Manage performance issues

ONTAP 9

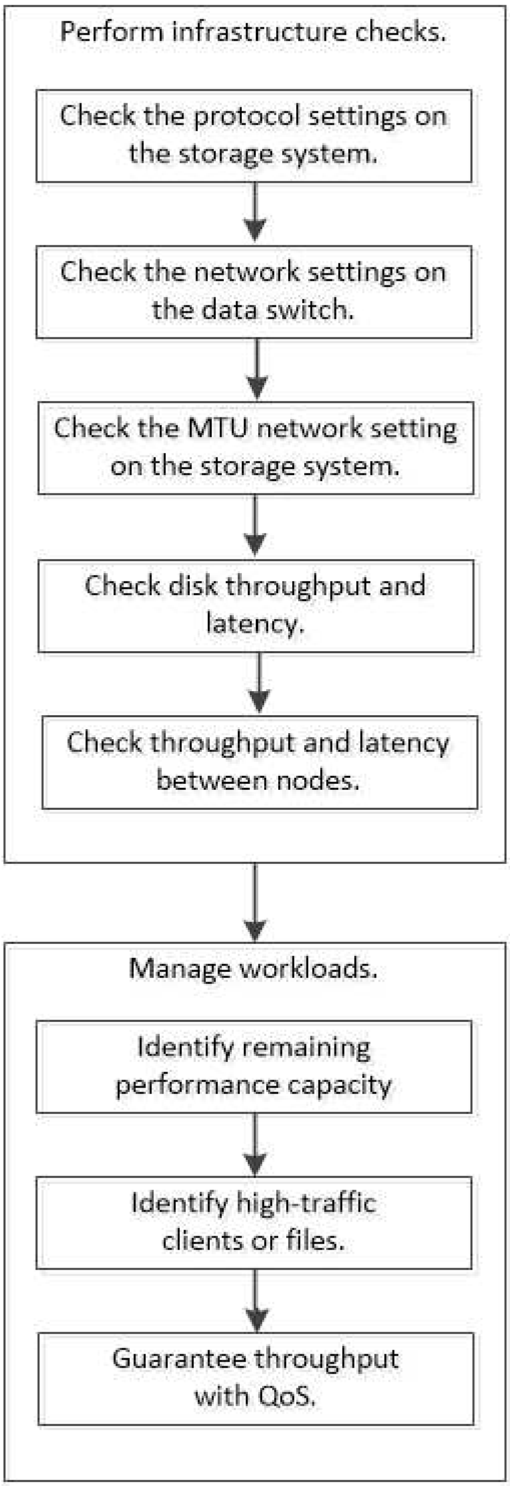
NetApp
February 28, 2022

Table of Contents

Manage performance issues

Performance management workflow

Once you have identified a performance issue, you can conduct some basic diagnostic checks of your infrastructure to rule out obvious configuration errors. If those don’t pinpoint the problem, you can start looking at workload management issues.



Perform basic infrastructure checks

Check protocol settings on the storage system

Check the NFS TCP maximum transfer size

For NFS, you can check whether the TCP maximum transfer size for reads and writes might be causing a performance issue. If you think the size is slowing performance, you can increase it.

What you'll need

- You must have cluster administrator privileges to perform this task.
- You must use advanced privilege level commands for this task.

Steps

1. Change to the advanced privilege level:

   `set -privilege advanced`

2. Check the TCP maximum transfer size:

   `vserver nfs show -vserver` *`vserver_name`* `-instance`

3. If the TCP maximum transfer size is too small, increase the size:

   `vserver nfs modify -vserver` *`vserver_name`* `-tcp-max-xfer-size` *`integer`*

4. Return to the administrative privilege level:

   `set -privilege admin`

Example

The following example changes the TCP maximum transfer size of `SVM1` to 1048576:

```
cluster1::*> vserver nfs modify -vserver SVM1 -tcp-max-xfer-size 1048576
```

Check the iSCSI TCP read/write size

For iSCSI, you can check the TCP read/write size to determine if the size setting is creating a performance issue. If the size is the source of an issue, you can correct it.

What you'll need

Advanced privilege level commands are required for this task.

Steps

1. Change to advanced privilege level:

   `set -privilege advanced`

2. Check the TCP window size setting:

   `vserver iscsi show -vserv,er vserver_name -instance`

3. Modify the TCP window size setting:

   `vserver iscsi modify -vserver vserver_name -tcp-window-size integer`

4. Return to administrative privilege:

   `set -privilege admin`

Example

The following example changes the TCP window size of `SVM1` to 131,400 bytes:

```
cluster1::*> vserver iscsi modify -vserver vs1 -tcp-window-size 131400
```

Check the CIFS multiplex settings

If slow CIFS network performance causes a performance issue, you can modify the multiplex settings to improve and correct it.

Steps

1. Check the CIFS multiplex setting:

   `vserver cifs options show -vserver -vserver_name -instance`

2. Modify the CIFS multiplex setting:

   `vserver cifs options modify –vserver -vserver_name –max-mpx integer`

Example

The following example changes the maximum multiplex count on `SVM1` to 255:

```
cluster1::> vserver cifs options modify -vserver SVM1 -max-mpx 255
```

Check the FC adapter port speed

The adapter target port speed should match the speed of the device to which it connects, to optimize performance. If the port is set to autonegotiation, it can take longer to reconnect after a takeover and giveback or other interruption.

