



Technical Report

NetApp Storage Performance Primer for Clustered Data ONTAP 8.2

Roy Scaife, Paul Updike, Chris Wilson, NetApp
July 2013 | TR-4211

Abstract

This paper describes the basic performance concepts as they relate to NetApp® storage systems and the clustered Data ONTAP® operating system. It also describes how operations are processed by the system, how different features in clustered Data ONTAP can affect performance, and how to observe the performance of a cluster.

TABLE OF CONTENTS

1	Introduction	4
1.1	Performance Fundamentals	4
1.2	Normal Performance Relationships	4
2	System Architecture Overview	6
2.1	Connectivity: NICs and HBAs	7
2.2	Controller Subsystem: Memory, CPU, NVRAM	7
2.3	Storage Subsystem: Disks, Flash Cache, and Flash Pool	7
3	Data Storage and Retrieval	8
3.1	Cluster Operations	8
3.2	Node Operations	9
4	Introduction to Storage Quality of Service	12
4.1	The Need for Storage QoS	12
4.2	Storage QoS Concepts	13
4.3	Examples of Using Storage QoS	17
5	Performance Management with Clustered Data ONTAP	19
5.1	Basic Workload Characterization	20
5.2	Observing and Monitoring Performance	20
5.3	Managing Workloads with Data Placement	26
6	A Performance Management Scenario	26
6.1	Establish Monitoring	27
6.2	Root Cause an Alert	27
6.3	Resolve Performance Problem	28
7	Conclusion	29
	Appendices	29
	Data ONTAP 8.2 Upgrade Recommendations	29
	Running and Interpreting perf report -t	30
	References	31
	Version History	31

LIST OF TABLES

Table 1)	QoS Limits	17
Table 2)	SLA Levels	19

Table 3) QoS Throughput Labels	24
Table 4) Data ONTAP 8.2 Upgrade Recommendations	29

LIST OF FIGURES

Figure 1) High-level System Architecture	6
Figure 2) Scaling of the Architecture	7
Figure 3) Direct Data Access.....	8
Figure 4) Indirect Data Access	9
Figure 5) Read from Disk	10
Figure 6) Read from Memory	10
Figure 7) Write to Flash.....	10
Figure 8) Read from Flash.....	11
Figure 9) NVRAM Segmenting – Standalone Node and HA Pair Configurations	11
Figure 10) Accepting a Write	12
Figure 11) Consistency Point	12
Figure 12) Storage QoS	17
Figure 13) Performance Advisor Dashboard	21
Figure 14) Volume Latency Summary	21
Figure 15) Viewing Volume Latency Statistics.....	22
Figure 16) Adding a Threshold	22
Figure 17) Performance Advisor Threshold Details.....	23
Figure 18) Performance Advisor Alarms Configuration	23
Figure 19) Performance Advisor Latency Monitoring	27
Figure 20) Performance Advisor Latency Alert.....	27

1 Introduction

The demand on IT departments for storage has been steadily increasing, but budgets have not expanded accordingly. Many departments are trying to squeeze more out of their storage infrastructures in both capacity and performance. This document provides performance groundwork as well as the architecture of a NetApp storage system and how the architecture works with clustered Data ONTAP to provide efficiently performing data storage.

This document is not intended to be a guide on tuning or a deep troubleshooting guide, but rather a general overview of the architecture and operation, the performance management principles following the normal performance management paradigm, and the capabilities of NetApp Data ONTAP and FAS systems. Before reading this guide, you should understand the basic concepts of NetApp clustered Data ONTAP. For an introduction to clustered Data ONTAP, read [TR-3982: Clustered Data ONTAP 8.2: An Introduction](#).

1.1 Performance Fundamentals

Many variables affect the performance of a storage system. Out of all the metrics that can be measured, two specifically give the most insight into the performance of the storage system. The first, throughput, describes how much work the system is doing. This is presented in a unit of work per fixed unit of time (for example MB/s or IO/s). The second, latency, describes how fast the system is doing the work and is presented as the amount of time needed to complete a single operation (for example, ms/op).

The Data ONTAP operating system and the storage system hardware work together to solve these problems in a way that enables data to be safe, consistent, and provided rapidly. The performance of a system depends a lot on the workload applied to the system. Workload characteristics that can affect performance include:

- Concurrency – The number of operations in flight at any point in time.
- Operation Size – The size of the operations requested of the storage system.
- Operation Type – The type of operation requested of the storage system (i.e. read, write, other).
- Randomness – The distribution of data access across a dataset.
- Working Set Size – The amount of data considered to be active and needed to complete work.

Modifying any of these workload characteristics ultimately ends up affecting the performance of the system and can be observed in either latency or throughput. Altering the workload requested of the storage system is generally not a task a storage administrator has the luxury of doing. Therefore, to meet performance requirements, the storage administrator must observe the performance of the system and adapt as necessary by making changes to the storage system configuration.

1.2 Normal Performance Relationships

For the purposes of day-to-day management, there are a few guiding principles behind performance. These can be stated as the relationships between the fundamental characteristics of a workload and the resulting performance.

- Throughput is a function of latency.
- Latency is a function of throughput.
- Throughput is a function of concurrency.
- Throughput is a function of operation size.
- Throughput is a function of randomness of operations.
- The host application controls the amount and type of operations.

Throughput and Latency

Workloads can be defined as either closed-loop or open-loop systems. In closed-loop systems, a feedback loop exists. Operation requests are dependent upon the completion of previous operations. In this scenario the number of concurrent requests is fixed and the rate that operations that can be completed depends on how long it took (latency) for previous operations to be completed. Simply put, in closed-loop systems throughput is a function of latency; if latency increases, throughput decreases.

In open-loop systems, operations are performed without relying on feedback from previous operations. This means that the response time from those operations doesn't affect when other operations will be requested. The requests will occur when necessary from the application. As throughput increases to the system, the utilization of the resources increases. As the resource utilization increases, so does operation latency. Because of this utilization increase, we can say that latency is a function of throughput in open-looped systems, although indirectly.

Concurrency

Storage systems are designed to handle many operations at the same time. In fact, peak efficiency of the system can never be reached until it is processing a large enough number of I/Os such that there is always an operation waiting to be processed behind another process. Concurrency, the number of outstanding operations in flight at the same time, allows the storage system to handle the workload in a more efficient manner. The effect can be dramatic in terms of throughput results.

Little's Law: A Relationship of Throughput, Latency, and Concurrency

Little's Law describes the observed relationship between throughput (arrival rate), latency (residence time) and concurrency (residents):

$$L = A \times W$$

This equation says that the concurrency of the system (L) is equal to the throughput (A) multiplied by latency (W). This would mean that for higher throughput, either concurrency would have to increase and/or latency to decrease. This explains why low concurrency workloads, even with low latencies can have lower than expected throughput.

Operation Size

A similar effect on concurrency is observed with the size of operations on a system. More work, when measured in megabytes per second, can be done with larger operations than can be done with smaller operations. Each operation has overhead associated with it at each point along the way in transfer and processing. By increasing the operation size, the ratio of overhead to data is decreased which allows more throughput in the same time. Similarly, when work depends on latency in low concurrency workloads, a larger operation size increases the efficiency of each individual operation.

Small operations might have a slightly better latency than large operations, so the operations per second could be potentially higher, but the throughput in megabytes will hold a general trend of being lower with smaller operations.

Randomness

Protocol operations sent to a storage system are assigned to a logical location within a data file or LUN. This logical address is subsequently translated into an actual physical location on the permanent storage media. The order of operations and the location of the data being accessed over time determine how random a workload is. If the logical addresses are in order (next to one another) they are considered sequential.

For read operations, performance improves on a NetApp storage system for sequential data. This is because fewer drive seeks and operations are required from one disk I/O operation to the next.

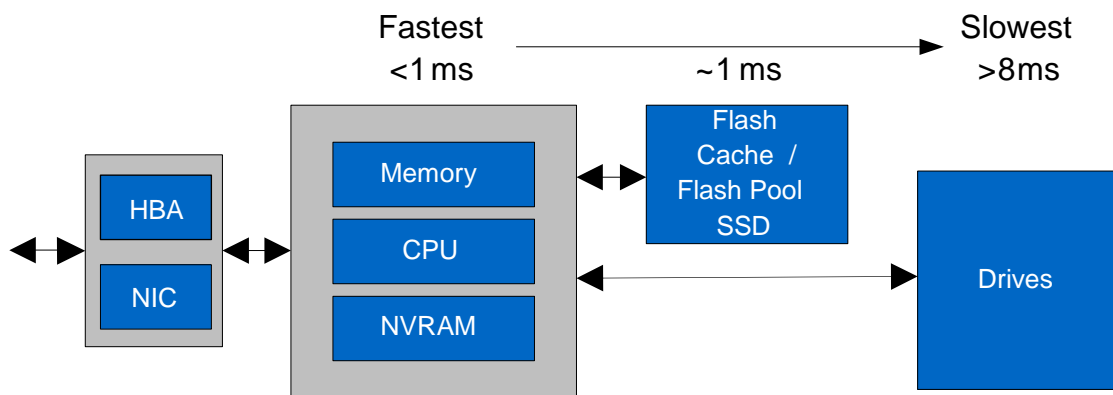
Data ONTAP is write-optimized. Due to the way writes are written to storage almost all writes behave as if they are sequential writes. Thus, we see less improvement in random versus sequential writes.

2 System Architecture Overview

Storage systems are designed to store and retrieve large amounts of data permanently, inexpensively, and quickly. Unfortunately, to store lots of data you need to use a slow medium: the mechanical disk drive. To access data quickly, you need a fast medium such as silicon-based Random Access Memory (RAM), which is neither persistent nor inexpensive. It is also important to remember that different workloads affect different parts of the system in different ways. This creates a problem as to how to optimize access to data to provide the best performance. NetApp does this by innovating in the way data is stored and accessed through the use of unique combinations of spinning disk, flash, and RAM.

