



Performance overview and validation in AWS - FSxN NetApp ONTAP

NetApp Solutions

NetApp
October 20, 2023

This PDF was generated from <https://docs.netapp.com/us-en/netapp-solutions/data-analytics/kafka-nfs-performance-overview-and-validation-in-aws-FSxN.html> on October 20, 2023. Always check docs.netapp.com for the latest.

Table of Contents

Performance overview and validation in AWS FSx for NetApp ONTAP	1
Apache Kafka in AWS FSx for NetApp ONTAP	1

Performance overview and validation in AWS FSx for NetApp ONTAP

[Previous: Performance overview and validation in AWS - Cloud Volume ONTAP.](#)

A Kafka cluster with the storage layer mounted on NetApp NFS was benchmarked for performance in the AWS FSx for NetApp ONTAP. The benchmarking examples are described in the following sections.

Apache Kafka in AWS FSx for NetApp ONTAP

Network File System (NFS) is a widely used network filesystem for storing large amounts of data. In most organizations data is increasingly being generated by streaming applications like Apache Kafka. These workloads require scalability, low latency, and a robust data ingestion architecture with modern storage capabilities. To enable real-time analytics and to provide actionable insights, a well designed and highly performant infrastructure is required.

Kafka by design works with POSIX compliant file system and relies on the file system to handle file operations, but when storing data on an NFSv3 file system, the Kafka broker NFS client can interpret file operations differently from a local file system like XFS or Ext4. A common example is the NFS Silly rename which caused Kafka brokers to fail when expanding clusters and re-allocating partitions. To deal with this challenge NetApp has updated the open-source Linux NFS client with changes now generally available in RHEL8.7, RHEL9.1, and supported from the current FSx for NetApp ONTAP release, ONTAP 9.12.1.

Amazon FSx for NetApp ONTAP provides a fully managed, scalable, and highly performance NFS file system in the cloud. Kafka data on FSx for NetApp ONTAP can scale to handle large amounts of data and ensure fault tolerance. NFS provides centralized storage management and data protection for critical and sensitive datasets.

These enhancements make it possible for AWS customer to take advantage of FSx for NetApp ONTAP when running Kafka workloads on AWS compute services. These benefits are:

- * Reducing CPU utilization to reduce the I/O wait time
- * Faster Kafka broker recovery time.
- * Reliability and efficiency.
- * Scalability and performance.
- * Multi-Availability Zone availability.
- * Data protection.

Performance overview and validation in AWS FSx for NetApp ONTAP

A Kafka cluster with the storage layer mounted on NetApp NFS was benchmarked for performance in the AWS cloud. The benchmarking examples are described in the following sections.

Kafka in AWS FSx for NetApp ONTAP

A Kafka cluster with AWS FSx for NetApp ONTAP was benchmarked for performance in the AWS cloud. This benchmarking is described in the following sections.

Architectural setup

The following table shows the environmental configuration for a Kafka cluster using AWS FSx for NetApp

ONTAP.

Platform component	Environment configuration
Kafka 3.2.3	<ul style="list-style-type: none">• 3 x zookeepers – t2.small• 3 x broker servers – i3en.2xlarge• 1 x Grafana – c5n.2xlarge• 4 x producer/consumer — c5n.2xlarge *
Operating system on all nodes	RHEL8.6
AWS FSx for NetApp ONTAP	Multi-AZ with 4GB/Sec throughput and 160000 IOPS

NetApp FSx for NetApp ONTAP setup

1. For our initial testing, we have created a FSx for NetApp ONTAP filesystem with 2TB of capacity and 40000 IOPs for 2GB/Sec throughput.

```
[root@ip-172-31-33-69 ~]# aws fsx create-file-system --region us-east-2
--storage-capacity 2048 --subnet-ids <desired subnet 1> subnet-<desired
subnet 2> --file-system-type ONTAP --ontap-configuration
DeploymentType=MULTI_AZ_HA_1,ThroughputCapacity=2048,PreferredSubnetId=<
desired primary subnet>,FsxAdminPassword=<new
password>,DiskIopsConfiguration="{Mode=USER_PROVISIONED,Iops=40000}" }
```

In our example, we are deploying FSx for NetApp ONTAP through the AWS CLI. You will need to customize the command further in your environment as needed. FSx for NetApp ONTAP can additionally be deployed and managed through the AWS Console for an easier and more streamlined deployment experience with less command line input.

