Volume Quality of Service Best Practices with Dell SC Series Storage

Dell Storage Engineering
April 2016

A Dell Best Practices Guide

Revisions

Date	Description
April 2016	Initial release

Acknowledgements

Author: Darin Schmitz

Table of contents

Executive summary

Dell Storage Center Operating System (SCOS) version 7 introduces the volume Quality of Service (QoS) feature, which lets storage administrators govern storage traffic for an individual volume or set of volumes. By using I/O balancing techniques, administrators can alleviate what is commonly known as the *noisy neighbor.* This effect is most notable in infrastructures containing shared resources in which a single tenant can monopolize bandwidth, I/O, or controller resources — essentially starving other applications.

This document explains basic configuration options, guidance, and best practices for implementing volume QoS within Storage Center.

1 Getting started

To use volume QoS, the following prerequisites must be met:

- SC Series system controllers must be running SCOS 7 or higher.
- User preferences must be changed to allow QoS profile selection.
- System storage settings must be modified to enable QoS limits and the server load equalizer.
- The user account must have admin-level or volume-manager-level permissions.

1.1 Configure QoS profile selection

Before enabling volume QoS on an array, the Storage Center user account must be allowed to view and administrate QoS profiles.

Within the Dell Storage Manager (DSM) client:

1. Choose **Edit Settings** for the system.
2. Navigate to the **Preferences** section.
3. Check the box, **Allow QoS Profile Selection**, and click **OK** when finished.

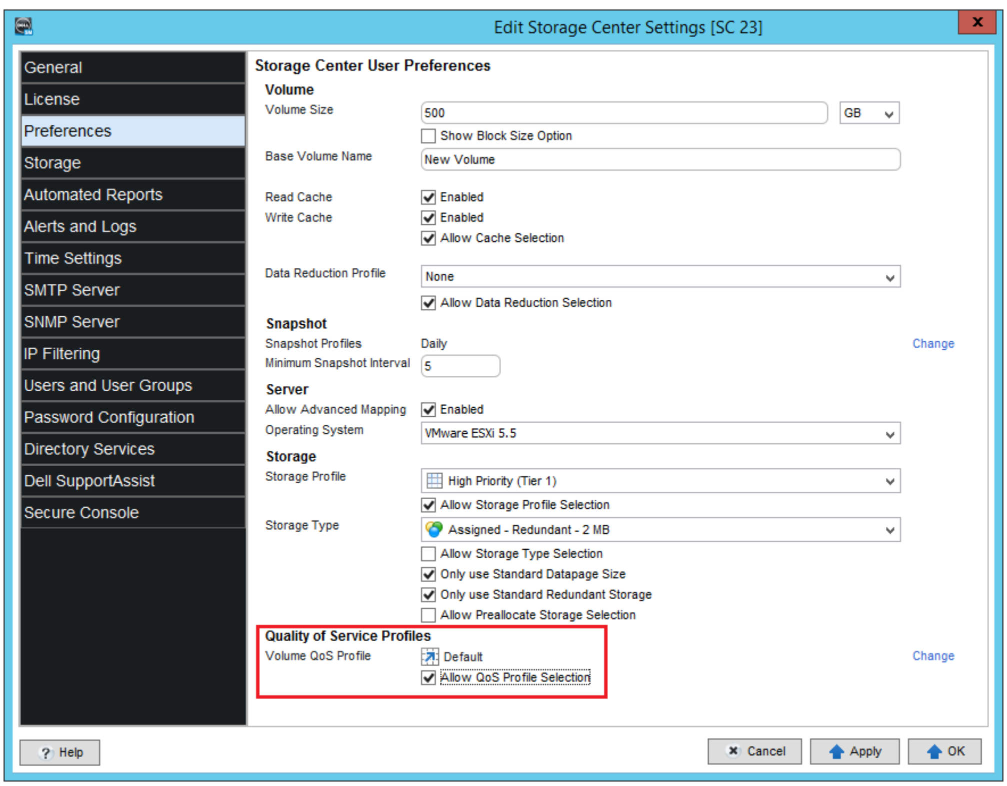


Figure 1 Enabling QoS profile selection

Note: The **Quality of Service Profiles** section also provides the option to select a default QoS policy for newly created volumes.

1.2 Enable Volume QoS on a system

Within the **Edit Storage Center Settings** window, the volume QoS settings are located in **Storage** > **Quality of Service Control**.

