

Ethernet SAN Storage Virtualization
EtherDrive® VSX3500

White Paper



December 2010



© 2010 Coraid, Inc.

Coraid, Inc.

The trademarks, logos, and service marks (collectively "Trademarks") appearing on the Coraid website are the property of Coraid and other parties. **ALL OF THE TRADEMARKS MENTIONED IN THIS DOCUMENT ARE THE PROPERTY OF THEIR RESPECTIVE OWNERS.** Nothing contained in this document should be construed as granting any license or right to use any Trademark without the prior written permission of the party that owns the Trademark. CORAID and EtherDrive are registered trademarks of Coraid. Coraid trademarks include RAIDShield and VirtualStorage.

Coraid, Incorporated
255 Shoreline Drive, Suite 650
Redwood City, California, 94065
United States of America

+1-706-548-7200
+1-877-548-7200

www.coraid.com

Table of Contents

Introduction.....	4
Storage Virtualization	4
VSX Logical Volume Management.....	4
Benefits of VSX Logical Volume Management	5
Mirroring Physical Volumes.....	6
Snapshot.....	7
Snapshot Use Case	7
Cloning Logical Volume	8
Clone Use Case	8
Remote Asynchronous Replication	8
High-Availability Configuration.....	9
Conclusion	10

Introduction

Enterprise computing continues to grow in complexity, driven by server virtualization and cloud computing environments. The networked storage needed to support these environments is growing in complexity too. Also, IT organizations are expected to manage these complex environments with the same or fewer resources. To keep up with these demands, IT administrators require the ability to allocate storage for any application based on the latest business needs, and to do so non-disruptively, quickly and easily. Moreover, they need to ensure the data is available 24 hours a day, 365 days a year. Coraid's VSX3500 (VSX) is a new class of storage appliance that provides advanced storage virtualization for an entire Ethernet SAN, allowing IT administrators to respond to complex storage needs quickly and reliably.

Storage Virtualization

Features like logical volume management, snapshots, and replication have become common in many enterprise class, monolithic, storage arrays; however, these features are constrained to operation within the chassis or control head of the "monolithic" storage box. In contrast, Coraid's VSX provides advanced virtualization functionality including, logical volume management (LVM), mirroring, snapshots, creating clones, and asynchronous replication, not just for a single array, but for all Coraid EtherDrive® storage on the SAN. VSX provides a more flexible and powerful approach to storage virtualization by introducing a SAN virtualization appliance that leverages the simplicity of Ethernet to create logical representation of resources not constrained by physical limitations. The SAN virtualization capabilities of the VSX increase flexibility and responsiveness to business requirements providing unique value for our customers.

VSX Logical Volume Management

Physical Volumes (PVs) are fixed size LUNs exported from Coraid EtherDrive storage arrays.

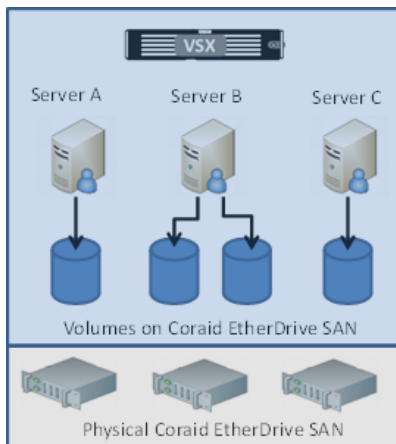


Figure 1 – VSX3500 and virtualized LUNs (LV)

As shown in Figure 1, Logical Volume Management is the process of grouping physical volumes (PVs) into storage pools, then one or more logical volumes (LVs) are created from each pool.

LVs are exported as storage virtual LUNs. Block storage devices that have been virtualized by LVM can be resized at any time rather than being constrained to the capacity of a single physical LUN.

The VSX administrator creates one or more storage pools, and adds physical LUNs from EtherDrive SR/SRX appliances to the pools. VSX pools are divided up into 4MByte storage extents. From these pools, extents are

allocated to create logical volumes (LVs) and exported to the SAN as VSX virtual LUNs. Virtual LUNs can be resized at any time by simply adding or subtracting extents. Servers read from and write to virtual LUNs, and the VSX takes care of forwarding server requests to the physical storage volumes who in turn respond directly to the initiating server.

