managing VMs from many customers on the same infrastructure, and there are many real-life examples where customer data was lost because of complex configuration and restore procedures.

In addition to this restructuring of the virtual environment, further measures are often needed, and this has led to recommendations for implementing specialized WAN optimization devices between sites, resulting in significant costs and complexity.

**Solution 3: Bandwidth Optimization**

By dealing with the task of replication at the much more granular VM level, you can optimize the transfer of data and spread it over longer periods, rather than sending bursts of bulk LUN-level data. In addition, hypervisors tend to track changes to individual VM disks at much smaller block size with Changed Block Tracking (CBT). Combining CBT with replication software optimized for WAN links provides optimal use of available bandwidth. Of course, with standard network connections, there is an optimal block size to use over TCP/IP; this is dealt with in the replication software itself.

In addition to using optimal block sizes for transferring incremental data to the DR site, replication software tools may also include WAN optimization capability, where multiple optimized TCP streams are used to transfer data. This can create upwards of a ten-fold increase in speed without the cost of a specialized device.
Source VM data can certainly be of a significant size, so some kind of seeding – where initial data is physically transferred to the DR site – is beneficial, but shipping around an entire SAN is excessive. In a software solution, it is possible to use a deduplicated and compressed backup file containing the source VMs as the seed for the whole replication operation. This creates a much smaller amount of data, which is much easier to physically transport. Think of a USB drive in a car, rather than a rack unit in a truck.

**Challenge 4: Failover and Failback Management**

During a disaster, a process must be used to failover to the DR site, where the VMs there are booted, brought on to the network and configured to make them accessible by the application users. SAN providers generally do not provide in-depth hypervisor integration to enable VMs to be registered with the DR site virtual infrastructure and controlled during failover. This requires the use of expensive additional third-party products such as Site Recovery Manager (SRM) for VMware. For Hyper-V, stretched clusters or other third-party solutions are shoehorned in to provide some of this capability. Installing, configuring and using some of these additional products can be complex, expensive and difficult to get in place. While they may be capable of staged orchestration of VM booting, they add little value over and above registering, re-IPing (changing the IP address configuration) and network mapping the VMs on the replicated LUNs. It is not uncommon for these extra products to require several months for implementation and, once installed, they do require additional maintenance.

Implementing failback (i.e., returning operation to the production site) can be equally as arduous with full copies of the DR site data being replicated back to production, even if the production site was rescued fully or in part just minutes after failover occurred.

**Solution 4: Simple Management for Failover and Failback**

As there are no hardware dependencies to cater to, a simple software solution to implement and track failover and failback procedures can easily resolve many of these issues. Is it really necessary to pay significantly more for extensive, staged boot-up processes? Why not just click “failover” to boot-up your list of VMs.

Software replication solutions can provide all of the must-have features mentioned above such as re-IPing machines, network mapping VMs and controlling the process of failover. In addition, some solutions can also provide easy failback to the production site should it be recovered in a short period of time. During the failback process, only changes that have occurred after failover was initiated are transferred back to the production site. This can save significant bandwidth, time, money and headaches.

In terms of overall implementation, a software solution can be installed, configured and replicating to the target site in a little as 30 minutes. This presents a stark contrast to the multi-week (or perhaps multi-month) process of implementing SAN replication for DR.
**Challenge 5: Replicating Corruption**

As with synchronous replication, should the “disaster” come in the form of massive data corruption, virus infection or malware, asynchronous SAN replication may put you in much the same position, where the corruption has already transferred to the DR site. Most SAN replication operations result in a single, “most recent” copy of the source LUNs being replicated. If that most recent copy is the only thing to failover to, and it’s infected, it’s pretty much game over for the DR plan.

**Solution 5: Multiple Restore Points**

VM-level software replication can address this problem. By implementing additional historical restore points on a per-VM basis, depending on the replication schedule, you can cycle back through historical copies of VMs until you get to a point before the corruption was caused or the problem began. Again, this represents a simple methodology to significantly decrease business risk in a granular, cost-effective manner.

