



Lenovo SAN Manager

Intelligent Real-Time Tiering & Read Cache

Gregg Shupe, Lenovo Enterprise Product Group

February 2013; Revised April 2015

©2015 Lenovo. All rights reserved.

Table of Contents

Introduction 3

Automated Sub-LUN Tiering 3

LUN-level tiering is inflexible and inefficient 4

SSD Read Cache..... 6

Figure 3- Add Disk Group Dialog..... 7

Key CLI Commands for Intelligent Real-time Tiering and Read Cache..... 7

Summary 8

Introduction

The Lenovo Storage S2200 and S3200 SAN products are differentiated offerings with Lenovo's Storage portfolio, providing both hardware and software feature advantages. These SAN products designed for simplicity, speed, and scalability and availability. The perfect blend of a high performing SAN product with a very easy-to-use interface.

With simplicity at their core, the S2200 and S3200 share the same software features and interface. The Lenovo SAN Manager, the operating software to manage the operational tasks and features of the SAN products, is an easy-to-use GUI that makes complex management 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:

- How to combine SSDs and HDDs to accelerate your application performance with the new automated Sub-LUN Tiering and Sub-LUN SSD Read Cache capabilities
- How to configure Lenovo SAN Manager for both Real-Time Tiering and Read Cache using the Storage Management Utility or CLI
- Key differences between Real-Time Tiering and SSD Read Cache and how they optimize performance

Automated Sub-LUN Tiering

Disk tiers are comprised of aggregating 1 or more Disk Groups (RAID groups) that exhibit similar characteristics in speed and density. There are 3 tiers:

1. A Performance tier with SSDs
2. A Standard tier with Enterprise SAS HDDs
3. An Archive tier with Nearline SAS HDDs, dense disks with rotational speed of 7.2K

Prior to Automated Sub-LUN Tiering, LUN-level tiers are manually created and managed by using dedicated VDisks and volumes. Applications requiring the highest performance are given volumes on top of the SSD-based VDisks while those requiring the lower performance, but greater capacity are typically dedicated to a VDisk comprised of Enterprise SAS or Nearline SAS HDDs.

LUN-level tiering is inflexible and inefficient

- Manual tiers act as static repositories, trapping data in one tier during dynamic workload changes
 - Cold data trapped in the performance tier limits the amount high-performing disks that can be brought to bear to a dynamic workload
 - Hot data trapped in the standard or archive tiers will underperform when needed the most; the controller cache usually isn't large enough to deal with this issue
- Migration between tiers requires external migration software or a cumbersome backup/restore cycle.
- Movement at the LUN-level is inefficient; the entire LUN must be migrated when typically only portions of the LUN are being referenced by the applications

Lenovo SAN Manager alleviates all of these issues because it seamlessly migrates data between tiers on-demand. Dynamic data workloads are based upon algorithms that reveal “hot” data – those which are consistently accessed with dynamic reads/writes. This feature is called Automated Tiering.

Automated Tiering is accomplished via a concept called “paging”. User volumes are logically broken down into small 4MB chunks called pages. Pages are ranked based upon a sophisticated algorithm. The page rank is used to very efficiently select good pages to move between tiers. The result is that pages can be migrated between tiers automatically such that I/O's are optimized in real-time. Figure 1 captures this concept. Note that 3 pages have been moved from the Enterprise SAS Tier to the SSD Tier, 1 page from the Nearline SAS Tier to the SSD Tier, and a single page has been moved down from the SSD Tier to the SAS Tier.