The VSX logically sits in front of EtherDrive Storage and all LVM configuration metadata is stored with the user data extents in each pool. Figure 2 below depicts a virtualized Ethernet SAN with two VMware ESX servers and three EtherDrive storage arrays.

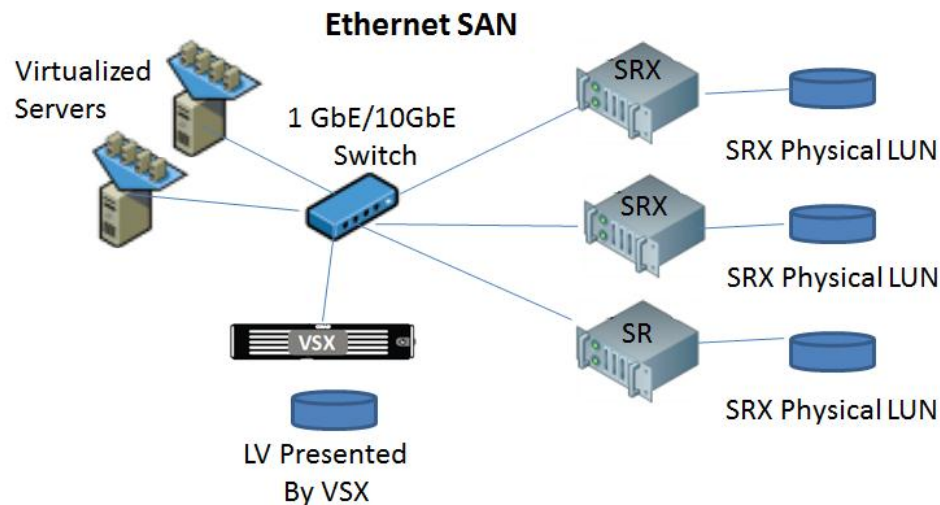


Figure 2 – An Ethernet SAN configuration with the VSX3500 and virtualized LUNs (LV).

A unique configuration feature of the VSX is it is not totally “in band” or inline, but is a hybrid virtualization solution designed to route read and write acknowledgement request from the initiator to the disk and back directly to the disk optimizing IO performance. The VSX’s use of logical block addressing translation technology avoids the pitfalls of traditional “in band” virtualization solutions of limited scalability and increased latency.

Benefits of VSX Logical Volume Management

One of the difficult decisions facing storage administrators is how to size LUN space for each server application and how much space to reserve for snapshots and clones. The need to estimate just how much space is likely to be needed for any one application makes the installation more complex than is necessary. Once the administrator has estimated (guessed) how much space is needed for an application then it is quite common for LUNs to fill up even if there is plenty of space in other LUNs or in other storage arrays on the SAN. With VSX logical volume management, virtual LUNs are dynamically resizable up to huge 500TB LUNs. Pools can easily be expanded by adding more physical LUNs to the pool from any available

storage appliance on the Ethernet SAN. This way if LUNs fill up, but there is unused space in other LUNs, it is possible to shrink a logical volume and reallocate that space. Another alternative would be to allocate minimal amounts of space for each logical volume and leave some of the disk unallocated. Then, when the LUNs start to fill up, they can be expanded as necessary.

The benefits of logical volume management are also advantageous for large cloud storage environments with many disk drives. Managing a large disk farm is a time-consuming job, made particularly complex if the system contains many disks of different sizes. Balancing the (often conflicting) storage requirements of various users can be a nightmare. User groups can be allocated to separate pools and logical volumes from these pools can be grown as required. It is possible for the system administrator to "hold back" disk storage until it is required. It can then be added to the pool (user group) that has the most pressing need.

When new drives (or complete disk arrays) are added to the Ethernet SAN, it is no longer necessary to move users files around to make the best use of the new storage; simply add the new disk into an existing pool and extend the logical volumes as necessary. It is also easy to take old drives out of service by moving the data from them onto newer drives - this can be done online, without disrupting user service.

Mirroring Physical Volumes

Mirroring is extremely important for High Availability (HA) environments. For applications that require high availability, LUNs (physical volumes) from EtherDrive storage appliances can be synchronously mirrored by the VSX to protect against SAN network and EtherDrive appliance failures. Figure 3 shows a synchronized mirror between two EtherDrive storage arrays.

