TRP NetApp Performance Analysis

Version 2.0
September 2011

Prepared by:
Michael Arndt
arndt@netapp.com

Abstract
This paper documents the methodologies used for NetApp storage system performance analysis within the Thomson Reuters Professional (TRP) environment. NetApp best practices, configuration specific to the TRP storage environment, and common use cases for a variety of tools used in doing performance analysis are covered.

TABLE OF CONTENTS

1 INTRODUCTION

This paper documents the methodologies used for NetApp storage system performance analysis within the Thomson Reuters Professional (TRP) environment. NetApp best practices, configuration specific to the TRP storage environment, and common use cases for a variety of tools used in doing performance analysis are covered. This document assumes that the reader is familiar with basic NetApp storage management tools such as Operations Manager (OM), NetApp Management Console (NMC), Performance Advisor (PA), and NetApp Perfstat. The document also assumes familiarity with basic NetApp storage components such as controllers, aggregates, flexible volumes (flexvols), protocol operations, throughput, and latency. The configurations are documented as they should be configured for the Thomson Reuters Professional (TRP) environment in Eagan, MN.

1.1 Conventions

The following conventions are used throughout this document:

> *NMC -> list-> list*
> The convention above indicates a list of links to click on in order to accomplish a task using the Performance Advisor interface in the Network Management Console software that is available in conjunction with Operations Manager.
>
> *om-cli#*
> The convention above indicates a command to run on an Operations Manager server, as the root user.
>
> *om-cli>*
> The convention above indicates a command to run on an Operations Manager server, as a non-root user

1.2 Storage performance analysis concepts

Most storage performance analysis exercises will be driven by one of two scenarios:

1. A performance or stability problem has been reported somewhere within an application stack, and storage is being investigated to determine if it is performing as designed.
2. There is a question about how much additional workload can be added to a particular storage controller without exceeding the performance capabilities of the system.

In scenario #1, the main metric of interest will be the response time of the storage system, also known as latency. If the latency is within the boundaries of what is typically deemed to be acceptable, then it is unlikely that the storage system is a root cause for the performance or stability problem being investigated. If the latency is higher than the boundary of what is typically deemed to be acceptable, then the storage system may be contributing to the performance or stability problem being investigated.

The table below shows some commonly accepted values for latency metrics. If these values are not breached, this typically indicates a properly performing storage system. Some applications may require lower latency values in order to operate as designed, and some applications may tolerate much large latencies, but latencies below these values will typically be indicative of a well performing storage system.

Metric Type	**Read/Other Latency**	**Write Latency**
Metric Value	**Under 20ms**	**Under 5ms**

Higher than expected values for other performance metrics such as the CPU utilization on the storage system, network throughput, disk utilization, etc, are typically not indicators of a storage performance problem unless they are also accompanied by high latency.

In scenario #2, the first step is to determine the typical I/O profile of the workloads that are seen on the storage system. An I/O profile can be summarized as consisting of the following components:

- The number of IOPs seen during most of the I/O peaks over a given period of time.
- The average I/O size. This can be calculated by dividing the throughtput (in units of kilobytes/second) by the number of IOPs.
- The ratio of reads versus writes.

Once the I/O profile is established, it can be compared against a known maximum desired IOPs value for the profile to determine how much additional workload can be added without exceeding the performance capabilities of the system. The maximum desired IOPs value for an I/O profile should be designed such that the system can handle the workload at an acceptable latency. The storage system may be able to handle higher IOPs values, but this will typically be done at a higher latency than desired.

Section 6.1 contains tables with maximum desired IOPS values based on I/O profile types and storage controller platforms.

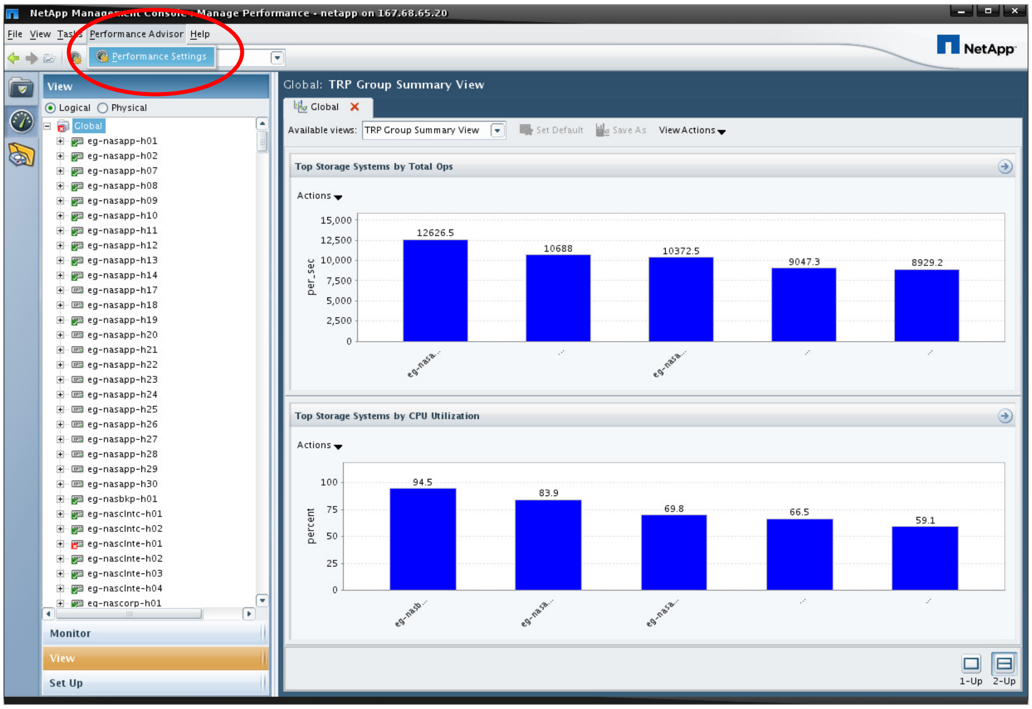
1.3 Storage performance analysis tools

The remainder of this document will focus on a variety of storage performance analysis tools available from NetApp, as well as the configurations in place in the TRP environment. The most commonly used tool for day to day performance analysis of NetApp storage systems will be the Performance Advisor (PA) component of NetApp Operations Manager (OM). PA is accessed via the NetApp Management Console (NMC) client that comes bundled with OM. Other tools such as NetApp Perfstat and the onboard performance archiver will also be discussed.

2 Using Performance Advisor (PA)

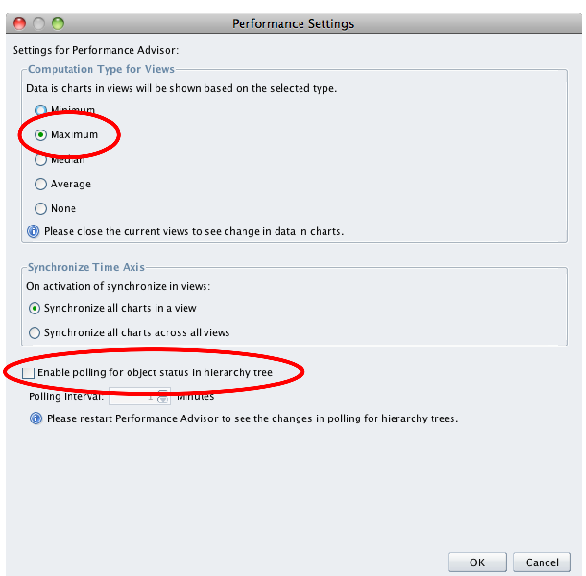
2.1 Recommended PA configuration settings

PA configuration settings can be accessed via *Performance Advisor -> Performance Settings*, as shown below.



