



Analyzing workload performance

OnCommand Unified Manager 9.5

NetApp

October 23, 2024

Table of Contents

- Analyzing workload performance 1
 - Determining whether a workload has a performance issue 1
 - Investigating a perceived slow response time for a workload 2
 - Identifying trends of I/O response time on cluster components 3
 - Analyzing the performance improvements achieved from moving a volume 4
 - Performance/Volume Details page 6

Analyzing workload performance

Unified Manager enables you to monitor and analyze I/O performance of volume workloads on your clusters. You can determine whether a performance issue is on the cluster and whether storage is the issue.



This chapter describes how to analyze workload performance using the Performance/Volume Details page and the Event details page.

Determining whether a workload has a performance issue

You can use Unified Manager to determine whether a detected performance event was truly caused by a performance issue on the cluster. The event might have been caused a spike in activity, for example, that drove up its response time, but now the response time has returned to its usual levels.

Before you begin

- You must have the Operator, OnCommand Administrator, or Storage Administrator role.
- You must have identified the name of the volume, or associated LUN, you want to analyze.
- Unified Manager must have collected and analyzed a minimum of five days of performance statistics from the cluster.

About this task

If you are viewing the Event details page, you can click the name link for a volume to go directly to the Performance/Volume Details page.

Steps

1. In the **Search** bar, type at least the first three characters of the volume name.

The name of the volume is displayed in the search results.

2. Click the name of the volume.

The volume is displayed on the Performance/Volume Details page.

3. In the **Historic data** chart, click **5d** to display the last five days of historical data.

4. Review the **Latency** chart to answer the following questions:

- Are there new performance events?
- Are there historic performance events, indicating that the volume has had issues in the past?
- Are there spikes in the response time, even if the spikes are within the expected range?
- Have there been configuration changes on the cluster that might have impacted performance? If the response time for the volume does not display performance events, spikes in activity, or recent configuration changes that might have impacted the response time, you can rule out the performance issue being caused by the cluster.

Investigating a perceived slow response time for a workload

You can use Unified Manager to determine whether operations on the cluster might have contributed to the slow response time (latency) for a volume workload.

Before you begin

- You must have the Operator, OnCommand Administrator, or Storage Administrator role.
- You must have identified the name of the volume, or associated LUN, you want to analyze.
- Unified Manager must have collected and analyzed a minimum of five days of performance statistics from the cluster.

About this task

If you are viewing the Event details page, you can click the name for a volume to go directly to the Performance/Volume Details page.

Steps

1. In the **Search** bar, type the name of the volume.

The name of the volume is displayed in the search results.

2. Click the name of the volume.

The volume is displayed on the Performance/Volume Details page.

3. On the Historic data chart, click **5d** to display the last five days of historical data.

4. Review the **IOPS** chart to answer the following questions:

- Are there dramatic spikes in the activity?
- Are there dramatic drops in the activity?
- Are there abnormal changes in the operations pattern? If the operations do not display dramatic spikes or drops in activity, and there were no changes to the cluster configuration during this time, the storage administrator can confirm that other workloads have not impacted volume performance.

5. On the **Break down data by** menu, under **IOPS**, select **Reads/writes/other**.

6. Click **Submit**.

The Reads/writes/other chart is displayed below the IOPS chart.

7. Review the **Reads/writes/other** chart to identify dramatic spikes or drops in the amount of reads or writes for the volume.

If there are no dramatic spikes or drops in reads or writes, the storage administrator can confirm that I/O on the cluster is operating normally. Any performance issues might be on the network or the connected clients.

Identifying trends of I/O response time on cluster components

You can use Unified Manager to view the performance trends for all monitored cluster components for a volume workload. You can see, over time, which components have the highest usage, whether the highest usage is from read or write requests, and how the usage has impacted the workload response time.

Before you begin

- You must have the Operator, OnCommand Administrator, or Storage Administrator role.
- You must have identified the name of the volume or associated LUN you want to analyze.
- To display 30 days of performance statistics, Unified Manager must have collected and analyzed a minimum of 30 days of performance statistics from the cluster.

