



SIMPLIFIED STORAGE OPTIMIZATION TECHNIQUES

Janarthanan Palanichamy
Senior Software Engineer
EMC Data Storage Systems
India Pvt. Ltd.
janarthanan.palanichamy@emc.com

Akilesh Nagarajan
Software Quality Engineer
EMC Data Storage Systems
India Pvt. Ltd.
akilesh.nagarajan@emc.com

Table of Contents

Introduction	3
VMware Environment (Host Environment)	5
VMware vSphere Infrastructure	5
VMware DRS.....	6
VMware Storage DRS	6
VMware vMotion.....	7
VMware Storage vMotion	8
VMware vCenter Server	9
Symmetrix Environment (Storage Array)	9
Virtual Provisioning in Symmetrix	9
FAST and FAST VP	10
VMware vCenter Server Plug-ins.....	11
Simplified Solutions for Storage Optimization.....	12
SIOC, Storage DRS, EMC FAST	12
Performance monitoring using Solutions Enabler	13
Conclusion	17
References	18

Disclaimer: The views, processes, or methodologies published in this article are those of the authors. They do not necessarily reflect EMC Corporation's views, processes, or methodologies.

Introduction

EMC Symmetrix® storage systems are best suited for VMware environments. The rapid adoption of VMware-Symmetrix deployments has created a need for easier ways to monitor and manage performance in these environments. The traditional way of looking at VMware performance statistics alone and trying to optimize storage is not the best method. As well, looking at performance statistics of Symmetrix and VMware using various third-party tools could prove counterproductive.

Knowing and applying simplified techniques to make decisions about storage is the need of the hour. This article explores various techniques to optimize storage performance in VMware-Symmetrix environments. These techniques can be employed to get a quick guide to make decisions to optimize storage. Though this article discusses storage optimization techniques for VMware-Symmetrix environments, it can be easily adapted to other environments with different virtualization software and arrays as well.

Let us begin with a few questions which are business critical, the very purpose which drove us to write this article.

- How to collect performance metrics using native tools and provide insights for making decisions to have a balanced environment?
- How to reduce I/O latencies between hosts and the storage arrays?
- How to identify I/O contention issues and what are the ways to address them?
- How to make storage load balancing decisions and prevent I/O boot storm situations?

EMC and VMware proposed numerous approaches and developed a vast set of tools and products to address each of the issues mentioned. The main theme of this article is to illustrate a few tools/products which help simplify storage operations such as resource provisioning, load balancing, centralized monitoring, and prevention of performance bottlenecks. Some of the tools can co-exist and work harmoniously to utilize storage space and related resources more efficiently.

In this article we tried to integrate aspects of VMware technology with EMC solutions for effective storage utilization. Figure 1 depicts a simple block illustration of the article.