What you'll need

All LIFs that use this adapter as their home port must be offline.

Steps

1. Take the adapter offline:

```
network fcp adapter modify -node nodename -adapter adapter -state down
```

2. Check the maximum speed of the port adapter:

```
fcp adapter show -instance
```

3. Change the port speed, if necessary:

```
network fcp adapter modify -node nodename -adapter adapter -speed
{1|2|4|8|10|16|auto}
```

4. Bring the adapter online:

```
network fcp adapter modify -node nodename -adapter adapter -state up
```

5. Bring all the LIFs on the adapter online:

```
network interface modify -vserver * -lif * { -home-node node1 -home-port e0c }
-status-admin up
```

Example

The following example changes the port speed of adapter `0d` on `node1` to 2 Gbps:

```
cluster1::> network fcp adapter modify -node node1 -adapter 0d -speed 2
```

Check the network settings on the data switches

Although you must maintain the same MTU settings on your clients, servers and storage systems (that is, network endpoints), intermediate network devices such as NICs and switches should be set to their maximum MTU values to ensure that performance is not impacted.

For best performance, all components in the network must be able to forward jumbo frames (9000 bytes IP, 9022 bytes including Ethernet). Data switches should be set to at least 9022 bytes, but a typical value of 9216 is possible with most switches.

Procedure

For data switches, check that the MTU size is set to 9022 or higher.

For more information, see the switch vendor documentation.

Check the MTU network setting on the storage system

You can change the network settings on the storage system if they are not the same as on the client or other network endpoints. Whereas the management network MTU setting is set to 1500, the data network MTU size should be 9000.

About this task

All ports within a broadcast-domain have the same MTU size, with the exception of the e0M port handling management traffic. If the port is part of a broadcast-domain, use the `broadcast-domain modify` command to change the MTU for all ports within the modified broadcast-domain.

Note that intermediate network devices such as NICs and data switches can be set to higher MTU sizes than network endpoints. For more information, see Check the network settings on the data switches.

Steps

1. Check the MTU port setting on the storage system:

   `network port show -instance`

2. Change the MTU port setting to 9000:

   `network port modify -node *nodename* -port *port* -mtu 9000`

Example

The following example changes the MTU port setting for `e0a` to 9000:

```
cluster1::*> network port modify -node node0 -port e0a -mtu 9000
```

Check disk throughput and latency

You can check the disk throughput and latency metrics for cluster nodes to assist you in troubleshooting.

About this task

Advanced privilege level commands are required for this task.

Steps

1. Change to advanced privilege level:

   `set -privilege advanced`

2. Check the disk throughput and latency metrics:

   `statistics disk show -sort-key latency`

Example

The following example displays the totals in each user read or write operation for `node2` on `cluster1`:

```
::*> statistics disk show -sort-key latency
cluster1 : 8/24/2015 12:44:15
                    Busy Total Read  Write  Read    Write   *Latency
  Disk        Node  (%)   Ops   Ops   Ops   (Bps)   (Bps)      (us)
------------ ----   ---- ----   ----- ----- ------   -----    -----

1.10.20      node2     4    5       3     2  95232  367616    23806
1.10.8       node2     4    5       3     2 138240  386048    22113

1.10.6       node2     3    4       2     2  48128  371712    19113
1.10.19      node2     4    6       3     2 102400  443392    19106

1.10.11      node2     4    4       2     2 122880  408576    17713
```

Check throughput and latency between nodes

You can use the `network test-path` command to identify network bottlenecks, or to prequalify network paths between nodes. You can run the command between intercluster nodes or intracluster nodes.

What you'll need

- You must be a cluster administrator to perform this task.
- Advanced privilege level commands are required for this task.
- For an intercluster path, the source and destination clusters must be peered.

About this task

Occasionally, network performance between nodes may not meet expectations for your path configuration. A 1 Gbps transmission rate for the kind of large data transfers seen in SnapMirror replication operations, for example, would not be consistent with a 10 GbE link between the source and destination clusters.

You can use the `network test-path` command to measure throughput and latency between nodes. You can run the command between intercluster nodes or intracluster nodes.

> The test saturates the network path with data, so you should run the command when the system is not busy and when network traffic between nodes is not excessive. The test times out after ten seconds. The command can be run only between ONTAP 9 nodes.

The `session-type` option identifies the type of operation you are running over the network path—for example, "AsyncMirrorRemote" for SnapMirror replication to a remote destination. The type dictates the amount of data used in the test. The following table defines the session types:

Session Type	Description
Default	SnapMirror replication between nodes in different clusters

Session Type	Description
AsyncMirrorLocal	SnapMirror replication between nodes in the same cluster
AsyncMirrorRemote	SnapMirror replication between nodes in different clusters
SyncMirrorRemote	SyncMirror replication between nodes in different clusters
RemoteDataTransfer	Data transfer between nodes in the same cluster (for example, an NFS request to a node for a file stored in a volume on a different node)

Steps

1. Change to advanced privilege level:

   `set -privilege advanced`

2. Measure throughput and latency between nodes:

   `network test-path -source-node` *`source_nodename`* `|local -destination-cluster` *`destination_clustername`* `-destination-node` *`destination_nodename`* `-session-type Default|AsyncMirrorLocal|AsyncMirrorRemote|SyncMirrorRemote|RemoteDataTransfer`

   The source node must be in the local cluster. The destination node can be in the local cluster or in a peered cluster. A value of "local" for `-source-node` specifies the node on which you are running the command.

