

Oracle Flash Storage System QoS Plus Operation and Best Practices

ORACLE WHITE PAPER | OCTOBER 2016





Table of Contents

Introduction	1
When to Use Auto-Tiering	1
Access Skews	1
Consistent Access	2
Recommendations	2
Basic Operation of Auto-Tiering	2
Progression Maps	2
Preferred Storage Classes	2
Initial Storage Class	3
Collecting the Statistics	3
Evaluating the Statistics	3
Migrating the Data	3
Storage Domains	3
Auto-Tiering Options	3
Enable Tier Reallocation	3
Enable Tier Statistics Collection	4
Performance Optimizing Scan Cycle	4
Allowable Storage Capacity for Auto-Tiered LUNs	4
Tier Storage Allocation	4
Tiers and Storage Classes	4
Determining QoS Plus Effectiveness	5
Summary	7

Introduction

An Oracle Flash Storage System uses QoS Plus, a technology that is patented by Oracle, to perform the following actions:

- » Choose the Storage Class on which to store portions of LUNs.
- » Choose the RAID level of certain portions of LUNs.
- » Prioritize access to the LUNs according to business needs.

QoS Plus provides a significant benefit by automatically configuring storage according to application data access patterns and to business needs. This automatic placement of data is referred to as auto-tiering.

This white paper explains the best practices for provisioning an Oracle FS System storage array.

NOTE: This information applies to software releases 6.0 through 6.2.08.

When to Use Auto-Tiering

Auto-tiering exploits “unfair” access of data in LUNs to migrate data blocks to an appropriate Storage Class. For auto-tiering to be effective, unfairness in the access to LUN data must exist. In addition, auto-tiering depends on being able to predict which parts of the LUNs are accessed more frequently. Therefore, LUNs with a consistent access pattern over time benefit the most from auto-tiering.

Access Skews

LUNs and Storage Domains are constructed from a collection of data blocks. Data access occurs across these data blocks and is characterized by a particular statistical pattern. The degree of asymmetry in the statistical distribution of the block accesses for a LUN is called *access skew*. The Oracle FS System calculates access skew in a range of 0–100 for comparative purposes.

When all the blocks in all the LUNs are accessed uniformly, auto-tiering has no hot data to place on high-performance storage, and no cold data to place on low-cost storage. **Fully random access is an example of zero skew because all data blocks are accessed equally. Fully sequential access is another example of zero skew as each data block is accessed only once.**

An access skew of 100 for a LUN indicates a limited number of data blocks accounting for all the data accesses. This situation can be characterized as an *unfair* data access pattern, because a limited number of data blocks account for virtually all the accesses.

Auto-tiering is at its best when access patterns are unfair resulting in a skew value that is rewarded with the highest performing storage class.

Consistent Access

The auto-tiering feature learns access patterns over time (typically over a day to a week), LUNs with access patterns that persist for many days or longer benefit from auto-tiering. A consistent access pattern, such as occurs when accounting updates complete every day at 15:00, is relative to the Performance Optimizing Scan cycle-time interval. Access patterns that change over many days are consistent relative to a 24-hour cycle-time interval. The migration of data through the tiers of storage takes time and consumes I/O of the system. Therefore, **avoid** setting a cycle-time interval to a low value.

Recommendations

- » LUNs that are good candidates for auto-tiering have significant access skew and access patterns that remain consistent for many days or longer.
- » LUNs with significant random I/O access, and LUNs that are accessed sequentially, should be configured as single-tier.
- » Use single-tier allocation on SSDs for LUNs that require guaranteed low latency.
- » You can change your mind about the LUN allocation parameters, both the choice of auto-tiering and the choice of a single-tier Storage Class. The Oracle FS System supports continuous access to LUNs during migration; however, migration affects system performance.