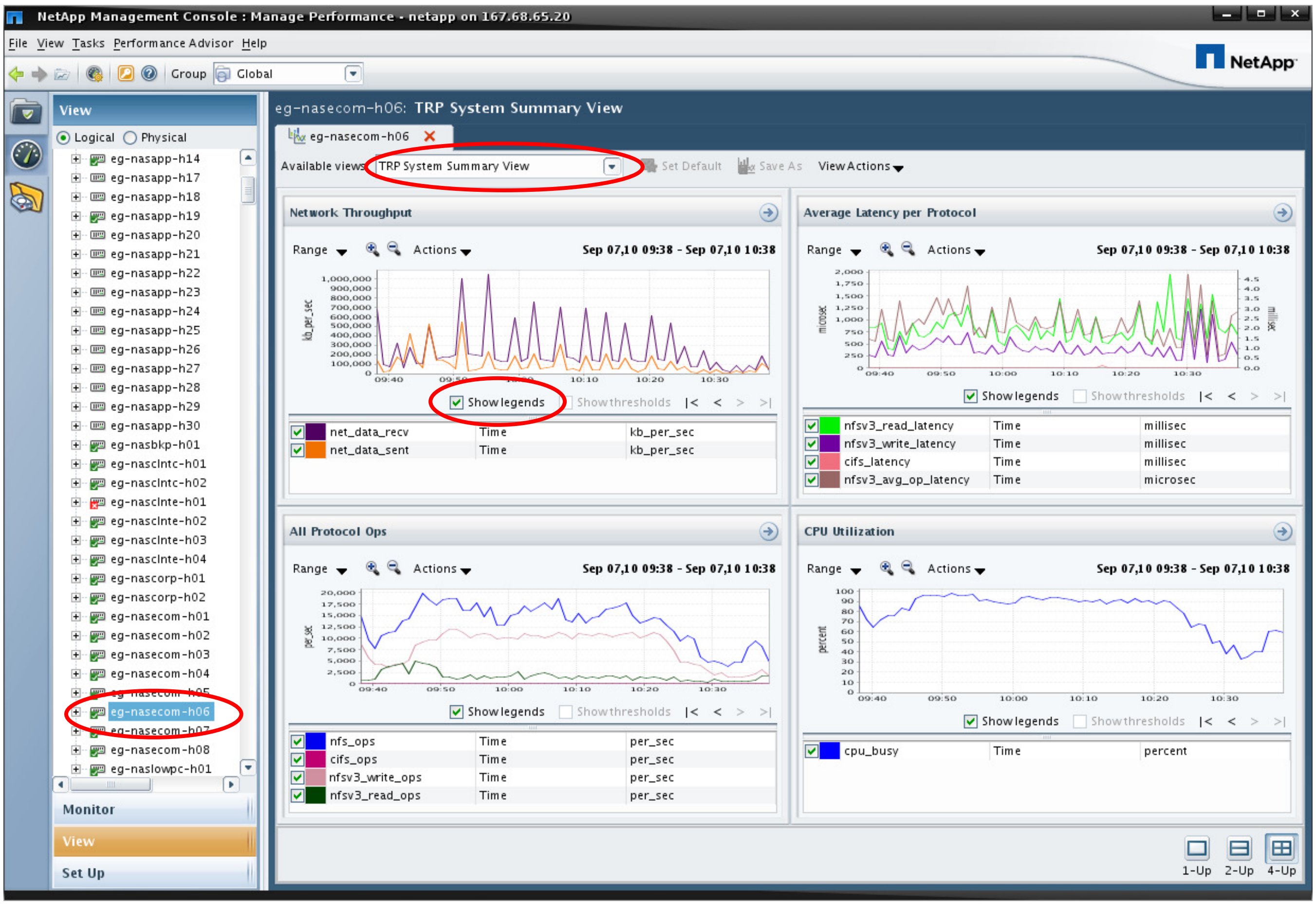
There are two configurations settings that we recommend to be modified for the TRP environment:

- Change the *Computation Type for Views* from *Average* to *Maximum*. This setting controls how PA will perform consolidation of data points in graphs when viewing data across longer periods of times, such as days or weeks. Since we typically want to see the peaks when doing troubleshooting and sizing with PA, it is highly recommended to change this setting to *Maximum*.
- Uncheck the option for *Enable polling for object status in hierarchy tree*. This is required to avoid NetApp bug 461282, which can cause significant performance degradation if polling is enabled on servers managing a large number of objects. This setting is only available with NMC 3.0D2.



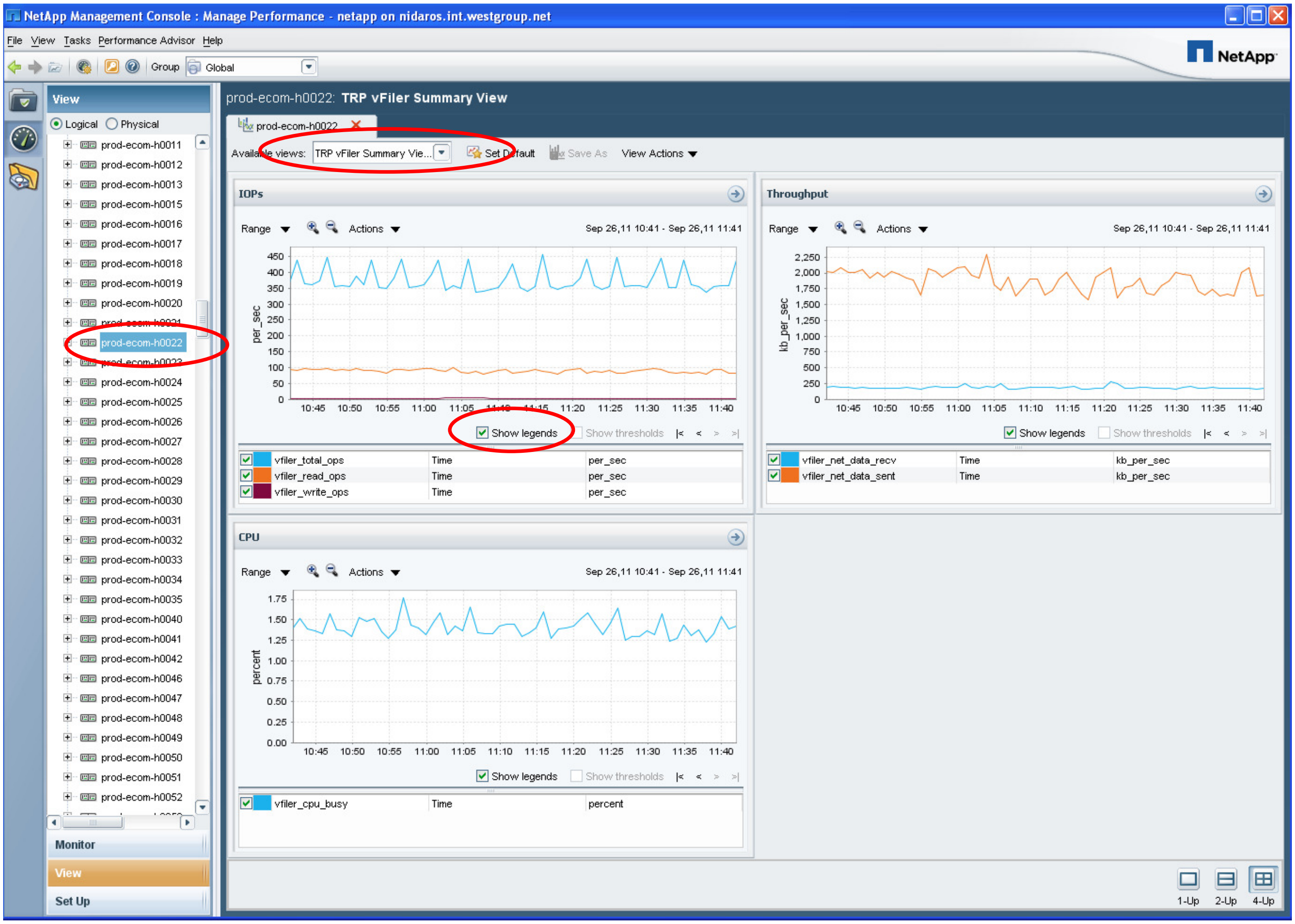
2.2 Analyzing overall storage controller performance

The most commonly used metrics at the storage controller level can be seen in the *TRP Storage System Summary View,* which is a custom performance view from within PA. This view will come up by default when you select a storage controller in the *NMC-> PA* interface. This view will allow to user to obtain the metrics required to calculate the I/O profile as described in section 1.2, view overall system latency metrics, and see system CPU utilization. Notice that by selecting the *Show legends* checkbox you are able to see the units for the counters in each graph, which is especially useful when there are counters with different units in the same graph. An example of this view is shown below:



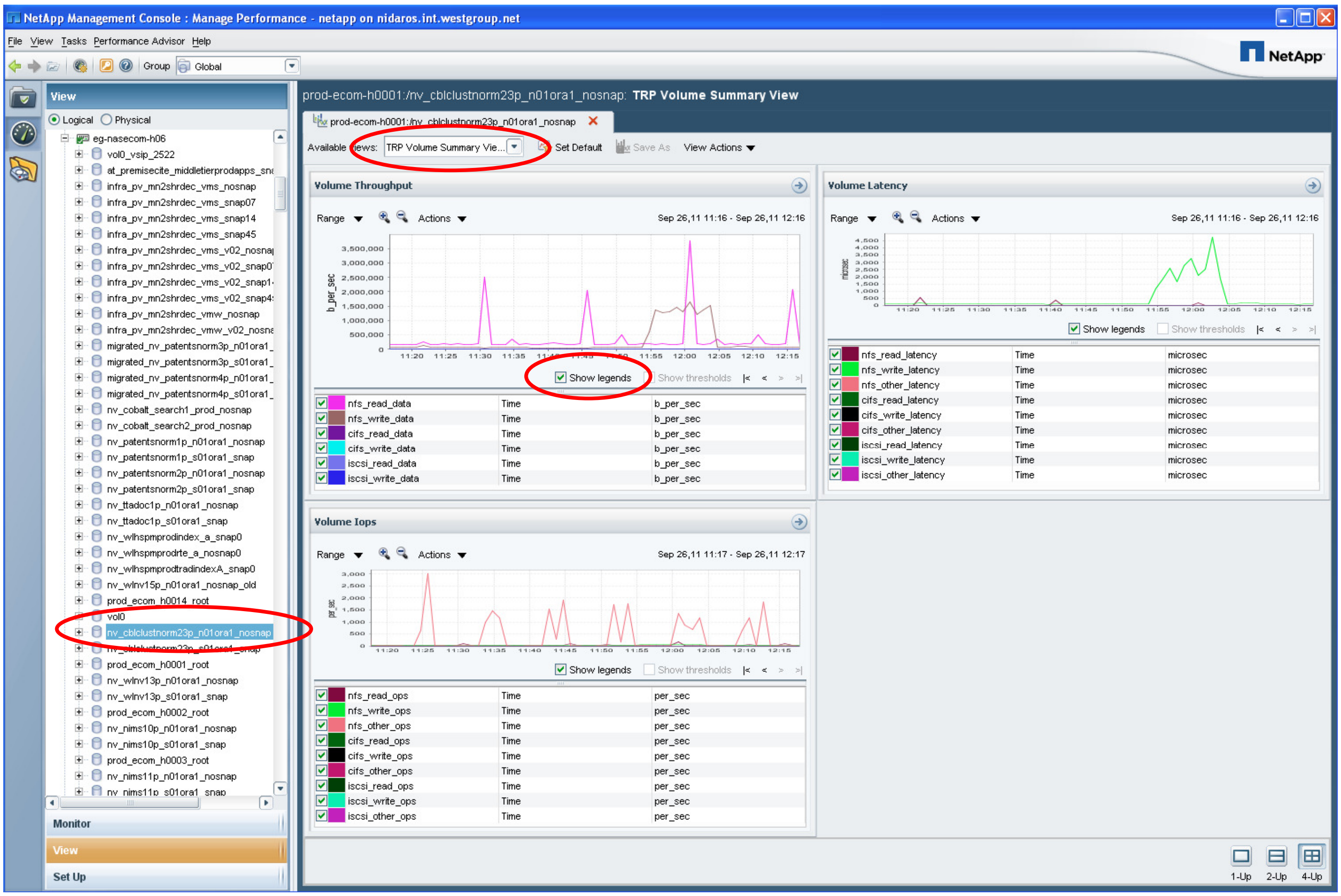
2.3 Analyzing vFiler level performance

The metrics required to analyze performance on a per vFiler level can be seen in the *TRP vFiler Summary View*, which is a custom performance view in PA. To see the vFiler level views, select the vFiler of interest from the list of vFilers in the left hand *View* pane, and then select the *TRP vFiler Summary View* from the *Available views* dropdown. An example of this view is shown below.



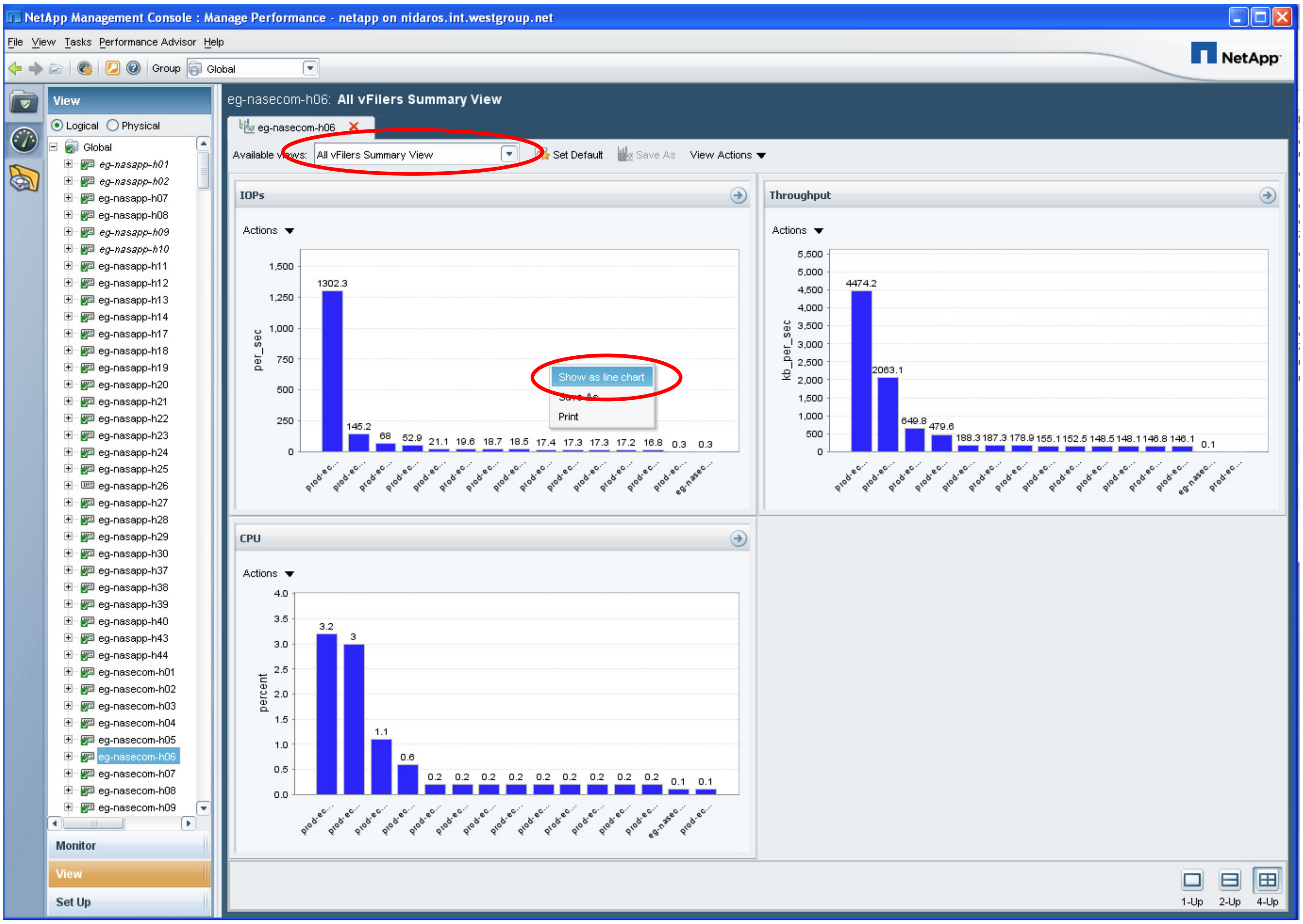
2.4 Analyzing volume level performance

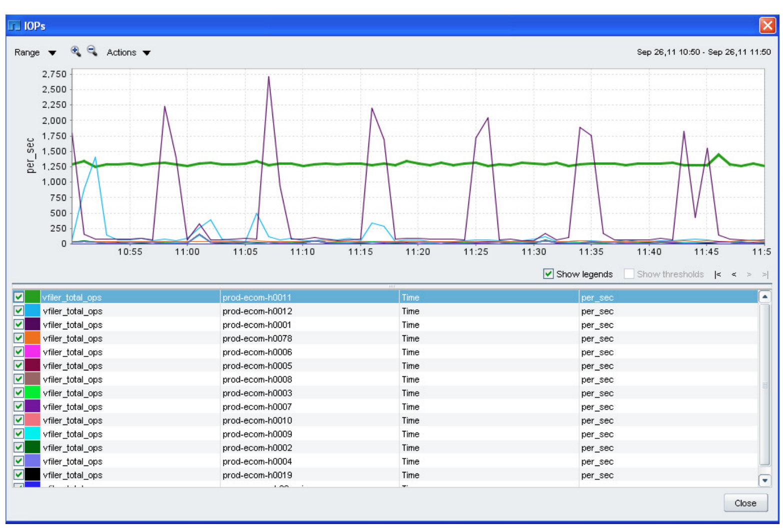
The metrics required to analyze performance on a per volume level can be seen in the *TRP Volume Summary View*, which is a custom performance view in PA. To see the volume level views, expand the list of volumes by clicking on the + sign next to the storage controller from within the PA interface and click on the volume of interest. Then select the *TRP Volume Summary View* from the *Available views* dropdown. This view will allow to user to obtain the metrics required to calculate the I/O profile on a per volume bases, as described in section 1.2, as well as view volume specific latency metrics. An example of this view is shown below. Note that while this view contains counters that are specific to each protocol, most volumes are only used with a single protocol and therefore can have a number of counters de-selected when viewing.