   The following command measures throughput and latency for SnapMirror-type replication operations between `node1` on the local cluster and `node3` on `cluster2`:

   ```
   cluster1::> network test-path -source-node node1 -destination-cluster
   cluster2 -destination-node node3 -session-type AsyncMirrorRemote
   Test Duration:      10.88 secs
   Send Throughput:    18.23 MB/sec
   Receive Throughput: 18.23 MB/sec
   MB sent:            198.31
   MB received:        198.31
   Avg latency in ms:  2301.47
   Min latency in ms:  61.14
   Max latency in ms:  3056.86
   ```

3. Return to administrative privilege:

   `set -privilege admin`

After you finish

If performance does not meet expectations for the path configuration, you should check node performance statistics, use available tools to isolate the problem in the network, check switch settings, and so forth.

Manage workloads

Identify remaining performance capacity

Performance capacity, or *headroom*, measures how much work you can place on a node or an aggregate before performance of workloads on the resource begins to be affected by latency. Knowing the available performance capacity on the cluster helps you provision and balance workloads.

What you'll need

Advanced privilege level commands are required for this task.

About this task

You can use the following values for the `-object` option to collect and display headroom statistics:

- For CPUs, `resource_headroom_cpu`.
- For aggregates, `resource_headroom_aggr`.

You can also complete this task using System Manager and Active IQ Unified Manager.

Steps

1. Change to advanced privilege level:

   ```
   set -privilege advanced
   ```

2. Start real-time headroom statistics collection:

   ```
   statistics start –object resource_headroom_cpu|aggr
   ```

   For complete command syntax, see the man page.

3. Display real-time headroom statistics information:

   ```
   statistics show -object resource_headroom_cpu|aggr
   ```

   For complete command syntax, see the man page.

4. Return to administrative privilege:

   ```
   set -privilege admin
   ```

Example

The following example displays the average hourly headroom statistics for cluster nodes.

You can compute the available performance capacity for a node by subtracting the `current_utilization` counter from the `optimal_point_utilization` counter. In this example, the utilization capacity for

`CPU_sti2520-213` is -14% (72%-86%), which suggests that the CPU has been overutilized on average for the past hour.

You could have specified `ewma_daily`, `ewma_weekly`, or `ewma_monthly` to get the same information averaged over longer periods of time.

```
sti2520-2131454963690::*> statistics show -object resource_headroom_cpu
-raw -counter ewma_hourly
  (statistics show)

Object: resource_headroom_cpu
Instance: CPU_sti2520-213
Start-time: 2/9/2016 16:06:27
End-time: 2/9/2016 16:06:27
Scope: sti2520-213

    Counter                                                     Value
    -------------------------------- --------------------------------
    ewma_hourly                                                     -
                         current_ops                             4376
                     current_latency                            37719
                 current_utilization                               86
                   optimal_point_ops                             2573
               optimal_point_latency                             3589
           optimal_point_utilization                               72
     optimal_point_confidence_factor                                1

Object: resource_headroom_cpu
Instance: CPU_sti2520-214
Start-time: 2/9/2016 16:06:27
End-time: 2/9/2016 16:06:27
Scope: sti2520-214

    Counter                                                     Value
    -------------------------------- --------------------------------
    ewma_hourly                                                     -
                         current_ops                                0
                     current_latency                                0
                 current_utilization                                0
                   optimal_point_ops                                0
               optimal_point_latency                                0
           optimal_point_utilization                               71
     optimal_point_confidence_factor                                1
2 entries were displayed.
```

Identify high-traffic clients or files

You can use ONTAP Active Objects technology to identify clients or files that are responsible for a disproportionately large amount of cluster traffic. Once you have identified these "top" clients or files, you can rebalance cluster workloads or take other steps to resolve the issue.

What you'll need

You must be a cluster administrator to perform this task.

Steps

1. View the top clients accessing the cluster:

   `statistics top client show -node` *`node_name`* `-sort-key` *`sort_column`* `-interval` *`seconds_between_updates`* `-iterations` *`iterations`* `-max` *`number_of_instances`*

   For complete command syntax, see the man page.

   The following command displays the top clients accessing `cluster1`:

```
cluster1::> statistics top client show

cluster1 : 3/23/2016 17:59:10


                                                       *Total
        Client Vserver           Node Protocol    Ops
-------------- ------- -------------- -------- ------
172.17.180.170     vs4 siderop1-vsim4      nfs    668
172.17.180.169     vs3 siderop1-vsim3      nfs    337
172.17.180.171     vs3 siderop1-vsim3      nfs    142
172.17.180.170     vs3 siderop1-vsim3      nfs    137
172.17.180.123     vs3 siderop1-vsim3      nfs    137
172.17.180.171     vs4 siderop1-vsim4      nfs     95
172.17.180.169     vs4 siderop1-vsim4      nfs     92
172.17.180.123     vs4 siderop1-vsim4      nfs     92
172.17.180.153     vs3 siderop1-vsim3      nfs      0
```

2. View the top files accessed on the cluster:

   `statistics top file show -node` *`node_name`* `-sort-key` *`sort_column`* `-interval` *`seconds_between_updates`* `-iterations` *`iterations`* `-max` *`number_of_instances`*

   For complete command syntax, see the man page.