***Best Practice:** Wait for migration I/O to complete before taking new measurements to gauge the effectiveness of a LUN allocation change. Do **not** make changes lightly. Although you can easily make a change using the Oracle FS System Manager (GUI) or the Oracle FS Command Line Interface (FSCLI), making a change queues up a potentially large amount of work for the system.*

Basic Operation of Auto-Tiering

Progression Maps

For each auto-tiered LUN, the Oracle FS System stores the metadata in an internally managed virtual LUN called the *progression map*. This particular LUN should be configured on the best available storage class. The metadata includes:

- » The physical location of the block
- » The access profile of the block, counting the read operations, the write operations, the sequential accesses, and the random accesses

Preferred Storage Classes

The preferred Storage Classes comprise an ordered list for each auto-tiered LUN. This list influences the order in which the Oracle FS System uses the Storage Classes that are available for ongoing allocations to the LUN.

Preferred Storage Classes are determined by selecting a particular Storage Profile for the auto-tiered LUN or by selecting a set of Storage Classes in the **Additional Options** dialog box. This dialog box is available when creating and modifying an auto-tiered LUN.

When space is not available in the selected Storage Classes, the Oracle FS System writes new data on any of the available Storage Classes that are available in the Storage Domain. This policy applies even when the system must use a Storage Class that has been deselected, or not, in the preferred list.

Initial Storage Class

The initial Storage Class determines the placement of the progression maps and **should be set only to SSD with thin provisioning**. Data blocks are also initially placed here.

Collecting the Statistics

The Oracle FS System bases data placement on LUN priority and block access statistics. The Oracle FS System constantly monitors and stores the 640-KB block access patterns in auto-tiered LUNs to identify the following properties:

- » Read or write bias
- » Random or sequential data patterns

The system stores the statistics in the progression map that is associated with the LUN.

Evaluating the Statistics

After collecting the statistical data for each allocated data block in the system, the Oracle FS System uses the data and the LUN QoS priorities to determine the data blocks that should be promoted or demoted.

Migrating the Data

Based on the analysis of the statistics and of the QoS Priority for a LUN, the Oracle FS System migrates the data to appropriate storage tiers. The Oracle FS System performs the data migration in the background with uninterrupted data access.

Data with the highest hit rates that show a bias toward write operations are placed on SSDs using RAID 10 technology. RAID 10 is used because that technology is the most efficient RAID level for write-biased data. However, if space becomes limited in the SSD tier, the auto-tiering algorithms automatically change the RAID level to RAID 5. Changing to RAID 5 technology saves space, thus permitting the system to store more data on SSDs.

Storage Domains

The physical Drive Enclosures for the various storage tiers are assigned to administrator-defined Storage Domains. An administrator can choose to place all Drive Enclosures in a single Storage Domain or to define multiple Storage Domains to isolate workloads.

Each LUN resides in a single Storage Domain. An auto-tiered LUN can migrate to different storage devices, but only within the Storage Domain in which the LUN resides.

Many QoS Plus options can be defined at the Storage Domain level. Some of these options can be overridden at the individual LUN level.

Auto-Tiering Options

Enable Tier Reallocation

The checkbox for **Enable Tier Reallocation** controls the migration of auto-tiered blocks from one storage tier to another storage tier. This option applies to all LUNs in the Storage Domain. In addition, this option can be set on a LUN-by-LUN basis in the **LUN Modify** dialog box.

Best Practice: *If you have a one-time event that you believe might affect tier reallocation, you need not disable this option. QoS Plus is designed to handle these transient events. However, if the disturbance is projected to last for a week or more and you want to hold the current allocations, disable tier reallocation.*

Enable Tier Statistics Collection

The checkbox for **Enable Tier Statistics Collection** controls the monitoring of block access patterns across the Storage Domain. This option can also be set on a LUN-by-LUN basis in the **LUN Modify** dialog box.

Best Practice: *If you have a one-time event that you believe might affect the statistics, you need not disable this option. QoS Plus is designed to handle these transient events. However, if the disturbance is projected to last for a week or more, disable statistics collection, in addition to reallocation, during the disturbance and re-enable statistics collection afterwards.*

Performance Optimizing Scan Cycle

The text box for a Storage Domain's **Performance-Optimizing Scan Cycle** sets the scan interval, in hours, **during which the frequency of data block access is analyzed to determine which blocks should be promoted** to a higher performing Storage Class. (If capacity is not available in the target Storage Class, only then is data block demotion initiated.) The default value for the scan interval is **24** (one full day).