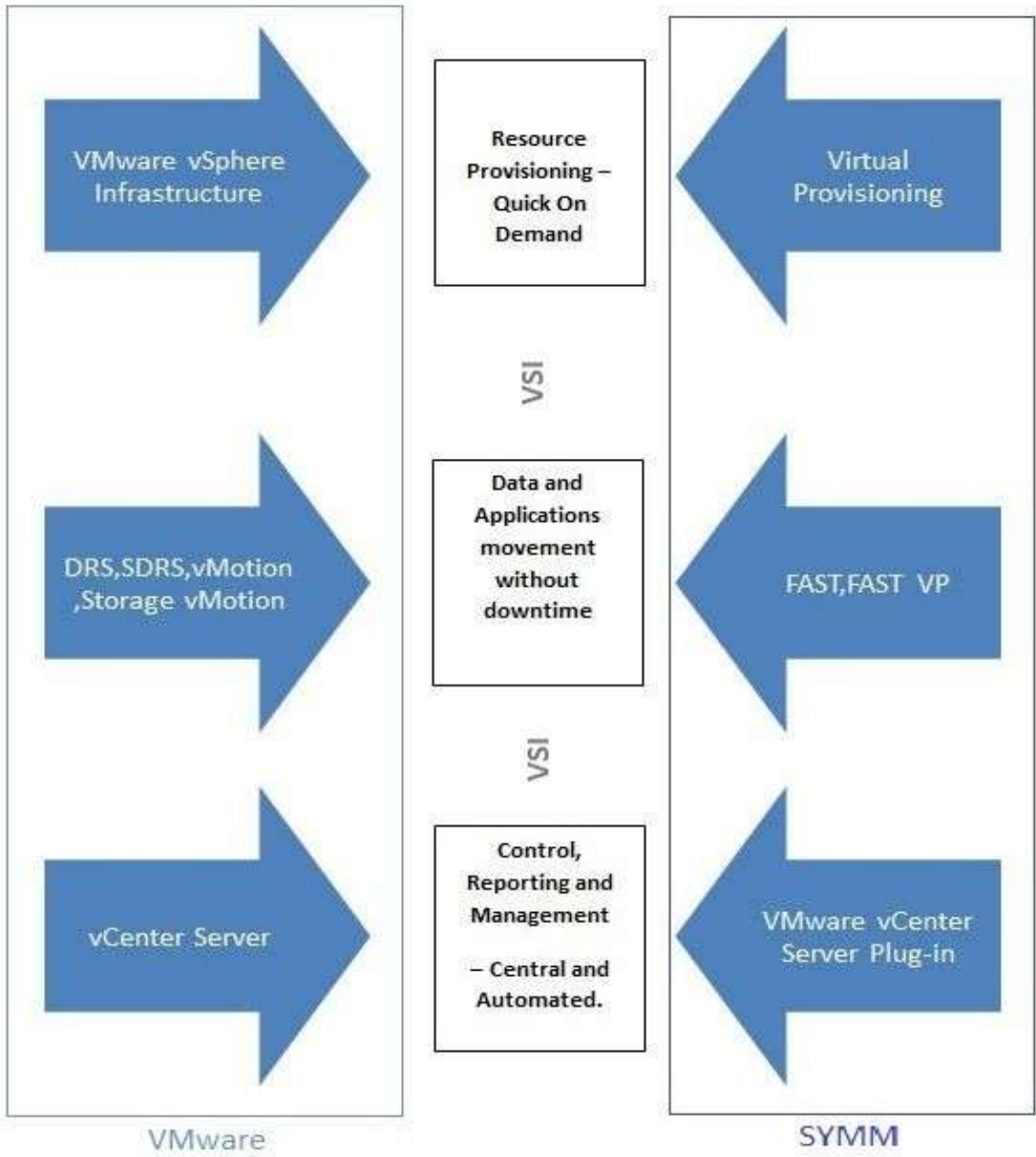


Fig 1) VMware - Symmetrix Environment

Each of these components will be discussed in detail along with some simplified solutions and best practices.

VMware Environment (Host Environment)

EMC Virtual Storage Integrator

EMC Virtual Storage Integrator (VSI) for VMware vSphere is a plug-in to the VMware vCenter.

VSI simplifies storage management by:

- Provisioning Storage
- Fastening Creation of VM Clones
- Mapping virtual machines with storage
- Monitoring data stores on storage arrays from vCenter
- Configuration of FAST-VP policies

VSI consists of four components

- Storage Pool Management
- Storage Viewer
- Centralized Path Management
- SRA Utilities

The Storage Pool Management component of VSI provides a UI that enables users to provision storage. The Storage Viewer provides monitoring and rich performance reporting facilities for all Storage Pool Management users. Data store I/O and LUN I/O can be seen in an easy view. Also, the latest version enables the user to create VM latency alarms which will automatically generate vCenter alerts.

VMware vSphere Infrastructure

vSphere is VMware's virtualization platform offering that helps virtualize x86 server resources and ensures high availability of server resources by decreasing downtime for maintenance activities of physical server and the storage. Known for its low operating cost, vSphere infrastructure also enables efficient storage utilization, simplified administration, and management of storage resources.

vSphere infrastructure provides multiple features, a few of which are explained in detail to make this article clearer.

- Distributed Resource Scheduler (DRS) and Storage DRS
- vMotion and Storage vMotion
- vCenter Server

VMware DRS

Distributed Resource Scheduler [DRS] is a utility used for load balancing based on requirements and resources in a virtualized environment. With DRS, users can describe rules for distribution of physical resources for the virtual machines. Isolation of resource pools for different business units is possible with the help of DRS. Some physical servers can be shut down based on the decrease in the incoming workload.