About this task

Identifying performance trends for the cluster components helps the administrator decide whether the cluster is being overused or underused.

If you are viewing the Event details page, you can click the name for a volume to go directly to the Performance/Volume Details page.

Steps

1. In the **Search** bar, type the name of the volume.

The name of the volume is displayed in the search results.

2. Click the name of the volume.

The volume is displayed on the Performance/Volume Details page.

3. On the Historic data chart, click **30d** to display the last 30 days of historical data.

4. Click **Break down data by**.

5. Under **Latency**, select **Cluster Components** and **Reads/writes latency**.

6. Click **Submit**.

Both charts are displayed below the Latency chart.

7. Review the **Cluster Components** chart.

The chart breaks down the total response time by cluster component. The response time at the aggregate is the highest.

8. Compare the **Cluster Components** chart to the **Latency** chart.

The Latency chart shows spikes in the total response time that are aligned with the spikes in response time for the aggregate. There are a few at the end of the 30-day time frame, where the performance threshold was crossed.

9. Review the **Reads/writes latency** chart.

The chart shows a higher response time for write requests than read requests, indicating that the client applications are waiting longer than usual to have their write requests fulfilled.

10. Compare the **Reads/writes latency** chart to the **Latency** chart.

The spikes in total response time that align with the aggregate in the Cluster Components chart also align with the writes in the Reads/writes latency chart. The administrator must decide whether the client applications using the workload must be addressed or whether the aggregate is being overused.

Analyzing the performance improvements achieved from moving a volume

You can use Unified Manager to investigate the impact of a volume move operation on the latency (response time) of other volumes on the cluster. Moving a high performing volume to a less busy aggregate or an aggregate with flash storage enabled allows the volume to perform more efficiently.

Before you begin

- You must have the Operator, OnCommand Administrator, or Storage Administrator role.
- You must have identified the name of the volume, or associated LUN, you want to analyze.
- Unified Manager must have collected and analyzed seven days of data.

About this task

Unified Manager identifies when a volume moves between aggregates. It can detect when the volume move is occurring, completed, or failed. The Performance/Volume Details page displays a change event icon for each state of the volume move, which helps you track when a move operation occurred and helps you determine whether it might have contributed to a performance event.

If you are viewing the Event details page, you can click the name of a volume to go directly to the Performance/Volume Details page.

Steps

1. In the **Search** bar, type the name of the volume.
2. Click the name of the volume.


The volume is displayed on the Performance/Volume Details page.

3. In the **Historic data** chart, adjust the sliders to show activity from the previous work week.
4. Analyze the **Latency** chart and the **IOPS** chart to see how the volume performed over the last few days.

Assume that you notice a consistent pattern of very high average response times of over 42 milliseconds per operation (ms/op), with performance events, each day of the week and decide to move the volume to a less busy aggregate to improve performance. Using OnCommand System Manager, you can move the volume to an aggregate with Flash Pool enabled for an increased performance boost. Approximately an hour after the volume move has been completed, you can return to Unified Manager to confirm that the

move operation was completed successfully and that the latency has improved.

5. If the **Performance/Volume Details** page is not displayed, search for the volume you want to view.
6. On the **Historic data** chart, click **1d** to view the activity from the last one day, a few hours since the volume move was completed.

At the bottom of the page, in the Events time line, a change event icon () is displayed to indicate the time that the volume move operation was completed. A black, vertical line is also displayed from the change event icon to the Latency chart.

7. Point your cursor to the change event icon to view details about the event in the **Events List**.

Because the volume moved to an aggregate with Flash Pool enabled, you can see the change in read and write I/O to cache.

8. On the **Break down data by** menu, under **MBps**, select **Cache hit ratio**.

The Cache hit ratio chart displays statistics about the reads and writes to cache.

The volume successfully moved to a less busy aggregate and the change event is highlighted in the Events List on the right. The average latency decreased significantly from over 42 ms/op to around 24 ms/op. The current latency is around 1.5 ms/op. In the Cache hit ratio chart, the amount of successful read and write hits to cache is now at 100% because the volume is now on an aggregate with Flash Pool enabled.

