Managing performance using QoS policy group information

Active IQ Unified Manager 9.13

NetApp
May 09, 2024

This PDF was generated from https://docs.netapp.com/us-en/active-iq-unified-manager-913/performance-checker/concept_how_storage_qos_can_control_workload_throughput.html on May 09, 2024. Always check docs.netapp.com for the latest.

Table of Contents

Managing performance using QoS policy group information . . . 1
How storage QoS can control workload throughput . . . 1
Viewing all QoS policy groups available on all clusters . . . 2
Viewing volumes or LUNs that are in the same QoS policy group . . . 2
Viewing the QoS policy group settings applied to specific volumes or LUNs . . . 3
Viewing performance charts to compare volumes or LUNs that are in the same QoS policy group . . . 4
How different types of QoS policies are displayed in the throughput charts . . . 4
Viewing workload QoS minimum and maximum settings in the Performance Explorer . . . 6

Managing performance using QoS policy group information

Unified Manager enables you to view the quality of service (QoS) policy groups that are available on all the clusters you are monitoring. The policies may have been defined using ONTAP software (System Manager or the ONTAP CLI) or by Unified Manager Performance Service Level policies. Unified Manager also displays which volumes and LUNs have a QoS policy group assigned.

For more information on adjusting QoS settings, see Performance management overview

How storage QoS can control workload throughput

You can create a Quality of Service (QoS) policy group to control the I/O per second (IOPS) or throughput (MB/s) limit for the workloads it contains. If the workloads are in a policy group with no set limit, such as the default policy group, or the set limit does not meet your needs, you can increase the limit or move the workloads to a new or existing policy group that has the desired limit.

“Traditional” QoS policy groups can be assigned to individual workloads; for example, a single volume or LUN. In this case the workload can use the full throughput limit. QoS policy groups also can be assigned to multiple workloads; in which case the throughput limit is “shared” among the workloads. For example, a QoS limit of 9,000 IOPS assigned to three workloads would restrict the combined IOPS from exceeding 9,000 IOPS.

“Adaptive” QoS policy groups can also be assigned to individual workloads or multiple workloads. However, even when assigned to multiple workloads, each workload gets the full throughput limit instead of sharing the throughput value with other workloads. Additionally, adaptive QoS policies automatically adjust the throughput setting based on the volume size, per workload, thereby maintaining the ratio of IOPS to terabytes as the size of the volume changes. For example, if the peak is set to 5,000 IOPS/TB in an adaptive QoS policy, a 10 TB volume will have a throughput maximum of 50,000 IOPS. If the volume is resized later to 20 TB, adaptive QoS adjusts the maximum to 100,000 IOPS.

Starting with ONTAP 9.5 you can include the block size when defining an adaptive QoS policy. This effectively converts the policy from an IOPS/TB threshold to a MB/s threshold for cases when workloads are using very large block sizes and ultimately using a large percentage of throughput.

For shared group QoS policies, when the IOPS or MB/s of all workloads in a policy group exceeds the set limit, the policy group throttles the workloads to restrict their activity, which can decrease the performance of all workloads in the policy group. If a dynamic performance event is generated by policy group throttling, the event description displays the name of the policy group involved.

In the Performance: All Volumes view, you can sort the affected volumes by IOPS and MB/s to see which workloads have the highest usage that might have contributed to the event. In the Performance/Volumes Explorer page, you can select other volumes, or LUNs on the volume, to compare to the affected workload IOPS or MBps throughput usage.

By assigning the workloads that are overusing the node resources to a more restrictive policy group setting, the policy group throttles the workloads to restrict their activity, which can reduce the use of the resources on that node. However, if you want the workload to be able to use more of the node resources, you can increase the value of the policy group.