2.5 Analyzing the performance of all vFilers on a controller

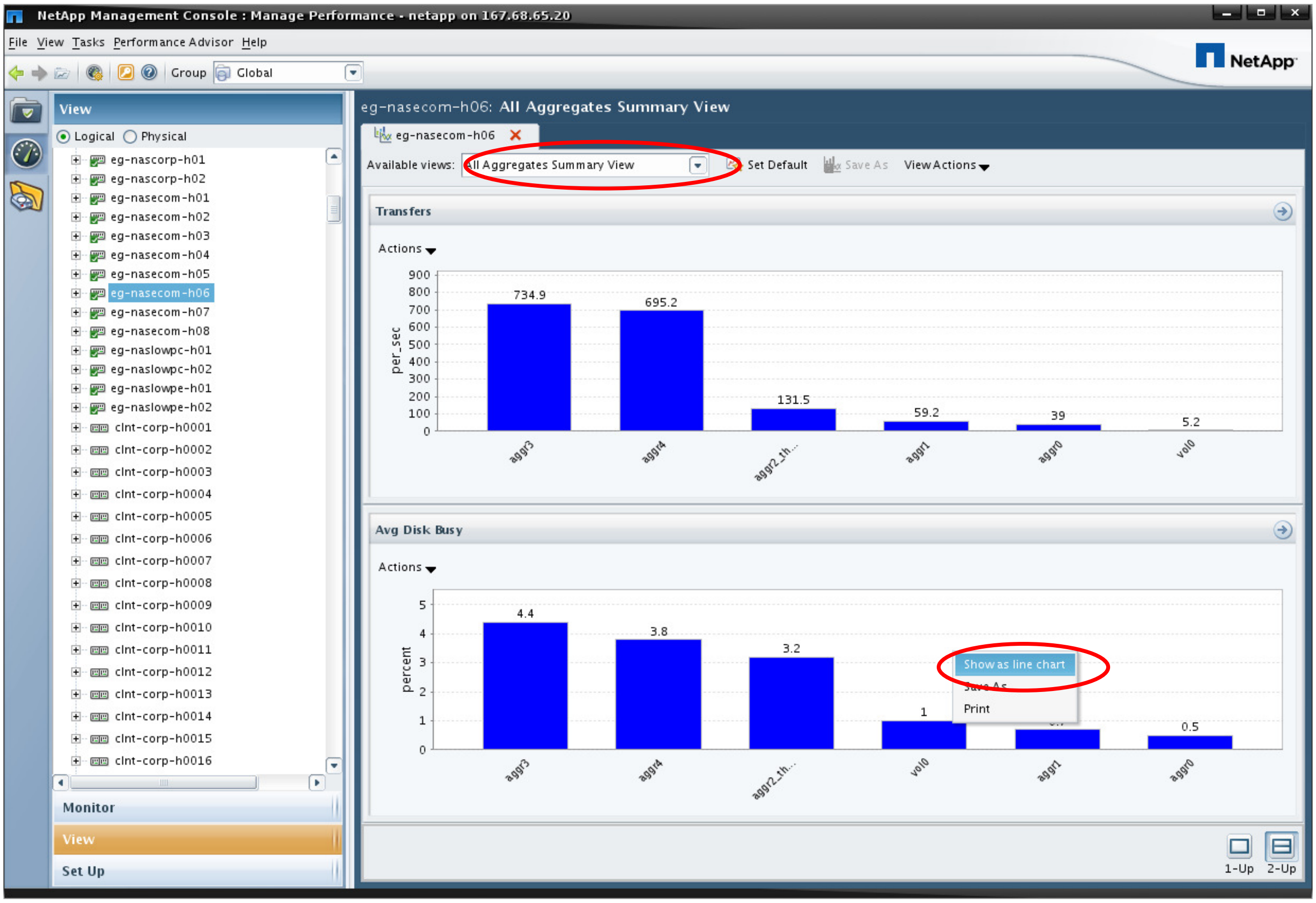
In some situations, it is useful to determine which vFilers on a storage controller are doing the most I/O at any given time. In this situation, the *All vFilers Summary View* can be selected from the *Available views* dropdown after selecting the storage controller of interest from the PA interface. The graphs in this view show a ranking of the busiest vFilers on the system over the past 10 minutes. You will typically want to see a longer history of this data, and that can be accomplished by right clicking on the bar chart and selecting the *show as line chart* option. An example of this view, and the corresponding line chart, is shown below.

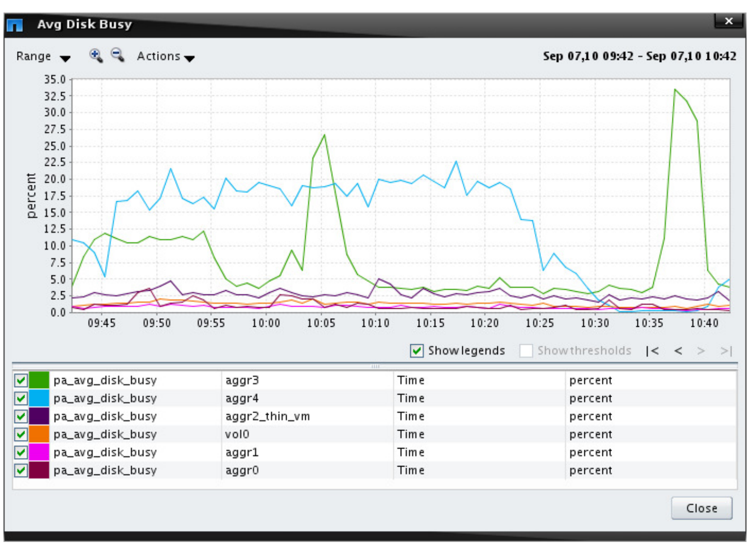




2.6 Analyzing the performance of all aggregates on a controller

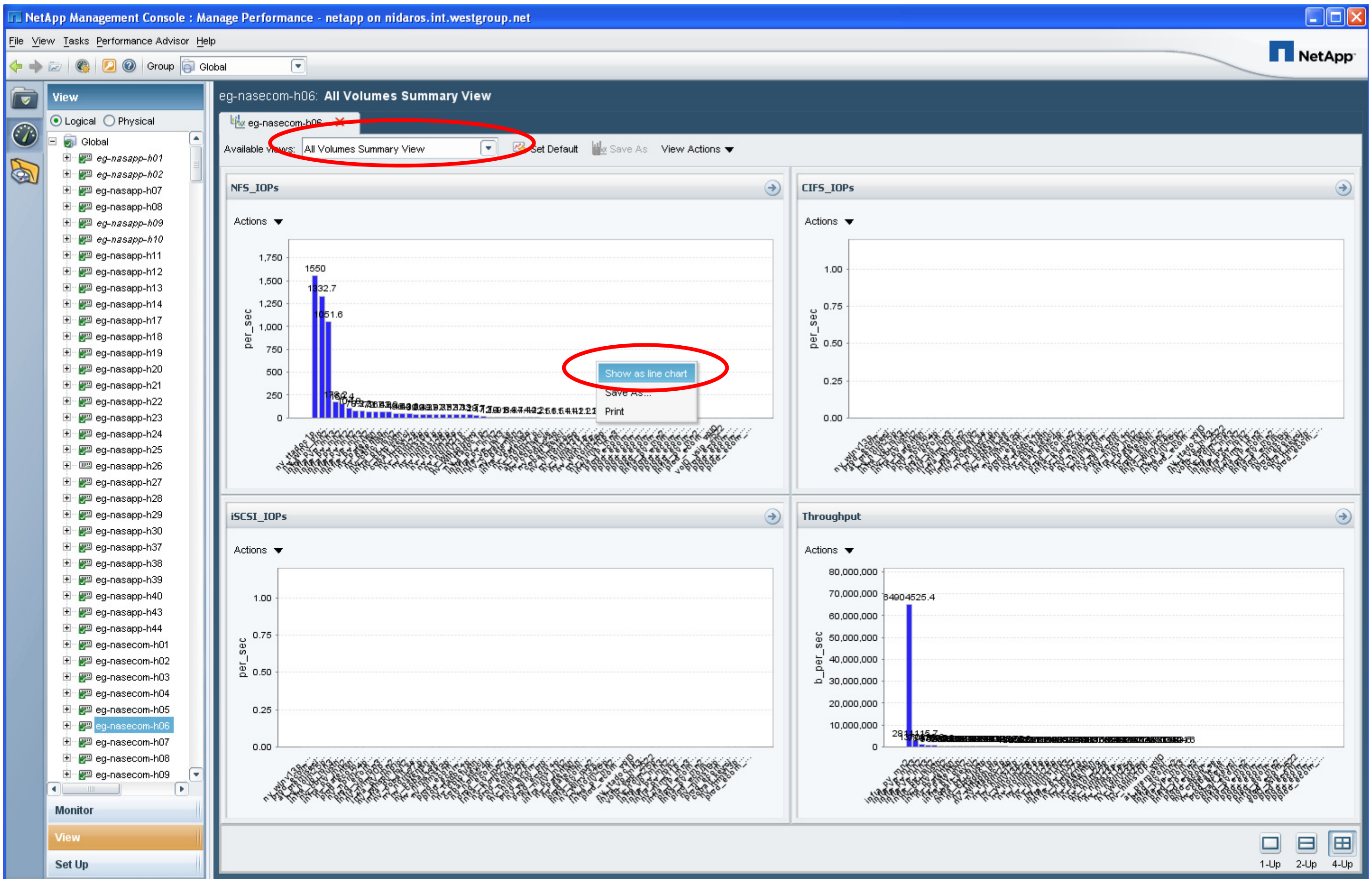
In some situations, it is useful to determine how busy a given aggregate of a storage system is, as compared to the other aggregates on that system. In this situation, the *All Aggregates Summary View* can be selected from the *Available views* dropdown after selecting the storage controller of interest from the PA interface. The graphs in this view show a ranking of the busiest aggregates on the system over the past 10 minutes. Typically, you will want to see a longer history of this data, and that can be accomplished by right clicking on the bar chart and selecting the *show as line chart* option. An example of this view, and the corresponding line chart, is shown below:

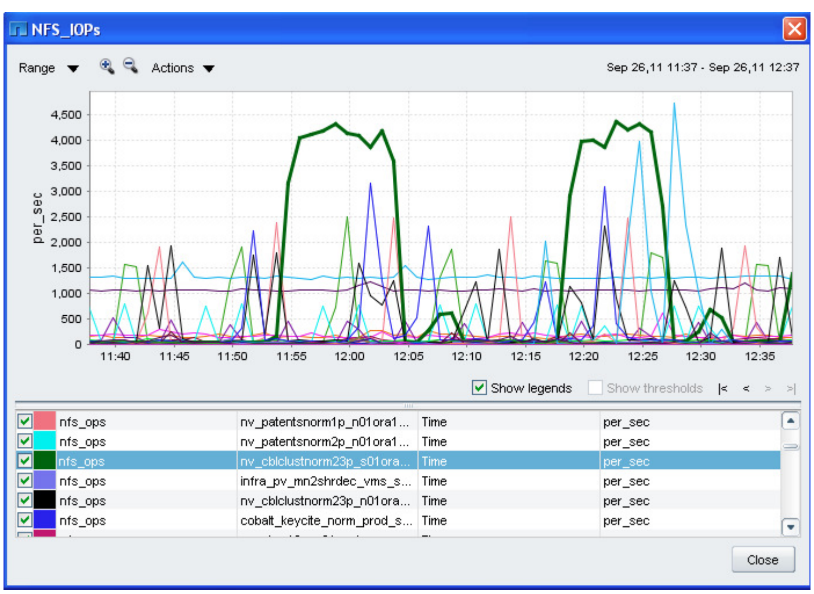




2.7 Analyzing the performance of all volumes on a controller

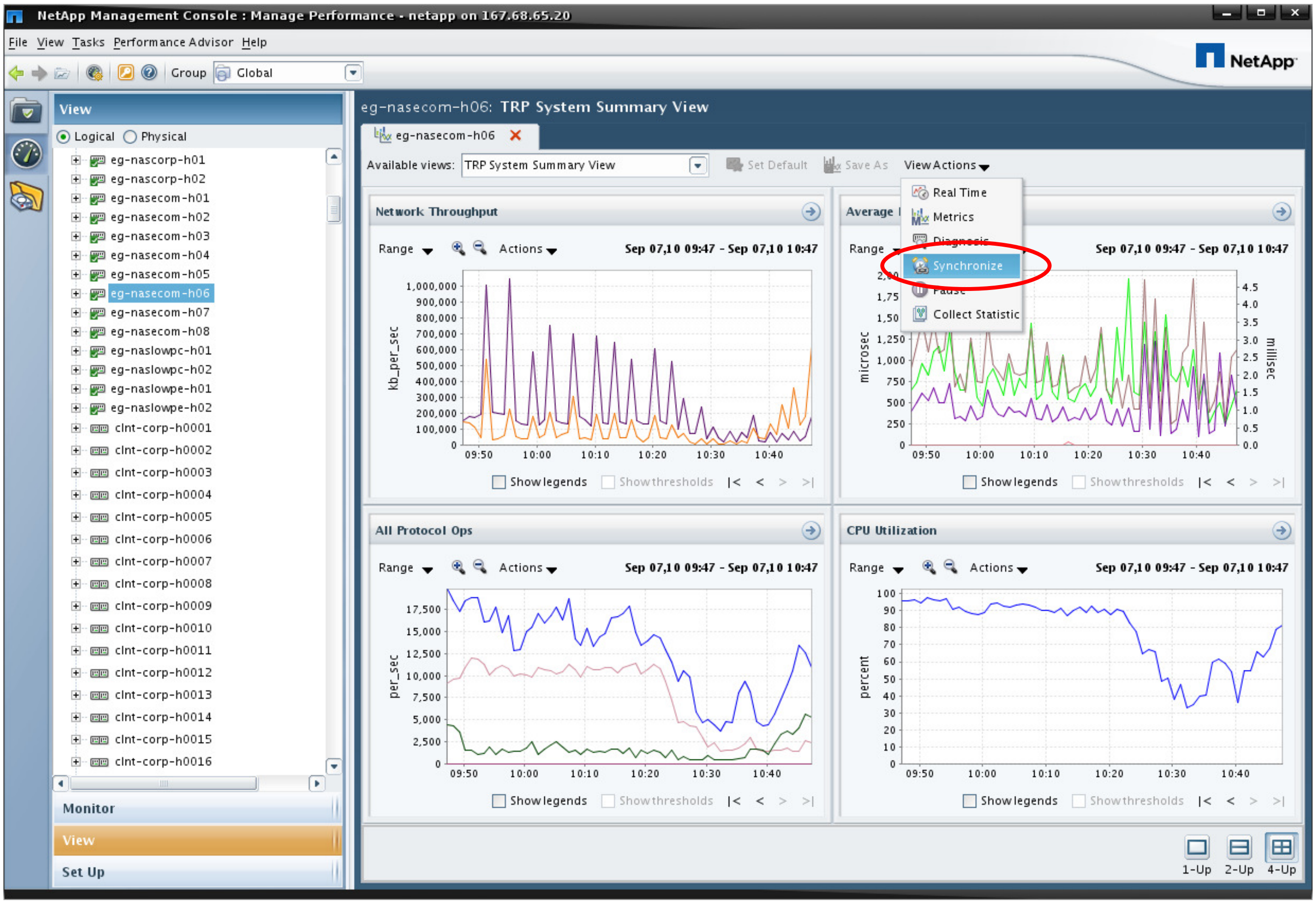
Similarly, it is often useful to determine which volumes on a storage controller are doing the most I/O at any given time. In this situation, the *All Volumes Summary View* can be selected from the *Available views* dropdown after selecting the storage controller of interest from the PA interface. The graphs in this view show a ranking of the busiest volumes on the system over the past 10 minutes, with the IOPs counters broken out by protocol. Again, you will typically want to see a longer history of this data, and that can be accomplished by right clicking on the bar chart and selecting the *show as line chart* option. An example of this view, and the corresponding line chart, is shown below. Notice that items selected in the legend get highlighted.