How moving a FlexVol volume works

Knowing how moving a FlexVol volume works helps you to determine whether the volume move satisfies service-level agreements and to understand where a volume move is in the volume move process.

FlexVol volumes are moved from one aggregate or node to another within the same storage virtual machine (SVM). A volume move does not disrupt client access during the move.

Moving a volume occurs in multiple phases:

- A new volume is made on the destination aggregate.
- The data from the original volume is copied to the new volume.

During this time, the original volume is intact and available for clients to access.

- At the end of the move process, client access is temporarily blocked.

During this time the system performs a final replication from the source volume to the destination volume, swaps the identities of the source and destination volumes, and changes the destination volume to the source volume.

- After completing the move, the system routes client traffic to the new source volume and resumes client access.

The move is not disruptive to client access because the time in which client access is blocked ends before clients notice a disruption and time out. Client access is blocked for 35 seconds by default. If the volume move operation cannot finish in the time that access is denied, the system aborts this final phase of the volume move

operation and allows client access. The system attempts the final phase three times by default. After the third attempt, the system waits an hour before attempting the final phase sequence again. The system runs the final phase of the volume move operation until the volume move is complete.

Performance/Volume Details page

This page provides detailed performance statistics for all I/O activity and operations for the selected FlexVol volume, FlexGroup volume, or FlexGroup constituent workload. You can select a specific time frame over which to view the statistics and events for the volume. The events identify performance events and changes that might be impacting I/O performance.

Historic data chart

Plots the historical performance analysis data for the selected volume. You can click and drag the sliders to specify a time frame. The sliders increase and decrease the time frame window. The data outside the time frame window is grayed out. You can use the slider at the bottom of the chart to move the time frame window across the historical data. The entire page, including the displayed charts and events, reflects the data available within the time frame window. Unified Manager retains a maximum of 30 days of historical data on this page.



On the historic data chart, if you select a time frame of more than 1 day, depending on your screen resolution, the charts display the maximum values for response time and IOPS across the number of days.

Options

• Time selector

Specifies the time range over which to view the volume performance statistics for the entire page. You can click 1 day (**1d**) through 30 days (**30d**), or click **Custom** to select a custom range. For a custom range, you can select a beginning and end date, and then click **Update** to update the entire page.



If you access the Performance/Volume Details page by clicking the name link of a volume on the Event details page, a time range, such as 1 day or 5 days prior to the current day, is automatically selected by default. When you move the slider in the historic data chart, the time range changes to a custom range, but the **Custom** time selector is not selected. The default time selector remains selected.

• Break down data by

Provides a list of charts you can add to the Performance/Volume Details page to display more detailed performance statistics for the selected volume.

Performance statistics displayed in the data breakdown charts

You can use the graphs to view performance trending for a volume. You can also view statistics for reads and writes, network protocol activity, the impact of QoS policy group throttling on latency, the ratio of reads and writes to cache storage, the total cluster CPU time used by a workload, and specific cluster components.

These views display a maximum of 30 days of statistics from the current day. On the historic data chart, if you select a time frame of more than 1 day, depending on your screen resolution, the charts display the maximum values for latency and IOPS across the number of days.



You can use the **Select All** check box to select, or deselect, all the listed chart options.

• Latency

The following charts detail the latency, or response time, information for the selected workload:

◦ Cluster Components

Displays a graph of the time spent at each cluster component used by the selected volume.

The chart helps you determine the latency impact by each component as it relates to the total latency. You can use the check box next to each component to show and hide its graph.

For QoS policy groups, data is only displayed for user-defined policy groups. Zeros are displayed for system-defined policy groups, such as default policy groups.

◦ Reads/writes latency

Displays a graph of the latencies of the successful read and write requests from the selected volume workload over the selected time frame.

Write requests are an orange line and read requests are a blue line. The requests are specific to the latency for the selected volume workload, not all workloads on the cluster.



The read and write statistics might not always add up to the total latency statistics displayed in the Latency chart. This is expected behavior based on how Unified Manager collects and analyzes read and write statistics for a workload.