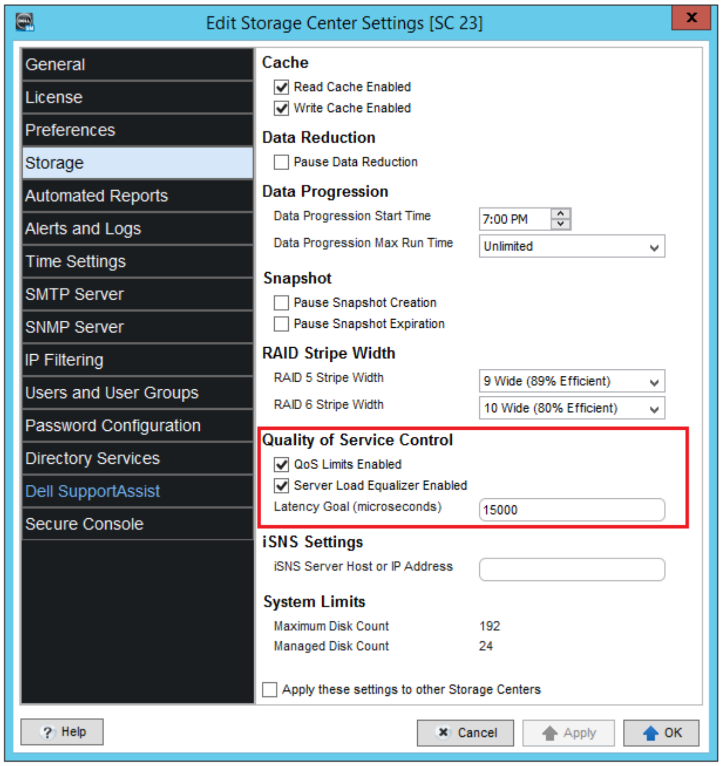


Figure 2 Enabling volume QoS

To fully implement the QoS feature set, it is recommended to check both **QoS Limits Enabled** and **Server Load Equalizer Enabled**. In addition, it is recommended to leave the **Latency Goal** setting at the default value until a performance baseline of the array has been established. Proper use of this setting is described in detail in section 2.

1.3 Default QoS profile

Once volume QoS has been enabled on a system, all of the volumes are assigned the default QoS profile. Effectively, the same relative priority is now applied to all volumes in the system, and each will have unfettered access to system resources until congestion occurs. Once congestion pushes the system latency past the latency goal, the load equalizer will begin to ensure fairness.

This default configuration has been designed to alleviate a majority of noisy-neighbor situations that may occur during normal operation of the array. Before deploying any advanced QoS profiles or strategies, it is recommended that the default configuration be monitored to see if it resolves contention adequately.

2 System latency goal

The latency goal setting within the system is a global value that helps the load equalization engine determine when to begin mitigating resource contention. By default, this setting is configured to 15,000 microseconds (15 milliseconds). While the load equalizer is enabled, it monitors a rolling average of the last 1,000 I/Os processed throughout the system to determine what is referred to as the *system latency*. If the load equalizer detects that the system latency has exceeded the user-provided latency goal, the load equalization engine will begin balancing the I/O stream in an attempt to keep the system latency at or below the latency goal specified.

Since the latency goal setting is a key component of configuring volume QoS, a proper baseline must be established through monitoring and testing for optimal operation.

2.1 Establish a baseline

To determine a proper latency goal, a baseline for a system must be established. If the latency goal is set too high, contention thresholds may never be reached, making it ineffective. However, if the latency goal is set too low, the system could inadvertently throttle loads that do not require throttling. Therefore, it is essential to diligently monitor the performance of volumes and their corresponding applications.

The first step to monitoring a system load is to determine an appropriate observation timeframe. The goal is to determine a period of time when the applications carry out their normal operation with acceptable performance. Depending on the applications on the system, this period could be a day or even a week in duration. The DSM data collector is an important component used to establish a baseline because it can gather and store up to a year's worth of performance statistics, which makes monitoring the performance trends in the array much easier.

Within the DSM client, each system has two distinct performance monitoring sections:

The **IO Usage** tab shows historical performance data.

The **Charting** tab displays up to an hour of real time statistics.

