Storage Resource Management

Introduction

The scope of Storage Industry, which is embracing all the new technological advancements, is getting wider and deeper everyday. With data growths per year averaging 50% to 100%, organizations are not only faced with rising hardware costs but also with the increased cost of managing the storage. According to ESG, the cost of storage management and administration is three times the cost of initial Hardware acquisition. SRM today is not just limited to device and quota management. All aspects of NAS and SAN management now come under the purview of SRM. The basic responsibilities of SRM are:

1. Discovery of Storage Nodes
2. Effective Storage provisioning and capacity management
3. File and Array level utilization
4. Performance monitoring and characterization
5. Backup and Recovery management
6. Workflow management
7. ILM
8. CSM
9. Intuitive Presentation and reporting mechanisms.

StorTrends -Data Collection, Analysis and Trending Module

This module lays the foundation of StorTrends SRM. The design is centered on a kernel level and an application level component that are glued into a very flexible and extensible framework. This makes it very easy to extend SRM to cover bigger and wider range of attributes without having to redesign from scratch. The key kernel component is the data collection agent that interfaces with other modules to collect instantaneous data about all the monitored resources. The other component, which is an application layer module, is called the 'data container '. This layer provides an interface for storing collected data in a structure that spans time to provide a context for trending data. Recent and instantaneous data along with statistical data going back into months and years are maintained.

Storage provisioning and capacity management

Capacity provisioning is probably one of the biggest challenges any storage administrator faces today. This challenge begins from Day 1, when the volumes are to be appropriately created and provisioned. iTX with its Thin-provisioning architecture takes this headache out. Once the initial hurdle is crossed, the next big and persistent problem an storage administrator faces is to take an inventory of what is present today and predicting how much of growth is anticipated over the next few months, year or even beyond. Just present moment inventory, which most of the storage stacks provide is just not enough to help out. Capacity planning is a very inexact science and it is very time consuming to administer without appropriate tools. StorTrends, with its statistically collected and maintained data along with extrapolation algorithms provide all that is required to predict capacity growth and enable strategic capacity planning to a reasonable degree of accuracy.

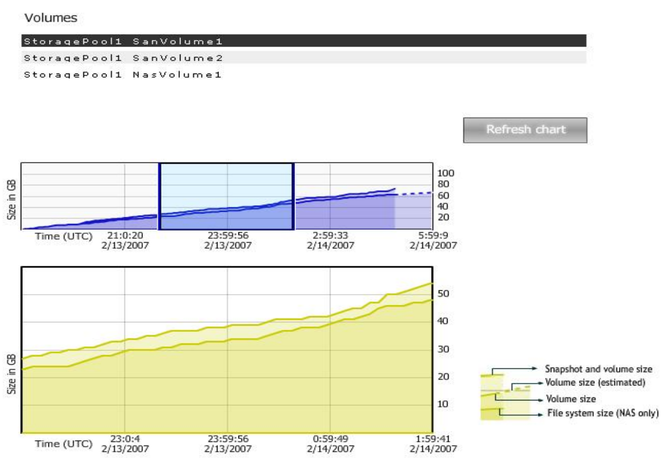

Figure 1: Storage provisioning & capacity management, showing long-term and short-term trends

Performance trending

Performance data at various levels of the I/O stack is collected and presented both for immediate analysis and statistical review. A few items worth special mention here are:

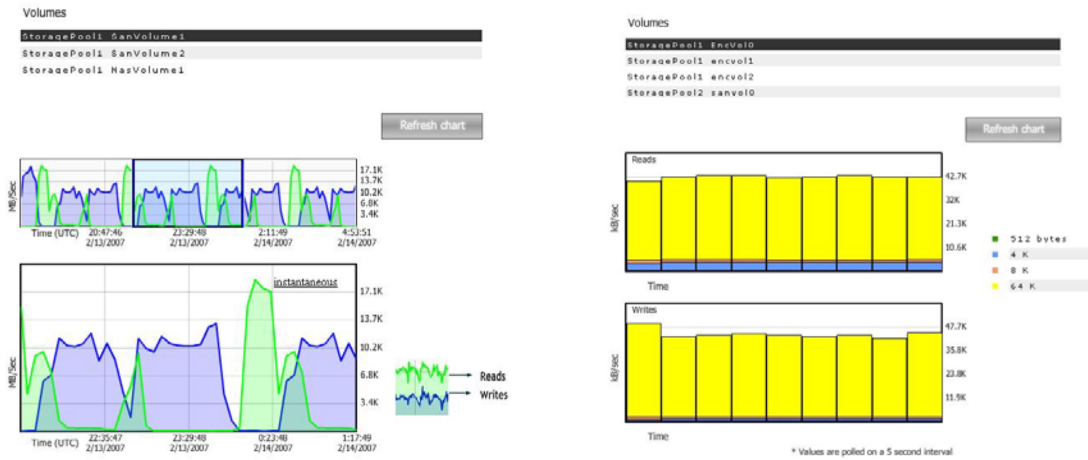

Figure 2: Performance monitoring – long term and short term trends

- Read & write performances of the volumes are also collected.
- In depth analysis of the instantaneous read and writes performance values. It gives the user information about the contributions of different I/O sizes towards the total I/O activity going on at the moment. The I/O sizes chosen are 512 byte, 4K, 8K and 64K reads or writes.
- I/O latency monitoring and analysis. Today with so many SATA drives populated into a single 2U or 3U enclosure, vibrational stability is very important. SATA drives with their dense track densities are very intolerant to vibrations and result in long recovery cycles and ultimately may be 'FAILED' by the RAID controller. Tracking I/O latencies is a very key attribute to detecting any such looming failures.
- H/W health monitoring and tracking S.M.A.R.T. indications.
  Various key parameters like enclosure temperature, Power supply, Fans etc. are continuously monitored

CPU Utilization and Memory Utilization

Figure 3: CPU and Memory Utilization

Workflow Management

This is probably the least understood yet most important aspect of a storage stack today. This forms the basis of background task scheduling and thereby ensuring the system maintains good I/O performance. System and I/O activities are continuously monitored and recorded. This information comes in very handy to pick up a period of relaxed activity to schedule background tasks like ILM and capacity expansion balancing.

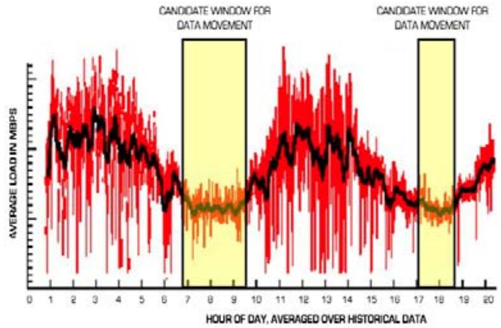

Figure 4: Workflow analysis

Backup and Restoration management

This falls under two categories:

First, to manage space effectively when numerous snapshots are created for backup and recovery. StorTrends with Tiered Storage and three level space mapping ensures that Snapshots are maintained in the most capacity-optimized manner.

Next, to ensure that the so called 'recovery images' are indeed consistent and good. Today, recovery management is more important than backup. Statistics show that a good fraction of recovery fail for various reasons. StortTrends offer various features to ensure the consistency and validity of 'recovery images'. At the NAS level, Snapshots are mounted and checked for correctness using 'fsck'. At SAN level, external agents perform scheduled 'chkdsk' chores on snapshots to assure correctness of snapshots.

In replication environments, background tasks are scheduled to check for the integrity of snapshots by creating and comparing CRCC checksums of images. This not method does not put any burden on the communication links.

Information Lifecycle Management

StorTrends provide excellent ILM features by tracking block level usage attributes. This is a vast subject and is outside the scope of this document.

Container Space Management

StorTrends provide CSM i.e. re-striping data across disks to increase performance. This is also a vast subject and is outside the scope of this document.

Presentation and Reporting Module

StorTrends offer a very elegant management interface for the various SRM chores.