◦ Policy Group Impact

Displays a graph of the percentage of the latency for the selected volume workload that is impacted by the throughput limit on its QoS policy group.

If the workload is throttled, the percentage indicates how much the throttling contributed to the latency at a specific point in time. The percentage values indicate the amount of throttling:

- 0% = no throttling
- > 0% = throttling
- > 20% = critical throttling If the cluster can handle more work, you can reduce throttling by increasing the policy group limit. Another option is to move the workload to a less busy aggregate.



The chart displays for workloads in a user-defined QoS policy group with a set throughput limit only. It does not display if the workloads are in a system-defined policy group, such as the default policy group, or a policy group that does not have a QoS limit. For a QoS policy group, you can point the cursor to the name of the policy group to display its throughput limit and the last time it was modified. If the policy group was modified before the associated cluster was added to Unified Manager, the last modified time is the date and time when Unified Manager first discovered the cluster.

- **IOPS**

The following charts detail the IOPS data for the selected workload:

- **Reads/writes/other**

Displays a graph showing the number of read and write IOPS and other IOPS, per second, over the selected time frame.

Other IOPS are protocol activities initiated by the client that are not reads or writes. For example, in an NFS environment, this could be metadata operations such as `getattr`, `setattr`, or `fsstat`. In a CIFS environment, this could be attribute lookups, directory listings, or antivirus scans. Write IOPS are an orange line and read requests are a blue line. The requests are specific to all operations for the selected volume workload, not all operations on the cluster.

- **MBps**

The following charts detail the throughput data for the selected workload:

- **Cache hit ratio**

Displays a graph of the percentage of read requests from client applications satisfied by cache over the selected time frame.

The cache could be on Flash Cache cards or solid state drives (SSDs) in Flash Pool aggregates. A cache hit, in blue, is a read from cache. A cache miss, in orange, is a read from a disk in the aggregate. The requests are specific to the selected volume workload, not all workloads on the cluster.

You can view more detailed information about volume cache usage in the Unified Manager Health pages and in OnCommand System Manager.

- **Components**

The following charts detail the data by cluster component used by the selected workload:

- **Cluster CPU Time**

Displays a graph of the CPU usage time, in ms, for all nodes in the cluster used by the selected workload.

The graph displays the combined CPU usage time for network processing and data processing. The CPU time for system-defined workloads that are associated to the selected workload, and are using the same nodes for data processing, is also included. You can use the chart to determine whether the workload is a high consumer of the CPU resources on the cluster. You can also use the chart, in combination with the Reads/writes latency chart under the Latency chart, or the Reads/writes/other chart under the IOPS chart, to determine how changes to workload activity over time impact cluster CPU utilization.

- **Disk Utilization**

Displays a graph showing the percentage of utilization on the data disks in the storage aggregate over the selected time frame.

The utilization includes disk read and write requests from the selected volume workload only. Reads from cache are not included. The utilization is specific to the selected volume workload, not all workloads on the disks. If a monitored volume is involved in a volume move, the utilization values in

this chart are for the target aggregate to which the volume moved.

How graphs of performance data work

Unified Manager uses graphs or charts to show you volume performance statistics and events over a specified period of time.

The graphs enable you to customize the range of time for which to view data. The data is displayed with the time frame on the horizontal axis of the graph and the counters on the vertical axis, with point intervals along the graph lines. The vertical axis is dynamic; the values adjust based on the peaks of the expected or actual values.

Selecting time frames

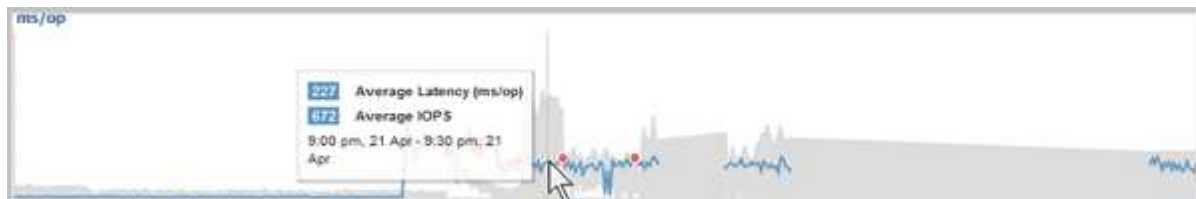
On the Performance/Volume Details page, the Historic data chart enables you to select a time frame for all graphs on the page. The 1d, 5d, 10d, and 30d buttons specify 1 day through 30 days (1 month) and the **Custom** button enables you to specify a custom time range within that 30 days. Each point on a graph represents a 5-minute collection interval, and a maximum of 30 days of historical performance data is retained. Note that intervals also account for network delays and other anomalies.



