`

**TR Consolidated NetApp CDOT Build Standards**

**Standards for Configuration and Build of NetApp CDOT Systems**

**Synopsis:** This document details the standards to be used when commissining or upgrading a new CDOT filer in a TR Data Center..

**Segment:** Data Centre Engineering and Managed Services – Storage

**Authors:** David Ng

**Contributors:** Ian Daniel

**Document Version:** V 5

**Date:** 23rd February 2016

Document Status: published

**CONFIDENTIAL INFORMATION**

This document contains information proprietary to Thomson Reuters and may not be reproduced, disclosed or used in whole or part without express permission of Thomson Reuters.

© Thomson Reuters 2013

Contents

[1 Introduction 4](#_Toc444034829)

[Management Summary 4](#_Toc444034830)

[Document Scope 4](#_Toc444034831)

[References 4](#_Toc444034832)

[Change History 4](#_Toc444034833)

[Distribution List 4](#_Toc444034834)

[Glossary 4](#_Toc444034835)

[2 Version Guidelines 4](#_Toc444034836)

[3 Max Nodes Per Cluster 4](#_Toc444034837)

[4 Performance Guidelines 5](#_Toc444034838)

[5 NetApp Controller and Storage Configurations 6](#_Toc444034839)

[FAS8040 (6RU) – Shared Primary 6](#_Toc444034840)

[DS2246 (2RU) Disk Shelves 6](#_Toc444034841)

[DS4246 (4RU) Disk Shelves 8](#_Toc444034842)

[DS4486 (4RU) Disk Shelves 8](#_Toc444034843)

[6 Aggregate Configurations 9](#_Toc444034844)

[5 Volume, Qtree and LUN Configurations 9](#_Toc444034845)

[Volume Guidelines 9](#_Toc444034846)

[Thin Provisioning Rules 9](#_Toc444034847)

[Volume Sizes 10](#_Toc444034848)

[Qtree Guidelines 10](#_Toc444034849)

[LUN Guidelines 10](#_Toc444034850)

[6 Network Configuration 12](#_Toc444034851)

[LACP – Interface Groups 12](#_Toc444034852)

[MTU 12](#_Toc444034853)

[Flowcontrol 12](#_Toc444034854)

[VLANs 13](#_Toc444034855)

[Management Protocols 13](#_Toc444034856)

[Firewalls 13](#_Toc444034857)

[7 Monitoring 14](#_Toc444034858)

[System Alerting 14](#_Toc444034859)

[Performance Monitoring 14](#_Toc444034860)

[8 Backups 14](#_Toc444034861)

[9 Disaster Recovery 16](#_Toc444034862)

[10 Options Settings 16](#_Toc444034863)

[10.1 NFS v3 Options 16](#_Toc444034864)

[10.2 NFS v4 Options (Standard for MQ only) 16](#_Toc444034865)

[10.3 CIFS Options 16](#_Toc444034866)

[10.4 ISCSI Options 17](#_Toc444034867)

[11 QOS 18](#_Toc444034868)

**Table of figures**

[Figure 1 - FAS2240-2 Rear 6](#_Toc349895388)

[Figure 2 - FAS2240-4 Rear (Backup) 6](#_Toc349895389)

[Figure 3 - FAS3220 Rear 6](#_Toc349895390)

[Figure 4 - FAS3220 Rear (Backup) 7](#_Toc349895391)

[Figure 5 - FAS3250 Rear 7](#_Toc349895392)

[Figure 6 - FAS3250 Rear (Backup) 7](#_Toc349895393)

[Figure 7 - FAS6210 Rear (Tier 1) 8](#_Toc349895394)

[Figure 8 - FAS6210 Rear (Tier 2) 8](#_Toc349895395)

[Figure 9 - FAS6210 Rear (Backup) 9](#_Toc349895396)

[Figure 10 - DS2246 Front 9](#_Toc349895397)

[Figure 11 - DS2246 Rear 9](#_Toc349895398)

[Figure 12 - DS4243 Front 10](#_Toc349895399)

[Figure 13 - DS4243 Rear 10](#_Toc349895400)

[Figure 14 - DS4486 Front 10](#_Toc349895401)

[Figure 15 - DS4486 Rear 10](#_Toc349895402)

# Introduction

## Management Summary

This document details the configuration and options to be used when commissioning or upgrading a Netapp CDOT cluster within a TR Data Center. It will provide standards for aggregates, networks, system configurations, connectivity, management and option settings.