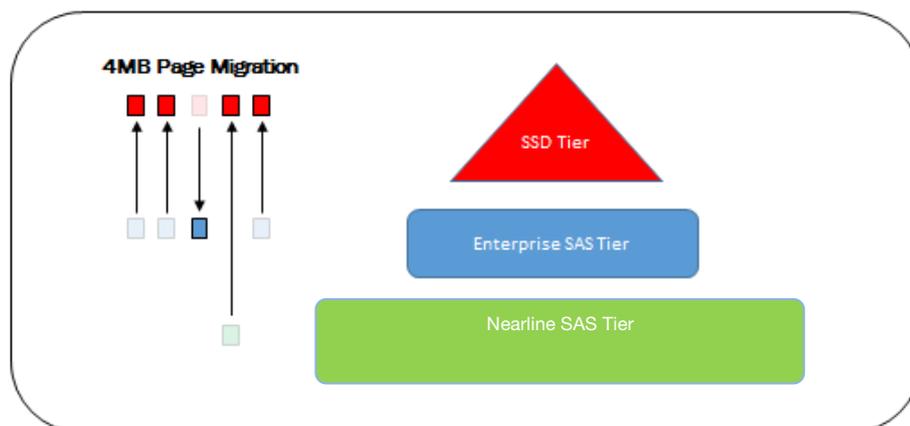


Figure 1 – Real-Time Tiering

In contrast to data movement at the LUN-level, the Sub-LUN nature of paging provides highly efficient data movement, a minimum of CPU and memory resources are needed to support data movement between tiers; therefore, movement can happen in real-time rather than in offline batch movements.

Key concepts: The tiering algorithm runs every 5 seconds. The algorithm ranks, scans, and migrates pages.

1. Pages are ranked according to access patterns
2. A scan looks for highly-ranked pages
3. Highly ranked pages are migrated up/down
 - a) Pages are only migrated down if room is needed for a highly-ranked page
 - b) Only 80MB migrated from any one 5-second interval to avoid degrading system throughput

One can monitor the migration between tiers from the Storage Management Console. Notice the IOPS status bar at the bottom Figure 2. This screen shot was captured 5 minutes after the start of a 100% random 100% write workload on a 10GB volume. The red portion of the status bar reflects those IOPS being served from the SSD tier, and the green portion of the status bar reflects those IOPS being served from the standard HDD tier.

Notice that only a small portion of the IOPS are coming from the SSD tier.