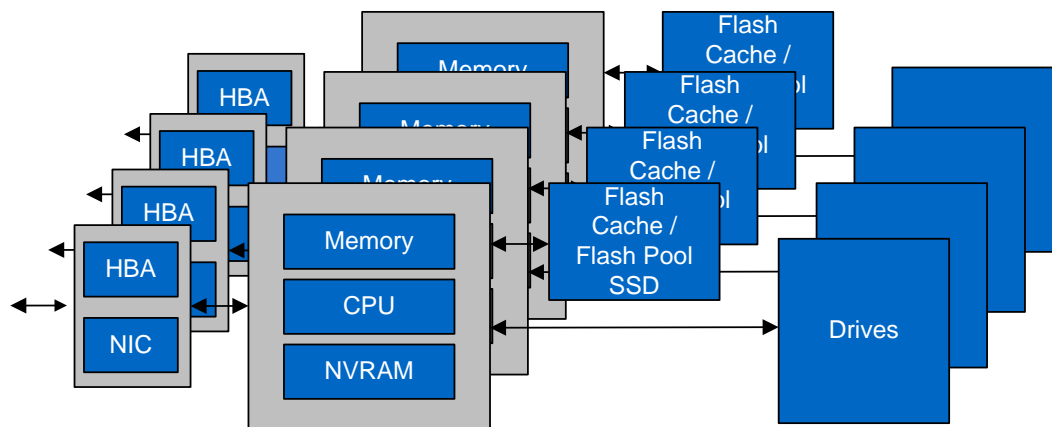
A NetApp storage system may be logically divided into three main areas when discussing performance. Those are connectivity, the system itself, and the storage subsystem. When we speak of connectivity, we refer to the network and HBA interfaces that connect the storage system to the clients and hosts. The system itself is the combination of CPU, memory, and NVRAM. Finally, the storage subsystem consists of the disks, and also Flash Cache™ and Flash Pool™ intelligent caching. The following picture logically represents a NetApp system.

Figure 1) High-level System Architecture



A system running clustered Data ONTAP consists of individual nodes joined together by the cluster interconnect. Every node in the cluster is capable of storing data on disks attached to it, essentially adding “copies” of the above architecture to the overall cluster. Clustered Data ONTAP has the capability to nondisruptively add additional nodes to the system to scale out both the performance and capacity of the system.

Figure 2) Scaling of the Architecture



2.1 Connectivity: NICs and HBAs

NICs and HBAs provide the connectivity to client, management, and cluster-interconnect networks. Adding more or increasing the speed of NICs or HBAs can scale client network bandwidth.

2.2 Controller Subsystem: Memory, CPU, NVRAM

Common to most systems, NetApp systems contain CPUs and some amount of memory, depending on the controller model. As with any computer, the CPUs serve as the processing power to complete operations for the system. Besides serving operating system functions for Data ONTAP, the memory in a NetApp controller also acts as a cache. Incoming writes are coalesced in main memory prior to being written to disk. Memory is also used as a read cache to provide extremely fast access time to recently read data.

NetApp systems also contain NVRAM. NVRAM is battery-backed memory that is used to protect in-bound writes as they arrive. This allows write operations to be committed safely without having to wait for a disk operation to complete and reduces latency significantly. High-availability (HA) pairs are created by mirroring NVRAM across two controllers.

Increasing the capacity of these components requires upgrading to a higher controller model. Clustered Data ONTAP allows nodes to be evacuated and upgraded nondisruptively to clients.

2.3 Storage Subsystem: Disks, Flash Cache, and Flash Pool

Spinning disk drives are the slowest components in the whole storage system. The typical response times for spinning disks are a few milliseconds. The performance of disk drives varies depending on the disk type and rotation speed: 7.2K RPM SATA disks have higher latency than 10K RPM SAS disks. Solid-state disks significantly reduce the latency at the storage subsystem. Ultimately, the type of disk needed for a specific application depends on capacity and performance requirements as well as the workload characteristics.

With the introduction of Flash Cache and Flash Pool, it is possible to combine the performance of solid-state flash technology with the capacity of spinning media. Flash Cache operates as an additional layer of read cache for the entire system. It caches recently read, or “hot,” data for future reads. Flash Pool serves as a read cache similar to Flash Cache at the aggregate level. Flash Pool is also capable of offloading random overwrites that are later destaged to disk to improve write performance.

3 Data Storage and Retrieval

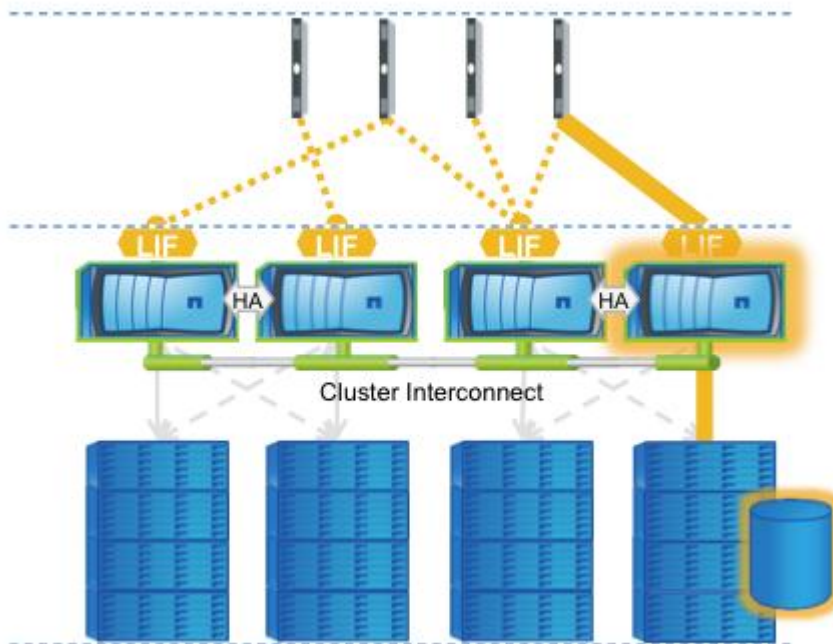
3.1 Cluster Operations

Data being stored or accessed does not need to reside on the node connected to the client. Data can be accessed directly or indirectly across the cluster.

Direct Data Access

Direct data access occurs when a client connected to a node accesses data stored directly on that node. When accessing data in this fashion there is no traversal of the cluster interconnect. Direct data access provides the lowest latency access. Therefore, having clients connect to LIFs on the same node as the data is beneficial to performance, but it is not necessary. Different protocols behave differently and have different features when it comes to clustered Data ONTAP. The section on protocol considerations discusses these differences.

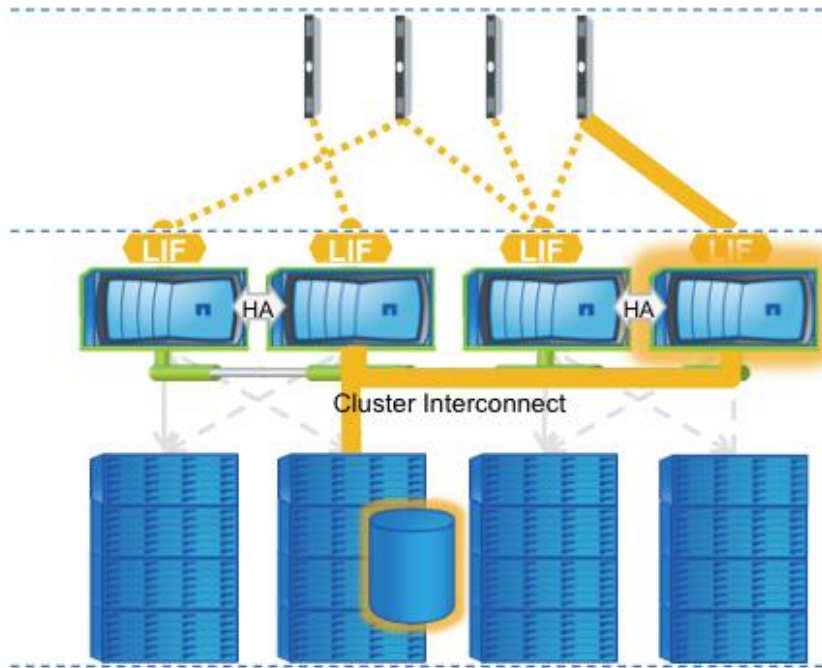
Figure 3) Direct Data Access



Indirect Data Access

Indirect data access occurs when a client accesses one node but the data is stored physically on another node. The node the client is communicating with identifies where the data is stored and accesses the other node via the cluster interconnect and then serves the data back to the client. Indirect data access allows data to live physically on any node without the need to force clients to mount more than a single location to access the data.

Figure 4) Indirect Data Access



Protocol Considerations

Accessing data directly on the node where it is stored reduces the amount of resources necessary to serve data and is ultimately the “shortest path” to where the data lives. Some of the protocols supported by clustered Data ONTAP have the ability to automatically provide direct data access. Independent of protocols, the management features of clustered Data ONTAP can be used to alter the data access path.

Certain protocols have the capability to automatically direct traffic to the node with direct data access. In the case of NAS protocols, NFS version 4 (NFSv4) can direct clients to local nodes through a variety of capabilities. NFSv4 referrals point the client to the directly attached node during mount. Another capability with NFSv4.1 is parallel NFS (pNFS). pNFS enables clients to connect to any node in the cluster for metadata work while performing direct data operations. To learn more about NFS capabilities in clustered Data ONTAP, read [TR-4067: NFSv3/v4 in Data ONTAP 8.1 Operating in Cluster-Mode Implementation Guide](#) and [TR-4063: Parallel Network File System Configuration and Best Practices for Data ONTAP Cluster-Mode](#). Similarly, the SMB 2.0 and 3.0 protocols support a feature called Auto Location. This capability automatically directs a client to the direct node when mounting a share. More information is available in the Data ONTAP documentation.

In SAN environments, the ALUA protocol enables optimal pathing to a LUN. Even if volumes are moved around in the cluster, the host will always access the LUN through the optimal path. To learn more about using SAN with clustered Data ONTAP, read [TR-4080: Best Practices for Scalable SAN in Data ONTAP 8.1 Cluster-Mode](#).