## Document Scope

An amount of NetApp knowledge is assumed and also knowledge of the TR Data Centers.

## References

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Document** | **Version** | **Date** | **Author** |
| 1 |  |  | N/A |  |

## Change History

|  |  |  |  |
| --- | --- | --- | --- |
| **Ver** | **Date** | **Author** | **Key Changes** |
| 0.1 | 10-Jun-2015 | David Ng | Initial version |
| 3 | 05-Oct-2015 | David Ng | Add note on volume level exports |
| 5 | 23-Feb-2016 | Ian Daniel | Added NFS v4 |

## Distribution List

|  |  |
| --- | --- |
| **Name** | **Role** |
|  | Storage Design and Engineering team |

## Glossary

|  |  |
| --- | --- |
| **Term** | **Definition** |
|  |  |

# Version Guidelines

|  |  |
| --- | --- |
| **TYPE** | **STANDARD** |
| Clustered Ontap | 8.2.3p5 |
| Firmware | Use latest |

# Max Nodes Per Cluster

|  |  |
| --- | --- |
| **TYPE** | **MAX NODE#** |
| High/Low Tier | 6 (2 can be added temporarily during tech refresh for a total of 8) |
| Backup | 1 |

# Performance Guidelines

The following table shows the peak performance capabilities of each TR NetApp storage solution.

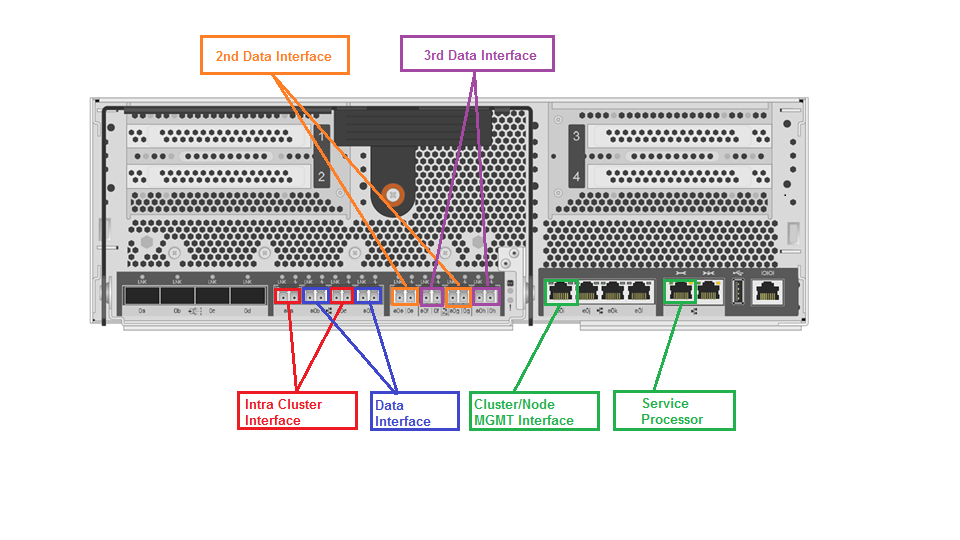
|  |  |  |
| --- | --- | --- |
| **Controller Configuration** | **Peak Throughput at 32K 50/50 Read/Write mix** | **Peak IOPS at 4K 50/50 Read/Write mix** |
| FAS8040 Shared HT Primary | 1376MB/sec per controller  43,000 IOPs per controller | 44,250 IOPs per controller  172MB/sec per controller |
| FAS8040 Backup | Backup Filer N/A | Backup Filer N/A |
| FAS8040 Shared LT Primary | 595MB/sec per controller  18,600 IOPs per controller | 15,500 IOPs per controller  62MB/sec per controller |

# NetApp Controller and Storage Configurations

We have a number of standard configurations that can be deployed into a TR data centre. These are as follows:

## FAS8040 (6RU) – Shared Primary

The standard FAS8040 network configuration is as follows.



NOTE: Additional 10Gbe cards can be inserted into slots 1, 3 and 4.

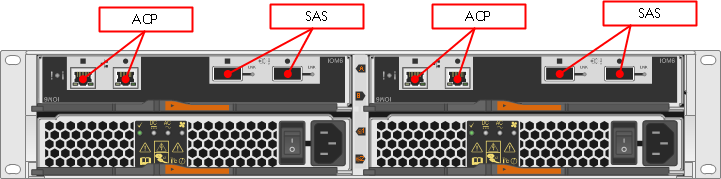
1. Figure 1 – FAS8040 Rear

## DS2246 (2RU) Disk Shelves

These shelves are the standard offering for FAS8040 HT storage in TR. Supported disk types are: 600GB SAS and 400GB SSD.