   The following command displays the top files accessed on `cluster1`:

```
cluster1::> statistics top file show

cluster1 : 3/23/2016 17:59:10

                                  *Total
                    File Volume Vserver          Node    Ops
------------------------ ------ ------- -------------- ------
/vol/vol1/vm170-read.dat   vol1     vs4 siderop1-vsim4     22
/vol/vol1/vm69-write.dat   vol1     vs3 siderop1-vsim3      6
     /vol/vol2/vm171.dat   vol2     vs3 siderop1-vsim3      2
     /vol/vol2/vm169.dat   vol2     vs3 siderop1-vsim3      2
      /vol/vol2/p123.dat   vol2     vs4 siderop1-vsim4      2
      /vol/vol2/p123.dat   vol2     vs3 siderop1-vsim3      2
     /vol/vol1/vm171.dat   vol1     vs4 siderop1-vsim4      2
     /vol/vol1/vm169.dat   vol1     vs4 siderop1-vsim4      2
     /vol/vol1/vm169.dat   vol1     vs4 siderop1-vsim3      2
      /vol/vol1/p123.dat   vol1     vs4 siderop1-vsim4      2
```

Guarantee throughput with QoS

Guarantee throughput with QoS overview

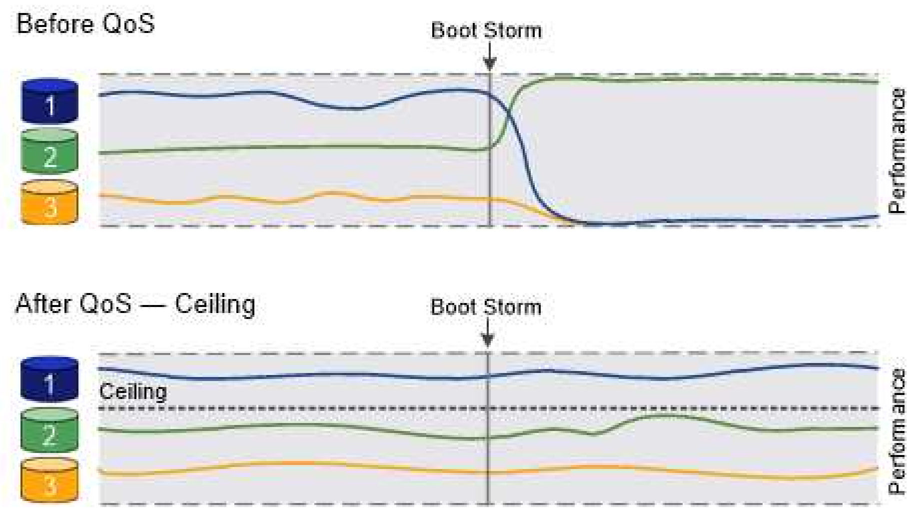
You can use storage quality of service (QoS) to guarantee that performance of critical workloads is not degraded by competing workloads. You can set a throughput *ceiling* on a competing workload to limit its impact on system resources, or set a throughput *floor* for a critical workload, ensuring that it meets minimum throughput targets, regardless of demand by competing workloads. You can even set a ceiling and floor for the same workload.

About throughput ceilings (QoS Max)

A throughput ceiling limits throughput for a workload to a maximum number of IOPS or MBps, or IOPS and MBps. In the figure below, the throughput ceiling for workload 2 ensures that it does not "bully" workloads 1 and 3.

A *policy group* defines the throughput ceiling for one or more workloads. A workload represents the I/O operations for a *storage object:* a volume, file, qtree or LUN, or all the volumes, files, qtrees, or LUNs in an SVM. You can specify the ceiling when you create the policy group, or you can wait until after you monitor workloads to specify it.

> ⓘ Throughput to workloads might exceed the specified ceiling by up to 10%, especially if a workload experiences rapid changes in throughput. The ceiling might be exceeded by up to 50% to handle bursts. Bursts occur on single nodes when tokens accumulate up to 150%



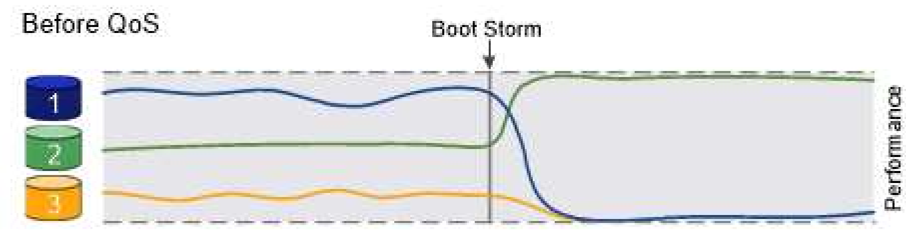
About throughput floors (QoS Min)

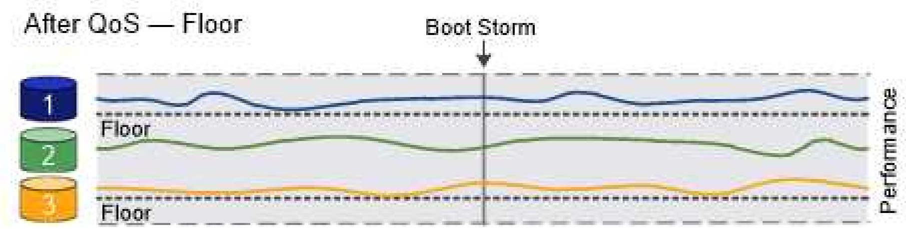
A throughput floor guarantees that throughput for a workload does not fall below a minimum number of IOPS or MBps, or IOPS and MBps.In the figure below, the throughput floors for workload 1 and workload 3 ensure that they meet minimum throughput targets, regardless of demand by workload 2.