Documentation In FSx for NetApp ONTAP, the max IOPS achievable for a 2GB/Sec throughput filesystem in our test region (US-East-1) is 80,000 iops. The total max iops for a FSx for NetApp ONTAP filesystem is 160,000 iops which requires a 4GB/Sec throughput deployment to achieve which we will demonstrate later in this document.

For more information on FSx for NetApp ONTAP performance specifications, please feel free to visit the AWS FSx for NetApp ONTAP documentation here: <https://docs.aws.amazon.com/fsx/latest/ONTAPGuide/performance.html> .

Detailed command line syntax for FSx “create-file-system” can be found here: <https://docs.aws.amazon.com/cli/latest/reference/fsx/create-file-system.html>

For instance, you can specify a specific KMS key as opposed to the default AWS FSx master key that is used when no KMS key is specified.

2. While creating the FSx for NetApp ONTAP filesystem, Wait till the “LifeCycle” status changes to “AVAILABLE” in your JSON return after describing your filesystem as follows:

```
[root@ip-172-31-33-69 ~]# aws fsx describe-file-systems --region us-east-1 --file-system-ids fs-02ff04bab5ce01c7c
```

3. Validate the credentials by login into Fsx for NetApp ONTAP SSH with the fsxadmin user:
Fsxadmin is the default admin account for FSx for NetApp ONTAP filesystems at creation. The password for fsxadmin is the password that was configured when first creating the filesystem either in the AWS Console or with the AWS CLI as we completed in Step 1.

```
[root@ip-172-31-33-69 ~]# ssh fsxadmin@198.19.250.244
The authenticity of host '198.19.250.244 (198.19.250.244)' can't be
established.
ED25519 key fingerprint is
SHA256:mgCyRXJfWRc2d/jOjFbMBsUcYOWjxoIky0ltHvVDL/Y.
This key is not known by any other names
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added '198.19.250.244' (ED25519) to the list of
known hosts.
(fsxadmin@198.19.250.244) Password:

This is your first recorded login.
```

4. Once your credentials have been validated, Create the storage Virtual Machine on the FSx for NetApp ONTAP filesystem

```
[root@ip-172-31-33-69 ~]# aws fsx --region us-east-1 create-storage-virtual-machine --name svmkafkatest --file-system-id fs-02ff04bab5ce01c7c
```

A Storage Virtual Machine (SVM) is an isolated file server with its own administrative credentials and endpoints for administering and accessing data in FSx for NetApp ONTAP volumes and provides FSx for NetApp ONTAP multi-tenancy.

5. Once you have configured your primary Storage Virtual Machine, SSH into the newly created FSx for NetApp ONTAP filesystem and create volumes in storage virtual machine using below sample command and similarly we create 6 volumes for this validation. Based on our validation, keep the default constituent (8) or less constituents which will provides better performance to kafka.

```
FsxId02ff04bab5ce01c7c:~> volume create -volume kafkafsxN1 -state
online -policy default -unix-permissions ---rwxr-xr-x -junction-active
true -type RW -snapshot-policy none -junction-path /kafkafsxN1 -aggr
-list aggr1
```