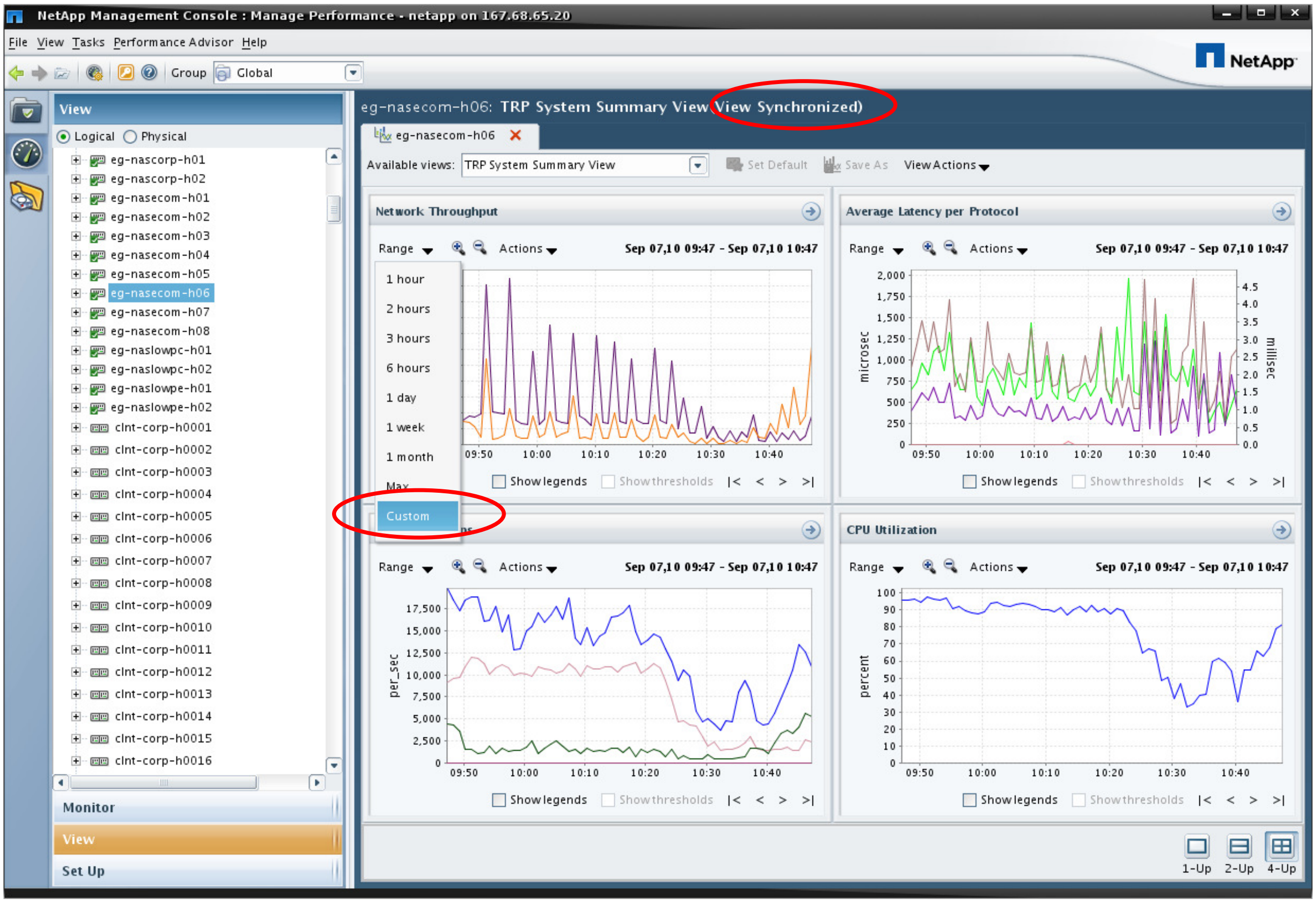


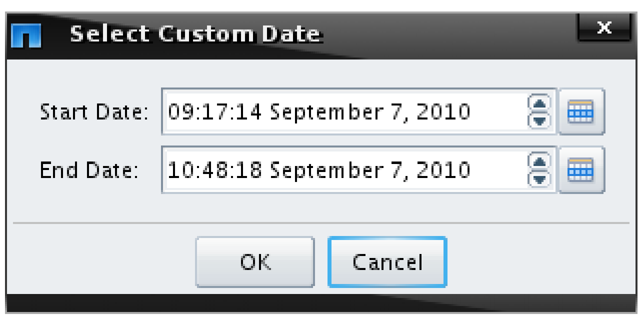
2.8 Helpful navigation tips for the PA interface

For PA views that show multiple line chart graphs in a single tab, it is often useful to synchronize the date range so that as you modify the range of one graph, the other graphs in the view update to the same range. This can be easily accomplished by clicking on the *Synchronize* option in the *View Actions* dropdown.



Whether or not you have your graphs synchronized, the range drop down can then be used to modify the date range displayed in the graph. While you can use any of the preset ranges, or use your mouse to select a portion of a graph already displayed, it is often useful to go directly to a date range of interest when you are troubleshooting an incident. In this case, select the custom option in the range dropdown, as shown below.





3 Performance related configurations in Operations Manager

3.1 Performance threshold settings

The TRP environment is configured with a number of performance threshold templates that are designed to track and alert on performance related concerns. The following tables lists the threshold settings in place on the TRP Operations Manager servers for read and write NFS latencies on Low Tier (LT) and High Tier (HT) storage systems.

Threshold Name	Threshold Value	Threshold Time	Automated Perfstat?
sys_nfs_read_latency_LT	40ms	240 seconds	No
sys_nfs_write_latency_LT	20ms	240 seconds	No
sys_nfs_read_RedAlert_LT	50ms	240 seconds	Yes
sys_nfs_write_RedAlert_LT	30ms	240 seconds	Yes
sys_nfs_read_latency_HT	30ms	180 seconds	No
sys_nfs_write_latency_HT	20ms	180 seconds	No
sys_nfs_read_RedAlert_HT	45ms	180 seconds	Yes
sys_nfs_write_RedAlert_HT	25ms	180 seconds	Yes

The "Automated Perfstat" column in the table above references the features implemented as described in section 3.3 of this document. These threshold settings can be modified in the PA interface by going to the *NMC->PA->Set Up->Threshold Templates* section.

3.2 Performance threshold breach summary report

In addition to obtaining email based alerts when performance thresholds are breached, it is useful to trend the number of threshold breaches over time. This trending data can be used to determine which systems are regularly breaching their performance thresholds, and therefore may require attention in order to reduce the workload on the systems. This data should also be consulted when provisioning new storage requests, so that new workload is not placed on systems that are already overloaded. The summary report can be found at http://nerstrand.int.westgroup.net/netapp/dfmperfeventreport/.

3.3 Automated NetApp Perfstat data gathering

While the data presented by PA can often be used to easily characterize and troubleshoot storage system performance, more detailed data is typically required when engaging the NetApp global support center for assistance with performance related incidents. The Netapp Perfstat tool is a Unix shell script that can be used to gather extremely detailed performance data that can be sent to the Netapp global support center to be analyzed in an easily repeatable fashion. Since gathering Perfstat data consumes system resources on the storage controllers, as well as a significant amount of disk space on the client systems that it is run from, it is recommended to typically gather Perfstat only as needed.

In the TRP environment, we have automated the Perfstat data gathering to start when the *RedAlert* performance thresholds are breached, as described in section 3.1. This is done via a shell script on the OM servers *(/dfm/netapp/perfstat/dfm_perfstat.sh)* that is attached to the alarms tied to the latency thresholds that are labeled with *RedAlert* in the threshold event name. See the comments in the *dfm_perfstat.sh* script for details on how to attach this script to an existing alarm that is triggered by a performance threshold breach.

The script is setup to gather Perfstat data only if a folder named after the storage system for which a threshold breach was detected is created in */dfm/netapp/perfstat/dfm_perfstat* on the OM server in question. For example, to enable this functionality for a controller named eg-nas-a01, simply create a directory named */dfm/netapp/perfstat/dfm_perfstat/eg-nas-a01* on the OM server that manages eg-nas-a01. Once this configuration is in place, administrators can check in this directory to see if a Perfstat was automatically gathered for the storage system of interest during an incident timeframe that they are investigating.

The Perfstat data gathered as a result of this process is taken only after a latency threshold is breached. The resulting Perfstat data will be gathered in 2 intervals of 2 minutes each. Since the latency thresholds must be breached for at least 3 minutes before an event is triggered, this method of gathering Perfstat data is typically only useful for performance incidents that persist for at least 5 minutes. For performance incidents that are of a shorter duration, manual Perfstat data collection must be configured per the instructions in section 5.3.

3.4 Custom PA views

The following views are not present in PA by default, but rather were created as custom views for the TRP environment: A description of the views, via a standard OM CLI command, is included.

TRP Storage System Summary View

om-cli> dfm perf view describe "TRP System Summary View"

```
View Name: TRP System Summary View
Applies To: Object type (filer)

Chart Details:

Chart Name: Network Throughput
Chart Type: simple chart
Counters in this Chart:
Counter: system:net_data_recv
Counter: system:net_data_sent

Chart Name: Average Latency per Protocol
Chart Type: simple chart
Counters in this Chart:
Counter: nfsv3:nfsv3_read_latency
Counter: nfsv3:nfsv3_write_latency
Counter: cifs:cifs_latency
Counter: nfsv3:nfsv3_avg_op_latency

Chart Name: All Protocol Ops
Chart Type: simple chart
Counters in this Chart:
Counter: system:nfs_ops
Counter: system:cifs_ops
Counter: nfsv3:nfsv3_write_ops
Counter: nfsv3:nfsv3_read_ops

Chart Name: CPU Utilization
Chart Type: simple chart
Counters in this Chart:
Counter: system:cpu_busy
```

TRP vFiler Summary View

om-cli> dfm perf view describe "TRP vFiler Summary View"

```
View Name: TRP vFiler Summary View
Applies To: Object type (vfiler)

Chart Details:

Chart Name: Throughput
Chart Type: simple chart
Counters in this Chart:
Counter: vfiler:vfiler_net_data_recv
Counter: vfiler:vfiler_net_data_sent

Chart Name: CPU
Chart Type: simple chart
Counters in this Chart:
Counter: vfiler:vfiler_cpu_busy

Chart Name: IOPs
Chart Type: simple chart
Counters in this Chart:
Counter: vfiler:vfiler_total_ops
Counter: vfiler:vfiler_read_ops
Counter: vfiler:vfiler_write_ops
```

TRP Volume Summary View

om-cli> dfm perf view describe "TRP Volume Summary View"

```
View Name: TRP Volume Summary View
Applies To: Object type (volume)

Chart Details:

Chart Name: Volume Throughput
Chart Type: simple chart
Counters in this Chart:
Counter: volume:nfs_read_data
Counter: volume:nfs_write_data
Counter: volume:cifs_read_data
Counter: volume:cifs_write_data
Counter: volume:iscsi_read_data
Counter: volume:iscsi_write_data

Chart Name: Volume Latency
Chart Type: simple chart
Counters in this Chart:
Counter: volume:nfs_read_latency
Counter: volume:nfs_write_latency
Counter: volume:nfs_other_latency
Counter: volume:cifs_read_latency
Counter: volume:cifs_write_latency
Counter: volume:cifs_other_latency
Counter: volume:iscsi_read_latency
Counter: volume:iscsi_write_latency
Counter: volume:iscsi_other_latency

Chart Name: Volume Iops
Chart Type: simple chart
Counters in this Chart:
Counter: volume:nfs_read_ops
Counter: volume:nfs_write_ops
Counter: volume:nfs_other_ops
Counter: volume:cifs_read_ops
Counter: volume:cifs_write_ops
Counter: volume:cifs_other_ops
Counter: volume:iscsi_read_ops
Counter: volume:iscsi_write_ops
Counter: volume:iscsi_other_ops
```