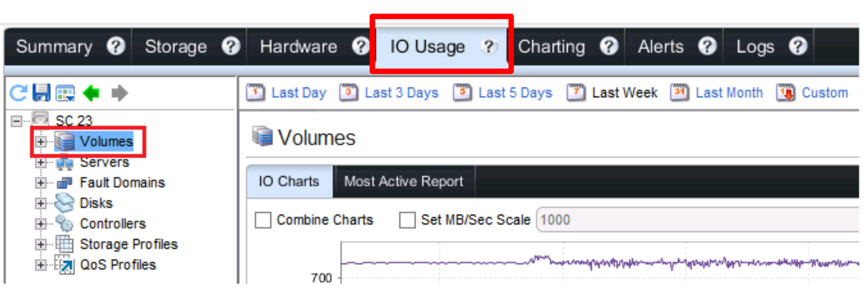


Figure 3 IO Usage and Charting tabs within the DSM client

Typically, the **IO Usage** tab offers the best aggregate of performance data to observe normalized latency data, because a full day's worth of data can be seen at once. In addition, since the system latency is calculated globally, the best place to watch latency is by highlighting the **Volumes** branch in the tree hierarchy. This provides the aggregate performance of all the volumes within the system. For example, Figure 4 shows a week's worth of performance data.

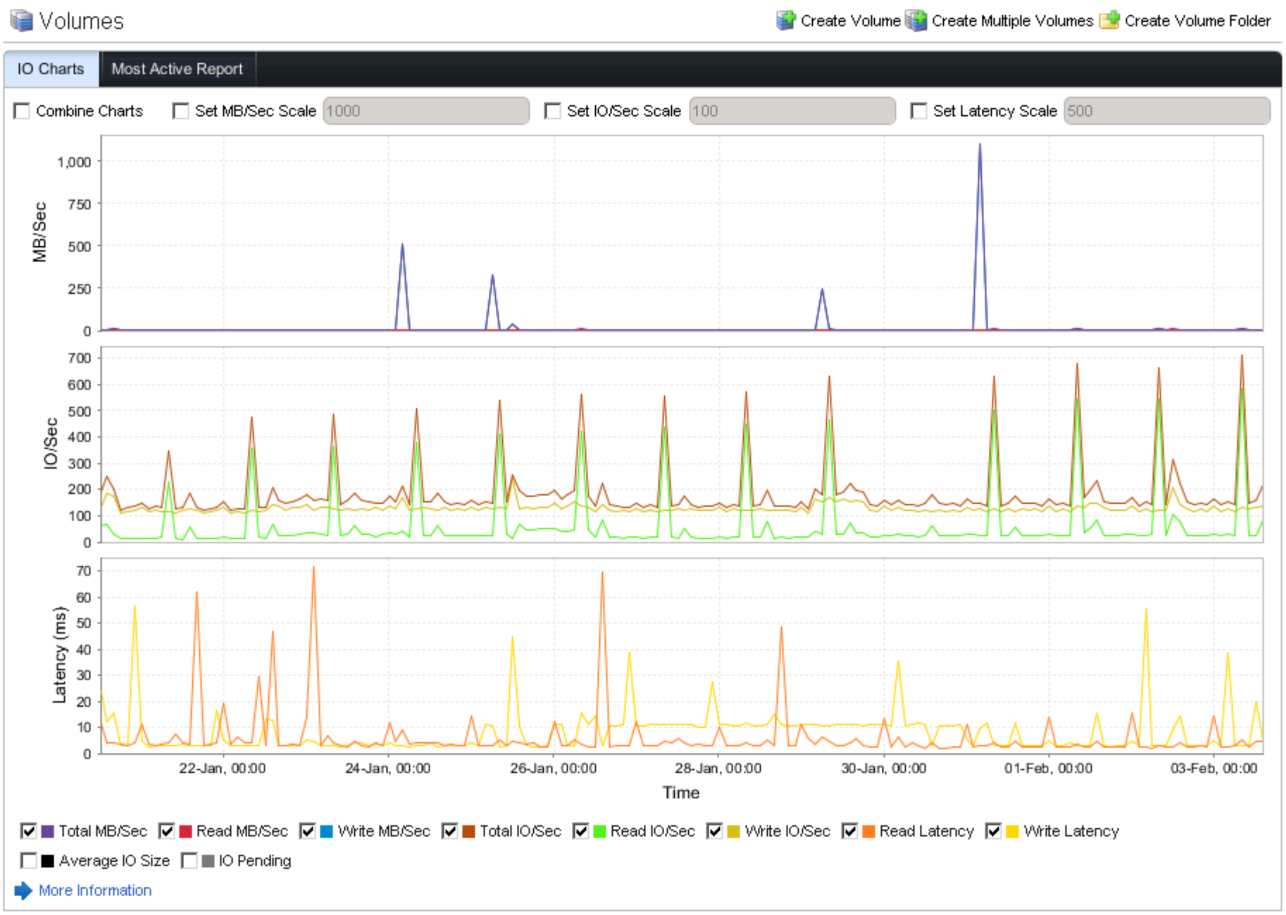


Figure 4 Volume performance monitoring within the DSM client

Select a specific I/O time range to zoom in and reveal more granular latency numbers.