1. Figure 10 - DS2246 Front

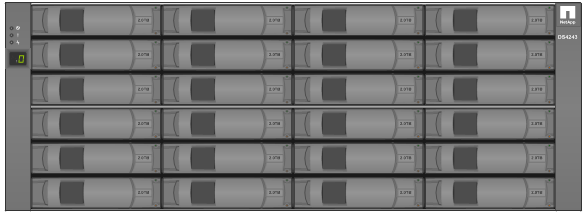


1. Figure 11 - DS2246 Rear

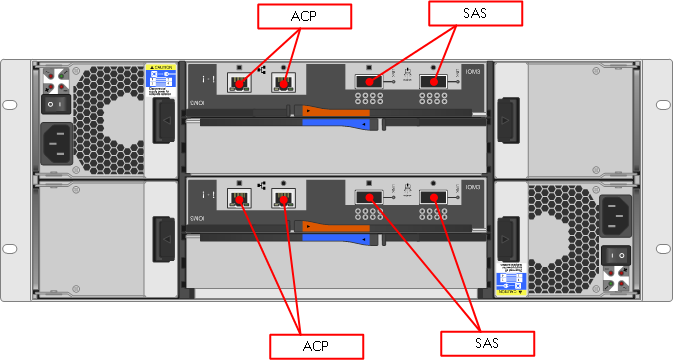
## 

## DS4246 (4RU) Disk Shelves

These shelves are the standard offering for FAS8040 LT storage attached to NetApp controllers in TR. Supported disk types are: 2TB SATA.



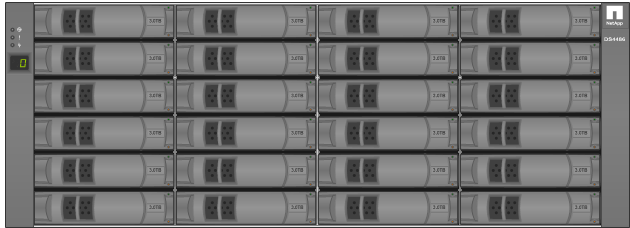
1. Figure 12 - DS4246 Front



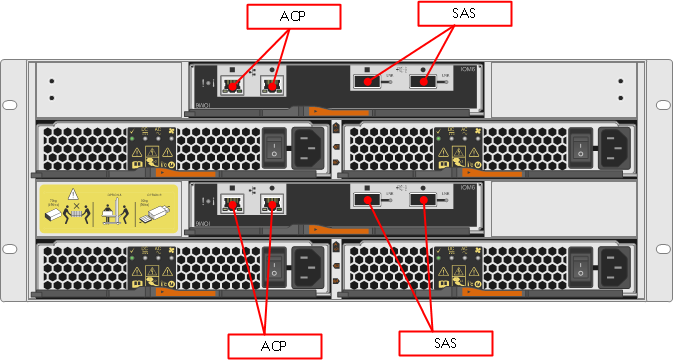
1. Figure 13 - DS4246 Rear

## DS4486 (4RU) Disk Shelves

These shelves are the standard offering for FAS8040 BKP storage attached to NetApp controllers in TR. Supported disk types are: 4TB SATA.



1. Figure 14 - DS4486 Front



1. Figure 15 - DS4486 Rear

# Aggregate Configurations

Aggregate configurations differ per device and disk type, the following table details the current configurations are shown below:

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **DOT** | **Tier** | **Model** | **Disk Size** | **SSD RG Size** | **HDD RG Size** | **Max Volume Count** | **Over Commit** | **Max RG per Aggr** | **Max Aggr per Ctlr** | **Aggr Size Max Usable (TB)** | **Spares (per Ctlr)** | **% Usable Threshold** |
| 8.2.x | 1 | FAS8040 | 600 | 8 | 19 | 500 (per cntlr) | 200% |  | 1 + root | 57 | 8 HDD / 1 SSD | 75 |
| 8.2.x | 2 | FAS8040 |  | NA | 16 | 500 (per cntlr) | 200% | 4 | 1 + root | 81.5 | 5 HDD | 75 |
| 8.2.x | BKP | FAS8040 |  | NA | 16 | 1000 (per cntlr) |  | 6 | 2 + root | 247.7 | 4 HDD |  |

The aggregate sizes and layouts above are based on the TR Architecture Groups specifications.