All vFilers Summary View

om-cli> dfm perf view describe "All vFilers Summary View"

```
View Name: All vFilers Summary View
Applies To: Object type (filer)

Chart Details:

Chart Name: IOPs
Chart Type: bar
Number of object instances: All
Top or Bottom Instances: Top
Counters in this Chart:
Counter: vfiler:vfiler_total_ops

Chart Name: Throughput
Chart Type: bar
Number of object instances: All
Top or Bottom Instances: Top
Counters in this Chart:
Counter: vfiler:vfiler_nw_throughput

Chart Name: CPU
Chart Type: bar
Number of object instances: All
Top or Bottom Instances: Top
Counters in this Chart:
Counter: vfiler:vfiler_cpu_busy
```

All Aggregates Summary View

om-cli> dfm perf view describe "All Aggregates Summary View"

```
View Name: All Aggregates Summary View
Applies To: Object type (filer)

Chart Details:

Chart Name: Transfers
Chart Type: bar
Number of object instances: All
Top or Bottom Instances: Top
Counters in this Chart:
Counter: aggregate:total_transfers

Chart Name: Avg Disk Busy
Chart Type: bar
Number of object instances: All
Top or Bottom Instances: Top
Counters in this Chart:
Counter: aggregate:pa_avg_disk_busy
```

All Volumes Summary View

om-cli> dfm perf view describe "All Volumes Summary View"

```
View Name: All Volumes Summary View
Applies To: Object type (filer)

Chart Details:

Chart Name: Throughput
Chart Type: bar
Number of object instances: All
Top or Bottom Instances: Top
Counters in this Chart:
Counter: volume:throughput

Chart Name: NFS_IOPs
Chart Type: bar
Number of object instances: All
Top or Bottom Instances: Top
Counters in this Chart:
Counter: volume:nfs_ops

Chart Name: CIFS_IOPs
Chart Type: bar
Number of object instances: All
Top or Bottom Instances: Top
Counters in this Chart:
Counter: volume:cifs_ops

Chart Name: iSCSI_IOPs
Chart Type: bar
Number of object instances: All
Top or Bottom Instances: Top
Counters in this Chart:
Counter: volume:san_ops
```

3.5 Configuration settings for custom views

Some of the counters used in the custom views described in this document require the *perfAdvisorShowDiagCounters* option to be enabled on the OM servers, which is done as follows:

om-cli> dfm options set perfAdvisorShowDiagCounters=Enabled

The following option was set on all OM servers in order to capture all volumes on a system in the custom *All Volumes Summary View*, rather than the default of the top 20 volumes:

om-cli> dfm options set perfMaxObjectInstancesInBarChart=500

4 Granular performance statistics

4.1 Configuration and use cases for the granular performance statistics

All storage systems running ONTAP 7.3 and above have a granular performance statistics data collection process enabled via a custom script on the OM servers. This data collection process runs via the following cron job entry on each OM server, and collects data in 5 second intervals. The second cron job listed below prunes the data files so that only 30 days worth of granular performance statistics are kept online:

```
15 * * * * netapp /dfm/netapp/scripts/stats_harvester.sh >> /dfm/netapp/scripts/stats_harvester.log 2>&1
5 5 * * 0 netapp find /dfm/netapp/scripts/stats -mtime +30 -exec rm { } \; >/dev/null 2>&1
```

At this time, the granular performance data is not meant to be used for full troubleshooting of performance cases with the NetApp global support center. This process has been created as a custom tool for TRP in order to aid with performance troubleshooting in use cases which require knowledge of a systems behavior during very short intervals of time (under 1 minute). Currently, NetApp Perfstat will continue to be the data gathering tool of choice when engaging the NetApp global support center to assist with troubleshooting performance issues.

4.2 Locating granular performance statistics

The granular performance statistics are stored in the /dfm/netapp/scripts/stats directory on each of the OM servers. Within this directory are subdirectories corresponding to each of the storage controllers being managed by the OM server. The files in these subdirectories each contain 1 hour of performance data each using 5 second granularity. The files are time/date stamped based on the hour in which they started recording data.

4.3 Processing granular performance statistics

The */dfm/netapp/scripts/stats_show.pl* script can be used to create CSV output from the granular performance statistics files located in */dfm/netapp/scripts/stats/<controller_name>*. The following command examples show how to use the stats_show.pl script.

- Show controller level statistics for eg-nasclnt-e01 on 2011-09-26 from 10:15am to 11:15am:
  > /dfm/netapp/scripts/stats_show.pl -f /dfm/netapp/scripts/stats/eg-nasclnt-e01/stats_eg-nasclnt-e01_20110926_1015.gz

- Show controller level statistics for eg-nasclnt-e01 on 2011-09-26 at 10:20am:
  > /dfm/netapp/scripts/stats_show.pl -f /dfm/netapp/scripts/stats/eg-nasclnt-e01/stats_eg-nasclnt-e01_20110926_1015.gz –t 10:20

- Show controller and volume level statistics for eg-nasclnt-e01 on 2011-09-26 from 10:15am to 11:15am:
  > /dfm/netapp/scripts/stats_show.pl -f /dfm/netapp/scripts/stats/eg-nasclnt-e01/stats_eg-nasclnt-e01_20110926_1015.gz -v

- Show controller and volume level statistics for controller eg-nasclnt-e01 on 2011-09-26 at 10:20am:
  > /dfm/netapp/scripts/stats_show.pl -f /dfm/netapp/scripts/stats/eg-nasclnt-e01/stats_eg-nasclnt-e01_20110926_1015.gz –t 10:20 -v

5 Performance incident management

5.1 Typical steps

The following steps should be used as a standard practice when analyzing performance incidents:

1. Use the *TRP Storage System Summary* view to determine if the overall system latency during the timeframe of interest is at an acceptable level.
    a. If the latency is at acceptable levels, and is not higher during the incident timeframe than it was at other times, then storage is most likely not the root cause of the incident.
    b. If the latency is not at acceptable levels during the incident timeframe, determine what the I/O profile is during the timeframe and compare the overall controller IOPs to the tables in section 6.1.
        i. Is this system already at over above the recommended peak values for the platform, based on the IOPs and I/O profile? If so, then the latency is likely high as a result of the system having too high of a workload placed on it.
        ii. If the system workload is below the recommended peak values for the platform, but latency is higher than expected, proceed to section 5.2
2. Use the *TRP Volume Summary View* to determine if the volume latency during the timeframe of interest was significantly different from the overall storage system latency.
3. Use the *All Volumes Summary View* to determine if the utilization of a certain volume is driving an unusual amount of workload on the system.

5.2 Opening a performance case with NetApp Global Support

In order to ensure a timely response from the NetApp global support center, it is recommended to only open a support case when you have the following information ready to share:

1. The storage controller hostname and serial number.
2. A concise description of the problem, including the metric of interest. Some good examples of the metric of interest are as follows:
    a. NFS write latency spiked to 50ms during a certain timeframe, and the I/O workload on the system was at a level that this type of storage controller should be able to handle.
    b. A certain batch processing job used to run in 10 minutes, and last night it took 60 minutes to complete. We can see elevated NFS read latency during last night's batch job timeframe, as compared to previous nights.
3. When did the issue start happening? Were there any other known changes in the environment that coincide with the start of this issue?
4. Which volumes on the storage controller are affected? All volumes, or only certain volumes?
5. What are the clients doing that are accessing the storage controller? For example, clients may be using the storage for storing Oracle data over NFS, storing VMware ESX data over NFS, or for storing user home directory data over CIFS.
6. What OS is being used by the clients of this storage controller? If there are any client side OS messages that are relevant, please include that information.
7. NetApp Perfstat output from the storage controller being investigated, which was gathered **during the time which this incident occurred**. See section 5.3 for the possible methods by which Perfstat data can be gathered. Opening a performance case with NetApp global support without having valid Perfstat output **during the timeframe of the incident** will significantly slow case resolution, and you will likely be asked to gather Perfstat data in order to continue.

5.3 Using NetApp Perfstat for data collection in special cases

In the course of troubleshooting performance incidents that require NetApp Perfstat data in order to engage the NetApp global support center, the order of preference for obtaining NetApp Perfstat data is as follows:

1. Determine if the problem can be caught by an automatically triggered Perfstat in response to a performance threshold breach, as described in section 3.3.
2. If the problem cannot be caught by an automatically triggered Perfstat, can a manual Perfstat job be setup to gather data? This is typically only possible for incidents that are ongoing or that occur at known timeframes. If there is a known timeframe for a performance problem, the following is an example of the recommended cron job entry to use. In this example, we have a cron job setup to gather perfstat for 10 iterations of 1 minute each on eg-nasclnt-e03, starting at 7:15PM each night:
   *15 19 * * *   netapp /dfm/netapp/perfstat/perfstat.sh -f eg-nasclnt-e03 -E ifstat -t 1 -i 10 -F -I -V -p -k > /dfm/netapp/perfstat/eg-nasclnt-e03/perfstat1_eg-nasclnt-e03.out 2>/dfm/netapp/perfstat/eg-nasclnt-e03/perfstat.log*
   Note that we have configured this perfstat command with arguments that are specifically setup for the TRP environment. In particular, the "*-E ifstat*" section is designed to work around bug 510126, which causes a short error in network performance counter values when resetting interface statistics via the ifstat command.
3. If the problem is too short to be caught by an automatically gathered Perfstat, and it occurs at unknown timeframes, then the use of the script described in this section is appropriate.