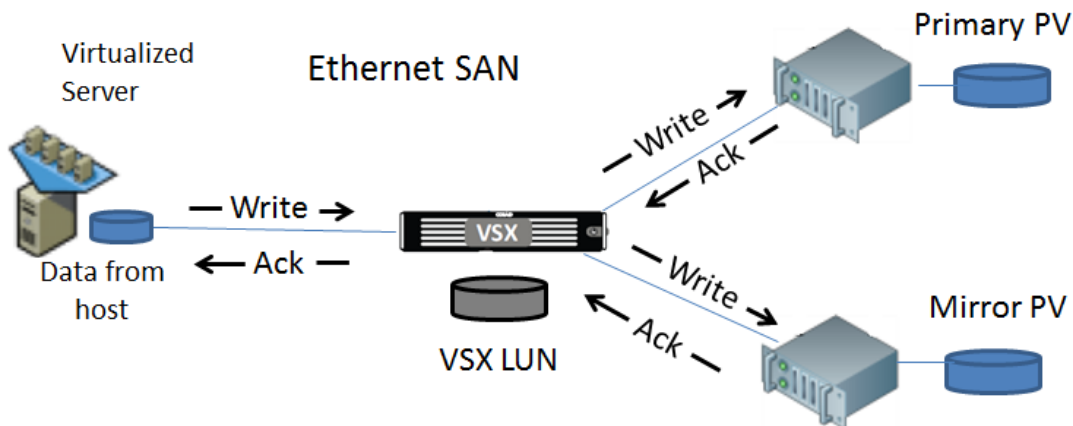


Figure 3 – A mirrored configuration with VSX3500 and two EtherDrive storage arrays.

Mirroring creates a simultaneous, exact duplicate of a PV to a designated EtherDrive LUN target. The original PV that is part of the pool is called the primary, and the second part of the mirrored relationship is referred to as the mirror.

VSX ensures synchronization by executing writes to both the primary and the mirror before responding to the Ethernet SAN initiator (server). Ethernet SAN requests for reads are always serviced from the primary. Figure 3 shows a synchronized mirror between two EtherDrive storage arrays.

Snapshot

VSX allows administrators to take snapshots of a logical volume to create fixed points-in-time views of storage volumes. It does this by creating a new COW (copy-on-write) logical volume with pointers to the original logical volume. So long as your data remains unchanged, the snapshot merely contains its metadata pointers and not the data itself. Whenever the original LV is modified, the changed extents are copied to snapshot LV extents before the original LV's extents are modified. As more extents in the original LV are changed, the storage space of the snapshot LV grows. As shown in Figure 4, a COW pointer matrix is used by the VSX to optimize snapshot storage space. This minimizes storage space when keeping multiple snapshots of the same LV.

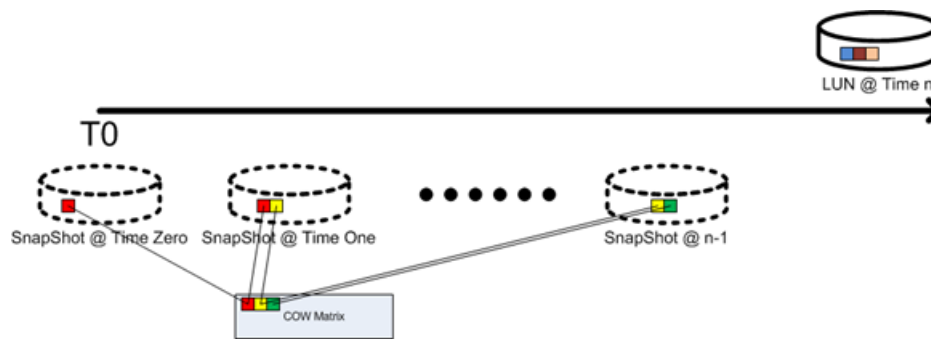


Figure 4 – Snapshots offer a point-in-time copy of a logical volume.

As with any SAN block storage device (FC or iSCSI), LUN snapshots should be synchronized with a momentary file system suspension to maintain snapshot file system consistency.

Snapshot Use Case

Snapshots allow system administrators to respond to and recover from storage loss and data corruption situations easier and faster than from other storage methods. A common scenario is a user has deleted a file and needs to recover the file. Instead of having to restore from a tape

backup that is a time consuming and resource intensive activity the administrator can recover the file quickly and easily from a snapshot of the data.