6. We will need additional capacity in our volumes for our testing. Extend the size of the volume to 2TB and mount on the junction path.

```

FsxId02ff04bab5ce01c7c:.*> volume size -volume kafkafsxN1 -new-size +2TB
vol size: Volume "svmkafkatest:kafkafsxN1" size set to 2.10t.

FsxId02ff04bab5ce01c7c:.*> volume size -volume kafkafsxN2 -new-size +2TB
vol size: Volume "svmkafkatest:kafkafsxN2" size set to 2.10t.

FsxId02ff04bab5ce01c7c:.*> volume size -volume kafkafsxN3 -new-size +2TB
vol size: Volume "svmkafkatest:kafkafsxN3" size set to 2.10t.

FsxId02ff04bab5ce01c7c:.*> volume size -volume kafkafsxN4 -new-size +2TB
vol size: Volume "svmkafkatest:kafkafsxN4" size set to 2.10t.

FsxId02ff04bab5ce01c7c:.*> volume size -volume kafkafsxN5 -new-size +2TB
vol size: Volume "svmkafkatest:kafkafsxN5" size set to 2.10t.

FsxId02ff04bab5ce01c7c:.*> volume size -volume kafkafsxN6 -new-size +2TB
vol size: Volume "svmkafkatest:kafkafsxN6" size set to 2.10t.

FsxId02ff04bab5ce01c7c:.*> volume show -vserver svmkafkatest -volume *
Vserver   Volume           Aggregate   State      Type        Size
Available Used%
-----
svmkafkatest
      kafkafsxN1   -           online     RW         2.10TB
1.99TB    0%
svmkafkatest
      kafkafsxN2   -           online     RW         2.10TB
1.99TB    0%
svmkafkatest
      kafkafsxN3   -           online     RW         2.10TB
1.99TB    0%
svmkafkatest
      kafkafsxN4   -           online     RW         2.10TB
1.99TB    0%
svmkafkatest
      kafkafsxN5   -           online     RW         2.10TB
1.99TB    0%
svmkafkatest
      kafkafsxN6   -           online     RW         2.10TB
1.99TB    0%
svmkafkatest
      svmkafkatest_root
                        aggr1      online     RW         1GB
968.1MB    0%
7 entries were displayed.

```

```

FsxId02ff04bab5ce01c7c::~*> volume mount -volume kafkafsxN1 -junction
-path /kafkafsxN1

FsxId02ff04bab5ce01c7c::~*> volume mount -volume kafkafsxN2 -junction
-path /kafkafsxN2

FsxId02ff04bab5ce01c7c::~*> volume mount -volume kafkafsxN3 -junction
-path /kafkafsxN3

FsxId02ff04bab5ce01c7c::~*> volume mount -volume kafkafsxN4 -junction
-path /kafkafsxN4

FsxId02ff04bab5ce01c7c::~*> volume mount -volume kafkafsxN5 -junction
-path /kafkafsxN5

FsxId02ff04bab5ce01c7c::~*> volume mount -volume kafkafsxN6 -junction
-path /kafkafsxN6

```

In FSx for NetApp ONTAP, volumes can be thin provisioned. In our example, the total extended volume capacity exceeds total filesystem capacity so we will need to extend the total filesystem capacity in order to unlock additional provisioned volume capacity which we will demonstrate in our next step.

7. Next, for additional performance and capacity, We extend the FSx for NetApp ONTAP throughput capacity from 2GB/Sec to 4GB/Sec and IOPS to 160000, and capacity to 5 TB

```

[root@ip-172-31-33-69 ~]# aws fsx update-file-system --region us-east-1
--storage-capacity 5120 --ontap-configuration
'ThroughputCapacity=4096,DiskIopsConfiguration={Mode=USER_PROVISIONED,Iops=160000}' --file-system-id fs-02ff04bab5ce01c7c

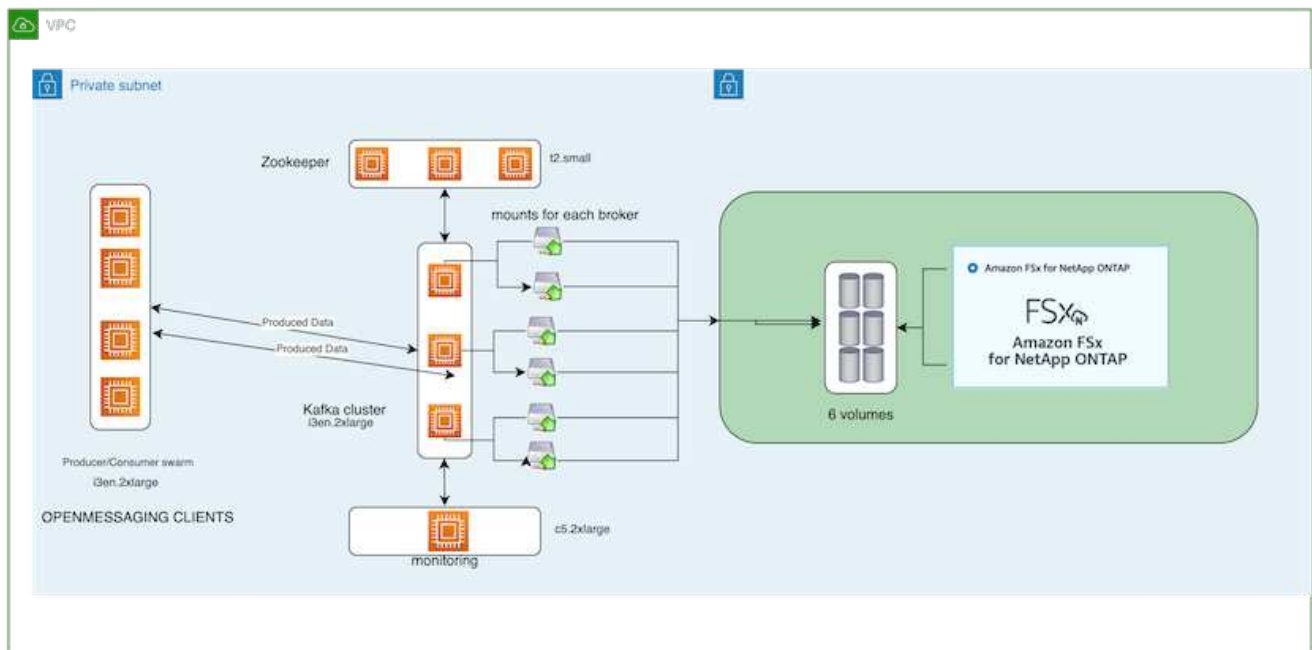
```