StorTrends Advantage

- **Built around an flexible and extendable framework**
- **Storage unit discovery**
- **Capacity trending and analysis tools**
- **Statistical and Instantaneous performance monitoring**
- **Backup and Recovery Management**
- **WorkFlow Management**
- **Tiered Storage Management**
- **Presentation and reporting tools**

Information Lifecycle Management &

Container Space Management

Introduction

Data, which is the lifeblood of any storage system, is now experiencing unprecedented growth. This has become the dominant-driving factor for the storage industry to bring forth a paradigm shift in the way data is handled. A NAS or SAN storage unit, which is the housekeeper of data, can now no longer afford to treat data as an opaque entity. Data thus has matured from being looked just as a binary stream of ones and zeros to an intelligent entity with values assigned. This has led to birth of new disciplines in storage industry called Information Life Cycle Management (ILM) and Container Space Management (CSM). It is a key aspect of Storage Resource Management (SRM), where a 'context' is assigned to the data based on some value that the block of data possesses. This context could be based on the pattern of data usage and/or on the content of the data. Data classification is the discipline where data is discovered (eDiscovery), analyzed and classified based on some policies. There are various products in the market today that do this job online with the incoming stream of data or offline with data-at-rest. Though the need is felt, the technology of data classification is still very fluid, since the perceived value of data may change over time and can not be definitive.

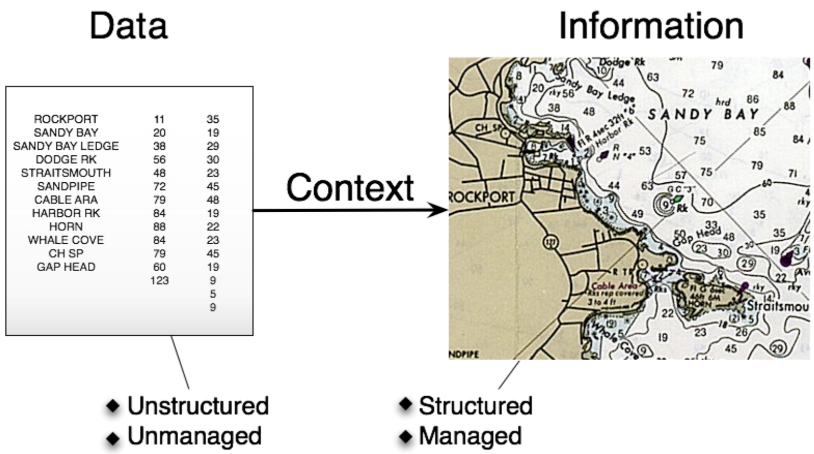

Figure 1: Adding structure and context to unmanaged data makes it information

Parallel to this the Disk Drive technology is also undergoing enhancements. SATA drives have proliferated the market today and will be there to stay for quite some time. They are very economical, but lack the performance and robustness called for by true enterprise class storage infrastructure. SAS drives are penetrating the storage arena with its enterprise class tag. SAS and SATA drives can coexist in the same SAN network ushering in the generation of affordable tiered-storage. To effectively utilize these storage devices that carry different price tags and performance characteristics, an efficient ILM is now almost mandatory. Scaling within the tiers is attained through CSM, which does the re-striping across disks.

According to researches conducted by the TheInfoPro Group, the interest in ILM over the last 18 months has heated up. The TIP heat index for ILM, which was 4 (in order of priority) in Spring 2005, has moved to 1 by Spring of 2006.

Basic Concepts

Logical Drive – A group of disks that form a RAID array (Redundant Array of Inexpensive Disks)

Container – A group of logical drives put together form a container.

Tier – A group of logical drives belonging to the same container but having same costs.

The cost is calculated based on RAID level, type of drives etc. A container may contain multiple tiers, and a tier may contain multiple LDs. The following diagram gives a pictorial description about the relationship between a Logical Drive, a Container and a Tier.

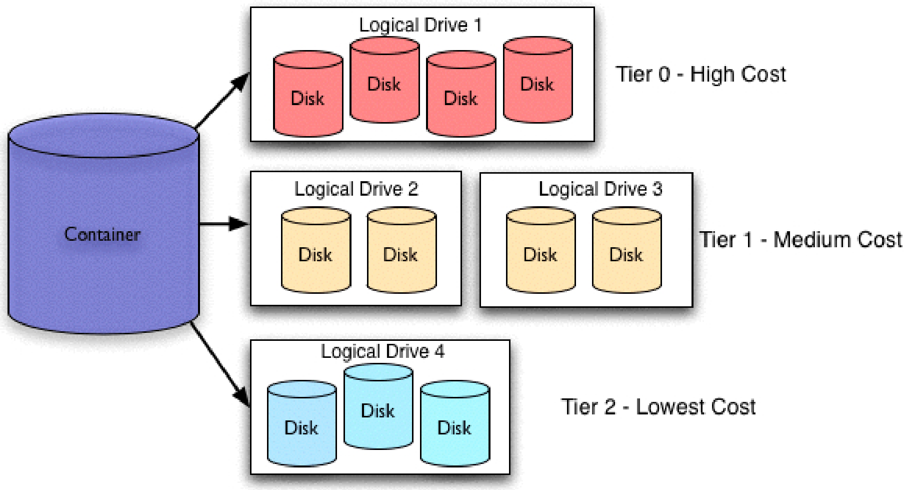

Figure 2: Relationship between Logical Drive, Container & Tier

StorTrends Tiered Storage and ILM (Scaling across tiers)

With StorTrends, even at block level, we handle 'information' and not just 'data'. Data when it is assigned a right context or intelligence becomes 'information'. The first challenge is to effectively segment out the data and assign appropriate vale. Next comes the policy aspect of ILM, where these chunks of 'information' are promoted or demoted according to certain strategies. Many block level storage vendors track the information at the data volumes level and, thereby, dilute the efficacy. StorTrends assigns and maintains relevant meta-data information at block level and parents this information from "Cradle to Grave".

Let us consider an example where a volume holds MS Exchange Database. The growth of email users, email data and corresponding regulatory push to retain emails for an extended period will result in messages and data to be scattered all over the space with various degrees of importance. The relevance of emails and corresponding references to them sharply tapers off with the passage of time. StorTrends has a powerful 'data collection and trending' agent that keeps track of near tem and long term block usage. It automatically demotes historically less referenced block form a more expensive storage tier to a less expensive one.

Access of cheaper tiers is also monitored and blocks are promoted back to faster tiers appropriately. In this way, by detecting 'stale' data and moving them out to lower tiers makes more room for newer and frequent data in faster tiers.

According to ESG: "…. The recommended users for Exchange Server increased by more than 100 simply by moving infrequently accessed blocks of data to a lower tier."

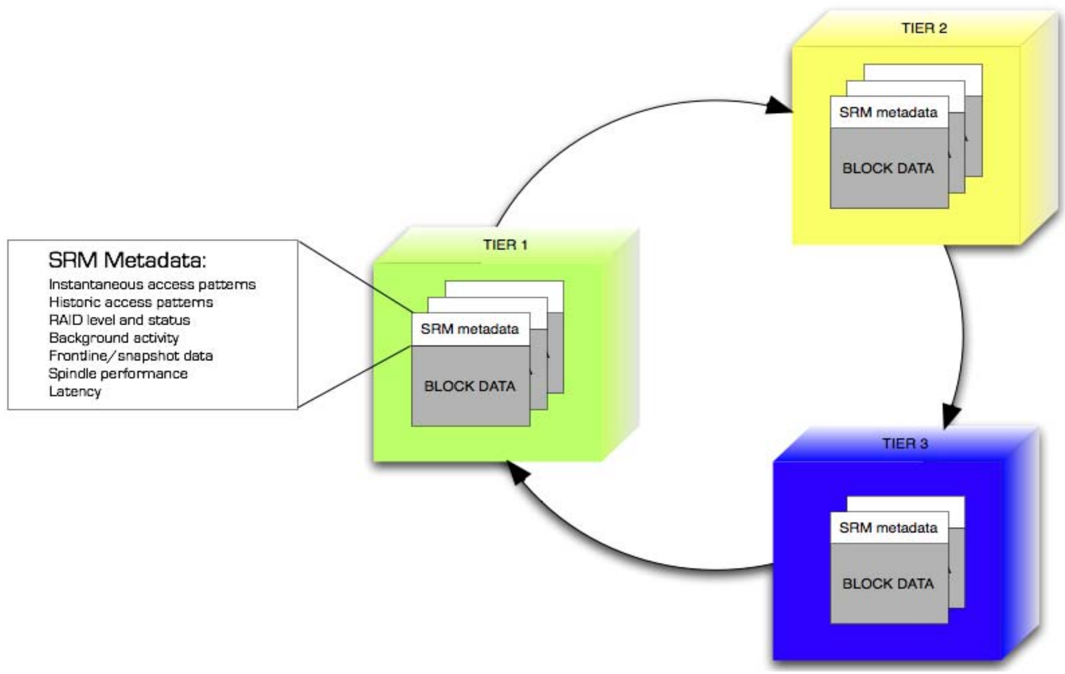

Figure 3:Information Lifecycle Management achieved by moving data based on SRM metadata