Snapshot also allows for data to be copied without taking the LUN offline. This means applications that are accessing the LUN can continue to access the LUN so users can continue to work. Another scenario is after installing a software upgrade a problem is discovered with the upgrade and the administrator wants to roll back to before the upgrade. Snapshot allows administrators to roll back before the upgrade to get the systems up and running to a prior state.

Cloning Logical Volume

The VSX supports creating a clone of a snapshot. A clone is a read/write logical volume similar to a snapshot. Creating the clone requires no new storage space, but after the clone is created, subsequent writes to the clone and the COW writes uses available extents from the pool and the clone size grows. A clone consumes less storage space and is provisioned instantaneously compared to the duplicate storage space and time required to make a copy of a physical volume.

Clone Use Case

Server or desktop virtualization deployments usually mean having to deploy the same disk image over and over again. With the clone feature a copy of the virtual machine (VM) image can be created and used to deploy to hundreds VMs faster than creating a separate copy for each VM. Another use case for using clones is with extensive testing of a database. In this scenario, IT administrators can use cloning to convert a snapshot to a point in time copy of the database and then use this clone of the database for different test scenarios.

Remote Asynchronous Replication

Remote replication is extremely important for disaster recovery as part of a business continuity plan. The VSX asynchronous replication feature –at no additional cost- enables LUNs to be shadow copied to remote location LUNs.

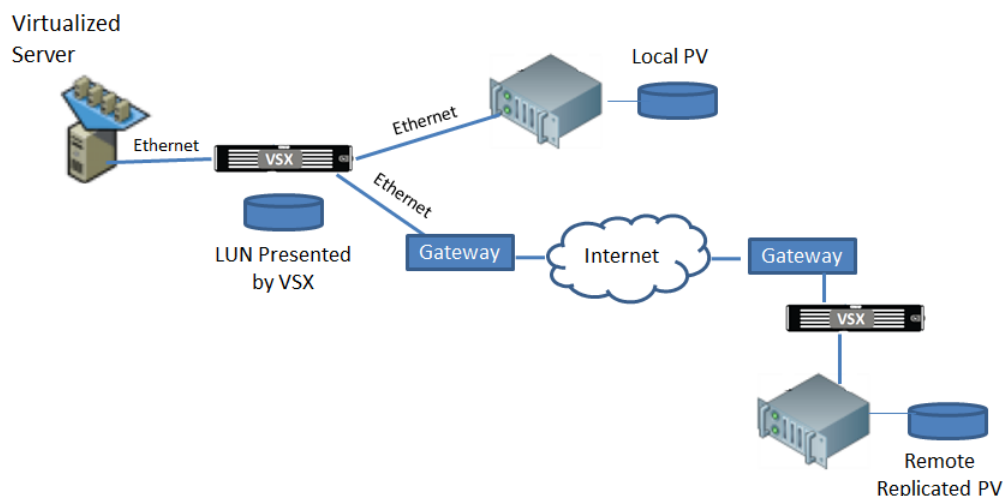


Figure 5 – Remote replication configuration for disaster recovery.

VSX allows users to establish a shadow relationship between LUNs. As shown in Figure 5, after a LUN has been created with the VSX a shadow LUN can be assigned to it. Shadow LUNs contain copies of all the used extents of the master LUN and all the snapshot information as well. Each time a snapshot is taken of the master LUN, the VSX automatically copies all updates since the previous snapshot to the shadow LUN. The shadow copy update speed is dependent upon the bandwidth of the interconnection link between the VSX and the remote shadow LUN and the amount of data changes to be updated. Therefore the shadow LUN may be asynchronous in time with the master LUN.

If the shadow LUN is located at a remote site like in Figure 6, an Ethernet SAN Gateway is required to tunnel the AoE protocol over an IP network link (WAN link). At the remote location, the shadow LUN is a virtual LUN created by a different VSX at the remote site. Shadow LUNs from one VSX can be shadowed by another VSX LUN to cascade data replication to multiple remote sites.