Other uses of VMware DRS include:

- Control of hardware parameters in a centralized way
- Assignment of hardware as a pool of resources to the virtual machines
- Zero-downtime server maintenance
- Uninterrupted monitoring of hardware consumption and optimization

VMware DRS consolidates physical resources across multiple servers into logical resource pools. In turn, these resources are fed to the virtual machines through rules based on the business needs. If a virtual machine is facing a heavy load, DRS checks for the priority of the workload against the rules. Resources are added if the need is justified. Resources are allocated either by moving the virtual machine to other servers having more resources or by making space on the current server by moving other virtual machines to a different server. The movement of virtual machines is accomplished through VMware vMotion.

VMware DRS can operate both in manual and automated mode. In manual mode, DRS recommends optimal placement of VM's and leaves the rest to the administrator. In automatic mode, VMware DRS makes automatic placement decisions, based on the rules and policies of a VM on specific host inside a cluster.

VMware Storage DRS

Storage DRS is mainly used to avoid resource contention and performance bottlenecks by balancing I/O load and space utilization. Storage DRS has rules and policies for allocation of storage resources to virtual machines. When a data store within the cluster enters maintenance mode, Storage DRS moves the vmdk [Virtual Machine Disk] files to a different data store. Until all vmdk files are moved, the data store from which the vmdk is moved will remain in the state "Entering Maintenance Mode" [1].

Other uses of Storage DRS

- Storage DRS checks on I/O Latency to recommend movement of vmdk's.
- Ensures prevention of data store running out of space. A threshold value is preset and if this value is reached the vmdk files are moved to a different data store which is underutilized.
- Integration of additional storage capacity to storage pools.
- Through Data Store Correlation Storage, DRS discovers if two different data stores are using the same spindles on an array and moves vmdk files accordingly to remove performance bottlenecks.

Storage DRS can operate both in automatic and manual mode. In manual mode, Storage DRS makes balancing recommendations and leaves the rest to the administrator. In automated mode, Storage DRS uses Storage vMotion to reduce I/O latency and to make all virtual machines run in a balanced way.

VMware vMotion

vMotion allows non-disruptive live movement of an active virtual machine from one physical server to another. The process involves movement of the memory and execution state of the virtual machine from one host to another using a high speed network. If using a gigabit Ethernet network, the entire process takes less than “two seconds.” [2]

Uses of vMotion

- Automated optimization of virtual machines within resource pools.
- Improving hardware utilization by scheduling maintenance without disruption and downtime.
- Move virtual machines out of a low performing or a failed server.
- Scheduled movements of virtual machines from one host to another without user intervention.
- More effective resource utilization for business operations.

There are three underlying technologies enabling this movement without service interruption.

1. Encapsulating the state of virtual machine by a set of files which is saved on a shared storage.
2. Transferring the active operational memory and state of execution through high speed networks from the host on which it currently runs to the destination host.
3. The network on ESX host used by the virtual machines is also virtualized to aid preserving the network connection and identity.

The result of using all three technologies is movement of VM's without downtime or service disruption.

VMware Storage vMotion

Storage vMotion ensures live migration of virtual machine disk files across and within hosts without service interruption. Movement of virtual machine disk files without disruption across different storage classes enables effective disk management of virtual disks at reduced costs.

Uses of Storage vMotion

- Effective array migrations and storage upgrades with complete integrity and zero downtime.
- Movement of virtual disks to and from any supported operating system and hardware.
- Live migration of vmdk files across "FC, FCoE, iSCSI and NFS" [3] storage systems.
- Movement of low priority VM's to slower storage, ensuring high performance VM's are reserved for high priority workloads.
- Used in conjunction with Storage DRS for automated management of storage performance.

If Storage vMotion encounters an out of space condition, the destination data store is cleaned by Storage vMotion and the virtual machine will continue to run on source. Storage vMotion initially starts the copy of metadata which contains the log and configuration files of the virtual machine. Later, vMotion starts copying the entire disk to a new location. This replication is done until all copies are in sync and the virtual machine is pointed to the new location.