# Volume, Qtree and LUN Configurations

## Volume Guidelines

The following lists standard volume sizes and configurations that should be used when deploying storage for customers. Volumes should be exported at the ‘**volume**’ level. All volumes are to be thin provisioned.

### Thin Provisioning Rules

|  |  |
| --- | --- |
| **Volume Attribute** | **Values** |
| Volume Style | flex |
| User ID | 0 |
| Group ID | 0 |
| Security Style | unix |
| UNIX Permissions | ---rwxr-xr-x |
| Junction Path Source | RW\_volume |
| Junction Active | true |
| Autosize Mode | off |
| Autosize Enabled (for flexvols only) | false |
| Space Guarantee Style | none |
| Space Guarantee in Effect | true |
| Snapshot Directory Access Enabled | true |
| Space Reserved for Snapshots | 20% |
| Language | En\_US |
| Is File System Size Fixed | false |
| Extent Option | off |
| Reserved Space for Overwrites | 0B |
| Fractional Reserve | 0% |
| **Primary Space Management Strategy** | **snap\_delete (primary volumes)** |
| **Primary Space Management Strategy** | **vol\_grow (secondary volumes)** |
| Read Reallocation Option | off |
| Block Type | 64-bit |
| Flash Pool Caching Eligibility | read-write |
| **Snapshot Management** |  |
| Enabled | true |
| Commitment | try |
| Delete Order | oldest\_first |
| Target Free Space | 20% |
| Trigger | snap\_reserve |

|  |  |
| --- | --- |
| **Deduplication** |  |
| Deduplication Enabled On Datastore | Yes |
| Deduplication Enabled On Swap Datastore | No |
| Deduplication Enabled On Pagefile Datastore | No |
| Deduplication Enabled On Application Datastore | Yes |
| DB Storage | **Disabled for DB storage** |
| Deduplication Enabled On NFS/CIFS Volume | Enabled for file shares and datastores |

### Export Policy Rules

|  |  |  |  |
| --- | --- | --- | --- |
|  | **RORULE** | **RWRULE** | **SuperUser** |
| **Default Policy** | sys | never | none |
| **Writable Volumes** | sys | sys | sys |
| **ReadOnly Volumes** | sys |  |  |

### Volume Sizes

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Volume Type** | **Size** | **Snap Reserve (%)** | **Local Snap Retention** | **Snapvault Retention** | **Snap AutoDelete** | **Auto Size** | **SnapVault**  **Backup** |
| VI Data | 3686 GB | 20 | 7 | 7 | Trigger on reserve | No | Yes |
| CIFS/NFS Share | 10TB (max) | 20 | 7 | 7, 14, 30, 45 | Trigger on reserve | No | Yes |
| LION Share | 10TB (max) | 20 | 7 day (2 snaps are vaulted each day) | 7, 14, 30, 45 days | Trigger on reserve | No | Yes |
| WISP Share | 10TB (max) | 50 | 7 | 7, 14, 30, 45 | Trigger on reserve | No | Yes |

## Qtree Guidelines

Qtrees should be used when creating shares. Quotas should be disabled if dedup is enabled on a volume.

## LUN Guidelines

LUNs are used for ISCSI deployments, mainly in WISP. All LUNs are to be thin provisioned and are to follow these guidelines.

|  |  |
| --- | --- |
| **Details** | **Primary Configuration** |
| SQL Server Volume/LUN limits | 6TB (max) |
| Lun limitation ( number) | 20 LUNS per system & 40 LUNS per cluster |
| **Multiple databases per volume** | <= 315 DB's per server |
|  | 35 DB's per volume |
|  | <= 200GB per database |
| **Single database per volume** | >200GB per database |
|  | 6TB volume limit |
| **Multiple databases per volume** | >315 DB's per server (Max DB's 1200) |
|  | 150 DB's per volume |
|  | Average user database size <=10GB |
|  | Maximum user database size <=50GB |
| **WISP V3** | Multiple databases per volume >315 DB's per server (Max DB's 1950) |
|  | 150 DB's per volume |
|  | Average user database size <=10GB |
|  | Maximum user database size <=50GB |
| **WISP STANDARD CONFIG FOR STANDALONE DB SERVER** | SNAP VOL1  **E:\ 50GB** --> data drive  **G:\ 25GB** --> log drive |
|  | SNAP VOL2   **S:\ 10GB** --> snapinfo drive |
|  | NOSNAP VOL1    **H:\ 25GB** --> sysdb lun |
| **WISP CLUSTERED DB** | SNAP VOL1  **E:\ 50GB  G:\ 25GB** |
|  | SNAP VOL2  **S:\ 10GB** |
|  | NOSNAP VOL1   **H:\ 25GB**  **Q:\ 1GB** --> Quorum lun |

# Network Configuration

## LACP – Interface Groups

The standard network configuration in the TR environment is a LACP interface group where networks will support it. But, VSIP IP is no longer required.