Remote asynchronous LUN replication maintains a complete copy data on a shadow LUN. Like having a second car if the first one gets damaged in a wreck you can switch to the backup car, if the source data becomes unavailable for some reason, a nearly current copy (since the last snapshot of data was transmitted) is available at the remote site to resume operations. This way, companies are protected from local and regional system failures or disasters like fire and floods.

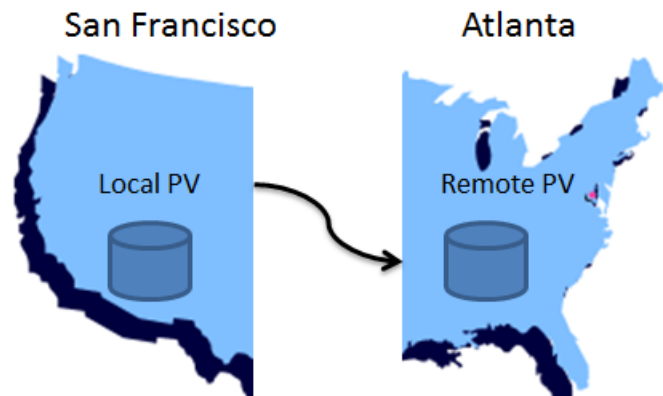


Figure 6 – Remote replication across the United States.

High-Availability Configuration

To ensure uninterrupted service in the event of system failure, you can employ a pair of VSX appliances in a high-availability (HA) configuration. See Figure 7 below for an example of a VSX3500 high availability/failover configuration with SR/SRX mirroring.

In a typical HA configuration like the one depicted below, one VSX functions as the primary, active appliance and the other VSX functions as the secondary, inactive appliance. The redundant secondary VSX defends the primary VSX against failure, serving the SAN storage if the connection to the primary VSX fails. If the secondary VSX detects a failed primary, the secondary VSX automatically becomes the active VSX. Coraid EtherDrive storage appliances, SR/SRX LUNs, VSX pools, PVs, LVs and VSX LUNs are then available from the new active VSX.

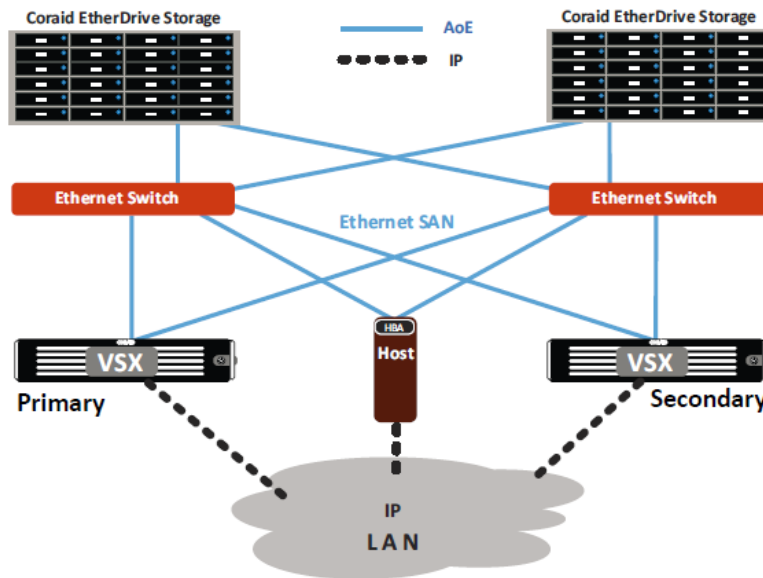


Figure 7 – Remote replication configuration for disaster recovery.

The VSX3500 SAN virtualization appliance incorporates HA to provide data availability above and beyond what hardware RAID can provide. While hardware RAID protects against disk drive failures, VSX3500 protect against a much broader array of failures including complete storage system failures. The ability to provide continuous availability in the event of a disk, controller, storage node, power, or network failure makes Coraid EtherDrive Ethernet SAN appliances an ideal match to the needs of HA environments.

Conclusion

This paper has described the principles underlying the design and operation of the VSX Ethernet SAN storage virtualization appliance. Traditional storage networking, shared storage provisioning and storage management methods are inefficient in today’s complex storage networks. Storage virtualization can take organizations to the next level eliminating inefficiencies and helping reduce cost.

With the advent of the EtherDrive VSX3500 Storage Virtualization Appliance, Ethernet SAN now can have features and benefits only found in more expensive Fibre Channel environments.