VMware vCenter Server

vCenter Server is a platform for centralized control and management which helps single console administration of VMware vSphere. vCenter helps in simplification and management of host profiles. It also aids in resource optimization by dynamic allocation of resources.

Features of vCenter server

- Allocates resources to virtual machines based on preset rules by continuously monitoring the resources utilization.
- Ensures high availability by providing failover of the virtual machines.
- During times of peak performance of the virtual machines, vCenter Server can modify resource allocations—such as processor and memory—while it is in running state.
- Ensures security by looking into the access control privileges of the virtual machines.

“vCenter server offers performance assessment, resource provisioning and centralized control and management.”[4]

Symmetrix Environment (Storage Array)

Virtual Provisioning in Symmetrix

Virtual Provisioning is a technology in which storage space is allocated with more capacity than its physical allocation. “Virtual provisioning simplifies storage administration, improves capacity utilization and performance”. [5]

Main uses of Virtual Provisioning

- Storage in Virtual Provisioned environments is provisioned independent of physical storage infrastructure.
- Unlike Standard Provisioning, because of the storage pool concept I/O spreads across multiple spindles which can, in turn, result in improved performance due to reduced disk contention.
- Ensure high availability of applications as there is additional storage, thus re-provisioning becomes less frequent.
- Multiple thin pools are created on the physical disks having many thin devices, reducing the amount of unused physical storage.

Virtual Provisioning in Symmetrix uses logical devices called Thin Devices. These thin devices do not have standard provisioning and are created over the thin pools. These thin pools are

comprised of data devices which are physically provisioned. Addition of new thin devices to the thin pool is possible.

FAST and FAST VP

Fully Automated Storage Tiering (FAST) was introduced with Enginuity release 5874. The first version of FAST is for disk-provisioned environments. Hence, it is named EMC FAST DP. Later, with Enginuity version 5875, support for virtual provisioning environments came in known as EMC FAST VP. FAST DP operates on standard Symmetrix disk-provisioned devices whereas FAST VP operates on thin devices. The support for different device types makes FAST DP and FAST VP operable in a single array in parallel.

FAST and FAST VP automates movement of application data across different tiers within an array for improved performance. The movement of data across these performance tiers is effected by policies and performance requirements of the applications. With FAST DP, data promotion or demotion occurs at full volume level whereas in FAST VP, data movement occurs at the sub-LUN level.

Benefits of FAST

- Automated movement of data between high-capacity SATA or high performance EFD's.
- No more manual tiering of applications when performance objectives change over time.
- Dynamic allocation of resources within a single array by optimizing and prioritizing business applications.
- Greater elasticity in meeting the price and performance needs throughout the lifecycle of the data stored.
- Reduced management effort, footprint, energy utilized, hardware and software cost.

FAST automates identification of which workload resides on what drives. Workload optimization based on drive technology is shown in Table 1.

Device Type	Back End Storage Density
EFD	High
SATA	Low
Fibre Channel / SAS Drives	Not constantly High or Low

Table 1: Workload Optimization based on Drive technology

Several other metrics are used for differentiating the type of workload, such as response time, Unit Storage Capacity Cost, and Unit Storage Request Processing Cost.

VMware vCenter Server Plug-ins

EMC vCenter Server plug-ins for VMware vCenter helps the virtual machine administrator to accelerate VM deployments by using the following set of functionalities

- Provision storage within vCenter.
- Provide storage mapping to ESX hosts.
- “Compress/ Decompress VMDK’s in real time to maximize storage efficiency.” [6]
- Add and modify block storage.

Each of the above mentioned components can integrate with each other as on Figure 1. For instance, FAST and FAST VP can be used in conjunction with vMotion and DRS to aid in virtual machine movement alongside sub-LUN tiering resulting in VM/application movement without downtime. Another example is Resource Provisioning VMware vSphere infrastructure with thin devices which aids in improved performance and efficient utilization of storage.

Following are examples of other proposals that simplifies storage optimization.