3.2 Node Operations

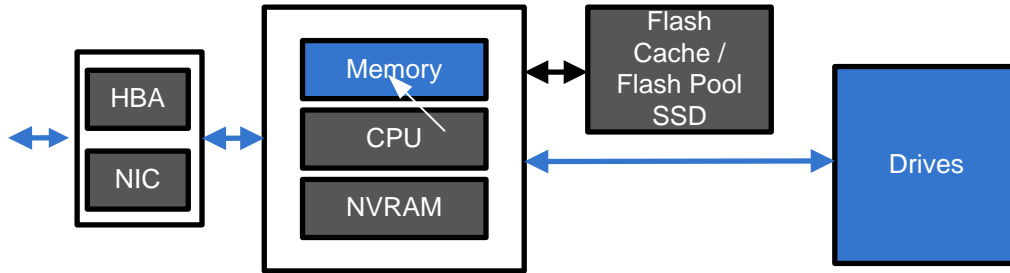
Once an operation has been directed to the proper node, that node becomes responsible for completing the read or write operation. In this section, we examine how reads and writes are completed on a node and how the components within the storage system are used.

Reads

Recall the storage system architecture presented in section 2; reads can be serviced from memory, flash-based cache, or spinning disk drives. The workload characteristics and capabilities of the system determine where reads are serviced and how fast. Knowing where reads are serviced can help set expectations as to the overall performance of the system. In the following diagrams, components and links in blue highlight the activity described.

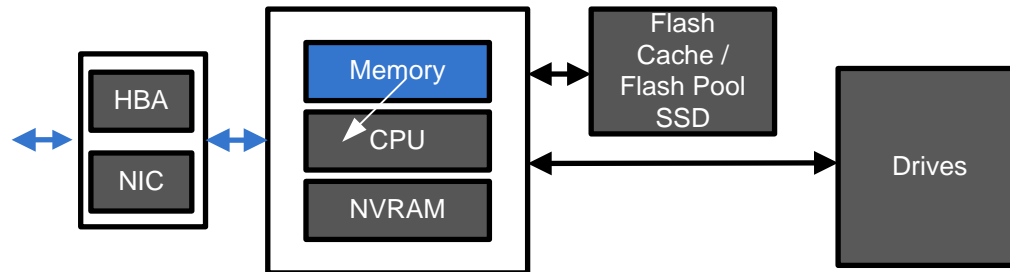
In the simple yet slowest case, read requests that are not cached anywhere are forced to come from disk. Once read from disk, the data is kept in main memory.

Figure 5) Read from Disk



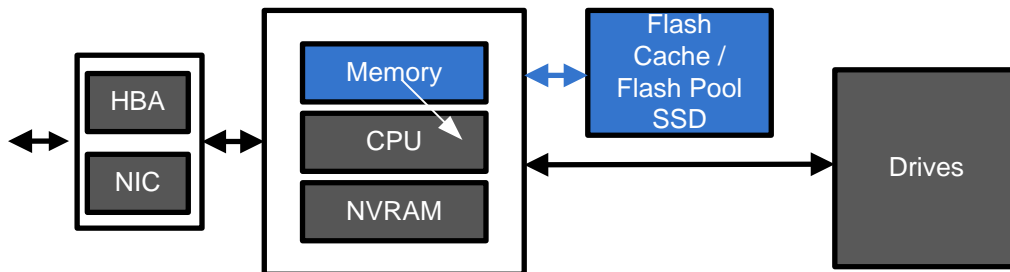
If this data is read again soon, it is possible for the data to be cached in main memory, making subsequent access extremely fast since no disk access would be required.

Figure 6) Read from Memory



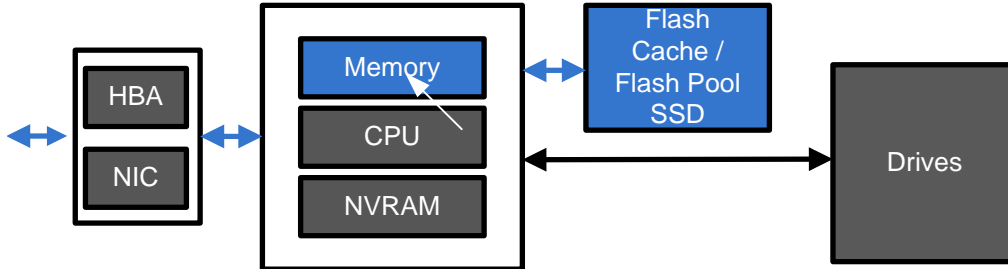
When more room is needed in the main memory cache, as is common with working sets larger than the buffer cache, the data is evicted. If Flash Cache or Flash Pool is in the system, that block could be inserted into the flash-based cache if it meets certain requirements. In general, only randomly read data and metadata are inserted into flash-based caches.

Figure 7) Write to Flash



Once inserted, subsequent reads of this block unable to be serviced from the buffer cache would be served from the flash-based cache until they are evicted from the flash-based cache. Flash access times are significantly faster than those of disk and adding cache in random read-intensive workloads can reduce read latency dramatically.

Figure 8) Read from Flash

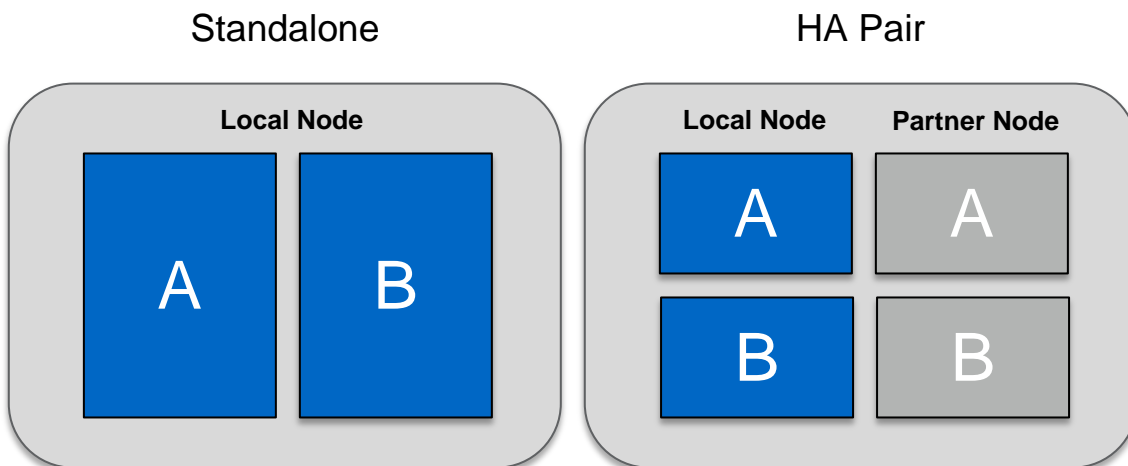


Incoming reads are continually being checked for patterns. For some data access patterns, such as sequential access, Data ONTAP has the capability to predict which blocks a client may want to access prior to the client ever requesting the read. This “read-ahead” mechanism preemptively reads blocks off disk and caches them in main memory. These read operations can be serviced at faster RAM speeds instead of waiting for disk when the read request is received.

Writes

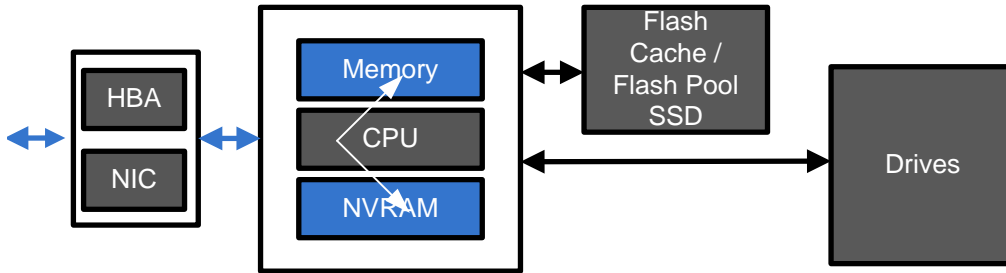
Next, consider how data is written to the storage system. For most storage systems, writes must be placed into a persistent and stable location prior to acknowledging to the client or host that the write was successful. As we know, drives are slow, and waiting for the storage system to write an operation to disk for every write could introduce significant latency. To solve this problem, NetApp storage systems use battery-backed RAM to create nonvolatile RAM to log incoming writes. NVRAM is divided in half and only one half is used at a time to log incoming writes. When controllers are in highly available pairs, half of the NVRAM is used to mirror the partner node’s log while the other half is used for logging writes. The part that is used for logging locally is still split in half just like a single node.

Figure 9) NVRAM Segmenting – Standalone Node and HA Pair Configurations



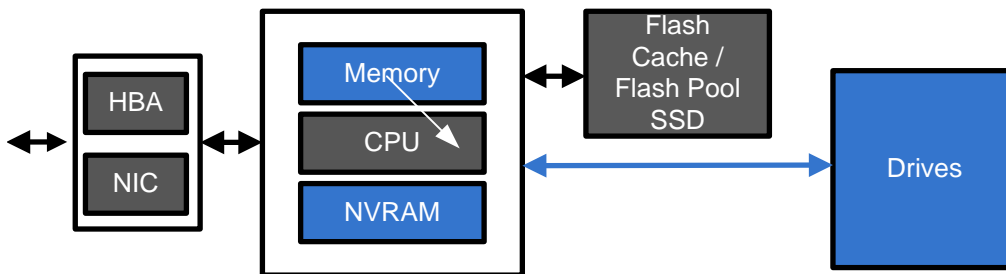
When a write enters a NetApp system, the write is logged into NVRAM and is buffered in main memory. Once the data is logged in persistent NVRAM, the client is acknowledged. NVRAM is accessed only in the event of a failure.

Figure 10) Accepting a Write



At a later point in time, called a consistency point, the data buffered in main memory is efficiently written to disk. Consistency points can be triggered for a number of reasons, including time passage, NVRAM fullness, or system-triggered events such as a snapshot.

Figure 11) Consistency Point



In general, writes take a minimal amount of time, on the order of low milliseconds to sub-milliseconds. If the disk subsystem is unable to keep up with the client workload and becomes too busy, write latency can begin to increase. When writes are coming in too fast for the back-end storage, both sides of the NVRAM can fill up and cause a scenario called a back-to-back CP. This means that both sides are filled up, a CP is occurring, and another CP will immediately follow on the current CP's completion. This scenario impacts performance because the system can't immediately acknowledge the write as persistent because NVRAM is full and must wait until the operation can be logged. The back-to-back CP scenario can most often be alleviated by improving the storage subsystem. Increasing the number of disks, moving some of the workload to other nodes, and considering flash-based caching can help solve write performance issues. Remember, only randomly overwritten data is written to the SSD portion of a Flash Pool and that Flash Cache is only a read cache but that off-loading any type of operation can reduce disk utilization.