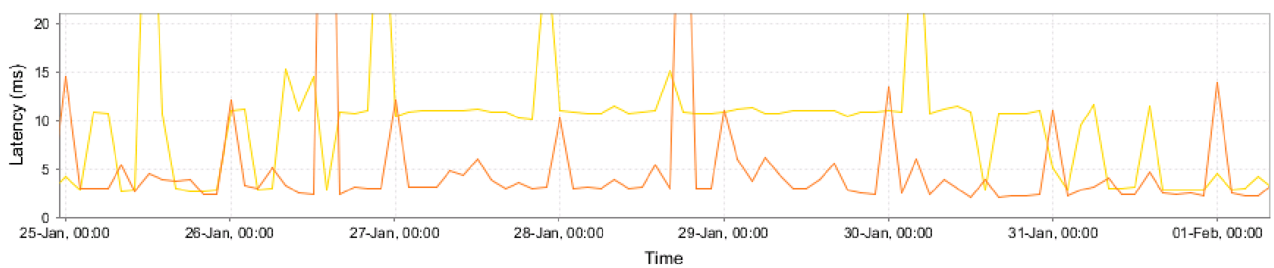


Figure 5 Latency timeframe enlarged to show detail

Looking at Figure 5, the read latency (orange) operates normally in the 4–7 ms range, and write latency (yellow) typically operates in the 4–12 ms range, meaning that the average system latency (read and write combined) will typically be in the 7 ms range. Using this data, a good baseline estimate for this system during the observation timeframe would be 7 ms.

2.2 Set and test the system latency goal

Once the baseline latency for a system has been identified, it can be used to determine a system latency goal. Using the previous example, if normal system operation has 7 ms of system latency, but users begin to complain when the system exceeds 15 ms of latency, a reasonable latency goal is between 9–14 ms.

To test this, choose a starting value within that range, such as 9 ms (9,000 microseconds). Open the system settings, and set the **Latency Goal** to **9000** microseconds.

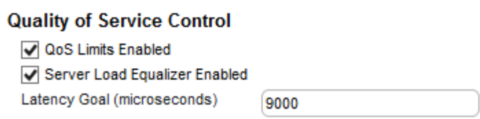


Figure 6 Setting the latency goal

Caution: Setting the latency goal below what the array is physically capable of delivering may have undesired results.

The key strategy to setting the latency goal is to set it, measure the outcome, then adjust the setting as necessary. Keep in mind that the load equalizer will attempt to keep the system latency at or below the defined threshold, and the goal may need to be raised or lowered depending on how the injected latency affects workloads. For example, when volume I/O consists of many small block transactions, even minute changes by injected latency can negatively magnify changes to the total IOPS.

Note: I/O patterns are highly variable and can change over time; it is expected that once a suitable value is found, it will need to be monitored and modified periodically.

2.3 Cross-disk tier, system-wide latency

Another key factor to consider when setting the latency goal variable is the types of disks in the system. Since the system latency is measured globally across all tiers of disks, extra care must be taken when setting the latency goal with multiple-tier systems. For example, in a two-tiered system that contains both SSD and 7K HDD drives, I/Os to each tier can have drastically faster or slower latency footprints. If the averaged system latency experiences an increase due to an influx of data being accessed from 7K drives, then the I/Os to the SSD drives may be unnecessarily throttled.

3 QoS profiles

QoS profiles provide a policy-based approach to assigning priorities or hard limits to volumes.

3.1 Create profiles

There are two types of profiles that can be created from the Dell Storage Manager client: volume QoS profiles and group QoS profiles.

3.1.1 Volume QoS profiles

Volume QoS profiles allow the administrator to set priority, alerts, and limits on volumes.