How it works

At the core of iTX stack there is a data collection and trending module that collects and categorizes both instantaneous and historical access patterns extending up to years. Storage blocks are also cataloged with respect to their RAID levels, class of disk drives (SAS, SATA etc) and rotational speed (15K RPM, 10K RPM, 7200 RPM etc.). Status and background activity of underlying logical drives (e.g. degraded or being rebuilt etc.) are also considered for the policy of migration. Along with this the extended knowledge of system activity is considered to determine the right time to migrate. Having chosen the right block to move out and the appropriate time to do so, iTX stack utilizes a very efficient algorithm to perform the migration. 'Data

Mover’ technology that works on advertisement of locks ensures that data blocks are migrated without corruption in the fastest possible time.

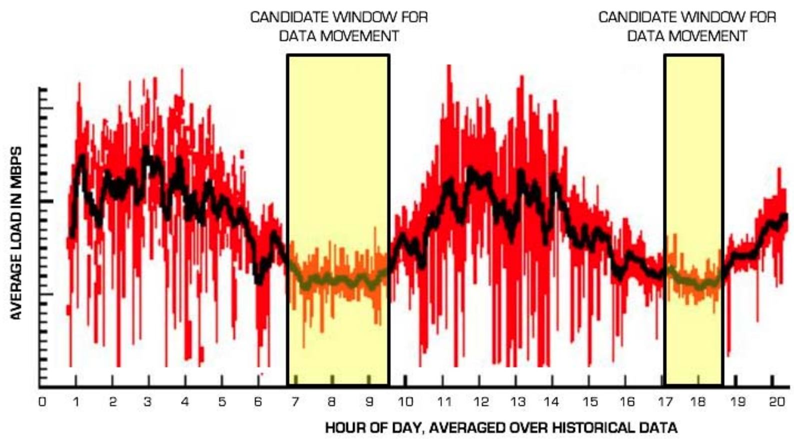

Figure 4: Choosing the best time to move depending on the average historic hourly load

In essence, even at block level, efficient ILM is implemented by maintaining block-level metadata. No content analysis is attempted, in managing blocks of data from conception to extinction.. According to the Taneja Group: “ If you are just interested in doing ILM and migrating between tiers, nine times out of ten you’ll be content just using metadata about the files that shows access patterns”.

StorTrends also moves out snapshots to more cost-effective storage tier. Depending on the access patterns, StorTrends not only demotes inactive data to less expensive storage tier, but also compresses data at rest.

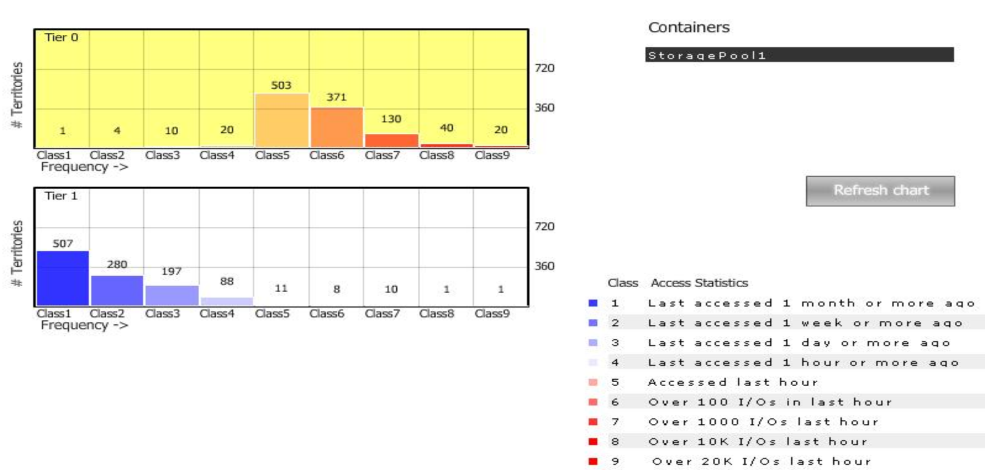

Figure 5: Data Access Patterns in ILM

StorTrends Tiered Storage and CSM (Scaling within tiers)

"Container Space Management can be defined as the technique which is used to equalize space utilization amongst all the logical disks of the same tier in a container, thereby enhancing the spindle performance"

There are various reasons for requiring multiple logical drives in a tier. Some of them are listed below.

Expansion on the logical drive level can take place by adding new physical disks in two ways:

Expansion of existing RAID arrays with the new disks; and

Creation of a new RAID array with the new disks.

Both methods have advantages, but it is often required to use the second method in order to avoid a lengthy re-striping process after expansion (during which time the I/O performance on the old data is much reduced.) This will result in multiple LDs in the same tier, and data has to be striped across them.

The restriction given by the RAID card in the number of disk that can be added to one logical drive or in brief maximum expansion support per logical drive. Some RAID cards may have restrictions on creating arrays across Enclosures. If a storage system contains multiple enclosures or JBODs, it may become necessary to have independent logical drives on each of them.

With a substantial spread in speeds and cost of physical drives available in the market today, it is conceivable that different logical drives may have different RAID levels but may still fall in the same tier. For example, a RAID-10 array with 7500RPM SATA drives and a RAID-50 array with 10000RPM SATA drives may be considered to belong to the same tier. Hence data will need to be stored using CSM on both of these entities. CSM makes sure that the maximum number of spindles may be made available to any I/O load, by distributing the data equally amongst all spindles, even if they fall in different logical drives.

How it works

The main challenges faced in re-striping data between LDs are:
"When and how often should data be re-striped?";
"How should the data be re-striped? ";

"What is the order in which the data should be re-striped?"

How often – whenever an existing container is expanded with a logical drive of the same cost, or every one-hour when an analysis of the data present in each container present in the system is done, only then should the re-striping begin.

Below is a pictorial representation of how the territories across the logical drives will be striped when a container is expanded with a logical drive of the same cost.

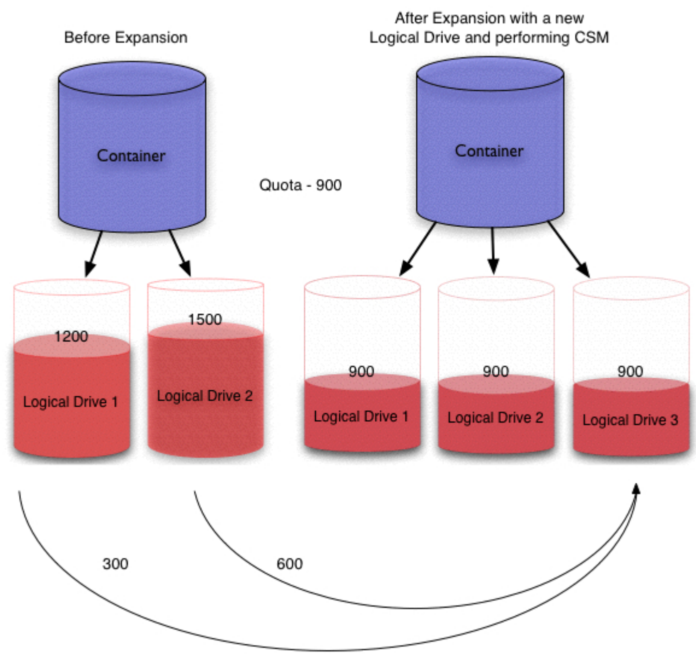

Figure 6: Pictorial representation of how CSM works

When should the data be re-striped: When the I/O load on the system is low it is the right time to re-stripe the data thereby it will not affect the bandwidth of the applications running above. The method of predicting the I/O load on the system at given point in time is done by the workflow analysis.