> As the examples suggest, a throughput ceiling throttles throughput directly. A throughput floor throttles throughput indirectly, by giving priority to the workloads for which the floor has been set.

A policy group that defines a throughput floor cannot be applied to an SVM. You can specify the floor when you create the policy group, or you can wait until after you monitor workloads to specify it.

> In releases before ONTAP 9.7, throughput floors are guaranteed when there is sufficient performance capacity available. In ONTAP 9.7 and later, throughput floors can be guaranteed even when there is insufficient performance capacity available. This new floor behavior is called floors v2. To meet the guarantees, floors v2 can result in higher latency on workloads without a throughput floor or on work that exceeds the floor settings. Floors v2 applies to both QoS and adaptive QoS. The option of enabling/disabling the new behavior of floors v2 is available in ONTAP 9.7P6 and later.A workload might fall below the specified floor during critical operations like `volume move trigger-cutover`. Even when sufficient capacity is available and critical operations are not taking place, throughput to a workload might fall below the specified floor by up to 5%. If floors are overprovisioned and there is no performance capacity, some workloads might fall below the specified floor.





About shared and non-shared QoS policy groups

Beginning with ONTAP 9.4, you can use a *non-shared* QoS policy group to specify that the defined throughput ceiling or floor applies to each member workload individually. Behavior of *shared* policy groups depends on the policy type:

- For throughput ceilings, the total throughput for the workloads assigned to the shared policy group cannot exceed the specified ceiling.
- For throughput floors, the shared policy group can be applied to a single workload only.

About adaptive QoS

Ordinarily, the value of the policy group you assign to a storage object is fixed. You need to change the value manually when the size of the storage object changes. An increase in the amount of space used on a volume, for example, usually requires a corresponding increase in the throughput ceiling specified for the volume.

Adaptive QoS automatically scales the policy group value to workload size, maintaining the ratio of IOPS to TBs|GBs as the size of the workload changes. That is a significant advantage when you are managing hundreds or thousands of workloads in a large deployment.

You typically use adaptive QoS to adjust throughput ceilings, but you can also use it to manage throughput floors (when workload size increases). Workload size is expressed as either the allocated space for the storage object or the space used by the storage object.

> ⓘ Used space is available for throughput floors in ONTAP 9.5 and later. It is not supported for throughput floors in ONTAP 9.4 and earlier.

- An *allocated space* policy maintains the IOPS/TB|GB ratio according to the nominal size of the storage object. If the ratio is 100 IOPS/GB, a 150 GB volume will have a throughput ceiling of 15,000 IOPS for as long as the volume remains that size. If the volume is resized to 300 GB, adaptive QoS adjusts the throughput ceiling to 30,000 IOPS.
- A *used space* policy (the default) maintains the IOPS/TB|GB ratio according to the amount of actual data stored before storage efficiencies. If the ratio is 100 IOPS/GB, a 150 GB volume that has 100 GB of data stored would have a throughput ceiling of 10,000 IOPS. As the amount of used space changes, adaptive QoS adjusts the throughput ceiling according to the ratio.

Beginning with ONTAP 9.5, you can specify an I/O block size for your application that enables a throughput limit to be expressed in both IOPS and MBps. The MBps limit is calculated from the block size multiplied by the

IOPS limit. For example, an I/O block size of 32K for an IOPS limit of 6144IOPS/TB yields an MBps limit of 192MBps.

You can expect the following behavior for both throughput ceilings and floors:

- When a workload is assigned to an adaptive QoS policy group, the ceiling or floor is updated immediately.
- When a workload in an adaptive QoS policy group is resized, the ceiling or floor is updated in approximately five minutes.

Throughput must increase by at least 10 IOPS before updates take place.

Adaptive QoS policy groups are always non-shared: the defined throughput ceiling or floor applies to each member workload individually.

Beginning with ONTAP 9.6, throughput floors is supported on ONTAP Select premium with SSD.

General support

The following table shows the differences in support for throughput ceilings, throughput floors, and adaptive QoS.

Resource or feature	Throughput ceiling	Throughput floor	Throughput floor v2	Adaptive QoS
ONTAP 9 version	All	9.2 and later	9.7 and later	9.3 and later
Platforms	All	• AFF • C190 * • ONTAP Select premium with SSD *	• AFF • C190 • ONTAP Select premium with SSD	All
Protocols	All	All	All	All
FabricPool	Yes	Yes, if the tiering policy is set to "none" and no blocks are in the cloud.	Yes, if the tiering policy is set to "none" and no blocks are in the cloud.	Yes
SnapMirror Synchronous	Yes	No	No	Yes

*C190 and ONTAP Select support started with the ONTAP 9.6 release.

Supported workloads for throughput ceilings

The following table shows workload support for throughput ceilings by ONTAP 9 version. Root volumes, load-sharing mirrors, and data protection mirrors are not supported.

