Lenovo SAN Manager – Rapid RAID Rebuilds and Performance Volume LUNs

Lenovo ThinkSystem DS2200, DS4200, DS6200

June 2017

David Vestal, WW Product Marketing

Lenovo.com/systems

Table of Contents

Introduction

The Lenovo ThinkSystem DS2200, DS4200, and DS6200 SAN products are differentiated offerings with Lenovo's Storage portfolio, providing both hardware and software feature advantages. These SAN products are designed for simplicity, speed, scalability and availability and are the perfect blend of a high performing SAN product with a very easy-to-use interface.

With simplicity at its core, the DS2200, DS4200, and DS6200 share the same software features and interface. The Lenovo SAN Manager, the operating software to manage the operational and features of the SAN products, is an easy-to-use GUI that makes complex administrative simple by automating configurations and operations. The Lenovo SAN Manager offers enterprise-class features like: Storage Tiering & Caching, Thin Provisioning, Rapid RAID Build, Snapshots, and more.

In this paper, you will learn:

- Learn how to create new performance Volumes with Lenovo SAN Manager using the Storage Management Utility or CLI.
- Understand how Virtual Disk Groups can rebuild parity quickly

Large LUNs

The Lenovo SAN Manager uses the concept of Virtual Disk Pools – which are comprised of multiple Virtual Disk Groups. This now decouples the LUN from being restricted to a single vdisk. A Volume's data on a given LUN can now expand across all disk drives in a pool. When capacity is added to a system, the user is also getting a performance benefit of the additional spindles – hence the term Performance Volume.

Today's data architectures benefit from large LUNs in several ways; general examples:

- General "Best Practice" for VMFS datastores is to use large LUNs for use by multiple VMs. Large LUNs give VMware vSphere users the most flexibility by not requiring storage administrators to provision new storage every time a new VM is created. Large LUNs also minimize rescans for new LUNs on the ESX server, which can impact VMFS I/O.
- Operating Systems now support up to 16TB file systems.
- As enterprise workloads for virtual environments grow in size and in performance demands, virtual hard disk (VHD) formats need to accommodate them. Hyper-V in Windows Server 2012 introduces a new version of the VHD format called VHDX which supports virtual disks up to 64TB in size.

The Lenovo SAN Manager supports large, flexible Volumes with sizes up to 128TB spanning up to 256 HDDs (architectural limit) providing seamless capacity expansion. Data automatically reflows to new drives and IOPs scales with additional storage.

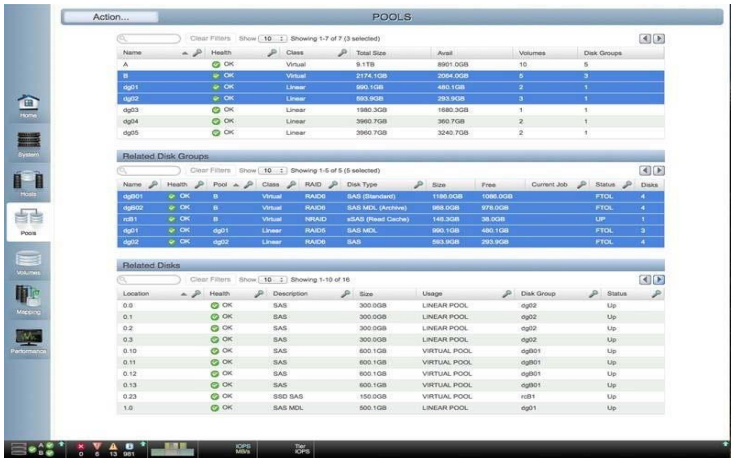

Figure 1- Creating Large Pools and Volumes

Thin Provisioning and Modern File Systems

Many modern operating systems become automatically aware of an array's ability to support thin provisioning. This is true of Windows 2012 / 2012 R2 and Linux with newer 6.x kernels. When using a non-thin provisioned storage array the deletion of a file does nothing more than rearrange the file system pointers.

If the array supports thin provisioning, as with Lenovo SAN Manager, thin-aware operating systems will notify the storage array to free the storage that had been allocated to the file that was deleted. This happens with the SCSI UNMAP command. This will allow for the re-use of the capacity that had been allocated for the deleted file.

Windows 2012 / 2012 R2 also issues notifications when the allocations reach a high percentage of the "advertised" capacity. This, along with the capacity usage notifications, signals the user that it may be time to add more physical storage.

Rapid RAID Rebuilds

Rebuild of conventional (Linear) storage requires that the entire vdisk is rebuilt prior to returning the vdisk (and thus the Volumes on the vdisk) to a fault tolerant state. The rebuild operates at the vdisk level and has no knowledge of Volumes or customer data contained therein.

A customer could have a single small Volume on a large vdisk. During the rebuild, even after it has completed the stripes where the Volume resides, the vdisk remains in a degraded or critical state. During this period, another drive failure would take the vdisk from degraded to critical, or from critical to down. In the latter case, the customer may have lost data. In the prior case, the customer’s data becomes at risk once again, even though that area of the vdisk had been rebuilt.

In another example, the customer may have created a Volume or Volumes consuming the entire vdisk space, but may have actually allocated very little. Since linear storage has no means to know what stripes contain customer data and which do not, that small amount of data will remain at risk until the entire vdisk has been rebuilt. Also, performance suffers during the entire rebuild.

Another exposure is encountering a media error. If the array is rebuilding two drives and encounters a media error on a third, the rebuild will fail. This leaves the vdisk in a critical state, exposed, and with no way to rebuild the vdisk. It is especially unfortunate if this media error resides in unallocated space. The outcome is the same, linear storage must assume all stripes contain customer data.

Lenovo SAN Manager has a set of technologies to improve these conditions. It enables Virtual Disk Pools which track how much data has been allocated for each Volume and where it resides in the Pool, at a granularity of 4 MB. The Virtual Pool algorithm decides where to allocate each new 4 MB page.

With this knowledge, the system is able to rebuild only those stripes containing allocated pages when a RAID rebuild is required. Once these stripes are rebuilt the Virtual Disk Group (and thus all the Volumes in the Pool) are returned to a Fault Tolerant State. This reduces the window of exposure and increases the MTTDL (Mean Time To Data Loss).

Advantages of Rapid RAID Rebuild

- LUNs are spread across many RAID sets. So one RAID set rebuilding only affects a fraction of all disk I/Os.
- Less rebuild work to do – so affected disks can return to full performance quickly.
- Volumes become fault tolerant more quickly – user data is restored to full protection more quickly.

- The rate of a RAID rebuild is directly proportional to the amount of unused disk space in a Virtual Disk Pool.