Below is a pictorial representation of when or at what point in time the task of re-striping should be performed.

How should the data be re-striped – The data should be re-striped in the manner that it is distributed equally across both the logical drives and with the minimum number of movements.

Order of data re-striping – It is beneficial to re-stripe the most frequently accessed data first in the system. This is because this data, when it is re-striped, will be given the performance boost due to extra spindles the earliest, thereby having a net positive impact on the system performance.

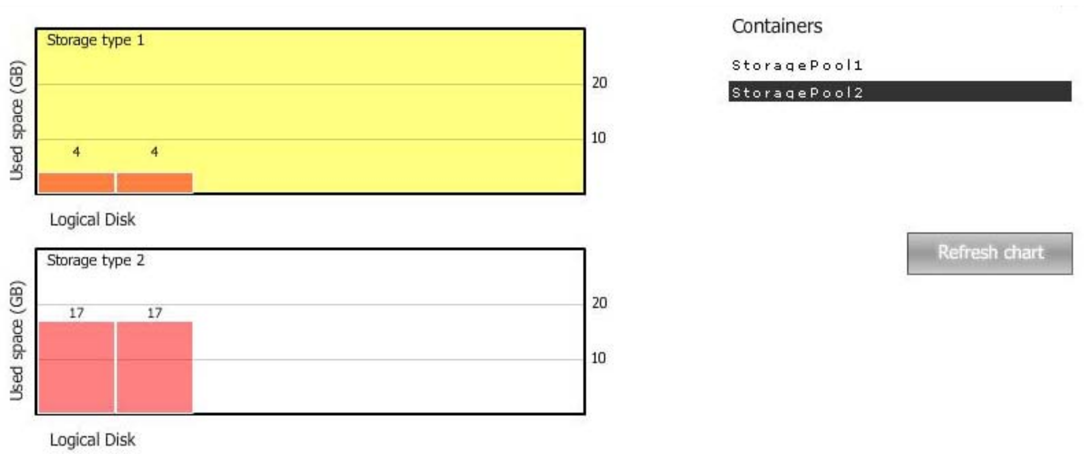

Figure 7: Container Space Management

StorTrends Advantage

- Granular block level tracking of access patterns
- Bi-directional migration between tiers and within tiers
- Efficiency of migration
- Automatic classification by type of storage, RPM, size of drives, RAID level and background activity
- Visual tools to monitor and analyze
- Optimizes and improves performance
- Reduces expenditure on disks by supporting multiple levels of tiered storage
- Improves continuity by conserving space for snapshots

Storage Resource Management User Guide

Navigating within ManageTrends

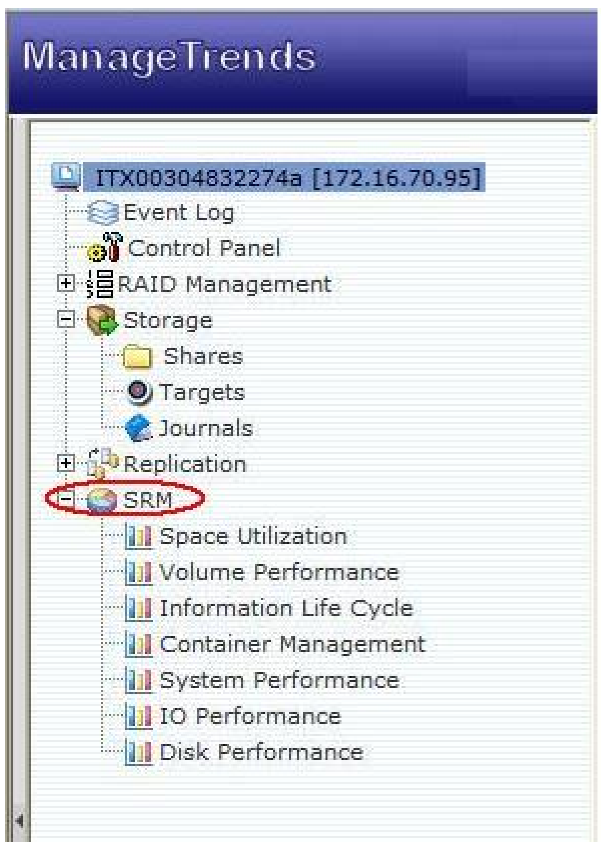

Figure 1: SRM in the ManageTrends Tree

Space Utilization

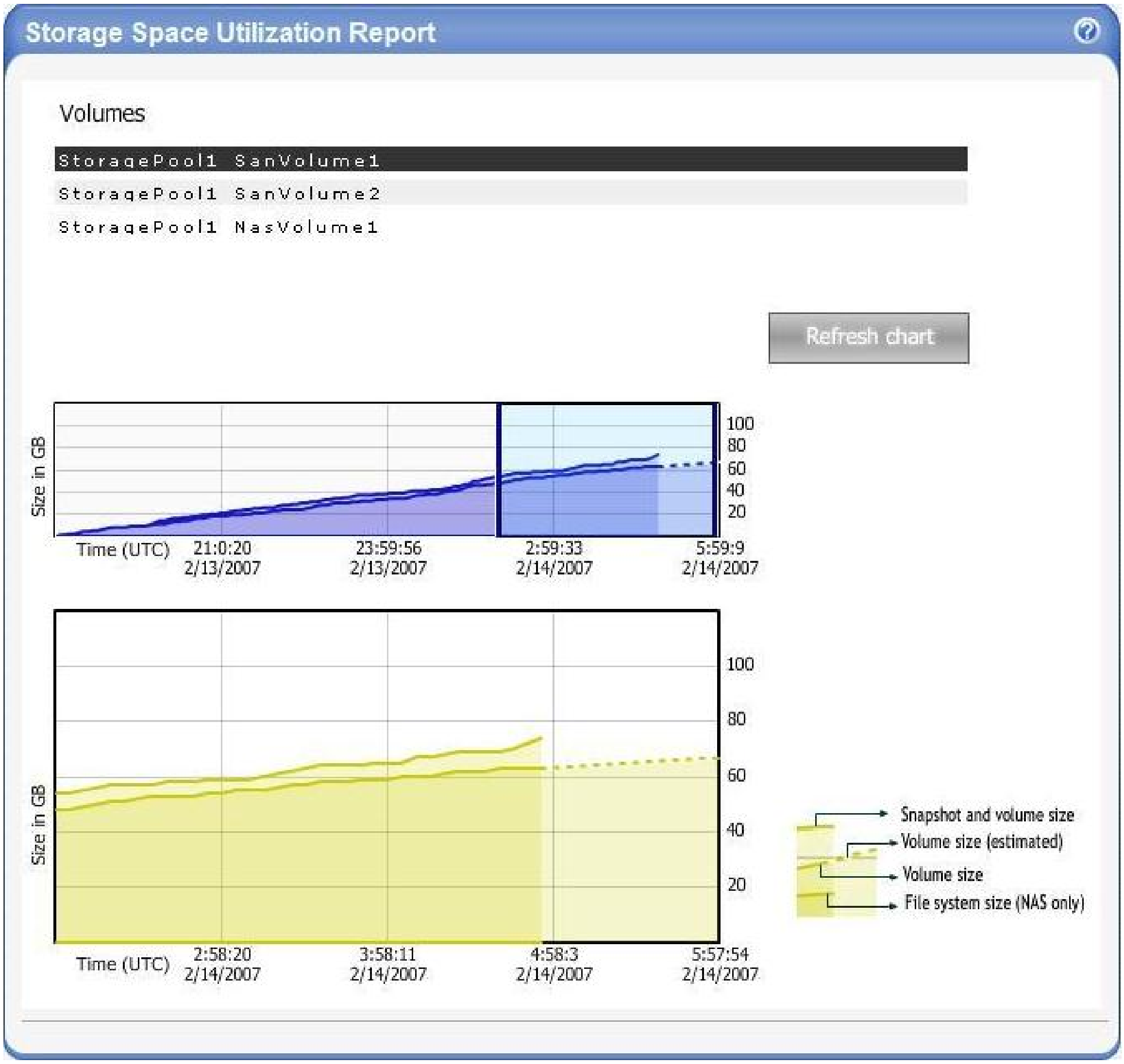

Figure 2: Space Utilization Report

The above chart shows how space is consumed over a period of time. The topmost line refers to the total space occupied by a volume and all its snapshots together. The middle line indicates total volume space occupied, and is also extrapolated (dotted line) to show the predicted growth. The bottom line is the space occupied by the file system (this is applicable for NAS Volumes only. In the above figure you can see that the volume selected is a SAN volume so the bottom line lies on the x-axis). The user can use the sliding window in the top chart to get a more detailed view in the bottom chart of how the space was consumed over a period of time. The size of the sliding window is also adjustable by the user.