Detailed command line syntax for FSx “update-file-system” can be found here:

<https://docs.aws.amazon.com/cli/latest/reference/fsx/update-file-system.html>

8. The FSx for NetApp ONTAP volumes are mounted with nconnect and default options in Kafka brokers

The following picture shows our final architecture of a our FSx for NetApp ONTAP based Kafka cluster:



- Compute. We used a three-node Kafka cluster with a three-node zookeeper ensemble running on dedicated servers. Each broker had six NFS mount points to a six volumes on the FSx for NetApp ONTAP instance.
- Monitoring. We used two nodes for a Prometheus-Grafana combination. For generating workloads, we used a separate three-node cluster that could produce and consume to this Kafka cluster.
- Storage. We used an FSx for NetApp ONTAP with six 2TB volumes mounted. The volume was then exported to the Kafka broker with an NFS mount. The FSx for NetApp ONTAP volumes are mounted with 16 nconnect sessions and default options in Kafka brokers.

OpenMessage Benchmarking configurations.

We used the same configuration used for the NetApp Cloud volumes ONTAP and their details are here - <https://docs.netapp.com/us-en/netapp-solutions/data-analytics/kafka-nfs-performance-overview-and-validation-in-aws.html#architectural-setup>

Methodology of testing

1. A Kafka cluster was provisioned as per the specification described above using terraform and ansible. Terraform is used to build the infrastructure using AWS instances for the Kafka cluster and ansible builds the Kafka cluster on them.
2. An OMB workload was triggered with the workload configuration described above and the Sync driver.

```
sudo bin/benchmark -drivers driver-kafka/kafka-sync.yaml workloads/1-
topic-100-partitions-1kb.yaml
```

3. Another workload was triggered with the Throughput driver with same workload configuration.


```
sudo bin/benchmark -drivers driver-kafka/kafka-throughput.yaml
workloads/1-topic-100-partitions-1kb.yaml
```

Observation

Two different types of drivers were used to generate workloads to benchmark the performance of a Kafka instance running on NFS. The difference between the drivers is the log flush property.

For a Kafka Replication factor 1 and the FSx for NetApp ONTAP:

- Total throughput generated consistently by the Sync driver: ~ 3218 MBps and peak performance in ~ 3652 MBps.
- Total throughput generated consistently by the Throughput driver: ~ 3679 MBps and peak performance in ~ 3908 MBps.