The text box for **Number of 24-hour Scan Cycles** defines the number of scan cycles that must execute before QoS Plus **performs calculated data block demotions in addition to promotions**. The default value is **7**, because a weekly cycle aligns with business operations.

Allowable Storage Capacity for Auto-Tiered LUNs

The section for **Allowable Storage Capacity for Auto-Tiered LUNs** allows you to set the percentage of each storage tier that is allocated to auto-tiering. These settings allow auto-tiered LUNs and single-tiered LUNs to share a Storage Domain.

Best Practice: *Allocate 100% of the capacity to auto-tiering. Reduce this allocation only when you are allocating single-tiered LUNs in the same Storage Domain.*

Tier Storage Allocation

The Initial Storage Class value for an auto-tiered LUN controls the placement of the progression map or metadata.

Best Practice: *When creating an auto-tiered LUN, always set the **Initial Storage Class** for the LUN as **SSD with thin provisioning**. This choice provides the following benefits:*

- » *Places the critical metadata for the LUN in the highest performing tier*
 - » *Allows for the highest performance of your auto-tiered LUN To prevent the first few LUNs that you create from claiming all the SSD capacity, set the **Allocated Logical Capacity** to **0**. That setting leaves the SSD capacity free until data blocks are written to the LUN.*
-

Tiers and Storage Classes

You can use the **Additional Options** dialog box to select the classes of data tiers to be used for a particular LUN.

Note: Under some circumstances, when you exclude a Storage Class from an auto-tiered LUN, the system might use that class anyway. The system uses the excluded class only when available capacity in the preferred Storage



Classes is depleted and capacity exists in the excluded Storage Class. This action prevents the LUN from becoming read only when a Storage Domain is heavily allocated.

Best Practice: For a particular LUN, you can select classes of data tiers that are not currently available to the LUN. When those tiers do become available, QoS Plus automatically begins using them for the LUN.

You can also use the **Additional Options** dialog box to disable automatic tier reallocation. By disabling tier reallocation, you can freeze the current storage allocations for a particular LUN.

Best Practice: Leave automatic tier reallocation enabled.

Determining QoS Plus Effectiveness

The chart shown in Figure 1 provides an example of the report that is available from the **Generate Auto-Tier Effectiveness Report** menu option for a Storage Domain. The report describes the decisions that QoS Plus has made, the benefits that QoS Plus has generated for two LUNs, and the overhead that QoS Plus has imposed.

Auto-Tier Effectiveness

July 2, 2015

Page 1

	Performance SSD	Capacity SSD	Performance Disk	Capacity Disk
Fixed Cost/GB	\$9.10	\$4.60	\$1.00	\$0.30
Fixed Cost/IOP	\$0.80	\$1.50	\$2.20	\$7.90
Fixed Relative Latency	0.1	0.12	0.64	1

AutoTier1	/SD1	Quality of Service: PREMIUM	
	Cost Averaged Latency:	0.696	Auto-Tier Cost/GB: \$2.12
Constrained Skew: 89%	Access Averaged Latency:	0.224	Auto-Tier Cost/IOP: \$0.03

	Auto-Tier	Performance SSD	Capacity SSD	Performance Disk	Capacity Disk
Peak IOPs	11,787	4,541	5,764	0	1,482
Average IOPs	6,981	2,672	3,423	0	886
Capacity (GB)	152	13	37	4	97
Cost for IOPs in Single Storage Class		\$9,430	\$17,680	\$25,931	\$93,117
Cost for Capacity in Single Storage Class		\$1,383	\$699	\$152	\$46

Summary:	Cost Using Auto-Tiering	\$322
	Best Cost Using Single Storage Class:	\$9,430 (Performance SSD)
	Cost Savings Using Auto-Tiering:	97%

The cost numbers in this report are approximate. To best interpret the cost-benefit represented in this report, compare the relative difference in costs between the storage classes. The relative difference in costs between storage classes provides a more important comparison than the absolute monetary amount.