Volume Performance

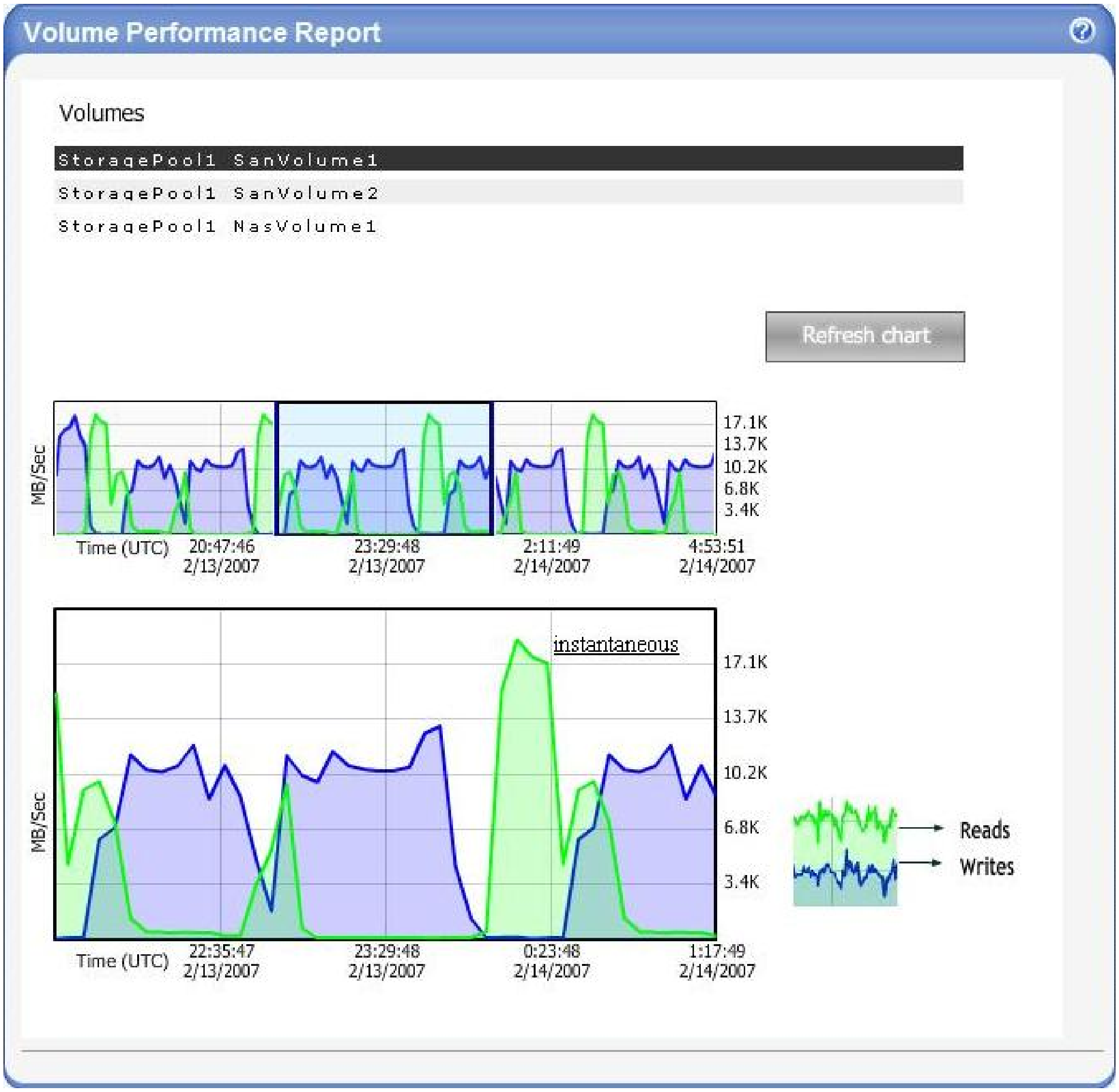

Figure 3: Volume Performance Report

The above chart shows the Volume performance over a period of time. The green colored line in the chart indicates the read performance and the blue colored line in the chart indicated the write performance. This chart has a sliding window, which can be dragged and adjusted for viewing detailed information for a particular period of time. The bottom chart also has a button to flip between the instantaneous performance and overall historical performance. For instantaneous performance the values are polled in a 5-second interval.

IO Performance

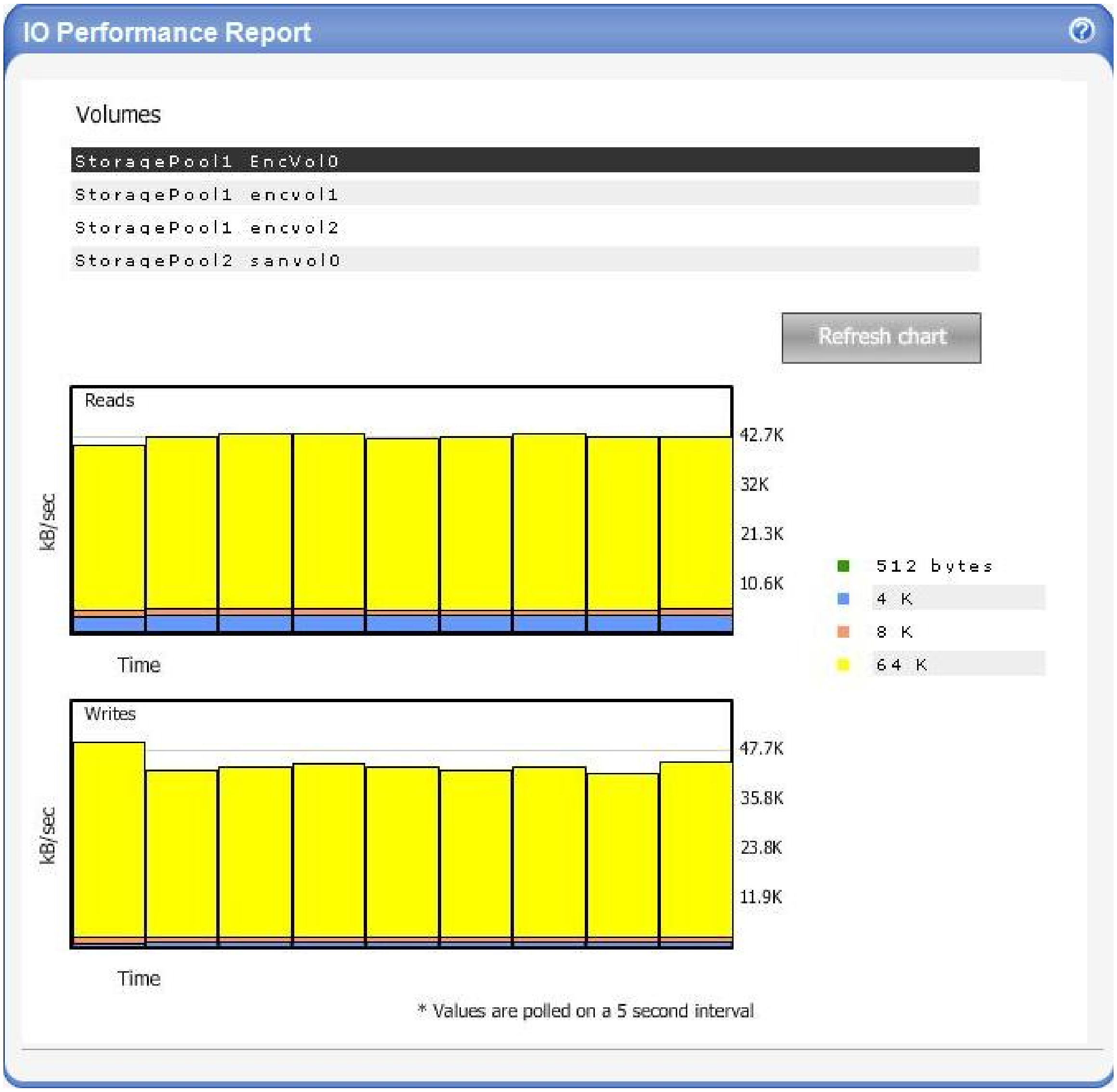

Figure 4: IO Performance Report

The above chart shows that the volume is undergoing more of 64K read and writes compared to the other segments. These values are polled in a 5-second interval.