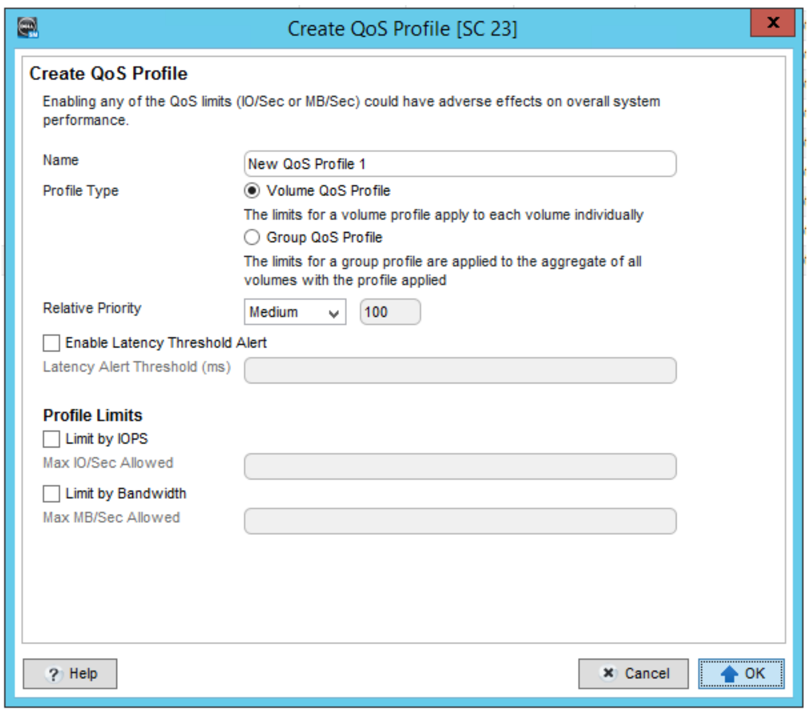


Figure 7 Creating a volume QoS profile

It is important to understand that all volumes placed into a volume profile will have the QoS attributes applied to each of them individually. For example, if a limit of 1,000 IOPS is set on a profile containing five volumes, each of them can use up to 1,000 IOPS at the same time.

Note: Guidelines on properly configuring profile settings are detailed in sections 3.1.3–3.4.

3.1.2 Group QoS profiles

Group profiles are different from volume profiles in that their sole purpose is setting aggregate limits on a group of volumes concurrently. For example, if an application has three volumes that can collectively never exceed a gigabit of bandwidth, but each volume should be able to use the full gigabit while the other volumes are idle, a group policy should be used.

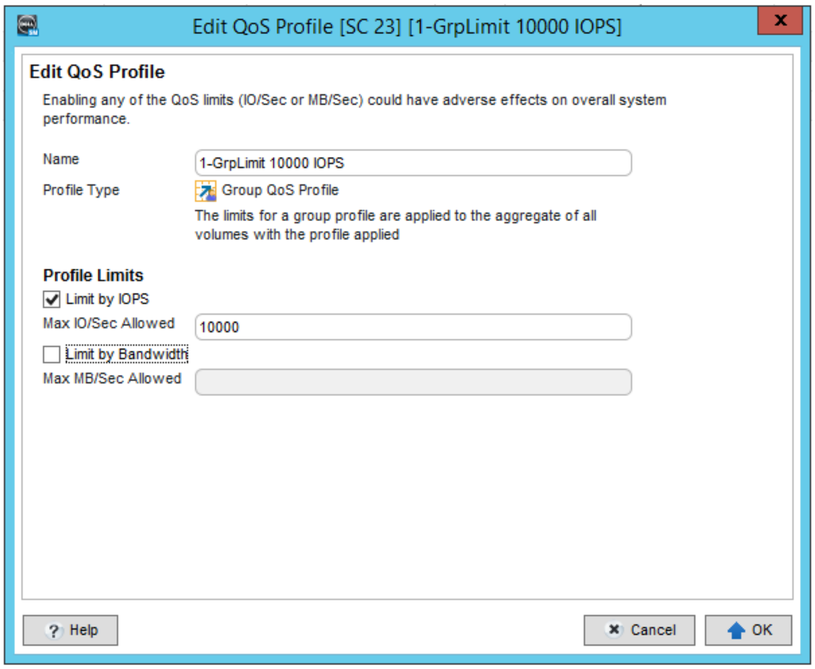


Figure 8 Group profile example

Note: Setting alerting and relative priority are capabilities only offered within volume QoS profiles.

3.1.3 Establish limits

Both volume QoS profiles and group QoS profiles can be applied to volumes to throttle the I/O performance.