You can use System Manager, the ONTAP commands, or Unified Manager Performance Service Levels to manage policy groups, including the following tasks:

- Creating a policy group
- Adding or removing workloads in a policy group
- Moving a workload between policy groups
- Changing the throughput limit of a policy group
- Moving a workload to a different aggregate and/or node

Viewing all QoS policy groups available on all clusters

You can display a list of all the QoS policy groups available on the clusters that Unified Manager is monitoring. This includes traditional QoS policies, adaptive QoS policies, and QoS policies managed by Unified Manager Performance Service Level policies.

Steps

1. In the left navigation pane, click **Storage** > **QoS Policy Groups**.

   The Performance: Traditional QoS Policy Groups view is displayed by default.

2. View the detailed configuration settings for each available traditional QoS policy group.
3. Click the expand button ( ) next to the QoS Policy Group name to view more details about the policy group.
4. In the View menu, select one of the additional options to view all the adaptive QoS policy groups or to view all the QoS policy groups that were created using Unified Manager Performance Service Levels.

Viewing volumes or LUNs that are in the same QoS policy group

You can display a list of the volumes and LUNs that have been assigned to the same QoS policy group.

In the case of traditional QoS policy groups that are “shared” among multiple volumes, this can be helpful to see if certain volumes are overusing the throughput defined for the policy group. It can also help you decide if you can add other volumes to the policy group without a negative affect to the other volumes.

In the case of adaptive QoS policies and Unified Manager Performance Service Levels policies, this can be helpful to view all the volumes or LUNs that are using a policy group so that you can see which objects would be affected if you changed the configuration settings for the QoS policy.

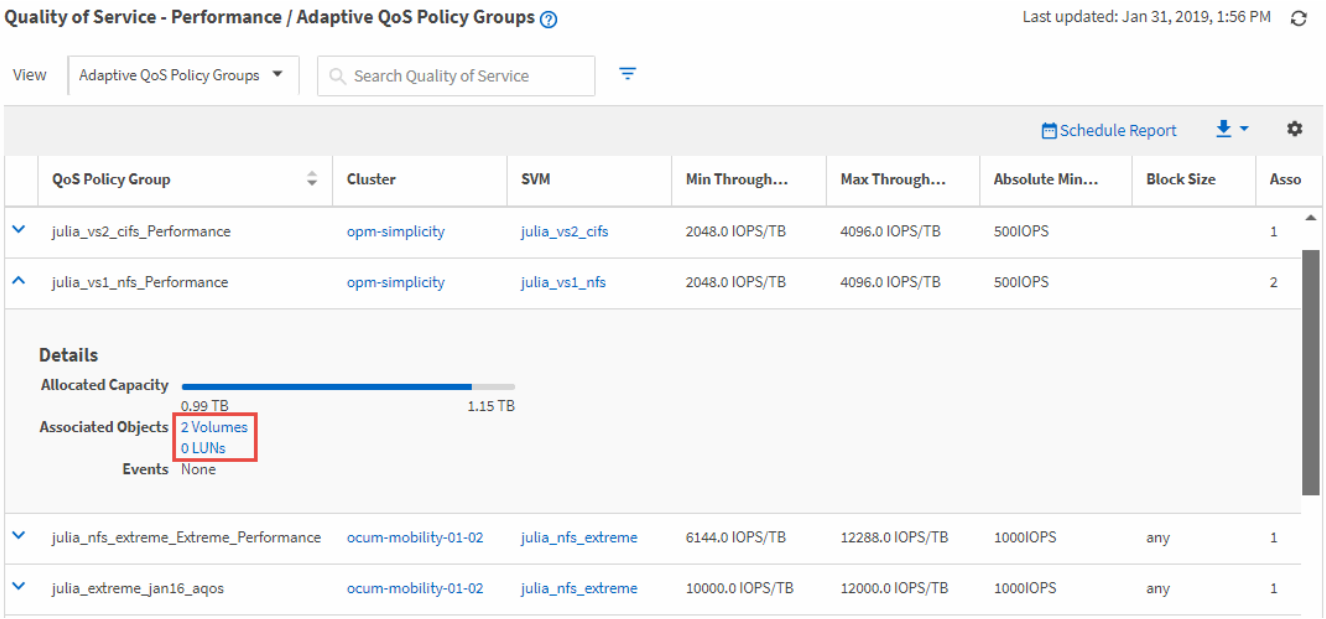
Steps

1. In the left navigation pane, click **Storage** > **QoS Policy Groups**.

   The Performance: Traditional QoS Policy Groups view is displayed by default.