Disk Performance

For each logical drive (LD) in the container, this chart tracks the Average I/O latencies against the Average Queue Depth on a historical basis to capture the past and present behavior. Average latency is the average amount of time that an I/O takes to complete once it has been dispatched to the logical drive, and is measured in microseconds. This value should be as low as possible. The average queue depth is the average number of outstanding I/Os on a particular logical drive. This number increases as the load on the system increases. If it is too low, the drive may be delivering less performance than it is capable of due to low load. If it is higher for one of many similar LDs, that one LD may be behaving as an I/O bottleneck.

Each line in the graph is indicative of the behavior of a different logical drive present in that container. The lines in the graph have a diminishing color gradient, which is indicative of the past history of the logical drive. By looking at the chart it is possible to detect drive failures in advance, since unusually high latency and outstanding I/Os indicate a failing drive. There is a tool tip at the end of line showing which logical drive it is and when did the last sampling of data happen. There is a chance that drives have same values and thereby the tool tips overlap. In this case we show only the tool tip of the logical drive which was plotted the last.

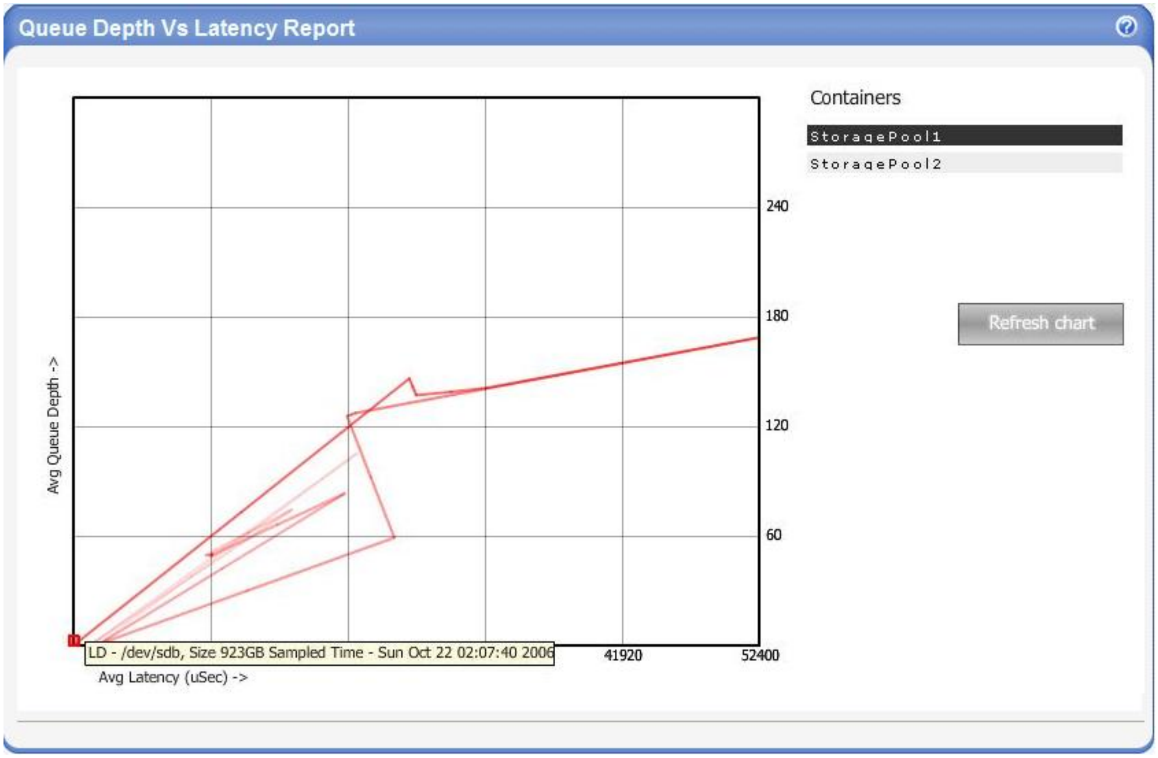

Figure 5: Queue Depth Vs Latency Report

The above chart shows the path traversed by a logical drive over a period of time.

System Performance

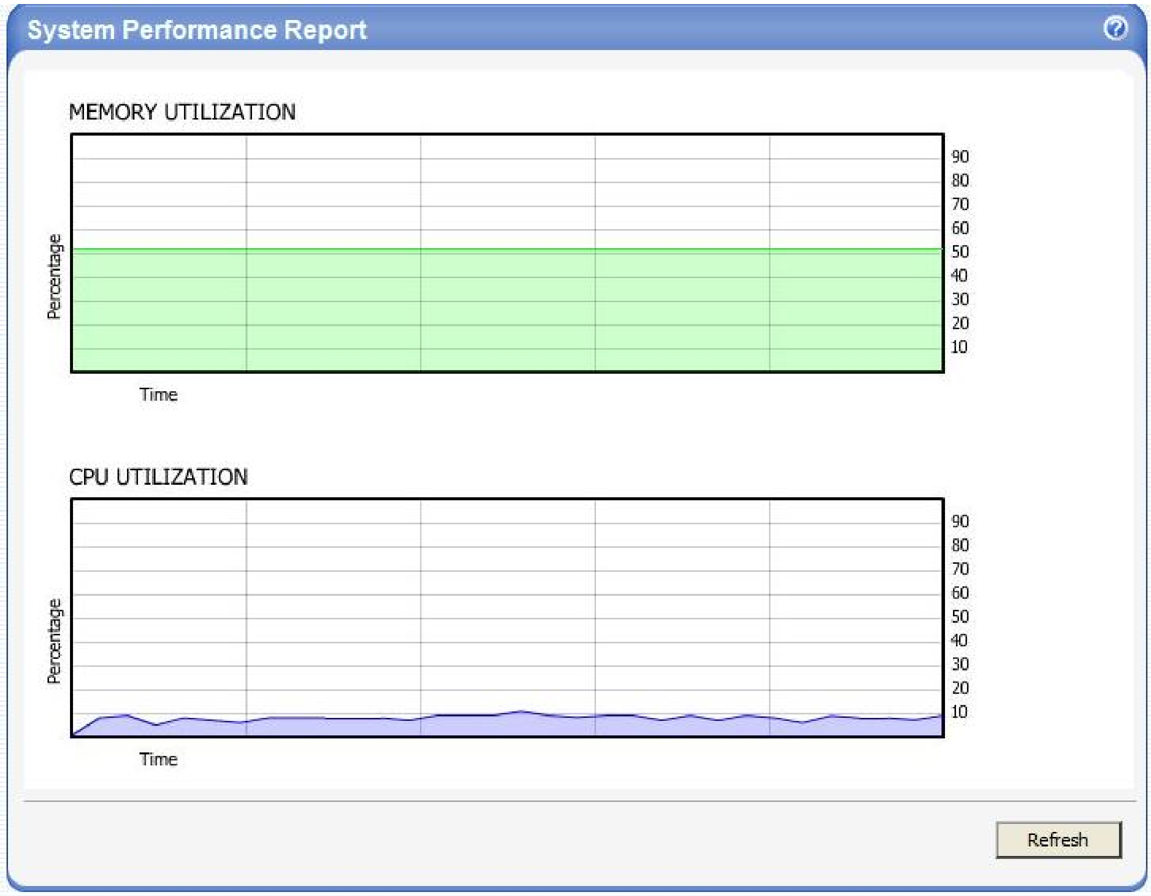

Figure 6: System Performance Report

System performance is tracked in the form of instantaneous CPU and Memory Utilization. These values are also polled on a 5-second interval. Above is the graph showing the CPU and Memory Utilization as a percentage