Simplified Solutions for Storage Optimization

SIOC, Storage DRS, EMC FAST

Storage I/O Control (SIOC) is used for prioritizing I/O for virtual machines that has access to a shared storage pool. SIOC monitors latency over the hosts when they are interacting with the data stores. When latency monitored exceeds desired levels, SIOC participates to relieve congestion.

Other uses of SIOC

- SIOC is used to configure rules and policies on Virtual Machines.
- In case of I/O latency or I/O congestion, if the values exceed set thresholds, resources are allocated to the virtual machines based on the preset rules.
- Increases the flexibility of storage infrastructure by reducing the need for storage volumes dedicated to a single application.
- Storage Priorities are enforced on virtual machines across a group of ESX hosts.

If a predefined threshold of latency is reached, SIOC brings it under control by restricting the number of I/O that a host can send to a data store.

Storage DRS allows pooling of data stores into single resource.

FAST, as discussed previously, is a sub-LUN tiering algorithm and moves applications across tiers based on preset policies. SIOC and FAST can work in conjunction with each other.

Storage DRS I/O metrics and FAST perform opposing roles, resulting in inefficient utilization of Storage Resources. Thus, Storage DRS I/O metrics needs to be disabled when used in conjunction with FAST VP.

Comparing the sampling window period for both FAST and SIOC, FAST uses a relatively longer sampling window and helps movement across tiers and lowering response times. Meanwhile, SIOC has a relatively smaller sampling period and is used to limit the guest latency to reduce I/O contention.

While SIOC excels in I/O burst scenarios due to short window period, FAST lowers the response time of the busiest I/Os and brings it under SIOC control. In this way, FAST and SIOC can be brought together to provide optimization and firefighting this is efficient and longer lasting.

Performance monitoring using Solutions Enabler

Solutions Enabler symstat command could be used for performance monitoring in Symmetrix-VMware environments where I/O contentions are seen on storage.

Performance statistics of the Symmetrix array as a whole could be gathered using the command mentioned in Figure 2.

```
[root@DLQA1126 ~]# symstat -sid 028 -type REQUESTS -i 5
```

Sym			IO/sec		KB/sec		% Hits		%Seq		Num WP
12:41:50	Dev	Physical	READ	WRITE	READ	WRITE	RD	WRT	RD	WRT	Tracks
12:42:51	00128	Not Visible	0	0	0	0	N/A	N/A	N/A	N/A	15
	00149	/dev/sdaf	82	0	329	230	96	N/A	100	N/A	102
	0014A	/dev/sdag	78	0	312	339	96	N/A	100	N/A	72
	0014B	/dev/sdah	81	0	327	276	98	N/A	100	N/A	100
	0014C	/dev/sdai	80	0	320	223	96	N/A	100	N/A	151
	0014D	/dev/sdaj	85	0	343	309	98	N/A	100	N/A	89
	0014E	/dev/sdak	84	0	338	336	98	N/A	100	N/A	93
	0014F	/dev/sdal	81	0	325	345	96	N/A	100	N/A	94
	00150	/dev/sdam	76	0	305	301	96	N/A	100	N/A	90
	00151	/dev/sdan	78	0	312	347	96	N/A	100	N/A	106
	00152	/dev/sdao	80	0	320	276	96	N/A	100	N/A	127
	00153	/dev/sdgt	63	0	255	284	97	N/A	100	N/A	106
	00154	/dev/sdgu	58	0	234	234	97	N/A	100	N/A	68
	00155	/dev/sdgv	65	0	260	237	97	N/A	100	N/A	47
	00156	/dev/sdgw	63	0	254	289	97	N/A	100	N/A	71
	00157	/dev/sdgx	57	0	230	274	96	N/A	100	N/A	64
	00158	/dev/sdgy	63	0	253	295	97	N/A	100	N/A	37
	00159	/dev/sdgz	57	0	231	220	96	N/A	100	N/A	39
	0015A	/dev/sdha	57	0	231	212	98	N/A	100	N/A	35
	0015B	/dev/sdhb	62	0	250	276	97	N/A	100	N/A	53
	0015C	/dev/sdhc	69	0	276	271	97	N/A	100	N/A	17
Total			1419	0	5705	5574	97	N/A	100	N/A	1576

Figure 2: Symmstat Array Level

Figure 3, depicting the data collected from the above symstat command, describes the average Read/Write IOPS on a few devices in the Symmetrix array.