2. If you are interested in traditional policy group, stay on this page. Otherwise, select one of the additional View options to display all the adaptive QoS policy groups or all the QoS policy groups that were created by Unified Manager Performance Service Levels.

3. In the QoS policy that you are interested in, click the expand button ( ) next to the QoS Policy Group name to view more
details.



4. Click the Volumes or the LUNs link to view the objects using this QoS policy.

   The Performance inventory page for Volumes or LUNs is displayed with the sorted list of objects that are using the QoS policy.

Viewing the QoS policy group settings applied to specific volumes or LUNs

You can view the QoS policy groups that have been applied to your volumes and LUNs, and you can link to the Performance/QoS Policy Groups view to display the detailed configuration settings for each QoS policy.

The steps to view the QoS policy that is applied to a volume are shown below. The steps to view this information for a LUN are similar.

Steps

1. In the left navigation pane, click **Storage** > **Volumes**.

   The Health: All Volumes view is displayed by default.

2. In the View menu, select **Performance: Volumes in QoS Policy Group**.
3. Locate the volume that you want to review and scroll to the right until you see the **QoS Policy Group** column.
4. Click the QoS Policy Group name.

   The corresponding Quality of Service page is displayed depending on whether it is a traditional QoS policy, an adaptive QoS policy, or a QoS policy that was created using Unified Manager Performance Service Levels.

5. View the detailed configuration settings for the QoS policy group.

6. Click the expand button ( ) next to the QoS Policy Group name to view more details about the policy group.

Viewing performance charts to compare volumes or LUNs that are in the same QoS policy group

You can view the volumes and LUNs that are in the same QoS policy groups and then compare the performance on a single IOPS, MB/s, or IOPS/TB chart to identify any issues.

The steps to compare the performance of volumes in the same QoS policy group are shown below. The steps to view this information for a LUN are similar.

Steps

1. In the left navigation pane, click **Storage** > **Volumes**.

   The Health: All Volumes view is displayed by default.

2. In the View menu, select **Performance: Volumes in QoS Policy Group**.
3. Click the name of the volume that you want to review.

   The Performance Explorer page is displayed for the volume.

4. In the View and Compare menu, select **Volumes in same QoS Policy Group**.

   The other volumes that share the same QoS policy are listed in the table below.

5. Click the **Add** button to add those volumes to the charts so that you can compare the IOPS, MB/s, IOPS/TB, and other performance counters for all the selected volumes in the charts.

   You can change the time range to view the performance over different time intervals other than the default of 72 hours.

How different types of QoS policies are displayed in the throughput charts

You can view the ONTAP-defined quality of service (QoS) policy settings that have been applied to a volume or LUN in the Performance Explorer and Workload Analysis IOPS, IOPS/TB, and MB/s charts. The information displayed in the charts is different depending on the type of QoS policy that has been applied to the workload.