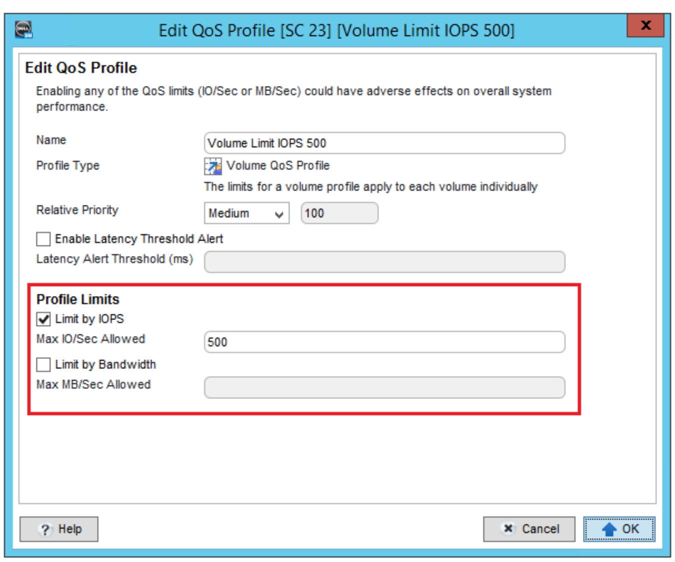


Figure 9 Example IOPS limiting profile

If the system needs to limit a volume, the transactions are throttled by artificially injecting latency into the I/O stream. For example, a volume is currently operating at 1,000 IOPS, but the user wishes to restrict that volume to 500 IOPS.

Since IOPS is defined as the number of input/output operations per second, and each second consists of 1,000 milliseconds, one can deduce that if each I/O takes 1 millisecond to process, the volume will be capable of 1,000 IOPS.

1 I/O x 1 ms each = 1,000 I/Os per sec (IOPS)

To slow that workload to 500, the limiter injects a millisecond of latency into each I/O.

1 I/O x 2 ms each = 500 I/Os per sec (1 ms_to_commit + 1 ms_injected = 2 ms_per_IO)

Note: Latency is not injected unless the server sends I/O faster than what is allowed by the profile established for that volume.

Similarly, for throttling bandwidth, the average block size is factored into the same equation.

1,000 IOPS x 4 KB block size each = 4,000 KBps (4 MBps)

If the administrator needs to limit the volume bandwidth usage to 2 MBps, the system needs to reduce the IOPS.

500 IOPS x 4 KB block size each = 2,000 KBps (2 MBps)

3.2 Configure profile priorities

The relative priority setting within each profile is used by the load equalizer to assign importance to each of the volume loads.

3.2.1 Built-in priorities

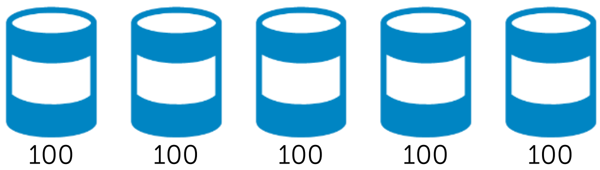
By default, volume QoS presents administrators with three built-in priority levels, as well as a custom priority that is user definable.

Table 1 Volume QoS priority levels

Name	Relative priority
Low	50
Medium (default)	100
High	200
Custom	User defined [1-1000]

When contention within the system occurs resulting in the goal latency being exceeded, the load equalizer begins its congestion management. For example, when all volumes are assigned the same priority, the load equalizer ensures fairness evenly across all the volumes.

In the following example, the default (medium) priority level is applied to all volumes:



100 + 100 + 100 + 100 + 100 = 500
(100/500 = 20%)

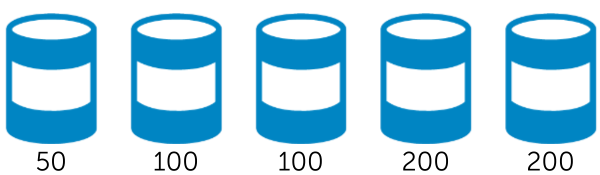
Each volume receives 20 percent of the total system queue depth.

3.2.2 Advanced priority calculations

In cases where certain volumes need higher or lower priority than others, the load equalizer uses the relative priority setting to distribute the total system queue depth. To understand how this works, one must first understand that behind the scenes, each array has a variable number of I/Os being processed at any given time in the system queue depth on each controller. Due to the burstable nature of disk I/O, the system queue depth can fluctuate from dozens to hundreds of I/Os in the system queue at any given time. The load equalizer polls repeatedly to determine the total number of I/Os currently in the queue, and uses the relative priority math to divide up the system queue slots accordingly.