The FAS8040 has sufficient ports for most scenarios. But, we may need to purchase additional network cards if a cluster require physical connectivity to more than 3 networks. Network port allocations are listed below.

|  |  |
| --- | --- |
| **FAS8040 Port Assignment** | |
| TYPE | PORT |
| IntraCluster (10Gbe interface) | e0a, e0c (do not use for data connections on backup filers) |
| SP (1Gbe interface) | SP |
| Node/Cluster MGMT (1Gbe interface) | e0i |
| 1st Data (10Gbe interface) | e0b, e0d |
| 2nd Data Interface (10Gbe interface) | e0e, e0g |
| 3rd Data Interface (10Gbe interface) | e0f, e0h |

## MTU

For all deployments the LACP 10GbE interfaces should be set with an MTU of 9000

InterCluster LIF are 1500 MTU.

### Flowcontrol

Flowcontrol should be disabled (None) for data and inter-cluster interfaces.

## 

## VLANs

VLANs are used to segregate traffic between vservers. The standard configuration uses three VLANs.

Link to VLANs in each datacenter:

https://theshare.thomsonreuters.com/sites/ie/storage/Lists/Storage%20VLANS%20and%20Addressing/Main\_Table.aspx

**Note:** In TR the default VLAN should be used to perform data migrations. In legacy markets environments the same applies, but VLAN IDs 199 and 200 are used by default in most locations.

## Management Protocols

All management protocols should be secure wherever possible. Use of SSH and HTTPS are mandatory.

## Firewalls

For standard non-Legacy deployments firewalls should not need consideration.

Firewalls are in use throughout the legacy environments and the following table shows ports that will need to be opened to ensure management connectivity to the NetApp controllers.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Number | Description | Source | Destination | Initiator (Source, Destination, Both) | Port/Protocol |
| 1 | Communication between Operations Manager Server and NetApp array | NetApp Array | Operations Manager Servers | Both | UDP/161  UDP/162  TCP/80  TCP/443  TCP/8443  TCP/8080  TCP/10000 |
| 2 | Communication between Operations Manager Server and NetApp array | Operations Manager Server | NetApp Array | Both | TCP/22  ICMP |
| 3 | NetApp Management Console communication to Operations Manager Database | Operations Manager Server | NetApp Management & NMC Workstation or Operations Manager Web Interface | Both | TCP/8443  TCP/8080  TCP/8088  TCP/8488  RDP/5616 |
| 4 | GMI Alerting | Operations Manager Server and WFA Server | GMI Netcool | Source | Standard GMI management ports  UDP/161  UDP/162 |
| 5 | SMTP for Management reporting | Operations Manager Server and WFA Server | Aurora SMTP gateway | Source | SMTP outbound  TCP/25 |
| 6 | Caching OPS Manager | WFA Server | OPS Manager Server | Destination | TCP/2638 |
| 7 | Command execution (ZAPI, PowerCLI) | WFA Server | NetApp Array | Source | TCP/22, 80, 443 |
| 8 | Desktop Access | NetApp Management & NMC Workstation or Operations Manager Web Interface | WFA Server | Destination | TCP/80, 443 |
| 9 | Caching VC | WFA Server | VC | Source | TCP/1433 |
| 10 | Snapmirror | Cluster | Cluster | both | TCP/11104 TCP/11105 |
| 11 | hw\_assist | HA pair MGMT | HA pair MGMT | both | TCP|UDP/4444 |

# Monitoring

## System Alerting

For all new deployments management is performed via OnCommand servers running on Linux. Alerts are sent to trap hosts which facilitate the automated logging of tickets. [Link](https://theshare.thomsonreuters.com/sites/DCO_Storage/Storage%20Services%20Documents/Forms/AllItems.aspx?RootFolder=%2Fsites%2FDCO%5FStorage%2FStorage%20Services%20Documents%2FAUTOMATION%20TOOLS%2FDFM) to Oncommand Architecture

Most of the DFM alerts are been converted into EMAT alert which will generate the tickets in various priority depends on the alert.