4 Introduction to Storage Quality of Service

4.1 The Need for Storage QoS

Storage Quality of Service (QoS) gives the storage administrator the ability to control storage object workloads to deliver consistent performance and meet performance goals for critical workloads.

Consistent Performance

Storage QoS is a feature within clustered Data ONTAP designed to help address the need for consistent workload performance. In environments without a quality-of-service (QoS) capability the dynamic effects

of an environment can cause utilization to rise to levels that are unacceptable, with some low-priority workloads taking more than their share of resources while higher-priority work is denied the resources it needs. The Storage QoS feature allows specific maximums to be set for groups of workloads, setting an upper boundary on the amount of throughput they may consume. With this capability, you can retain resources for important work by restricting access to the less important workloads.

Workload Isolation

Similarly, some workloads may need to be isolated from other workloads in the cluster. Rogue workloads can consume an enormous amount of resources and reduce the available amount for others. The ability to monitor and then isolate rogue workloads creates value for the administrator in dynamic environments. Storage QoS accomplishes these tasks with a simple command line interface that allows settings to be configured for the cluster on the fly, without requiring an extended planning process or a complicated interface.

4.2 Storage QoS Concepts

Before discussing examples and use cases for Storage QoS it is important to understand some basic QoS concepts and terminology.

Workload

A workload is the set of I/O requests sent to one or more storage objects. In clustered Data ONTAP 8.2, QoS workloads include I/O operations and data throughput, and they are measured in IO/s and MB/s, respectively. IO/s (IOPS) workload measurement includes all client I/O, including metadata and disk I/O, regardless of I/O block size. I/O related to system processes is not counted in the IOPS measurement.

Storage Objects

A storage object is the entity on the controller to be assigned to a QoS policy group for monitoring and control. QoS storage objects can be any of the following:

- Storage Virtual Machines (SVMs), formerly called Vservers.
- FlexVol® volumes
- LUNs
- Files

Policies

QoS policies are behaviors to apply to a QoS policy group and its storage objects. In clustered Data ONTAP 8.2, you can define a QoS policy to impose a throughput limit on the storage objects in the QoS policy group. This throughput limit is applied collectively to the group. QoS policies may be configured to control IO/s or MB/s throughput. In addition, the QoS policy may be configured to `none` to allow the storage administrator to monitor the workload throughput of the storage objects in the QoS policy group without limiting the workload throughput.

Limits

As previously discussed, the storage administrator can control the workload throughput using IO/s or MB/s limits. When the workload throughput exceeds the QoS policy limit, the workload is reduced at the protocol layer. The storage administrator should expect the response time for I/O requests to increase while QoS throttles the workload. Occasionally, some applications may time out. This behavior is no different than when a system runs out of performance headroom. Throttling a workload in the protocol stack prevents it from consuming incremental cluster resources, thus freeing up resources for the other workloads deployed on the cluster.

When the QoS policy is configured to throttle IOPS or MBPS, the specified policy value is a hard limit. The storage administrator should be aware that the workload IOPS or MBPS throughput may exceed the value set in the QoS policy by up to 10% while the I/O operations are queued and throttled. As a general rule, the lower the QoS policy limit the higher the deviation from the limit while the policy takes effect. I/O operations queued as a result of hitting the QoS policy limit do not impact cluster resources. QoS policies are applicable to all supported protocols, including NFS, SMB, SAN, iSCSI, and FCoE, except NFS 4.1.

Note: QoS is not compatible with NFS 4.1 in clustered Data ONTAP 8.2.

When to Use MB/s

For large block IO workloads NetApp recommends configuring the QoS policy using MB/s.

When to Use IO/s

For transactional workloads NetApp recommends configuring the QoS policy using IO/s.

Policy Groups

QoS policy groups are collections of storage objects (that is, SVMs, volumes, LUNs, or files) to enable the storage administrator to monitor and control workload throughput. One QoS policy (behavior) can be assigned to a QoS policy group. The storage administrator can monitor storage object workloads by assigning the storage objects to a policy group without applying a QoS policy.

Note: Only one QoS policy may be applied to a QoS policy group.

Storage objects assigned to QoS policy groups are SVM (Vserver) scoped. This means that each QoS policy group may have only storage objects assigned to it from a single SVM. QoS policy groups support assignment of several FlexVol volumes, LUNs, and files within the same SVM. The IO limits are applied collectively across the storage objects in a policy group, and are not applied at an individual storage object-level. Individual storage objects within a policy group are expected to consume resources using a fair-share methodology.

Note: The QoS policy throughput limit is applied to the *aggregate* throughput of all storage object workloads assigned to the policy group.

Nested storage objects cannot be assigned to the same or a different QoS policy group. For example, a VMDK file and its parent volume may not both be assigned to a QoS policy group.

Note: Nested storage objects may not both be assigned to the same or a different QoS policy group in clustered Data ONTAP 8.2.

QoS policy group membership remains unchanged as storage objects are moved within the cluster. However, as previously discussed, storage objects cannot be nested. For example, if a VMDK file which is part of a policy group is moved to a different datastore (volume) which is already part of a policy group, then the VMDK file will no longer be assigned to the policy group.

Some environments may utilize NetApp FlexCache[®] functionality to enhance performance. When a FlexVol volume leveraging FlexCache is assigned to a QoS policy group the FlexCache volume workload is included in the QoS policy group workload.

Monitor

Assigning storage objects to a QoS policy group without a QoS policy, or modifying an existing QoS policy limit to `none`, gives the storage administrator the ability to monitor the workload placed on those storage objects without limiting the workload throughput. In this configuration the storage administrator can monitor workload latency, IOPS, and data throughput. Storage QoS measures latency from the network interface to and from the disk subsystem.

Creating a QoS policy group to monitor workload latency and throughput:

```
Cluster1::> qos policy-group create -policy-group monitor_workload -vserver vserver1
```

Assigning volumes to a QoS policy group:

```
Cluster1::> vol modify -vserver vserver1 -volume vol1 -qos-policy-group monitor_workload  
(volume modify)
```

Volume modify successful on volume: vol1

```
Cluster1::> vol modify -vserver vserver1 -volume vol2 -qos-policy-group monitor_workload  
(volume modify)
```

Volume modify successful on volume: vol2

Assigning a SVM to a QoS policy group:

```
Cluster1::> vserver modify -vserver vserver2 -qos-policy-group vserver2_qos_policy_group
```

Displaying the QoS policy group configuration and the number of workloads assigned to the policy group:

```
Cluster1::> qos policy-group show  
Name           Vserver      Class           Wklds  Throughput  
-----  
monitor_workload vserver1     user-defined    2      0-INF  
vol1_qos_policy vserver1     user-defined    0      0-500IOPS  
vol2_qos_policy vserver1     user-defined    0      0-100MB/S  
3 entries were displayed.
```

Note: QoS policy groups that do not have a throughput limit are shown with 0-INF. This represents an infinite QoS policy limit.

Viewing the QoS policy group latency statistics:

```
cluster1::> qos statistics latency show  
Policy Group  Latency  Network  Cluster  Data  Disk  QoS  
-----  
-total-      16ms    6ms      2ms     3ms   4ms   1ms
```

Viewing the QoS policy group performance statistics:

```
cluster1::> qos statistics performance show  
Policy Group  IOPS      Throughput  Latency  
-----  
-total-      12224     47.75MB/s  512.45us  
rogue_policy  7216     28.19MB/s  420.00us  
prevent_policy 5008     19.56MB/s  92.45us
```

For more information on the QoS policy group monitoring commands, see the “Clustered Data ONTAP 8.2 Commands: Manual Page Reference.”

Control

QoS policy groups with a policy can control and limit the workloads of the storage objects assigned to the policy group. This capability gives the storage administrator the ability to manage and, when appropriate, throttle storage object workloads. In clustered Data ONTAP 8.2, the storage administrator can control I/O and data throughput. When a policy is configured the storage administrator can continue to monitor the latency, IOPS, and data throughput workloads of storage objects.

Creating a QoS policy group to control workload IOPS:

```
Cluster1::> qos policy-group create -policy-group vol1_qos_policy_group -max-throughput 500iops  
-vserver vserver1
```

Creating a QoS policy group to control workload data throughput:

```
Cluster1::> qos policy-group create -policy-group vol2_qos_policy_group -max-throughput 1000MBPS
-vserver vserver1
```

Assigning volumes to QoS policy groups:

```
Cluster1::> vol modify -vserver vserver1 -volume vol1 -qos-policy-group vol1_qos_policy_group
(volume modify)

Volume modify successful on volume: vol1

Cluster1::> vol modify -vserver vserver1 -volume vol2 -qos-policy-group vol2_qos_policy_group
(volume modify)

Volume modify successful on volume: vol2
```

Assigning LUNs to QoS policy groups:

```
Cluster1::> lun modify -vserver vserver1 -lun lun1 -vol vol2
-qos-policy-group lun_qos_policy_group
```

Assigning files to QoS policy groups:

```
Cluster1::> volume file modify -vserver vserver1 -vol vol2 -file log.txt
-qos-policy-group file_qos_policy_group
```

Displaying the QoS policy group configuration and the number of workloads assigned to the policy group:

```
Cluster1::> qos policy-group show
Name          Vserver      Class      Wklds  Throughput
-----
monitor_workload vserver1    user-defined 0      0-INF
vol1_qos_policy_group
              vserver1    user-defined 1      0-500IOPS
vol2_qos_policy_group
              vserver1    user-defined 1      0-100MB/S
3 entries were displayed.
```

For more information on the QoS policy group monitoring commands, see the “Clustered Data ONTAP 8.2 Commands: Manual Page Reference.”

Storage QoS Summary

The Storage QoS capability in NetApp clustered Data ONTAP 8.2 enables customers to increase utilization of storage resources by consolidating multiple workloads in a single shared storage infrastructure, while minimizing the risk of workloads impacting each other’s performance. Administrators can prevent tenants and applications from consuming all available resources in the storage infrastructure, improving the end-user experience and application uptime. In addition, pre-defining service level objectives allows IT to provide different levels of service to different stakeholders and applications, ensuring that the storage infrastructure continues to meet the business needs.