For example, when volumes assigned with varying levels of priority are intermixed, the relative priority is adjusted proportionately.

The following example illustrates distributed priority levels:



50 + 100 + 100 +200 + 200 = 650

Low = 8% (50/650)

Medium = 15% (100/650)

High = 31% (200/650)

Each volume receives respective portions of the total system queue depth.

This applies to only the volumes that are contending for the system queue at that given moment. If any volume is not using its allotted share of the queue, the balancer will redistribute the excess between the battling members. This ensures that the full performance capabilities of the array can be secured by high performance volumes, and also ensures that the lower performance volumes still have a small amount of the system queue to process their I/O.

3.3 Configure QoS profile alerting

Each volume QoS profile can have a latency threshold defined to alert array administrators if the set latency value is exceeded.

As shown in Figure 10, if volumes assigned to this profile exceed the defined latency threshold, an alert will be sent to all array administrators. This can used as a troubleshooting tool or an advanced warning system to alert array administrators at the onset of potential problems.

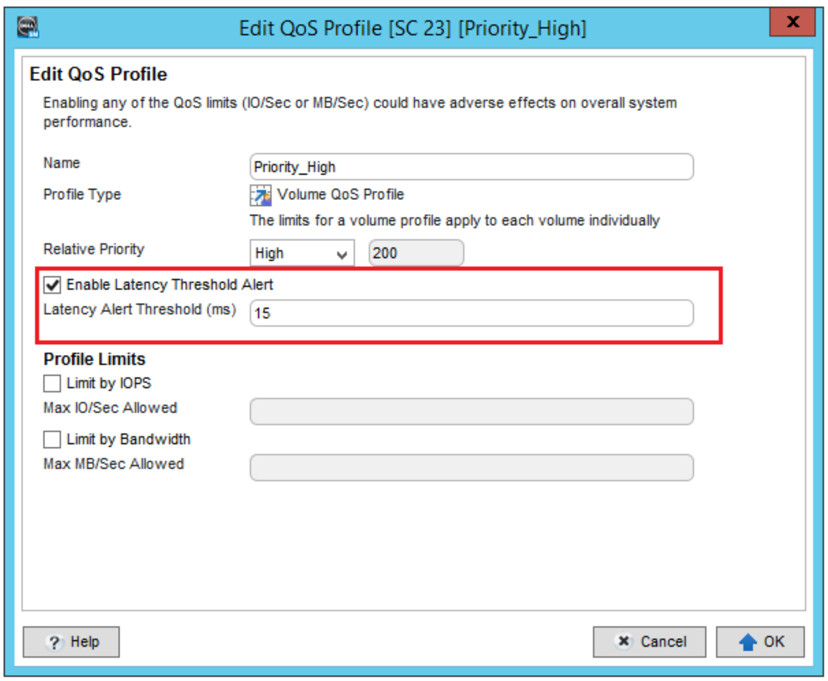


Figure 10 Example of setting the latency threshold alert

3.4 Schedule and automate QoS tasks

Using the Dell Storage APIs, such as the PowerShell SDK, advanced manipulation and scheduling of tasks can be accomplished. For example, if the environment contains a reporting server which routinely become a noisy neighbor on the last day of every month, a PowerShell script can be scheduled to run to apply a restrictive or limiting profile to that volume, and then remove the profile once the reports are finished being generated. In addition, administrators and developers have methods and cmdlets available to create, modify, and delete volume and group QoS profiles.

4 Workload evaluation

When testing volume QoS in an environment, it is important to evaluate each workload to make sure a particular priority or limit will not affect the volumes negatively.

4.1 Monitor QoS metrics

To monitor performance within the DSM client, the QoS latency metric was added. This new value shows how much latency is being injected by the load equalizer into a volume in order to balance the load.