In this example, the Historic data chart has a time frame set to the beginning and the end of the month of March. In the selected time frame, all historic data before March is grayed out.

Viewing data point information

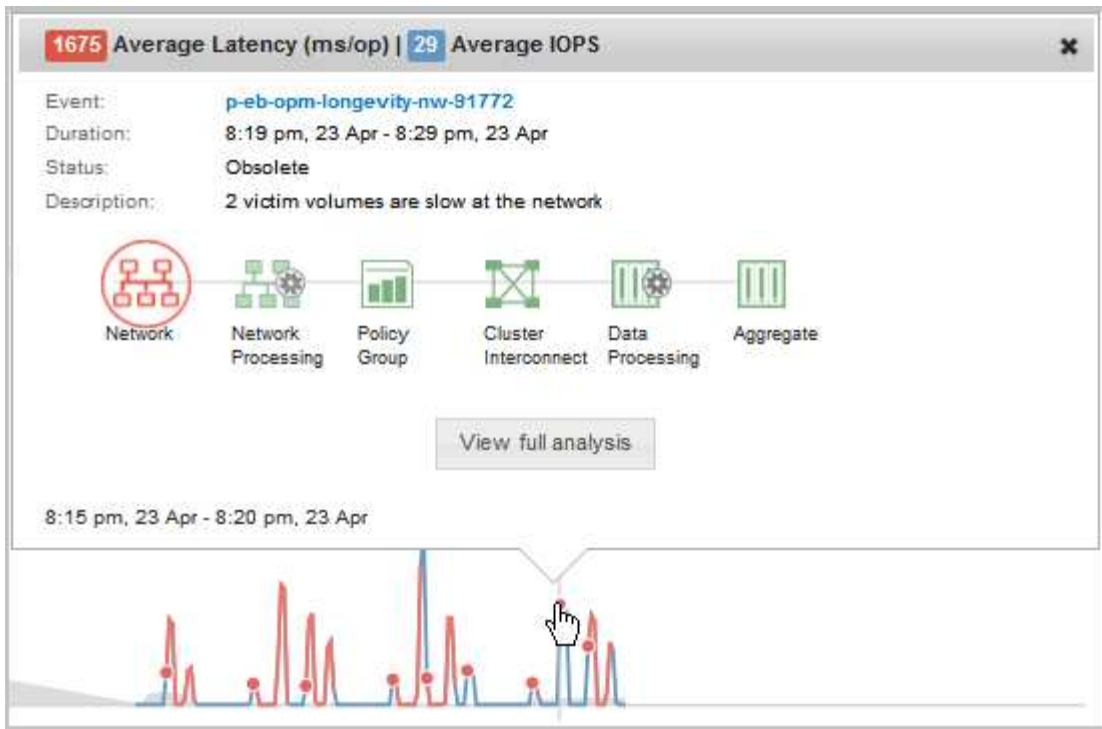
To view data point information on a graph, you can position the cursor over a specific point within the graph, and a pop-up box displays listing the value and date and time information.



In this example, positioning the cursor over the IOPS chart on the Performance/Volume Details page displays the response time and operations values between 3:50 a.m. and 3:55 a.m. on October 20th.

Viewing performance event information

To view event information on a graph, you can position the cursor over an event icon to view summary information in a pop-up box, or you can click the event icon for more detailed information.



In this example, on the Performance/Volume Details page, clicking an event icon on the Latency chart displays detailed information about the event in a pop-up box. The event is also highlighted in the Events List.

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.