Storage QoS adds new capabilities for the storage administrator to monitor and control user workloads. Following is a brief summary of the QoS functionality delivered in clustered Data ONTAP 8.2.

- Monitor and manage storage object workloads
- Control I/O and data throughput workloads on SVMs, volumes, LUNs, and files
- Multiprotocol support, including SMB, SAN, iSCSI, FCoE, NFS (except NFS 4.1)
- Provisioning of policy groups in Workflow Automation (WFA) 2.1
- QoS support for V-Series

However, there are a few caveats to remember when you consider QoS:

- QoS is not supported on Infinite Volumes

- Alerts may not be configured for QoS
- QoS does not provide workload guarantees
- QoS is not supported with NFS v4.1

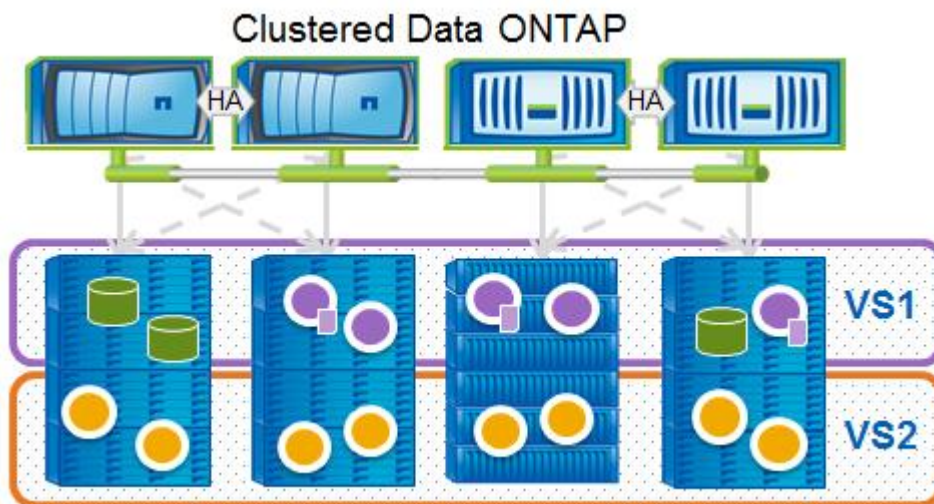
Table 1) QoS Limits

QoS Feature Area	Maximum	
	Per Node	Per Cluster
QoS policy groups supported	3,500	3,500
Number of controllers supported by QoS	1	8
Storage objects assigned to a QoS policy group	10,000	10,000

4.3 Examples of Using Storage QoS

Storage QoS will have many applications for the storage administrator. Below are a few scenarios that illustrate QoS capabilities. The first use case is an example in which the storage administrator throttles a “rogue” workload that is impacting other workloads. The second scenario describes how a storage administrator may prevent runaway (or rogue) workloads by proactively setting QoS policies. The final use case looks at managing workloads so that service providers can meet their service-level agreements (SLAs).

Figure 12) Storage QoS



Reactively Respond

In this scenario the storage administrator has not applied any storage objects to a QoS policy group. By default, Data ONTAP treats all storage objects on a best-effort basis. However, one of the storage objects (that is, a volume) has a rogue workload impacting the performance of other workloads on the system. Using Data ONTAP `statistics` and `qos statistics` commands the storage administrator can identify the rogue workload. Once the rogue workload is identified the storage administrator can use Storage QoS to isolate the workload by assigning it to a QoS policy group and applying a throughput limit.

Throttle Rogue Workloads

After identifying the rogue workload the storage administrator creates a new QoS policy group and sets the I/O throughput limit to 1,000 IOPS.

```
Cluster1::> qos policy-group create voll_rogue_qos_policy -max-throughput 1000iops
                -vserver vserver1
```

To view the QoS policy group configuration the storage administrator may use the `qos policy-group show` command.

```
Cluster1::> qos policy-group show -policy-group voll_rogue_qos_policy

    Policy Group Name: voll_rogue_qos_policy
                Vserver: vserver1
                Uuid: a20df2c2-c19a-11e2-b0e1-123478563412
    Policy group class: user-defined
                Policy Group ID: 102
    Maximum Throughput: 1000IOPS
    Number of Workloads: 0
    Throughput Policy: 0-1000IOPS
```

Next, the offending volume is assigned to the QoS policy group to begin throttling the rogue workload. The `modify` option of storage objects is used to assign an existing storage object to a QoS policy group.

```
Cluster1::> volume modify -vserver vserver1 -volume voll -qos-policy-group voll_rogue_qos_policy
```

The storage administrator can verify that the volume has been assigned to the QoS policy group using the `volume show` command.

```
Cluster1::> volume show -vserver vserver1 -volume voll -fields qos-policy-group
vserver  volume qos-policy-group
-----  -----
vserver1 voll    voll_rogue_qos_policy
```

Proactively Prevent Runaway Workloads

This is a scenario in which the storage administrator proactively sets a QoS policy group for the storage objects to prevent the impact of new, and possibly runaway, workloads. This situation may arise in a large virtualized environment in which the storage administrator needs to prevent a development or test application from impacting other production applications.

Apply Limits Before a Problem Occurs

The first step is to create a QoS policy group and apply a throughput limit.

```
Cluster1::> qos policy-group create -policy-group vmdk_13_qos_policy_group
                -max-throughput 100iops -vserver vserver1
```

Once the QoS policy group has been created the storage objects are assigned to the policy group. It is important to remember that the QoS limit is applied to the aggregate throughput of all storage objects in the QoS policy group.

```
Cluster1::> volume file modify -vserver vserver1 -vol vol2 -file vmdk-13.vmdk
                -qos-policy-group vmdk_13_qos_policy_group
```

Lastly, the storage administrator should monitor the storage object workload and adjust the policy as needed. Changes to the policy throughput limit can be completed quickly without impacting other workloads.

```
Cluster1::> qos statistics performance show
Policy Group      IOPS      Throughput  Latency
-----
-----
```

-total-	867	47.75MB/s	512.45us
voll_rogue_qos_policy	769	28.19MB/s	420.00us
vmdk_13_qos_policy_group	98	19.56MB/s	92.45us

After reviewing the required IOPS resources for vmdk-13 with engineering, the storage administrator agrees to increase the policy throughput limit to 200 IOPS.

```
Cluster1::> qos policy-group modify -policy-group vmdk_13_qos_policy_group
-max-throughput 200iops
```

Isolate Tenants with per-SVM Throughput Limits

In our final use case we look at a service provider who needs to isolate customer workloads to meet the service-level agreements. A new SVM is created in the cluster for each customer and the service provider must enable workloads to be controlled based on the SLA level. This service provider has three SLA levels - Bronze, Silver, and Gold - corresponding to the maximum data throughput allowed.

Table 2) SLA Levels

SLA Level	Data Throughput
Bronze	100MBPS
Silver	200MBPS
Gold	400MBPS

Once the service provider determines the SLA throughput limits the storage administrator can create the Storage QoS policy group with the appropriate limit and assign the SVM storage object to the policy group. For this example, we use three fictional service provider customers - Acme, Bravos, and Trolley - who have purchased the Bronze, Silver, and Gold service levels, respectively.

Create a policy group with the appropriate throughput limit (determined by service-level) and assign the SVM for each customer to the policy group:

```
Cluster1::> qos policy-group create -policy-group acme_svm_bronze -max-throughput 100MBPS
-vserver acme_svm

Cluster1::> qos policy-group create -policy-group bravos_svm_silver -max-throughput 200MBPS
-vserver bravos_svm

Cluster1::> qos policy-group create -policy-group trolley_svm_gold -max-throughput 400MBPS
-vserver trolley_svm
```

Apply the SVM for each customer to the QoS policy group:

```
Cluster1::> vservers modify -vserver acme_svm -qos-policy-group acme_svm_bronze

Cluster1::> vservers modify -vserver bravos_svm -qos-policy-group bravos_svm_silver

Cluster1::> vservers modify -vserver trolley_svm -qos-policy-group trolley_svm_gold
```

5 Performance Management with Clustered Data ONTAP

Ensuring the performance of a system is essential throughout the lifecycle. Some applications have more stringent requirements on performance than others, but performance is never a non-requirement. Performance management starts with the first request for workload characteristics and continues until the system is decommissioned. Being able to understand workloads, identify problems, and relate them back to the system's operation is essential to achieving performance goals.

This section introduces the capabilities of Data ONTAP and other NetApp software to complete performance management functions, including looking at statistics and using features to alter the performance of the system or workloads.

5.1 Basic Workload Characterization

As mentioned earlier, the workload characteristics and system architecture ultimately define the performance of the system. Also mentioned earlier were the Storage QoS capabilities available in clustered Data ONTAP 8.2. You can use the statistics generated by QoS to understand the characteristics of workloads in the system. These characteristics can then be used when sizing or setting expectations for performance. When reviewing this data, keep in mind the relationships introduced in section 1.2.

The following command shows the workload characteristics of the busiest workloads on the system:

```
Cluster1::> qos statistics workload characteristics show
```

Workload	ID	IOPS	Throughput	Request size	Read	Concurrency
-total-	-	5076	37.88MB/s	7825B	65%	2
volume_a	14368	4843	37.82MB/s	8189B	68%	2
...						

Although this is just an example, the data above shows that the system is processing about 5,000 IOPS, roughly 8kB in size, with 65% being reads.

Entire policy group characteristics can be viewed by eliminating the workload part of the command, as in the following:

```
Cluster1::> qos statistics characteristics show
```

These examples are basic, and more statistics than are presented here are available in Data ONTAP.

5.2 Observing and Monitoring Performance

Keeping an eye on performance can help quickly resolve performance problems and assist in deciding if additional load to the system is advisable. Latency should be considered the primary indicator of performance. Corroborating metrics for latency include throughput and resource utilizations. This means that if high latency is observed alongside increased throughput, a workload may have grown and is causing the saturation of a resource in the system, confirmed by resource utilization. Low throughput is not necessarily a problem, since clients may simply not be requesting that work be done. Ideally, workload objects are the best to monitor because they likely correspond to an application or tenant. Other abstractions are also useful to monitor, including the volume abstraction.