Figure 3: Read/Write IOPS Array Level

Disk-level Performance statistics could be gathered using the command shown in Figure 4.

```
[root@DLQA1126 ~]# symstat -i 5 -type DISK -sid 028
```

DISK	IO/sec		KB/sec	
	READ	WRITE	READ	WRITE
01C:C0	0	0	32	0
01C:C1	0	0	32	0
01C:C3	0	0	32	0
01C:C4	0	0	32	0
01C:C6	0	0	32	0
01C:C7	0	0	32	0
01C:C9	0	0	32	0
01C:Ca	0	0	32	0
01C:Ce	0	0	32	0
01C:C10	0	0	32	0
01C:C11	0	0	32	0
01C:C13	0	0	32	0
01C:C14	0	0	32	0
01C:C16	0	0	32	0
01C:C17	0	0	32	0
01C:C18	0	0	32	0
01C:C1d	0	0	32	0
01C:C20	0	0	32	0
01C:C23	0	0	32	0
01C:C24	0	0	32	0
01C:C25	0	0	32	0
01C:C27	0	0	32	0
01C:C2a	0	0	32	0
01C:C2d	0	0	32	0
01C:C32	0	107	34	13708
01C:C33	0	107	34	13708
01C:C36	0	107	34	13708

Figure 4: Symmstat Disk Level

Figure 5, created with the data collected from the above symstat command, describes the average Read/Write IOPS on the Symmetrix Disks.

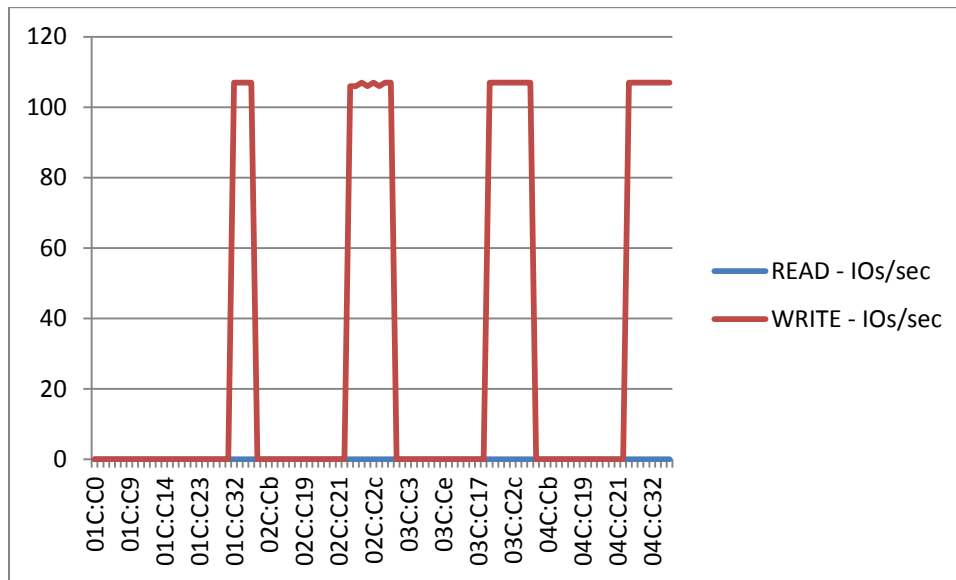


Figure 5: Read/Write IOPS Disks Level

LUN-level Performance statistics could be gathered using the command shown in Figure 6.

```
[root@DLQA1126 SYMSTAT]# symstat -i 5 -type REQUESTS -dev 0014D -sid 028
```

Sym	IO/sec	KB/sec	% Hits	%Seq	Num WP				
Dev	READ	WRITE	READ	WRITE	RD	WRT	RD	WRT	Tracks
23:06:41	30	0	6865	0	100	N/A	100	N/A	69030
23:07:41	30	0	6897	0	100	N/A	100	N/A	69030
23:08:44	30	0	6897	0	100	N/A	100	N/A	69030
23:09:44	28	0	6668	0	100	N/A	100	N/A	69030
23:10:44	23	0	5497	0	100	N/A	100	N/A	69030
23:11:45	19	0	4556	0	100	N/A	100	N/A	69030
23:12:45	20	0	4793	0	100	N/A	100	N/A	69030
23:13:45	18	0	4294	0	100	N/A	100	N/A	69030
23:14:45	19	0	4514	0	100	N/A	100	N/A	69030

Figure 6: Symmstat LUN Level

Figure 7, created with the data collected from the above symstat command, describes the IOPS on device 0014D over time.