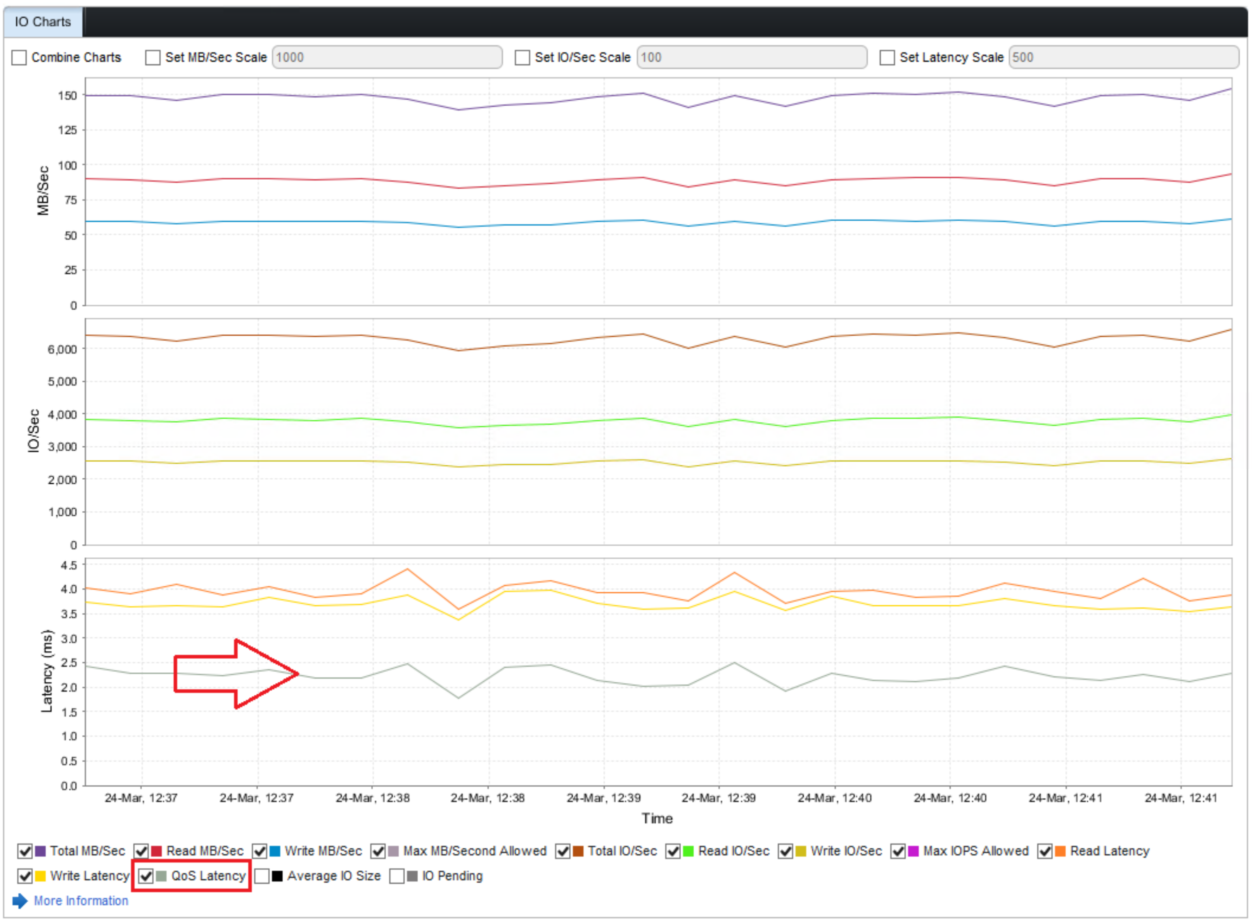


Figure 11 Example of the QoS latency statistic

Caution: When monitoring the performance of an entire folder of volumes within the DSM client, the aggregated performance for latency statistics will always display the highest latency measurement within the polling period. This means that the aggregated latency values will always show the worst-case scenario. For example, if there are five volumes in a folder, with four volumes displaying 1 ms of latency, but one volume is experiencing 6 ms of latency, the folder level statistics pane will report 6 ms. If high latency is witnessed at the folder level, it is recommended to drill-down into the individual volumes to check the specific latencies of each volume.

A Technical support and resources

Dell.com/support is focused on meeting customer needs with proven services and support.

For additional support information on specific array models, see the following table.

Dell Storage	Online support	Email	Phone support (US only)
SC Series and Compellent	https://customer.compellent.com	support@compellent.com	866-EZ-STORE (866-397-8673)
SCv Series XC Series	http://www.dell.com/support	Specific to service tag	800-945-3355
PS Series (EqualLogic)	http://eqlsupport.dell.com	eqlx-customer-service@dell.com	800-945-3355

For SC Series technical content on Dell TechCenter, visit SC Series Technical Documents.