Workload support - ceiling	9.0	9.1	9.2	9.3	9.4 and later	9.8 and later
Volume	yes	yes	yes	yes	yes	yes
File	yes	yes	yes	yes	yes	yes
LUN	yes	yes	yes	yes	yes	yes
SVM	yes	yes	yes	yes	yes	yes
FlexGroup volume	no	no	no	yes	yes	yes
qtrees*	no	no	no	no	no	yes
Multiple workloads per policy group	yes	yes	yes	yes	yes	yes
Non-shared policy groups	no	no	no	no	yes	yes

*Beginning with ONTAP 9.8, NFS access is supported in qtrees in FlexVol and FlexGroup volumes with NFS enabled. Beginning with ONTAP 9.9.1, SMB access is also supported in qtrees in FlexVol and FlexGroup volumes with SMB enabled.

Supported workloads for throughput floors

The following table shows workload support for throughput floors by ONTAP 9 version. Root volumes, load-sharing mirrors, and data protection mirrors are not supported.

Workload support - floor	9.2	9.3	9.4 and later	9.8 and later
Volume	yes	yes	yes	yes
File	no	yes	yes	yes
LUN	yes	yes	yes	yes
SVM	no	no	no	yes
FlexGroup volume	no	no	yes	yes
qtrees *	no	no	no	yes

Multiple workloads per policy group	no	no	yes	yes
Non-shared policy groups	no	no	yes	yes

*Beginning with ONTAP 9.8, NFS access is supported in qtrees in FlexVol and FlexGroup volumes with NFS enabled. Beginning with ONTAP 9.9.1, SMB access is also supported in qtrees in FlexVol and FlexGroup volumes with SMB enabled.

Supported workloads for adaptive QoS

The following table shows workload support for adaptive QoS by ONTAP 9 version. Root volumes, load-sharing mirrors, and data protection mirrors are not supported.

Workload support - adaptive QoS	9.3	9.4 and later
Volume	yes	yes
File	no	yes
LUN	no	yes
SVM	no	no
FlexGroup volume	no	yes
Multiple workloads per policy group	yes	yes
Non-shared policy groups	yes	yes

Maximum number of workloads and policy groups

The following table shows the maximum number of workloads and policy groups by ONTAP 9 version.

Workload support	9.3 and earlier	9.4 and later
Maximum workloads per cluster	12,000	40,000
Maximum workloads per node	12,000	40,000
Maximum policy groups	12,000	12,000

Enable or disable throughput floors v2

You can enable or disable throughput floors v2 on AFF. The default is enabled. With floors v2 enabled, throughput floors can be met when controllers are heavily used at the expense of higher latency on other workloads. Floors v2 applies to both QoS and Adaptive QoS.

Steps

1. Change to advanced privilege level:

   `set -privilege advanced`

2. Enter one of the following commands:

If you want to…	Use this command:
Disable floors v2	`qos settings throughput-floors-v2 -enable false`
Enable floors v2	`qos settings throughput-floors-v2 -enable true`

> To disable throughput floors v2 in an MetroCluster cluster, you must run the
>
> `qos settings throughput-floors-v2 -enable false`
>
> command on both the source and destination clusters.

```
cluster1::*> qos settings throughput-floors-v2 -enable false
```

Storage QoS workflow

If you already know the performance requirements for the workloads you want to manage with QoS, you can specify the throughput limit when you create the policy group. Otherwise, you can wait until after you monitor the workloads to specify the limit.

Set a throughput ceiling with QoS

You can use the `max-throughput` field for a policy group to define a throughput ceiling for storage object workloads (QoS Max). You can apply the policy group when you create or modify the storage object.

What you'll need

- You must be a cluster administrator to create a policy group.
- You must be a cluster administrator to apply a policy group to an SVM.

About this task

- Beginning with ONTAP 9.4, you can use a *non-shared* QoS policy group to specify that the defined throughput ceiling applies to each member workload individually. Otherwise, the policy group is *shared:* the total throughput for the workloads assigned to the policy group cannot exceed the specified ceiling.

Set `-is-shared=false` for the `qos policy-group create` command to specify a non-shared policygroup.

- You can specify the throughput limit for the ceiling in IOPS, MB/s, or IOPS, MB/s. If you specify both IOPS and MB/s, whichever limit is reached first is enforced.

> If you set a ceiling and a floor for the same workload, you can specify the throughput limit for the ceiling in IOPS only.

- A storage object that is subject to a QoS limit must be contained by the SVM to which the policy group belongs. Multiple policy groups can belong to the same SVM.
- You cannot assign a storage object to a policy group if its containing object or its child objects belong to the policy group.
- It is a QoS best practice to apply a policy group to the same type of storage objects.

Steps

1. Create a policy group:

   `qos policy-group create -policy-group` *`policy_group`* `-vserver` *`SVM`* `-max -throughput` *`number_of_iops`*`|Mb/S|iops,Mb/S -is-shared true|false`

   For complete command syntax, see the man page. You can use the `qos policy-group modify` command to adjust throughput ceilings.

   The following command creates the shared policy group `pg-vs1` with a maximum throughput of 5,000 IOPS:

   ```
   cluster1::> qos policy-group create -policy-group pg-vs1 -vserver vs1
   -max-throughput 5000iops -is-shared true
   ```

   The following command creates the non-shared policy group `pg-vs3` with a maximum throughput of 100 IOPS and 400 Kb/S:

   ```
   cluster1::> qos policy-group create -policy-group pg-vs3 -vserver vs3
   -max-throughput 100iops,400KB/s -is-shared false
   ```

   The following command creates the non-shared policy group `pg-vs4` without a throughput limit:

   ```
   cluster1::> qos policy-group create -policy-group pg-vs4 -vserver vs4
   -is-shared false
   ```

2. Apply a policy group to an SVM, file, volume, or LUN:

   `storage_object` `create -vserver` *`SVM`* `-qos-policy-group` *`policy_group`*

   For complete command syntax, see the man pages. You can use the *`storage_object`* `modify` command to apply a different policy group to the storage object.