Data ONTAP provides statistics that can be monitored through graphical tools as well as through the cluster CLI and APIs. The following subsections introduce ways to monitor performance metrics using NetApp tools and on-box features. In advanced or more complex environments, these CLI commands and related APIs from the NetApp SDK can be used to create custom monitoring tools.

Performance Advisor

OnCommand Performance Advisor is a graphical performance tool for clustered Data ONTAP and is part of the NetApp Management Console, bundled with OnCommand® Unified Manager. It provides the ability to establish thresholds for important metrics (like latency) on monitored objects as well as produce graphs showing the performance of both the physical and logical objects in the storage system. These features make it useful for monitoring numerous systems and helping identify problems as soon as they occur.

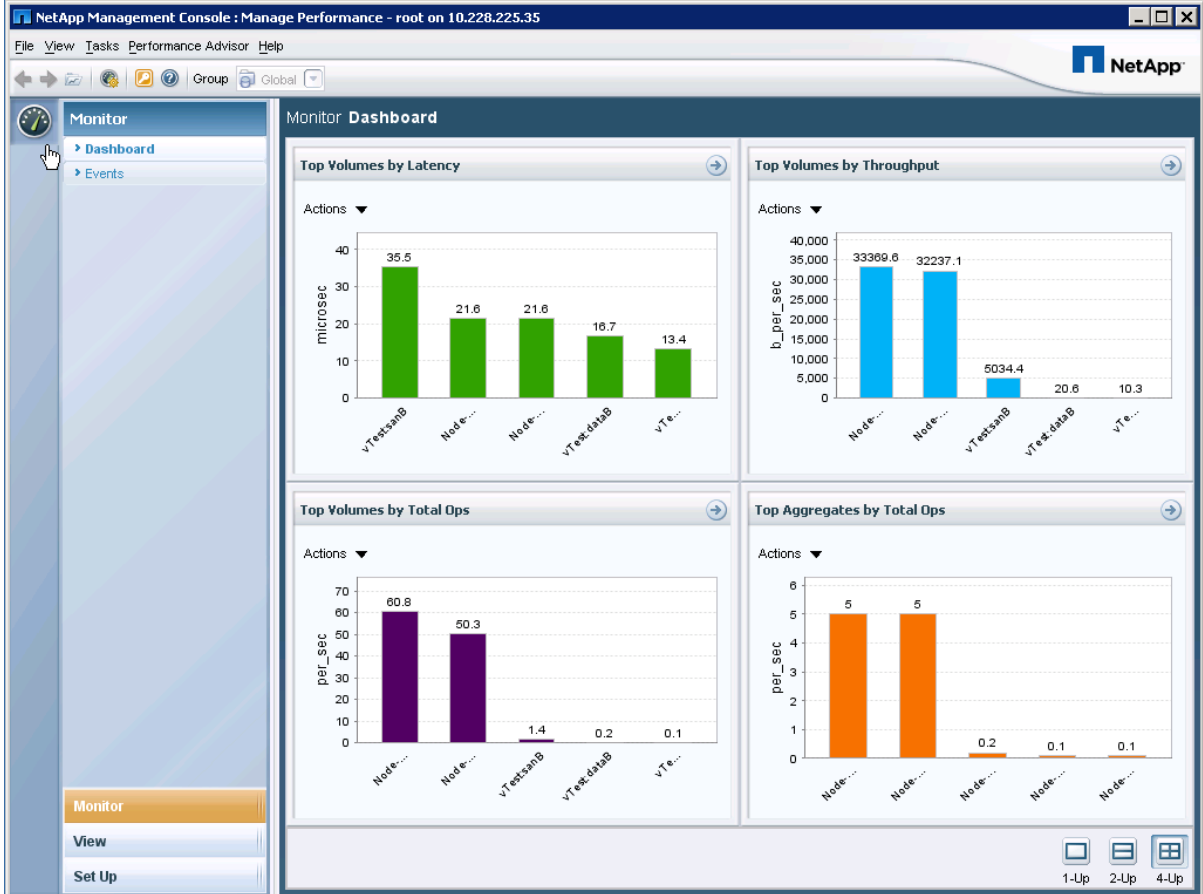
The following assumes that OnCommand Unified Manager is installed and data is being collected from the storage systems targeted for monitoring. OnCommand Unified Manager 5.2+ is required for clustered Data ONTAP 8.2. For more information about OnCommand Unified Manager and Performance Advisor,

review the product documentation and [TR-4090: Performance Advisor Features and Diagnosis: OnCommand Unified Manager 5.0/5.1.](#)

Basic Features

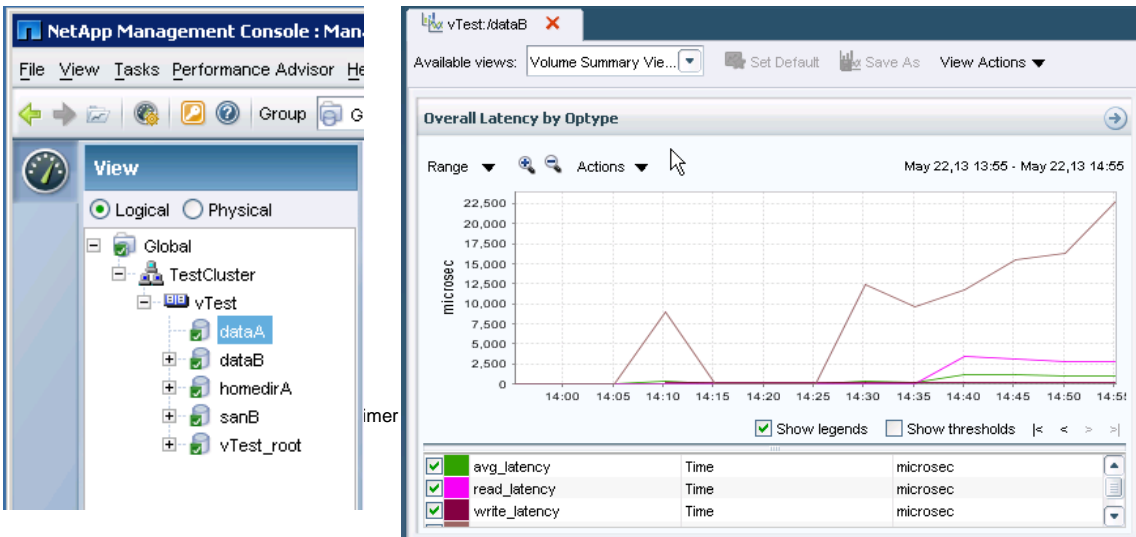
Without any configuration, Performance Advisor will collect data at regular fixed intervals and present this data in bar graphs. By default the busiest, or top, objects are presented in the dashboard.

Figure 13) Performance Advisor Dashboard



Besides the dashboard, individual objects, both logical and physical, are observable using the View tab. For instance, if the latency of a specific volume is interesting, use the view capability and drill down into the volume in the logical object tree.

Figure 14) Volume Latency Summary



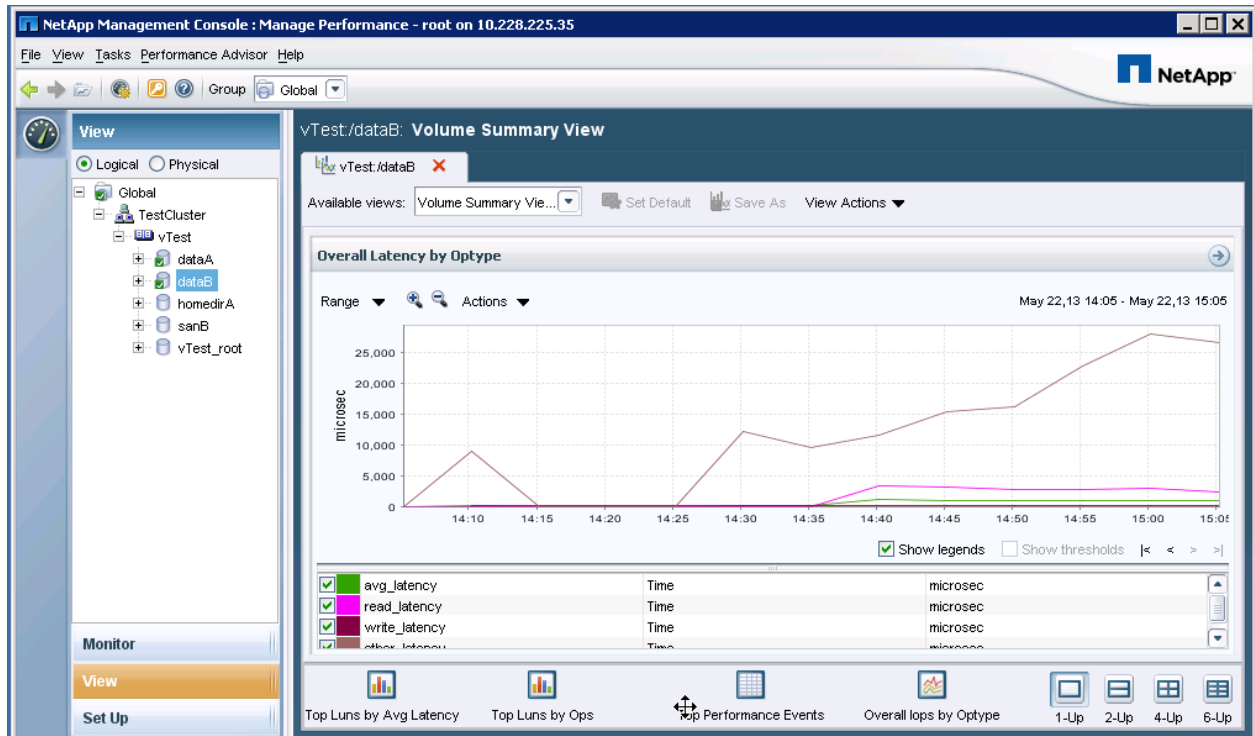
Establishing Thresholds and Alarms

Managing a large number of systems and systems with many objects is very difficult if done manually. Creating thresholds and setting alarms will automate alerting for key performance metrics and can help quickly resolve performance issues before they become worse.