This is all depends on the EMAT configuration. All basic issues like disk, shelf, snapshot are initially dealt with by the DCO L1 team.

## Performance Monitoring

Performance monitoring limited in OCUM 5x is limited. So, the ‘wlstats’ script will be used to capture performance metrics when issues arise. Acceptible latencies are listed below:

Read\_latency <= 20ms

Write\_latency <= 5ms

# Backups

SnapVault is the preferred backup mechanism and should be used without exception. CIFS and NFS protocols will share backup vservers. Separate vservers will be created for the ISCSI protocol. The table below outline retention policies wich should be used in TR:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Volume Type** | **Backup Methodoligy** | **Local Snap Retention** | **Snapvault Retention** | **Snap AutoDelete** | **Auto Size** | **SnapVault**  **Backup** |
| VI Data | Snapshot schedule | 7 snapshots | 7 snapshots | Trigger on reserve | No | Yes |
| CIFS/NFS Share | Snapshot schedule | 7 snapshots | 7, 14, 30, 45 snapshots | Trigger on reserve | No | Yes |
| LION Share | Run\_hotbackup\_netapp.sh | 7 days (2 snaps are vaulted each day) | 7, 14, 30, 45 days | Trigger on reserve | No | Yes |
| WISP Share | Snap manager | 7 snapshots | 7, 14, 30, 45 snapshots | Trigger on reserve | No | Yes |

# Disaster Recovery

The Snapmirror technology is used to replicate data for DR purposes. It can be scheduled to run nightly. Environments which require an RPO < 24 will need to be reviewed.

# Options Settings

The default volume option provide sufficient optimization in CDOT, so tunning is not required. But, i will list 7mode option along with CDOT mapping for reference.

## 10.1 NFS v3 Options

|  |  |  |
| --- | --- | --- |
| **Option** | **7Mode Value** | **CDOT Value** |
| nfs.tcp.recvwindowsize | 262144 | Auto-tuned in cDOT |
| nfs.tcp.xfersize | 65536 | Auto-tuned in cDOT for NFSv3 |
| nfs.v3.enable | On | Vserver nfs modify –vserver <vserver>   -v3 enabled |
| nfs.v4.enable | Off | Vserver nfs modify –vserver <vserver >   -v4 enabled |
| nfs.tcp.enable | On | Vserver nfs modify –vserver <vserver >   -tcp enabled |

## 10.2 NFS v4 Options (Standard for MQ only)

|  |  |  |
| --- | --- | --- |
| **Option** | **7Mode Value** | **CDOT Value** |
| nfs.v4.id.domain | <domain-name> | vserver nfs modify –vserver <vserver >  -v4-id-domain <NFS\_V4\_DOMAIN> |
| nfs.v4.enable | on | vserver nfs modify –vserver <vserver >   -v4 enabled |
| nfs.v4.id.allow\_numerics | on | nfs server modify -vserver nfsv4-testvs01 -v4-numeric-ids  enabled |
| nfs.tcp.enable | on | vserver nfs modify –vserver <vserver >   -tcp enabled |

## 10.3 CIFS Options

|  |  |  |
| --- | --- | --- |
| **Option** | **7Mode Value** | **CDOT Value** |
| cifs.max\_mpx | 255 | vserver cifs options modify -max-mpx (default is 255) |
| cifs.neg\_buf\_size | 65340 | Auto-tuned in cDOT |
| cifs.tcp\_window\_size | 262144 | Auto-tuned in cDOT |
| cifs.smb2.enable | On | vserver cifs options modify -smb2-enabled |
| cifs.smb2.signing.required | Off | vserver cifs security modify -is-signing-required |
| cifs.oplocks.enable | On | vserver cifs share modify -share-properties |
| cifs.smb2.durable\_handle.enable | Off | NA in cDOT |

## 10.4 ISCSI Options

|  |  |  |
| --- | --- | --- |
| **Option** | **7Mode Value** | **CDOT Value** |
| iscsi.enable | On | vserver iscsi start |
| iscsi.ip\_based\_tpgroup | Off | N/A |
| iscsi.max\_ios\_per\_session | 128 | vserver iscsi modify –max\_ios\_per\_session (default is 128) |
| iscsi.tcp\_window\_size | 131400 | vserver iscsi modify –tcp-window-size  (default is 131400) |

# QOS

Each data volume in a vserver should have QOS policy set to 6000 IOPS