The following command applies policy group `pg-vs1` to SVM `vs1`:

```
cluster1::> vserver create -vserver vs1 -qos-policy-group pg-vs1
```

The following commands apply policy group `pg-app` to the volumes `app1` and `app2`:

```
cluster1::> volume create -vserver vs2 -volume app1 -aggregate aggr1
-qos-policy-group pg-app
```

```
cluster1::> volume create -vserver vs2 -volume app2 -aggregate aggr1
-qos-policy-group pg-app
```

3. Monitor policy group performance:

   `qos statistics performance show`

   For complete command syntax, see the man page.

   > Monitor performance from the cluster. Do not use a tool on the host to monitor performance.

   The following command shows policy group performance:

```
cluster1::> qos statistics performance show
Policy Group           IOPS      Throughput   Latency
-------------------- -------- --------------- ----------
-total-                 12316       47.76MB/s  1264.00us
pg_vs1                   5008       19.56MB/s     2.45ms
_System-Best-Effort        62       13.36KB/s     4.13ms
_System-Background         30           0KB/s        0ms
```

4. Monitor workload performance:

   `qos statistics workload performance show`

   For complete command syntax, see the man page.

   > Monitor performance from the cluster. Do not use a tool on the host to monitor performance.

   The following command shows workload performance:

```
cluster1::> qos statistics workload performance show
Workload            ID     IOPS       Throughput     Latency
--------------- ------ -------- ---------------- ----------
-total-              -    12320        47.84MB/s  1215.00us
app1-wid7967      7967     7219        28.20MB/s   319.00us
vs1-wid12279     12279     5026        19.63MB/s     2.52ms
_USERSPACE_APPS     14       55        10.92KB/s   236.00us
_Scan_Backgro..   5688       20            0KB/s        0ms
```

> You can use the `qos statistics workload latency show` command to view detailed latency statistics for QoS workloads.

Set a throughput floor with QoS

You can use the `min-throughput` field for a policy group to define a throughput floor for storage object workloads (QoS Min). You can apply the policy group when you create or modify the storage object. Beginning with ONTAP 9.8, you can specify the throughput floor in IOPS or MBps, or IOPS and MBps.

What you'll need

- You must be running ONTAP 9.2 or later. Throughput floors are available beginning with ONTAP 9.2.
- You must be a cluster administrator to create a policy group.

About this task

- Beginning with ONTAP 9.4, you can use a *non-shared* QoS policy group to specify that the defined throughput floor be applied to each member workload individually. This is the only condition in which a policy group for a throughput floor can be applied to multiple workloads.

  Set `-is-shared=false` for the `qos policy-group create` command to specify a non-shared policy group.
- Throughput to a workload might fall below the specified floor if there is insufficient performance capacity (headroom) on the node or aggregate.
- A storage object that is subject to a QoS limit must be contained by the SVM to which the policy group belongs. Multiple policy groups can belong to the same SVM.
- It is a QoS best practice to apply a policy group to the same type of storage objects.
- A policy group that defines a throughput floor cannot be applied to an SVM.

Steps

1. Check for adequate performance capacity on the node or aggregate, as described in permalink :identify-remaining-performance-capacity-task.html[Identifying remaining performance capacity].
2. Create a policy group:

   `qos policy-group create -policy group` *`policy_group`* `-vserver` *`SVM`* `-min -throughput` *`qos_target`* `-is-shared true|false`

For complete command syntax, see the man page for your ONTAP release. You can use the `qos policy-group modify` command to adjust throughput floors.

The following command creates the shared policy group `pg-vs2` with a minimum throughput of 1,000 IOPS:

```
cluster1::> qos policy-group create -policy group pg-vs2 -vserver vs2
-min-throughput 1000iops -is-shared true
```

The following command creates the non-shared policy group `pg-vs4` without a throughput limit:

```
cluster1::> qos policy-group create -policy group pg-vs4 -vserver vs4
-is-shared false
```

3. Apply a policy group to a volume or LUN:

   `*storage_object* create -vserver *SVM* -qos-policy-group *policy_group*`

   For complete command syntax, see the man pages. You can use the `_storage_object_modify` command to apply a different policy group to the storage object.

   The following command applies policy group `pg-app2` to the volume `app2`:

```
cluster1::> volume create -vserver vs2 -volume app2 -aggregate aggr1
-qos-policy-group pg-app2
```

4. Monitor policy group performance:

   `qos statistics performance show`

   For complete command syntax, see the man page.

   > Monitor performance from the cluster. Do not use a tool on the host to monitor performance.

   The following command shows policy group performance:

```
cluster1::> qos statistics performance show
Policy Group           IOPS      Throughput   Latency
-------------------- -------- --------------- ----------
-total-                 12316       47.76MB/s  1264.00us
pg_app2                  7216       28.19MB/s   420.00us
_System-Best-Effort        62       13.36KB/s     4.13ms
_System-Background         30           0KB/s        0ms
```

5. Monitor workload performance:

```
qos statistics workload performance show
```

For complete command syntax, see the man page.