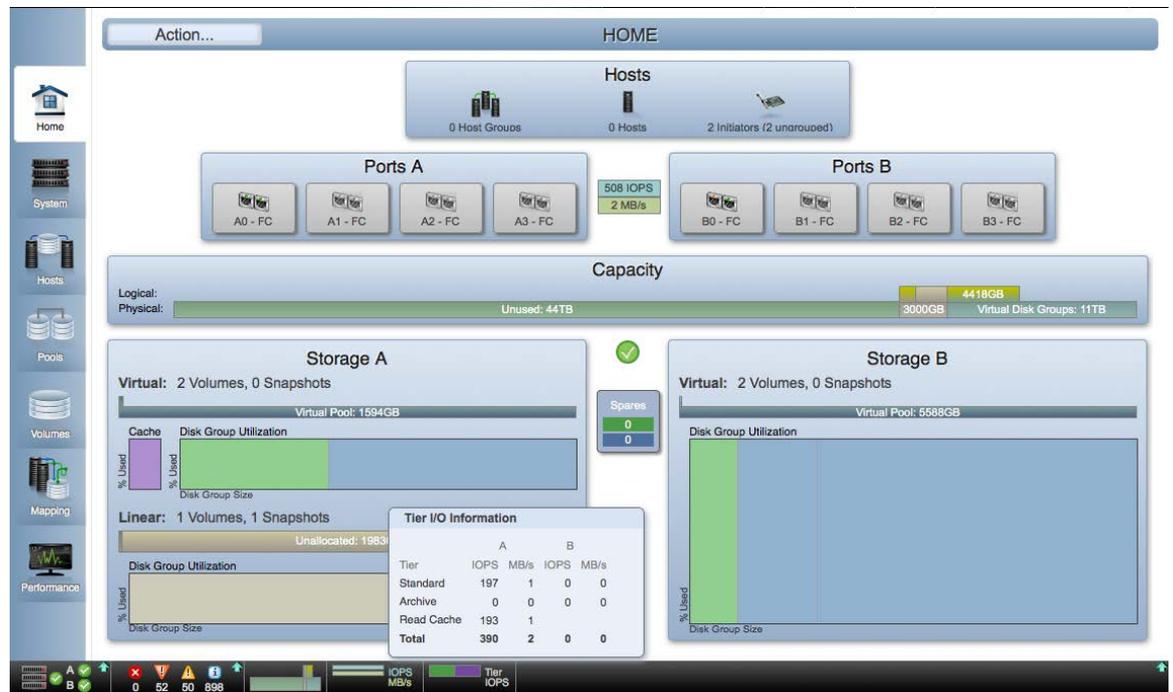


Figure 2 – Percent of IO from the Two Tiers

Volumes must be created from virtual storage pools to benefit from automated tiering. Volumes created from Linear Storage Pools (previously called VDisks) cannot take advantage of the benefits of paging and Sub-LUN tiering.

Automated SSD Tiering eliminates rigid hardware silos, cumbersome manual migration between tiers and can improve \$/IOPs metrics.

SSD Read Cache

A Read Cache is a special-purpose Disk Group dedicated to hold volatile copies of data within one or more SSDs. An SSD Read Cache is a new concept in the Lenovo SAN Manager.

Some advantages of a Read Cache are:

- Improved performance in read-centric workloads where maximum IOPS is crucial
- Requires fewer SSDs when compared to a dedicated SSD tier
 - A single SSD disk can be used as a Read Cache for a storage pool; whereas, an SSD tier requires multiple disk to provide fault tolerance
- Takes advantage of tiering-like algorithms so that only active pages are copied into the Read Cache
 - Active pages can be from any volume within the pool

The Read Cache becomes active when the controller cache is saturated. The controller cache software gives “hints” to the system software about which pages it would like to keep, but can’t. The system then copies these pages into the Read Cache. Reads are served from the SSD, and writes are written to both SSD and HDD.

Key concept: The read cache keeps copies of data that already exist in the hard disk tier; therefore, it doesn't require more SSD disks to be fault-tolerant. The SSD disk(s) provide the performance and the HDD disks provide the fault tolerance. This is a cost-effective method of providing performance.

A Read Cache is created using the Add Disk Group dialog box. From the SMU:

- Click on the “Pools” icon in the navigation menu to the left of the SMU
- Click the “Action” menu, and select the “Add Disk Group” menu item



Figure 3 – Add Disk Group

Figure 3- Add Disk Group Dialog

An example can be seen in figure 3. Notice that two SSDs are selected to act as the Read Cache for storage Pool A. Up to 2 SSDs may be dedicated to the Read Cache Disk Group; however, only one Disk Group may serve as a Read Cache for a pool.

Once active, the Read Cache for a pool can contain pages from any volume within the pool that is experiencing a read-intensive workload when the controller cache is exhausted.

Key CLI Commands for Tiering and Read Cache

The CLI may be used to create disks groups for Automated Tiers and Read Cache. Figures 4-6 demonstrate the creation of a Read Cache, the viewing of Disk Groups, and the deletion of a Disk Group respectively. The examples highlight a Read Cache, but the commands may be used for any type of tier.

Figure 4 - Creating a Disk Group

```
# add Disk-Group type read-cache pool a disks 1.21-22
```

Figure 5 - Viewing Disk Groups

```
# show Disk-Groups
```

Figure 6 - Removing a Disk Group

```
# remove Disk-Groups MyReadCache
```

Summary

The power of Virtual Disk Groups and paging introduces great flexibility around the use of disk resources. Intelligent Real-time Tiering and SSD Read Cache are two high-value features built using this malleable architecture. They provide efficient, self-managed use of disk resources with little maintenance.



LENOVO PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. This information could include technical inaccuracies or typographical errors. Changes may be made to the information herein; these changes will be incorporated in new editions of the publication. Lenovo may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any performance data contained herein was determined in a controlled environment; therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems, and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Any references in this publication to non-Lenovo Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this Lenovo product, and use of those Web sites is at your own risk.

The following terms are trademarks of Lenovo in the United States, other countries, or both: Lenovo