Given that latency is the most important indicator, establishing latency thresholds with alerts per protocol, and possibly even on one or more volumes, is a good idea. The following process is the quickest way to establish a threshold on an important metric.

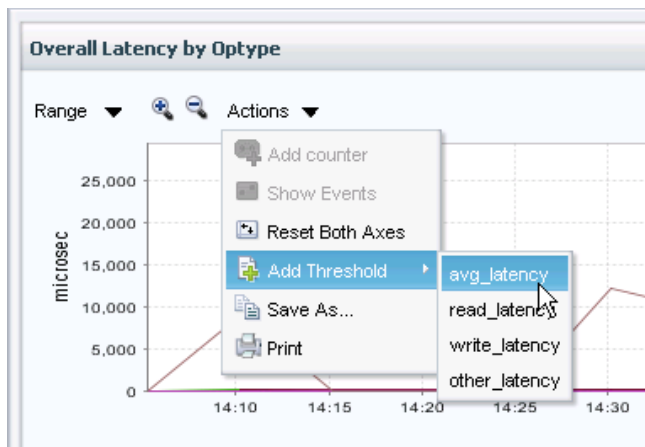
In this case, data is already being collected and latency is being viewed for a volume in the cluster.

Figure 15) Viewing Volume Latency Statistics



To set a threshold, simply go to the actions dropdown and select to add a threshold.

Figure 16) Adding a Threshold



This dialog allows you to specify a threshold if you know a value you would like to maintain or you can use Performance Advisor to suggest a value based on the history of the monitored object.

Figure 17) Performance Advisor Threshold Details

The dialog box is titled "Threshold details" and contains the following fields:

- Counter: volume:avg_latency
- Type: Upper Lower
- Value (microsec): 30000 (with a "Suggest value" button)
- Interval (seconds): 300 (with a spinner)
- Event name: volume:avg_latency

After the thresholds are established, you can add alarms to the desired objects to receive an alert if any thresholds are exceeded for a period of time. Adding alarms is done by going to the Set Up menu, reviewing the established thresholds, and opting to add an alarm.

Figure 18) Performance Advisor Alarms Configuration

The screenshot shows the "Set Up Thresholds" interface in the NetApp Management Console. The interface includes a left-hand navigation menu with "Set Up" selected, and a main content area with a "Summary" tab. The "Summary" tab displays a table of thresholds:

Object	Interval (seconds)	Enabled	Template name	Event name
vTest:/dataB	300	<input checked="" type="checkbox"/>		volume:avg_latency

Below the table, there are two sections: "Counter Details" and "Resource Properties".

Counter Details:

Counter	Type	Value	Unit
volume:avg_la...	Upper	30000	microsec

Resource Properties:

Name	Value
------	-------

There are other ways to add thresholds and alarms within Performance Advisor. To learn more about Performance Advisor and the other capabilities it offers, read [TR-4090: Performance Advisor Features and Diagnosis: OnCommand Unified Manager 5.0/5.1](#).

Storage QoS CLI

In clustered Data ONTAP 8.2, the QoS throughput policy is configured by setting either an IO limit (IO/s) or a data throughput limit (B/s). The table below provides a list of available QoS CLI labels.

Table 3) QoS Throughput Labels

QoS Storage unit	QoS CLI labels
IOPS	IOPS, iops, IO/s, io/s
Bytes/second	Mb/s, MB/s, mb/s, MB/S, MBPS, mbps, B/s, B/S, b/s, bps

Note: The data throughput limit can only be specified in bytes/second, including megabytes/second.

Observing Throughput and Latency

Workload-level statistics are available once storage objects are assigned to a QoS policy group. These statistics can be displayed using the QoS statistics CLI commands.

As discussed earlier, throughput and latency are important observable metrics. Similar to workload characteristics, the throughput and latency of the system, policy group, or workloads can be determined by using the following command:

```
Cluster1::> qos statistics workload performance show
Workload      ID      IOPS      Throughput  Latency
-----
-total-      -       5060      37.97MB/s   492.00us
volume_a     14368   4847      37.86MB/s   510.00us
...
```

More detailed latency information is also available by looking at the output from the following command:

```
Cluster1::> qos statistics workload latency show
Workload      ID      Latency    Network     Cluster     Data       Disk       QoS
-----
-total-      -       608.00us  270.00us    0ms        148.00us  190.00us   0ms
volume_a     14368   611.00us  270.00us    0ms        149.00us  192.00us   0ms
```

The output describes the latency encountered at the various components in the system discussed in previous sections. Using this output, it's possible to observe where most of the latency is coming from for a specific workload.

- Latency – Refers to the total latency observed
- Network – The amount of latency introduced by the network-level processing in Data ONTAP
- Cluster – The amount of latency introduced by the cluster interconnect
- Data – The amount of latency introduced by the system, except latency from the disk subsystem
- Disk – The amount of latency introduced by the disk subsystem. Note, any reads that were serviced by the WAFL[®] cache will not have a disk latency component because those operations did not go to disk.
- QoS – The amount of latency introduced by queuing by QoS if throughput limits have been established

Observing Resource Utilizations

QoS also enables users to view disk and CPU utilizations for a policy group or workload. This output can help indicate which workloads are utilizing resources the most, and can aid in identifying the workloads that could be considered bullies. For instance, if workload A is a very important workload and has high latencies from the previously introduced output, and in the resource utilizations output you notice workload B is using a lot of a systems resource, the contention between workload A and workload B for

the resource could be the source of the latency. Setting a limit or moving workload B could help alleviate workload A's latency issue. Resource utilizations are provided on a per-node basis.

```
Cluster1::> qos statistics workload resource cpu show -node Node-01
Workload          ID  CPU
-----
-total- (100%)    -   29%
volume_a         14368  12%
System-Default   1    10%
...
```

```
Cluster1::> qos statistics workload resource disk show -node Node-01
Workload          ID  Disk  No. of Disks
-----
-total-           -    4%      29
volume_a         14368  5%      22
System-Default   1     1%      23
...
```

Viewing Cluster-Level and Node-Level Periodic Statistics

Clustered Data ONTAP includes statistics beyond those presented in the QoS CLI statistics. One command to get an overall view into the cluster's state is `statistics show-periodic`. The output from this command provides details about the number of operations being serviced and additional cluster-wide resource utilizations. Looking at an individual node's state is also possible.

Note: Use CTRL-C to stop scrolling statistics and print a summary.

This command should be run in "advanced" privilege to get more information:

```
TestCluster::> set -privilege advanced
```

```
Warning: These advanced commands are potentially dangerous; use them only when directed to do so
by NetApp personnel.
Do you want to continue? {y|n}: y
```

The following example shows cluster-wide performance. The output is too wide to fit in the document, so it is divided between two output blocks.

```
TestCluster::*> statistics show-periodic
cluster:summary: cluster.cluster: 6/7/2013 18:27:39
cpu  cpu      total          fcache      total      total data      data      data cluster ...
avg busy      ops  nfs-ops  cifs-ops      ops      recv      sent busy      recv      sent      busy ...
-----
58%  91%    17687    17687      0          0      156MB    133MB  59%    68.7MB    47.3MB    4% ...
65%  92%    18905    18905      0          0      199MB    184MB  84%    103MB    74.9MB    6% ...
54%  86%    17705    17705      0          0      152MB    132MB  58%    68.9MB    47.2MB    4% ...
cluster:summary: cluster.cluster: 6/7/2013 18:27:47
cpu  cpu      total          fcache      total      total data      data      data cluster ...
avg busy      ops  nfs-ops  cifs-ops      ops      recv      sent busy      recv      sent      busy ...
-----
Minimums:
54%  86%    17687    17687      0          0      152MB    132MB  58%    68.7MB    47.2MB    4% ...
Averages for 3 samples:
59%  89%    18099    18099      0          0      169MB    150MB  67%    80.3MB    56.5MB    4% ...
Maximums:
65%  92%    18905    18905      0          0      199MB    184MB  84%    103MB    74.9MB    6% ...
```

```
... continued
```

```
... cluster cluster      disk      disk      pkts      pkts
...   recv   sent      read     write     recv     sent
... -----
```

...	87.6MB	86.3MB	96.4MB	139MB	87861	75081
...	96.2MB	109MB	108MB	261MB	127944	111190
...	84.0MB	85.4MB	69.6MB	101MB	87563	75402
...	cluster	cluster	disk	disk	pkts	pkts
...	recv	sent	read	write	recv	sent
...	-----	-----	-----	-----	-----	-----
...	84.0MB	85.4MB	69.6MB	101MB	87563	75081
...	89.3MB	93.8MB	91.6MB	167MB	101122	87224
...	96.2MB	109MB	108MB	261MB	127944	111190

The following command shows a single node. The output is similar to the previous example, but it is for a single node.

```
TestCluster::*> statistics show-periodic -object node -instance node -node Node-01
```

Note: When reviewing CPU information in Data ONTAP 8.2, CPU AVG is a better indicator of overall CPU utilization compared to CPU BUSY.

5.3 Managing Workloads with Data Placement

QoS is a very valuable tool to manage workloads within the cluster; however, the location and access path of data in the cluster can also play a role in performance, as was mentioned earlier. Clustered Data ONTAP has features that allow data to be moved, cached, and duplicated across nodes in the cluster to help manage performance.

DataMotion for Volumes

Independent of protocol, volumes can be moved and mirrored in the storage layer. Using volume move (vol move), volumes can be moved to the node handling the most client access to increase direct access. Using the same method, volumes can be moved to different disk types or nodes with different hardware to achieve different performance characteristics. Volume moves should be used to proactively manage performance and not when encountering performance problems, since volume move requires resources to perform the movement.

Intracluster FlexCache

For NAS protocols, such as NFS and SMB, volumes can be cached on other nodes in the cluster using intracluster FlexCache. With FlexCache, reads to a single volume can be distributed across the cluster while still allowing writes to be “written through” to the original volume location. Caches are not complete copies of data because they contain only read data that clients actually requested and will always contain the most recent data.