> Monitor performance from the cluster. Do not use a tool on the host to monitor performance.

The following command shows workload performance:

```
cluster1::> qos statistics workload performance show
Workload          ID     IOPS      Throughput     Latency
--------------- ------ -------- ---------------- ----------
-total-              -    12320        47.84MB/s  1215.00us
app2-wid7967      7967     7219        28.20MB/s   319.00us
vs1-wid12279     12279     5026        19.63MB/s     2.52ms
_USERSPACE_APPS     14       55        10.92KB/s   236.00us
_Scan_Backgro..   5688       20            0KB/s        0ms
```

> You can use the `qos statistics workload latency show` command to view detailed latency statistics for QoS workloads.

Use adaptive QoS policy groups

You can use an *adaptive QoS* policy group to automatically scale a throughput ceiling or floor to volume size, maintaining the ratio of IOPS to TBs|GBs as the size of the volume changes. That is a significant advantage when you are managing hundreds or thousands of workloads in a large deployment.

What you'll need

- You must be running ONTAP 9.3. Adaptive QoS policy groups are available beginning with ONTAP 9.3.
- You must be a cluster administrator to create a policy group.

About this task

A storage object can be a member of an adaptive policy group or a non-adaptive policy group, but not both. The SVM of the storage object and the policy must be the same. The storage object must be online.

Adaptive QoS policy groups are always non-shared: the defined throughput ceiling or floor applies to each member workload individually.

The ratio of throughput limits to storage object size is determined by the interaction of the following fields:

- `expected-iops` is the minimum expected IOPS per allocated TB|GB.

> `expected-iops` is guaranteed on AFF platforms only. `expected-iops` is guaranteed for FabricPool only if the tiering policy is set to "none" and no blocks are in the cloud. `expected-iops` is guaranteed for volumes that are not in a SnapMirror Synchronous relationship.

- `peak-iops` is the maximum possible IOPS per allocated or used TB|GB.
- `expected-iops-allocation` specifies whether allocated space (the default) or used space is used for expected-iops.

> `expected-iops-allocation` is available in ONTAP 9.5 and later. It is not supported in ONTAP 9.4 and earlier.

- `peak-iops-allocation` specifies whether allocated space or used space (the default) is used for `peak-iops`.
- `absolute-min-iops` is the absolute minimum number of IOPS. You can use this field with very small storage objects. It overrides both `peak-iops` and/or `expected-iops` when `absolute-min-iops` is greater than the calculated `expected-iops`.

  For example, if you set `expected-iops` to 1,000 IOPS/TB, and the volume size is less than 1 GB, the calculated `expected-iops` will be a fractional IOP. The calculated `peak-iops` will be an even smaller fraction. You can avoid this by setting `absolute-min-iops` to a realistic value.

- `block-size` specifies the application I/O block size. The default is 32K. Valid values are 8K, 16K, 32K, 64K, ANY. ANY means that the block size is not enforced.

Three default adaptive QoS policy groups are available, as shown in the following table. You can apply these policy groups directly to a volume.

Default policy group	Expected IOPS/TB	Peak IOPS/TB	Absolute Min IOPS
`extreme`	6,144	12,288	1000
`performance`	2,048	4,096	500
`value`	128	512	75

Steps

1. Create an adaptive QoS policy group:

   qos adaptive-policy-group create -policy group *policy_group* -vserver *SVM* -expected-iops *number_of_iops*/TB|GB -peak-iops *number_of_iops*/TB|GB -expected -iops-allocation-space|used-space -peak-iops-allocation allocated-space|used-space -absolute-min-iops *number_of_iops* -block-size 8K|16K|32K|64K|ANY

   For complete command syntax, see the man page.

   > `-expected-iops-allocation` and `-block-size` is available in ONTAP 9.5 and later. These options are not supported in ONTAP 9.4 and earlier.

   The following command creates adaptive QoS policy group `adpg-app1` with `-expected-iops` set to 300 IOPS/TB, `-peak-iops` set to 1,000 IOPS/TB, `-peak-iops-allocation` set to `used-space`, and `-absolute-min-iops` set to 50 IOPS:

```
cluster1::> qos adaptive-policy-group create -policy group adpg-app1
-vserver vs2 -expected-iops 300iops/tb -peak-iops 1000iops/TB -peak-iops
-allocation used-space -absolute-min-iops 50iops
```

2. Apply an adaptive QoS policy group to a volume:

   `volume create -vserver SVM -volume` *`volume`* `-aggregate` *`aggregate`* `-size` *`number_of`* `TB|GB -qos-adaptive-policy-group` *`policy_group`*

   For complete command syntax, see the man pages.

   The following command applies adaptive QoS policy group `adpg-app1` to volume `app1`:

```
cluster1::> volume create -vserver vs1 -volume app1 -aggregate aggr1
-size 2TB -qos-adaptive-policy-group adpg-app1
```

   The following commands apply the default adaptive QoS policy group `extreme` to the new volume `app4` and to the existing volume `app5`. The throughput ceiling defined for the policy group applies to volumes `app4` and `app5` individually:

```
cluster1::> volume create -vserver vs4 -volume app4 -aggregate aggr4
-size 2TB -qos-adaptive-policy-group extreme
```

```
cluster1::> volume modify -vserver vs5 -volume app5 -qos-adaptive-policy
-group extreme
```