Information Lifecycle Management

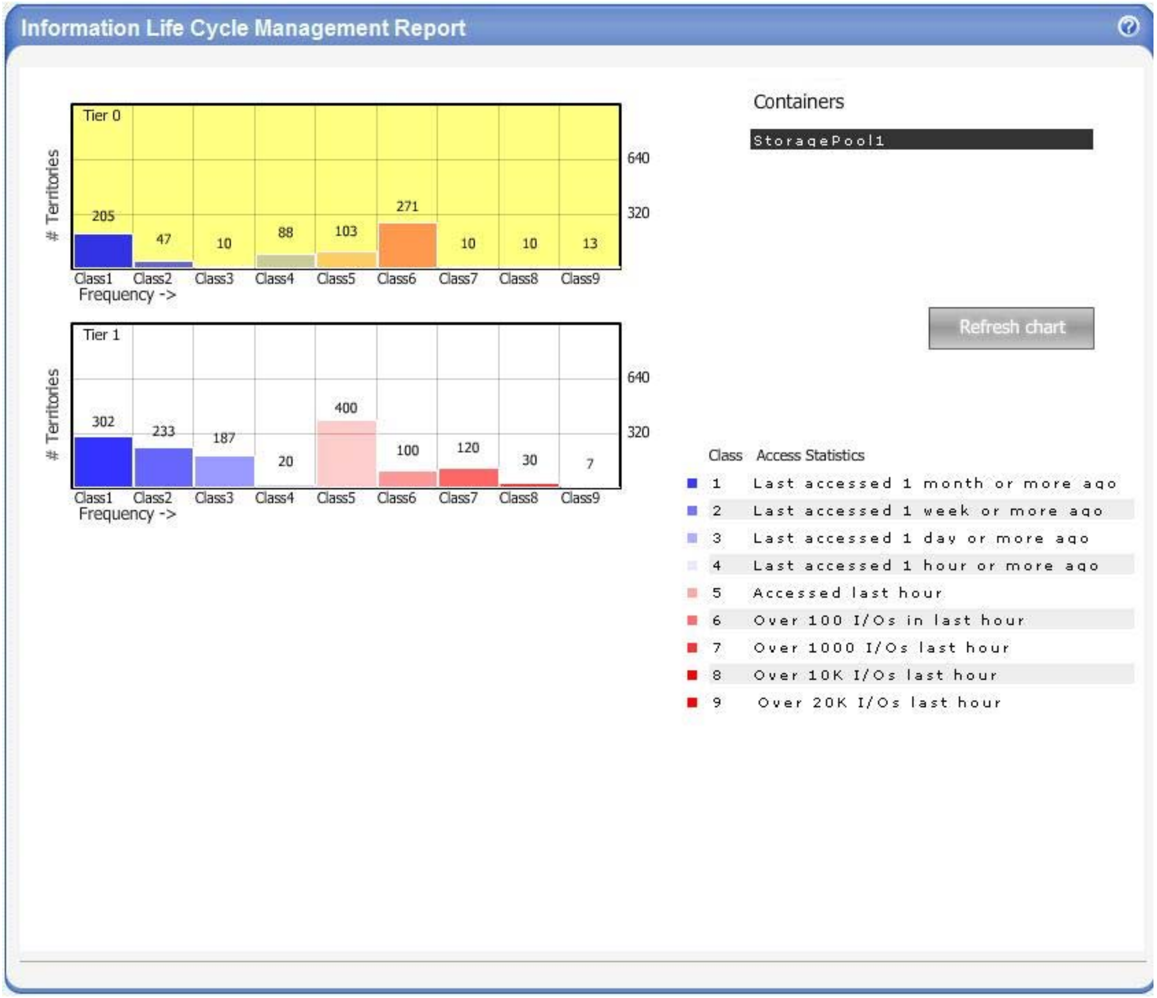

Figure 7: Random data access patterns

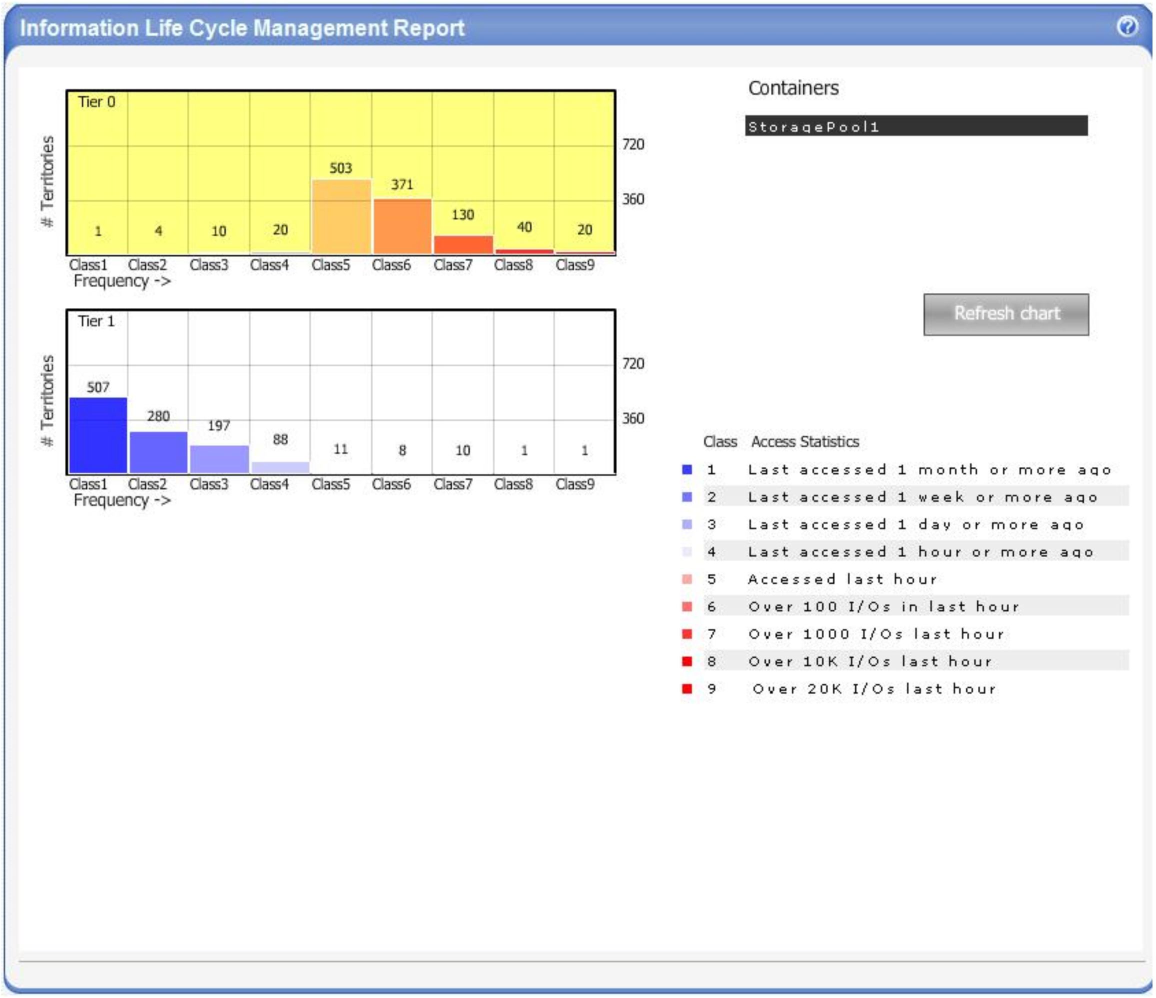

Figure 8: Information Lifecycle Management

The above charts show the access patterns of the data and how they are spread across the tiers. The first chart shows the random access patterns of data. The ideal situation (Figure 8) would be for the most frequently accessed data (red) be present in the high cost tier (yellow background) and least accessed data (blue) present in the low cost tier (white background). ILM takes care of the migrating data across tiers to reach this ideal situation.