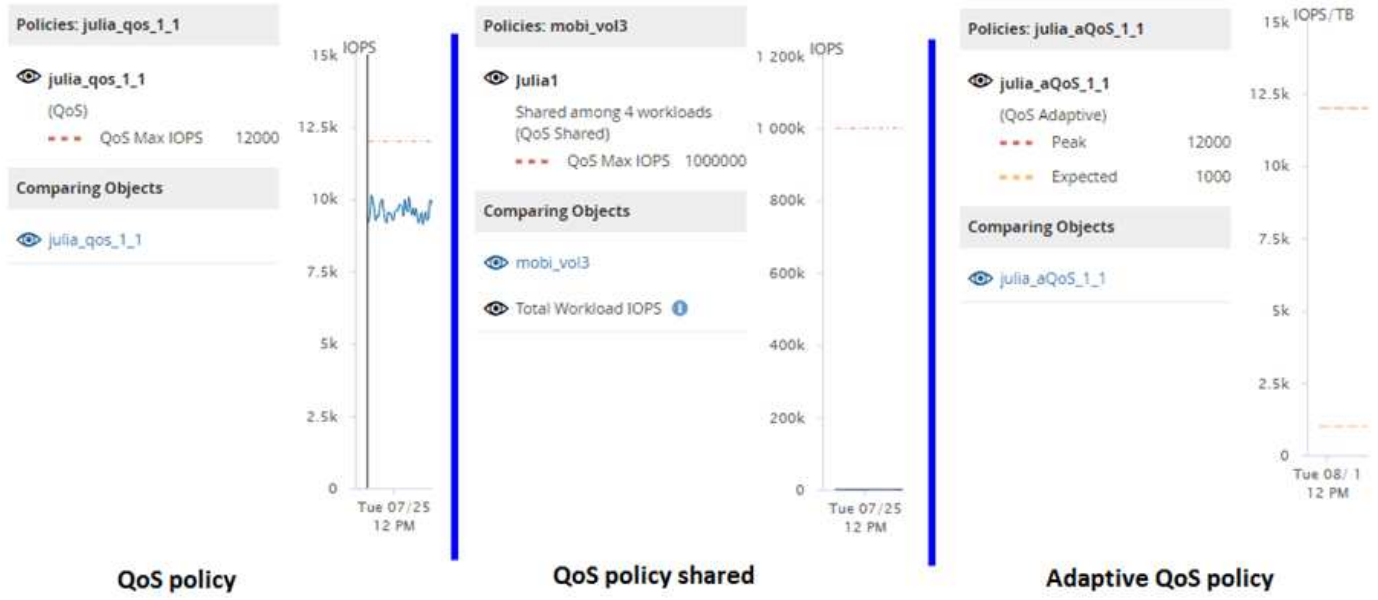
A throughput maximum (or “peak”) setting defines the maximum throughput that the workload can consume, and thereby limits the impact on competing workloads for system resources. A throughput minimum (or “expected”) setting defines the minimum throughput that must be available to the workload so that a critical workload meets minimum throughput targets regardless of demand by competing workloads.

Shared and non-shared QoS policies for IOPS and MB/s use the terms “minimum” and “maximum” to define the floor and ceiling. Adaptive QoS policies for IOPS/TB, which were introduced in ONTAP 9.3, use the terms “expected” and “peak” to define the floor and ceiling.

While ONTAP enables you to create these two types of QoS policies, depending on how they are applied to workloads there are three ways that the QoS policy will be displayed in the performance charts.

Type of policy	Functionality	Indicator in Unified Manager interface
QoS shared policy assigned to a single workload, or QoS non-shared policy assigned to a single workload or multiple workloads	Each workload can consume the specified throughput setting	Displays “(QoS)”
QoS shared policy assigned to multiple workloads	All workloads share the specified throughput setting	Displays “(QoS Shared)”
Adaptive QoS policy assigned to a single workload or multiple workloads	Each workload can consume the specified throughput setting	Displays “(QoS Adaptive)”

The following figure shows an example of how the three options are shown in the counter charts.



When a normal QoS policy that has been defined in IOPS appears in the IOPS/TB chart for a workload, ONTAP converts the IOPS value to an IOPS/TB value and Unified Manager displays that policy in the IOPS/TB chart along with the text “QoS, defined in IOPS”.