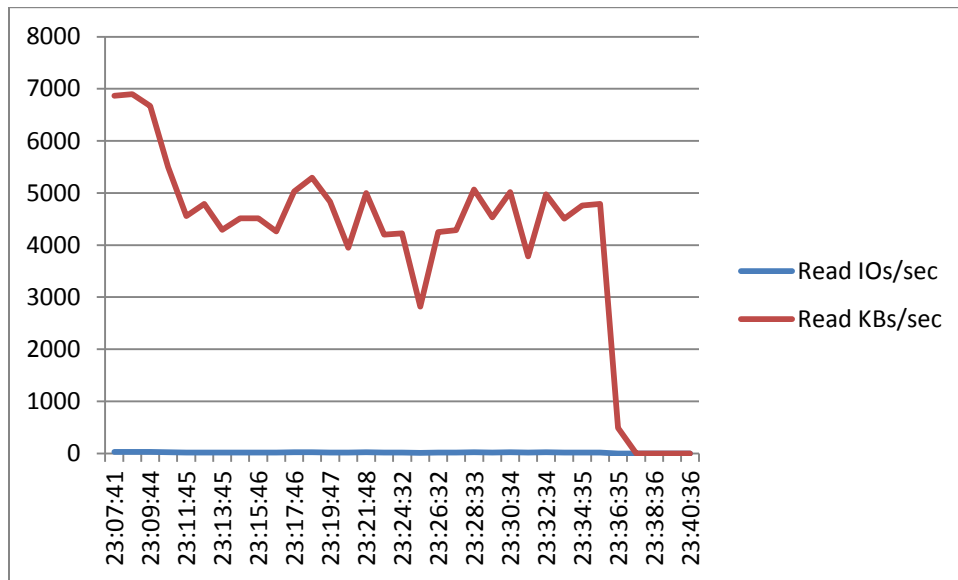


Figure 7: Read/Write IOPS LUN Level

Once we see I/O contentions, the symstat command and the graphs created with the data collected are helpful in identifying the LUN/DISK where the contention originates. Administrators should begin creating mitigation plans after carefully going through the Symmetrix and VMware performance statistics along.

Conclusion

VMware/Symmetrix environments hold a great deal of potential. Simple tools such as FAST and symstat could be used to gain quick insight into performance statistics on Symmetrix arrays and should be compared with VMware performance statistics accumulated by vCenter Server to provide a complete set of performance metrics to optimize the entire environment. No doubt, seamless integration of VMware and EMC products will result in a high level of optimization for the future of storage and virtualization.

References

[1] <http://www.vmware.com/products/vsphere/features/storage-drs.html>

[2] <http://www.vmware.com/products/vsphere/features/vmotion.html>

[3] <http://www.vmware.com/products/vsphere/features/storage-vmotion.html>

[4] <http://www.vmware.com/products/vcenter-server/features.html>

[5] <http://india.emc.com/storage/symmetrix-vmax/virtual-provisioning.htm>

<http://searchitchannel.techtarget.com/tip/How-Storage-vMotion-works-and-what-it-can-be-used-for>

<http://www.techopedia.com/definition/26818/vmware-vcenter-server>

<http://www.techopedia.com/definition/16785/virtual-provisioning>

<https://community.emc.com/thread/180492>

<http://www.emc.com/collateral/hardware/white-papers/h4139-emc-symmetrix-virtual-provisioning-wp.pdf>

[6] <http://www.emc.com/collateral/software/solution-overview/h7229-plug-ins-vmware-so.pdf>

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED "AS IS." EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.