Container Space Management

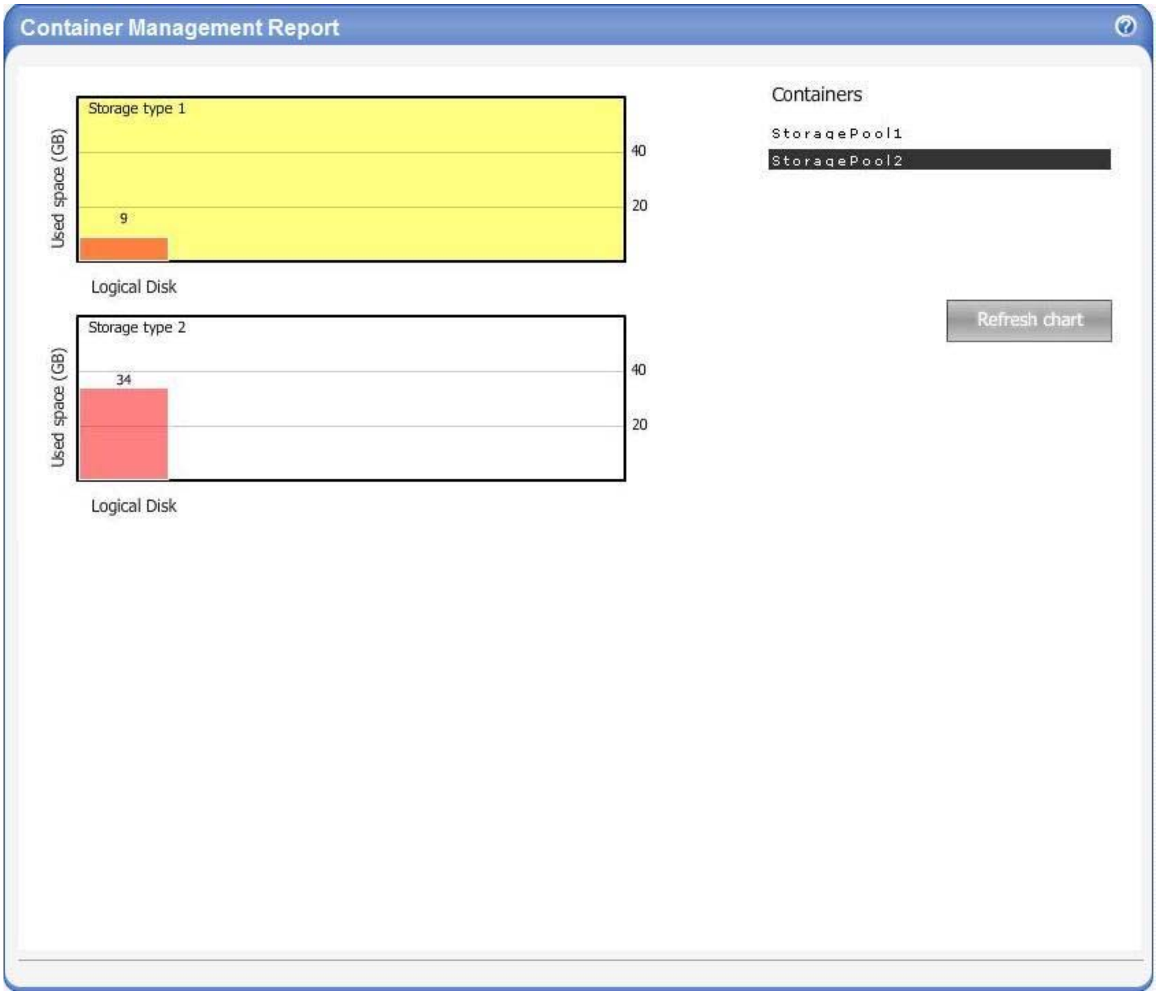

Figure 9: Before Logical Drive Expansion and CSM

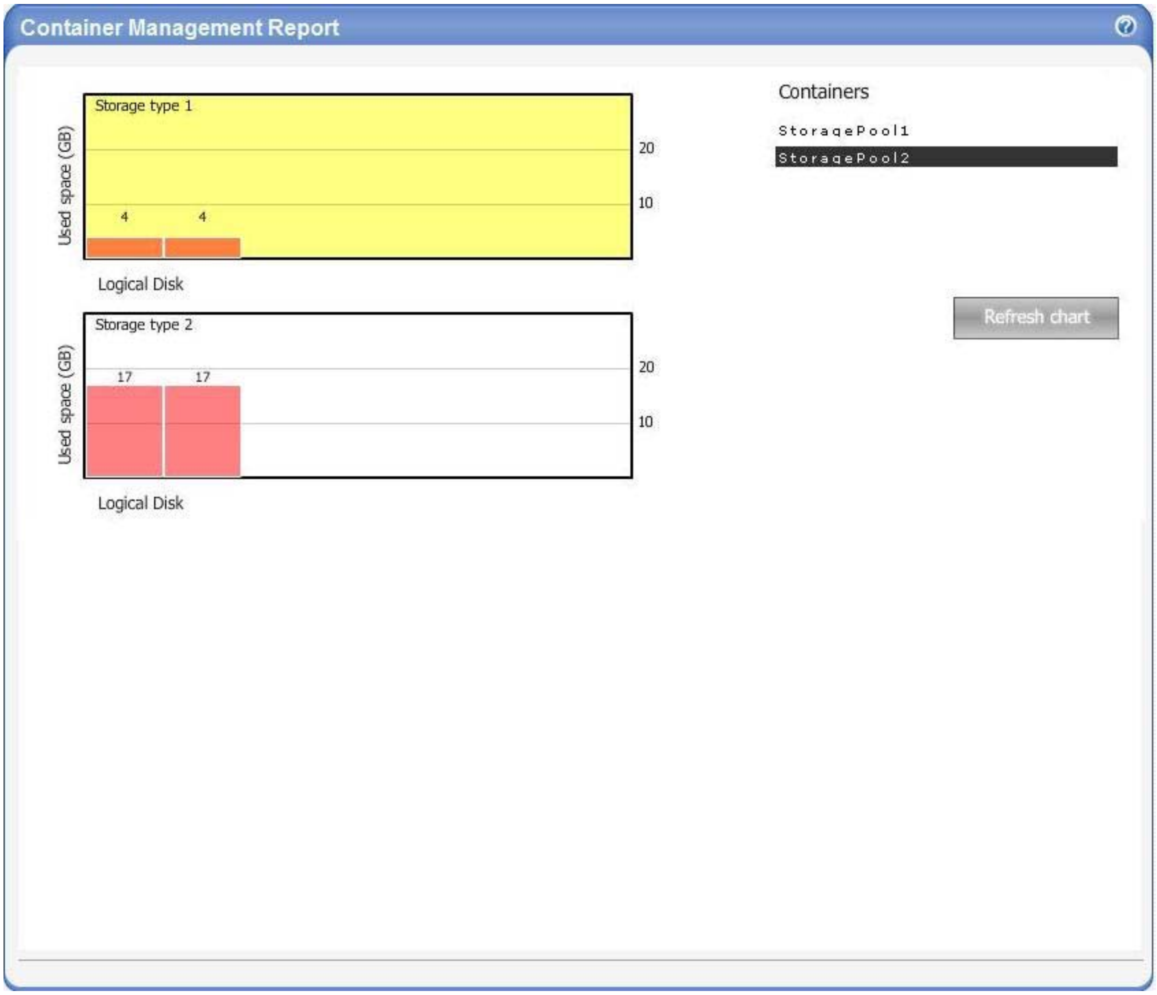

Figure 10: After Logical Drive Expansion and CSM (re-striping of data)

The above charts show how a container was before expansion and how it is now after expansion and re-striping. The x-axis represents the Logical Drives and the y-axis represents the Storage space used in each logical drive.

Reference: CLI User Guide for CLI Commands related to SRM