**Conclusions**

As you have seen from the previous discussion, implementing SAN replication to provide disaster recovery for virtual environments is a complex, expensive and high-risk strategy. In order to achieve basic results, you are often forced to fundamentally break some of the core tenets of virtualization. You are forced to redesign your virtual infrastructure to support the capabilities of the SAN, rather than the structuring your infrastructure in a way that best supports your business needs. Hardware independence, encapsulation and manageability must all be sacrificed to achieve even basic levels of performance and capability. In addition, many supplementary third-party products are required to increase performance and introduce features that are simply not provided by SAN devices. Add to this a lack of protection against replication of corrupt data and the risk profile becomes unviable.

A software-based replication solution that works at the VM level does, however, provide much more flexibility. Veeam® Backup & Replication™ is such a solution.

Veeam Backup & Replication provides both business continuity and disaster recovery capability in a single cohesive product. To facilitate disaster recovery strategies, Veeam Backup & Replication provides VM-level replication features that can overcome all five of these SAN replication challenges:

- **Hardware independence**: The ability to replicate VMs between dissimilar hardware. Different hosts, SANs and network configuration can be used in the DR site, therefore avoiding vendor lock-in and vastly reducing costs.

- **Granular application consistency**: Processing of application consistency on a per-VM basis, spreading the load of tasks to avoid resource-grinding flash points as well as providing a simplified management process.
• **Bandwidth optimization:** Granular per-VM processing at an optimized block size with built-in WAN optimization. This includes the ability to automatically deduplicate, compress, remove whitespace and swap files and in-guest page files from data streams before transmission. Add to this the ability to use multiple, optimized TCP streams for transmission and the result is that we can massively reduce costs while also increasing stability and throughput.

• **Fully managed failover and failback:** The ability to manage staged and controlled failover and failback processes. This includes advanced initial seeding and sending back only incremental changes during failback procedures. Additionally, this includes the ability to re-IP and re-map network settings during failover.

• **Replica restore points:** The ability to store multiple historical restore points on each replica VM, which in turn provides the opportunity to failover to a pre-corruption good copy of the relevant VM.

Veeam Backup & Replication can drastically simplify the whole process of off-site replication. The solution will work with your current virtual infrastructure and there is no requirement to restructure it in order to support the replication process. As a result of this simple and effective approach, what would take several months of implementation using SAN replication can literally be installed, configured and replicating in less than 30 minutes.

In addition to providing the above capabilities, Veeam Backup & Replication is also a full-featured backup product that provides:

• **Faster backups:** 77 times faster than traditional agent-based backup solutions.

• **Instant VM Recovery:** Restore “any” VM regardless of size in under three minutes.

• **Recovery verification:** Automated restore testing moves customers from testing an average of 1-2% of backups to 100%.

• **1-Click File Restore:** Files restored in a single click directly to their original location.

• **Universal Application Item Recovery (U-AIR®):** Restore “any” item from “any” application, with no requirement for agents or unsupported application internal interrogation.

Using Veeam Backup & Replication can remove the following significant costs:

• Duplicate SAN hardware

• Third-party management tools

• WAN optimization devices

• Backup solution

• Associated management, training and maintenance costs for the above.

If you would like to know more about Veeam Backup & Replication or would like to see it working in your environment, please register for a trial at [www.veeam.com](http://www.veeam.com)
About the Author

Hani El-Qasem is a System Engineer at Veeam Software. For 15 years he has been working with a broad range of organizations to provide management, data protection and disaster recovery solutions. These organizations include corporate, government, defence and intelligence agencies.

About Veeam Software

Veeam® Software develops innovative solutions for VMware backup, Hyper-V backup, and virtualization management. Veeam Backup & Replication™ is the #1 VM Backup solution. Veeam ONE™ is a single solution for real-time monitoring, resource optimization, documentation and management reporting for VMware and Hyper-V. Veeam extends deep VMware monitoring to Microsoft System Center with Veeam Management Pack™ (MP) and to HP Operations Manager with Veeam Smart Plug-In™ (SPI). Veeam also provides free virtualization tools. Learn more by visiting www.veeam.com.
Extending the lead in VM backup with these new capabilities:

- **Enterprise scalability:** new distributed architecture streamlines deployment and maintenance of remote office/branch office (ROBO) and large installations.
- **Advanced replication:** accelerates replication by 10x, streamlines failover, and provides real failback with delta sync.
- **Multi-hypervisor support:** brings Veeam’s award-winning data protection to Microsoft Hyper-V and lets you protect all your VMs—VMware and Hyper-V—from one console.
- **and more!**

To learn more, visit http://www.veeam.com/backup