In order to catch performance incidents that have a short duration, are intermittent, and unpredictable in terms of their timing, we have a script that will continuously gather Perfstat data in 1 minute intervals. This script is located at */dfm/netapp/scripts/perfstat_loop.sh* on all OM servers. In order to run this script, simply supply the storage system name as the first argument, and the path to an output directory as the second argument. Since this script runs until it is cancelled with a *<CTRL-C>*, it is recommend to run this via a *screen* session. For example:

om-cli# screen bash
om-cli# /dfm/netapp/scripts/perfstat_loop.sh eg-nasecom-h06 /dfm/netapp/perfstat/eg-nasecom-h06
<CTRL-A> <CTRL-D>

The *<CTRL-A> <CTRL-D>* key sequence will disconnect your screen session and the script will continue to run. In order to view the list of running screen sessions, use the *screen –list* command. Use *screen –r <socketid>* to reconnect to your screen session when you want to kill the script with the *<CTRL-C>* command, and exit the screen session with the *exit* command. The output file(s) to upload to the NetApp global support center will have a date/time stamp and there will be one file per hour.

Notes:

1. While this script gives us a high probability of catching short lived performance issues, there is a chance that it will miss them due to the startup/shutdown time required for each Perfstat iteration.
2. If you use this method to gather Perfstat data, make sure to disable the automated threshold breach Perfstata data collection as described in section 3.3 by renaming or removing the */dfm/netapp/perfstat/dfm_perfstat/<filename>* directory. This directory can be recreated to re-enable automated Perfstat data collection once you have finished using this method.

6 Additional resources and configuration settings

6.1 NetApp Peak Performance by platform

The following tables list the maximum desirable IOPs values for the commonly used platforms in TRP, depending on the I/O profile.

FAS6210 controllers with 450GB SAS drives

Max IOPs	Read/Write Mix	100/0	75/25	50/50	25/75	0/100
Avg IO Size						
32k		56,000	36,600	27,200	21,600	18,000
24k		57,200	40,400	31,400	25,600	21,600
16k		58,600	45,800	38,000	32,600	28,400
8k		60,600	53,000	49,000	45,000	42,000
4k		61,600	58,000	57,000	56,000	55,400

FAS6210 controllers with 2TB SATA drives

Max IOPs	Read/Write Mix	100/0	75/25	50/50	25/75	0/100
Avg IO Size						
32k		42000	27450	20400	16200	13500
24k		42900	30300	23550	19200	16200
16k		43950	34350	28500	24450	21300
8k		45450	39750	36750	33750	31500
4k		46200	43500	42750	42000	41550

FAS6080 controllers

Max IOPs	Read/Write Mix	100/0	75/25	50/50	25/75	0/100
Avg IO Size						
32k		18,300	15,200	13,000	11,300	10,000
24k		21,800	17,800	15,600	13,500	12,000
16k		26,800	22,400	19,400	17,400	15,600
8k		35,000	30,400	28,000	25,200	23,400
4k		41,000	37,200	35,200	33,400	31,800

FAS3170 controllers

Max IOPs	Read/Write Mix	100/0	75/25	50/50	25/75	0/100
Avg IO Size						
32k		18,000	14,300	11,900	10,100	8,800
24k		21,600	17,100	14,100	12,000	10,500
16k		27,200	22,000	18,200	16,000	14,000
8k		36,800	30,800	27,000	23,000	21,000
4k		44,600	39,800	34,400	30,200	27,200

FAS980 controllers

Max IOPs	Read/Write Mix	100/0	75/25	50/50	25/75	0/100
Avg IO Size						
32k		5,600	4,600	3,900	3,400	3,000
24k		6,800	5,500	4,700	3,900	3,500
16k		8,500	7,300	6,100	5,200	4,400
8k		11,500	9,100	8,400	7,500	6,600
4k		18,200	13,800	11,400	9,800	8,600

Notes on peak performance max IOPs values:

- All sizing was done with sub 20ms response time and 20% system CPU headroom in mind.
- Peak IOP values are per controller and can be doubled for per cluster numbers.
- Use this only as a guide when determining controller headroom on existing systems.
- The values listed in this document assume adequate spindles are in place.
- The values do not account for additional system loads such as deduplication, disk scrubs, etc.

6.2 PA custom counter configuration

The */dfm/netapp/scripts/pa_counter_setup.sh* script resides on all OM servers in the TRP environment, and should be run on all new storage systems to configure the appropriate performance counter configuration. This script will disable a number of performance counters that are not used in the TRP environment, as well as extend the retention period for volume based counters. The following is the list of counter groups that are disabled by this script for the TRP environment, which is done to reduce the load placed on the OM servers:

- disk (enabled on systems running 7.3.3 or higher due to a bug in earlier versions of ONTAP)
- qtree
- processor
- target
- lun
- fcp
- priorityqueue
- prisched

In addition to disabling the above counter groups, the script makes the following changes:

- Increases the retention period of the "volume" counter group to be 12 weeks instead of the default 1.
- Increases the retention period of the "vfiler" counter group to be 12 weeks instead of the default 1.
- Enables the following diag mode counters:
  volume:nfs_read_data
  volume:nfs_write_data
  volume:cifs_read_data
  volume:cifs_write_data
  volume:iscsi_read_ops
  volume:iscsi_write_ops
  volume:iscsi_other_ops
  volume:iscsi_read_data
  volume:iscsi_write_data
  volume:iscsi_read_latency
  volume:iscsi_write_latency
  volume:iscsi_other_latency

6.3 Extracting PA data via the CLI for special use cases

In some cases, it can be useful to get the raw performance data from the OM servers. This is useful if you want to consolidate data from many systems on different OM servers into a single graph, or if you want to feed the data into another tool for analysis. This can be accomplished use the OM cli, with the *dfm perf data retrieve* command. For more details on using the *dfm perf data retrieve* command see the OM manual pages, which are available via the Help -> General Help section of the OM web interface.

Since the output of the "dfm perf data retrieve" command can be difficult to directly import into Excel for graphing, we have written a wrapper script around this command to make that task easier. This script is installed as */dfm/netapp/scripts/dfmperfdatacsv.pl* on all the OM servers. The usage of this command can be seen by running the script with no arguments, and the output is a simple CSV file.

6.4 Troubleshooting the loss of historical performance data in OM

It is important to understand that there are certain circumstances in which performance data will not be gathered by the OM servers. The most common cause for this is when the *perfdata* directory on the OM server, which is */dfm/dfminst/perfdata* in the TRP environment, goes over 90% utilization. The gathering of performance data is automatically re-enabled by OM once the utilization of the *perfdata* directory goes back under 90%. If there are large gaps of data missing from the graphs of PA, this is typically the first item to check. The OM servers in the TRP environment will log instances of this occurring to */dfm/dfminst/logs/dfmserver.log*.

Another way that OM can throttle the gathering of performance data is due to the load on the OM server itself. This loss of data is more likely to be seen when gathering performance data via the CLI, since having a small percentage of data missing will not be visible in the PA graphs. Again, instances of this occurring are logged to the *dfmserver.log* file. For example:

om-cli> grep 'server load' /dfm/dfminst/log/dfmserver.log*

```
dfmserver-0.log:Jun 27 16:22:13 [dfmserver: WARN]: Thread 0xeaffaba0: The DataFabric
Manager server load is 97%; slowing monitoring to 90%.
dfmserver-0.log:Jun 28 09:44:20 [dfmserver: INFO]: Thread 0xf53edba0: The DataFabric
Manager server load is 41%; resuming full monitoring.
```

6.5 Bugs related to performance analysis and performance data

The following is a list of known bugs related to performance analysis and performance data, including the software releases in which they are present. This is useful to know so that these issues do not impact your analysis when viewing data impacted by these bugs.

Bug #	Description	Fixed In
356178	Performance Advisor shows double the actual network throughput numbers on systems with VLAN tagged interfaces.	ONTAP versions 7.3.2.P2, 7.3.3
461282	DFM(Data Fabric Manager) performance degrade when multiple NMC instances are open and a large number of objects are being monitored	DFM 4.0D15 and NMC 3.0D2
510126	ifstat -z -a messes up network in/out counters	None

6.6 Setting the default performance view for systems

The PA interface allows you to set the default performance view for any object you may be viewing. Since the most common object to view in PA is the storage controller itself, we have a script in place to set the default view on controllers to our "TRP System Summary View". This script is */dfm/netapp/scripts/setdefaultperfviews.pl*, and it is set to run once per week via an entry in */etc/crontab* on all OM servers.

6.7 The Operations Manager online documentation

https://now.netapp.com/NOW/knowledge/docs/DFM_win/dfm_index.shtml