When an adaptive QoS policy that has been defined in IOPS/TB appears in the IOPS chart for a workload, ONTAP converts the IOPS/TB value to an IOPS value and Unified Manager displays that policy in the IOPS chart along with the text “QoS Adaptive - Used, defined in IOPS/TB” or “QoS Adaptive - Allocated, defined in IOPS/TB” depending on how the peak IOPS allocation setting is configured. When the allocation setting is set to “allocated-space”, the peak IOPS is calculated based on the size of the volume. When the allocation setting is set to “used-space”, the peak IOPS is calculated based on the amount of data stored in the volume, taking into account storage efficiencies.

The IOPS/TB chart displays performance data only when the logical capacity used by the volume is greater than or equal to 128 GB. Gaps are displayed in the chart when the used capacity falls below 128 GB during the selected timeframe.

Viewing workload QoS minimum and maximum settings in the Performance Explorer

You can view the ONTAP-defined quality of service (QoS) policy settings on a volume or LUN in the Performance Explorer charts. A throughput maximum setting limits the impact of competing workloads on system resources. A throughput minimum setting ensures that a critical workload meets minimum throughput targets regardless of demand by competing workloads.

QoS throughput “minimum” and “maximum” IOPS and MB/s settings are displayed in the counter charts only if they have been configured in ONTAP. Throughput minimum settings are available only on systems running ONTAP 9.2 or later software, only on AFF systems, and they can be set only for IOPS at this time.

Adaptive QoS policies are available starting with ONTAP 9.3 and are expressed using IOPS/TB instead of IOPS. These policies automatically adjust the QoS policy value based on the volume size, per workload, thereby maintaining the ratio of IOPS to terabytes as the size of the volume changes. You can apply an adaptive QoS policy group to volumes only. The QoS terminology “expected” and “peak” are used for adaptive QoS policies instead of minimum and maximum.

Unified Manager generates warning events for QoS policy breaches when workload throughput has exceeded the defined QoS maximum policy setting during each performance collection period for the previous hour. Workload throughput may exceed the QoS threshold for only a short period of time during each collection period, but Unified Manager displays the “average” throughput during the collection period on the chart. For this reason you may see QoS events while the throughput for a workload might not have crossed the policy threshold shown in the chart.

Steps

1. In the **Performance Explorer** page for your selected volume or LUN, perform the following actions to view the QoS ceiling and floor settings:

If you want to…	Do this…
View the IOPS ceiling (the QoS max)	In the IOPS Total or Breakdown chart, click **Zoom View**.
View the MB/s ceiling (the QoS max)	In the MB/s Total or Breakdown chart, click **Zoom View**.
View the IOPS floor (the QoS min)	In the IOPS Total or Breakdown chart, click **Zoom View**.
View the IOPS/TB ceiling (the QoS peak)	For volumes, in the IOPS/TB chart, click **Zoom View**.

If you want to…	Do this…
View the IOPS/TB floor (the QoS expected)	For volumes, in the IOPS/TB chart, click **Zoom View**.

The dashed, horizontal line indicates the maximum or minimum throughput value set in ONTAP. You can also view when changes to the QoS values were implemented.

2. To view the specific IOPS and MB/s values compared to the QoS setting, move your cursor into the chart area to see the popup window.

If you notice that certain volumes or LUNs have very high IOPS or MB/s and are stressing system resources, you can use System Manager or the ONTAP CLI to adjust the QoS settings so that these workloads do not affect the performance of other workloads.

For more information on adjusting QoS settings, see Performance management overview

Copyright information

Copyright © 2024 NetApp, Inc. All Rights Reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP “AS IS” AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

LIMITED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (b)(3) of the Rights in Technical Data -Noncommercial Items at DFARS 252.227-7013 (FEB 2014) and FAR 52.227-19 (DEC 2007).

Data contained herein pertains to a commercial product and/or commercial service (as defined in FAR 2.101) and is proprietary to NetApp, Inc. All NetApp technical data and computer software provided under this Agreement is commercial in nature and developed solely at private expense. The U.S. Government has a non-exclusive, non-transferrable, nonsublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b) (FEB 2014).

Trademark information

NETAPP, the NETAPP logo, and the marks listed at http://www.netapp.com/TM are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.