6 A Performance Management Scenario

This scenario describes how to use the tools introduced in the previous section to manage the performance of a clustered Data ONTAP system. In this example, let’s assume we have a single cluster with two nodes that has been appropriately sized for the expected workloads. The cluster is responsible for multiple workloads including home directories and a database with development and test volume clones. The database is used by an essential business application and must maintain a response time below 20ms.

This is a simple example, but the process can be used with systems that might be more complicated.

6.1 Establish Monitoring

For day-to-day monitoring, Performance Advisor has been set up to monitor the volume latencies of each node and a threshold of 20ms with alarms has been set.

Figure 19) Performance Advisor Latency Monitoring

Counter	Type	Value	Unit
volume:avg_latency	Upper	20000	microsec

We also want to be ready to react to any performance problems so we place all volumes into a Storage QoS policy group on the cluster by doing the following:


```
TestCluster::> qos policy-group create -policy-group testPG -vserver vTest
TestCluster::> volume modify -vserver vTest -volume * -qos-policy-group
                    default_vol_pg
```

Note: This command places ALL volumes in a SVM under a single policy group. In environments with multiple policy groups or limits are already in place this command may be used to target only volumes that are not already in a policy group.

6.2 Root Cause an Alert

One afternoon, an email alert is received from Performance Advisor saying that latency on a node has exceeded 20ms.

Figure 20) Performance Advisor Latency Alert

Event:	perf:volume:avg_latency:breached
Source:	vTest:/dataA →
Source Type:	Volume
Severity:	 Error
About:	Status of counters:volume:avg_latency
Triggered:	10 Jul 10:09
Acknowledged:	No
Deleted:	No
Condition:	The following counter value(s) have breached the specified threshold(s): Current value of counter volume:avg_latency is 22411.2 microsec, which is higher than the threshold value of 20000 microsec.

A quick review of the system tells us that nothing has failed and no configuration changes have been made recently, so we rule out hardware problems and configuration issues. To troubleshoot this problem, we can review additional details in Performance Advisor by looking for changes in workloads; in addition, the Storage QoS CLI statistics can help quickly identify contention if it exists. First, look at the workload latencies:

```
TestCluster::> qos statistics latency show
Policy Group      Latency    Network    Cluster    Data       Disk       QoS
-----
-total-          25.02ms   37.00us   199.00us   176.00us   24.61ms   0ms
default_vol_pg   25.02ms   37.00us   199.00us   176.00us   24.61ms   0ms
```

```
TestCluster::> qos statistics workload latency show
```

Workload	ID	Latency	Network	Cluster	Data	Disk	QoS
-total-	-	23.68ms	36.00us	202.00us	121.00us	23.33ms	0ms
dataA_testClo..	2750	27.99ms	61.00us	463.00us	121.00us	27.35ms	0ms
dataA-widl0808	10808	21.91ms	7.00us	0ms	120.00us	21.78ms	0ms
dataB-wid6762	6762	1.71ms	169.00us	135.00us	114.00us	1288.00us	0ms

From this, we can see that most of the latency for our important volume is coming from disk. If we look at the disk resources for the node we can see the resources are split between the important volume workload and its clone, where we really want the important volume to be able to use more to satisfy the performance requirements:

```
TestCluster::> qos statistics workload resource disk show -node Node-01
```

Workload	ID	Disk	No. of Disks
-total-	-	36%	27
dataA-widl0808	10808	50%	20
dataA_testClo..	2750	49%	20

We can also look at the workload performance output to see how much the clone is doing compared to the important volume:

```
TestCluster::> qos statistics workload performance show
```

Workload	ID	IOPS	Throughput	Latency
-total-	-	10390	78.44MB/s	23.87ms
dataA-widl0808	10808	5470	42.73MB/s	22.82ms
dataA_testClo..	2750	4346	33.95MB/s	28.20ms
dataB-wid6762	6762	413	1.61MB/s	1452.00us
_USERSPACE_APPS	14	89	140.15KB/s	131.00us
_Scan_Backgro..	8630	70	0KB/s	0ms
_Scan_Besteff..	8336	2	0KB/s	0ms

6.3 Resolve Performance Problem

Since we have identified that one of the test volume clones has increased load, possibly from a runaway test, and is impacting our other important volume on the same aggregate there are a few things we could do to solve this problem. The simplest way is to set a Storage QoS limit on this volume test clone to reduce how many operations it is allowed.

```
TestCluster::> qos statistics policy-group create -policy-group test_clone_limit -vserver vTest -max-throughput 500iops
```

```
TestCluster::> volume modify -vserver vTest -volume dataA_testClone1 -qos-policy-group test_clone_limit
```

After applying the limit, the latencies for the important volume are now improved, since contention is reduced.

```
TestCluster::> qos statistics workload latency show
```

Workload	ID	Latency	Network	Cluster	Data	Disk	QoS
-total-	-	26.05ms	30.00us	15.00us	256.00us	13.66ms	12.09ms
dataA_testClo..	2750	258.86ms	225.00us	172.00us	244.00us	26.16ms	232.06ms
dataA-widl0808	10808	13.81ms	6.00us	0ms	266.00us	13.54ms	0ms
dataB-wid6762	6762	1.66ms	294.00us	152.00us	80.00us	1129.00us	0ms

7 Conclusion

This paper introduces the fundamental performance concepts related to the workload characteristics, measurement, and architecture of NetApp storage systems. For more information about specific features or the capabilities of clustered Data ONTAP consider reviewing the collateral mentioned in the following References section or the documentation available on the NetApp Support site.

Appendices

Data ONTAP 8.2 Upgrade Recommendations

When considering upgrading an existing system to Data ONTAP 8.2, NetApp recommends first reviewing the performance of the system to enable acceptable performance post-upgrade.

Table 4) Data ONTAP 8.2 Upgrade Recommendations

Historical CPU Utilization	Upgrade Recommendation
Greater than 60% CPU utilization for more than 10% of the time	Do not upgrade without additional evaluation of the system
Less than 60% but greater than 40% CPU utilization for more than 10% of the time	Upgrade with caution
Less than 40%	OK to upgrade

The eligibility for upgrading following the guidance above can be determined by:

- Using NetApp Upgrade Advisor on the NetApp Support site <http://support.netapp.com/NOW/asuphome/>
- Using the 8.2 Upgrade Check tool on the NetApp Support site ToolChest for customers who do not provide AutoSupport™ data to NetApp. https://support.netapp.com/NOW/download/tools/82_upgrade_check/
- Execute and review the output of `perf report -t` as described below. Both the Upgrade Advisor and the 8.2 Upgrade Check tool automatically perform this interpretation.

Running and Interpreting perf report -t

The CLI command `perf report -t` is a node-level, advanced privilege command. It provides historical CPU and disk utilization samples collected at one second sample periods since the last reboot of the node. Samples collected for `perf report -t` are the values of the busiest CPU and disk utilizations during a sample period. The data is presented in a table in which the columns are disk utilization sample counts in 10% increments and the rows are CPU utilization sample counts in 10% increments.

```
Cluster::> node run -node Node-01 -command "priv set advanced; perf report -t"
```

```
Node: Node-01
```

```
Warning: These advanced commands are potentially dangerous; use
         them only when directed to do so by NetApp
         personnel.
```

```
Perf Report Version 1
```

```
Samples: 85429
```

```
Frequency: 1
```

cdnum1:	79001	309	17	10	3	2	7	10	9	110
cdnum2:	4062	411	130	21	1	2	1	4	5	30
cdnum3:	410	106	187	56	11	0	2	1	1	13
cdnum4:	210	23	37							
cdnum5:	29	5	19	39	13	2	1	0	0	4
cdnum6:	31	4	3	13	12	2	0	0	0	0
cdnum7:	4	1	0	2	0	1	1	0	0	0
cdnum8:	3	1	1	0	0	0	1	0	0	0
cdnum9:	0	0	0	0	0	0	0	0	0	0
cdnum10:	0	0	0	0	0	0	0	0	2	0

All disk utilization samples at 0–10% CPU utilized

Number of samples collected when system was 31–40% CPU utilized and 21–30% disk utilized

When considering upgrading to 8.2, the sum of these samples should be less than 10% of all the samples less the idle samples in the first row and column (79001 in this example)

```
disk KB: 63356 45773588 95382456
          90568 2678756
```

All CPU utilization samples at 0–10% disk utilized

References

The following references were used for this report:

- [TR-4015: SnapMirror Configuration and Best Practices Guide for Clustered Data ONTAP 8.2](#)
- [TR-4090: Performance Advisor Features and Diagnosis: OnCommand Unified Manager 5.0/5.1](#)
- [TR-4063: Parallel NFS File System Configuration and Best Practices for Data ONTAP Cluster-Mode](#)
- [TR-4067: Clustered Data ONTAP NFS Implementation Guide](#)
- [TR-3982: NetApp Clustered Data ONTAP 8.2: An Introduction](#)
- [TR-4080: Best Practices for Scalable SAN in Clustered Data ONTAP 8.2](#)
- [TR-3832: Flash Cache Best Practices Guide](#)
- [TR-4070: Flash Pool Design and Implementation Guide](#)

Version History

Version	Date	Document Version History
1.0	July 2013	Initial version

Refer to the [Interoperability Matrix Tool \(IMT\)](#) on the NetApp Support site to validate that the exact product and feature versions described in this document are supported for your specific environment. The NetApp IMT defines the product components and versions that can be used to construct configurations that are supported by NetApp. Specific results depend on each customer's installation in accordance with published specifications.

NetApp provides no representations or warranties regarding the accuracy, reliability, or serviceability of any information or recommendations provided in this publication, or with respect to any results that may be obtained by the use of the information or observance of any recommendations provided herein. The information in this document is distributed AS IS, and the use of this information or the implementation of any recommendations or techniques herein is a customer's responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. This document and the information contained herein may be used solely in connection with the NetApp products discussed in this document.

Go further, faster®



www.netapp.com

© 2013 NetApp, Inc. All rights reserved. All rights reserved. No portions of this document may be reproduced without prior written consent of NetApp, Inc. Specifications are subject to change without notice. NetApp, the NetApp logo, Go further, faster, ASUP, AutoSupport, DataMotion, Data ONTAP, Flash Cache, Flash Pool, FlexCache, FlexVol, OnCommand, and WAFL are trademarks or registered trademarks of NetApp, Inc. in the United States and/or other countries. TR-4211-0713