AutoTier4	/SD1	Quality of Service: PREMIUM	
	Cost Averaged Latency:	0.565	Auto-Tier Cost/GB: \$3.04
Constrained Skew: 90%	Access Averaged Latency:	0.257	Auto-Tier Cost/IOP: \$0.03

	Auto-Tier	Performance SSD	Capacity SSD	Performance Disk	Capacity Disk
Peak IOPs	17,658	8,159	6,561	1	2,937
Average IOPs	12,417	5,734	4,621	0	2,062
Capacity (GB)	170	27	52	8	81
Cost for IOPs in Single Storage Class		\$14,126	\$26,487	\$38,848	\$139,498
Cost for Capacity in Single Storage Class		\$1,547	\$782	\$170	\$51

Summary:	Cost Using Auto-Tiering	\$517
	Best Cost Using Single Storage Class:	\$14,126 (Performance SSD)
	Cost Savings Using Auto-Tiering:	96%

The cost numbers in this report are approximate. To best interpret the cost-benefit represented in this report, compare the relative difference in costs between the storage classes. The relative difference in costs between storage classes provides a more important comparison than the absolute monetary amount.

Figure 1. Auto-tier effectiveness.

The Effectiveness Report illustrates the effectiveness of auto-tiering when compared to all-flash single-tiering. The report includes both the IOPS cost and the capacity cost for each alternative. The Oracle FS System shows the following information in the report:

- » The theoretical maximum performance (in terms of IOPS) and the storage cost for the LUN, as allocated by QoS Plus
- » The performance (IOPS) and the storage cost for the LUN when it is allocated to a particular Storage Class and RAID level

For each LUN, the report shows the access skew and the cost savings resulting from the data placement.

Note: The values in the Capacity (GB) row reflect capacity **usage** (not allocation) per Storage Class, which enables the calculation and comparison of costs per GB.



Summary

QoS Plus intelligently places data on storage tiers to improve the price and the performance of your Oracle FS system.

QoS Plus is best used when a large skew exists in the data access pattern for a LUN. For highly randomly accessed data or for sequential data, single-tiered LUNs are the best option. However, for most LUNs with a typical access skew, QoS Plus is extremely effective in learning where to place your data to optimize the performance of your storage configuration.

QoS Plus requires multiple storage tiers so that the data blocks with varying access patterns can be favorably located. QoS Plus places the data blocks on the Storage Class that best matches the system performance requirements. When a LUN is initially configured, you most likely will not immediately observe performance improvements. After several Storage Optimization Scans have been performed, however, you should observe significant performance improvements when there is enough access skew to work with. These improvements result from the QoS Plus algorithms learning the data access patterns of the LUN.

In summary, QoS Plus can greatly reduce the need to configure your storage array exclusively with all flash storage. The reduction in required flash capacity is accomplished through intelligently placing the data in the optimal Storage Class. In the absence of effective QoS Plus, the effective but more costly approach is to use an all-flash, non-auto-tiered system. The Oracle FS product offers both options.



Oracle Corporation, World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065, USA

Worldwide Inquiries
Phone: +1.650.506.7000
Fax: +1.650.506.7200

CONNECT WITH US



blogs.oracle.com/oraclepartners



facebook.com/oraclepartners



twitter.com/oraclepartners



oracle.com/partners

Hardware and Software, Engineered to Work Together

Copyright © 2016, Oracle and/or its affiliates. All rights reserved. This document is provided for information purposes only, and the contents hereof are subject to change without notice. This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document, and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Intel and Intel Xeon are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark of The Open Group. 1016

Oracle Flash System QoS Plus Operation and Best Practices
October 2016



Oracle is committed to developing practices and products that help protect the environment