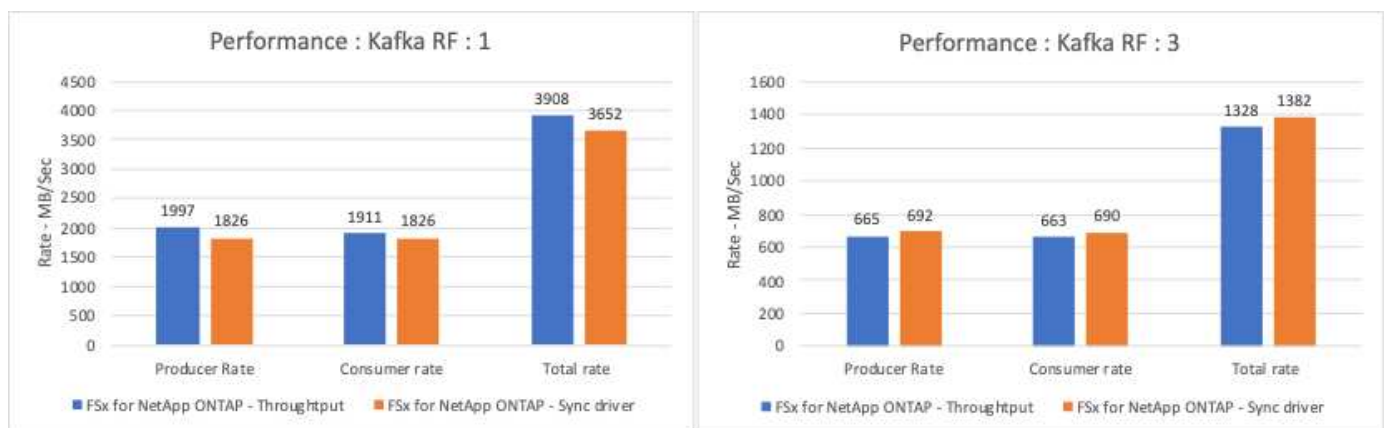
For Kafka with replication factor 3 and the FSx for NetApp ONTAP :

- Total throughput generated consistently by the Sync driver: ~ 1252 MBps and peak performance in ~ 1382 MBps.
- Total throughput generated consistently by the Throughput driver: ~ 1218 MBps and peak performance in ~ 1328 MBps.

In Kafka replication factor 3, the read and write operation happened three times on the FSx for NetApp ONTAP, In Kafka replication factor 1, the read and write operation is one time on the FSx for NetApp ONTAP, so in both validation, we able to reach the maximum throughput of 4GB/Sec.

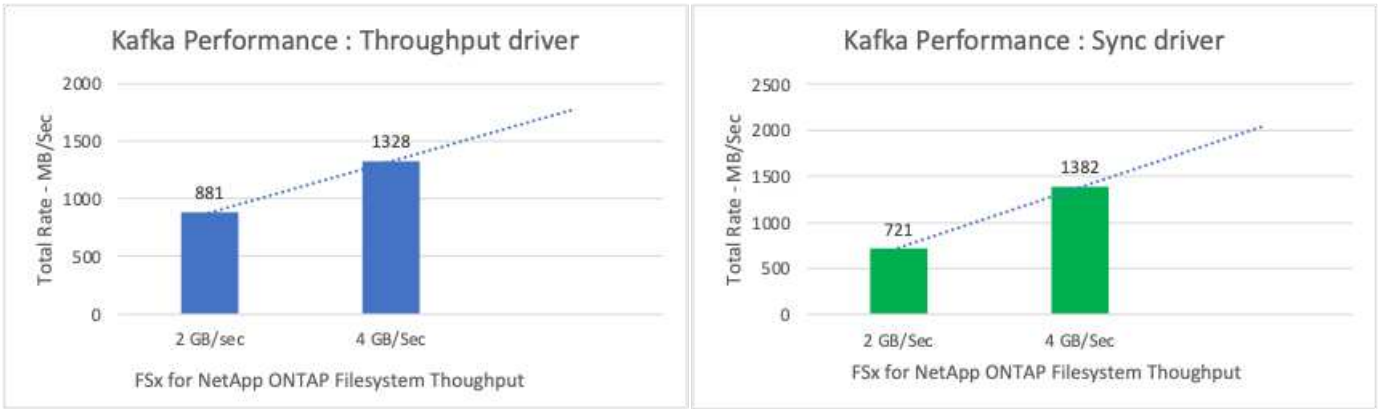
The Sync driver can generate consistent throughput as logs are flushed to the disk instantly, whereas the Throughput driver generates bursts of throughput as logs are committed to disk in bulk.

These throughput numbers are generated for the given AWS configuration. For higher performance requirements, the instance types can be scaled up and tuned further for better throughput numbers. The total throughput or total rate is the combination of both producer and consumer rate.

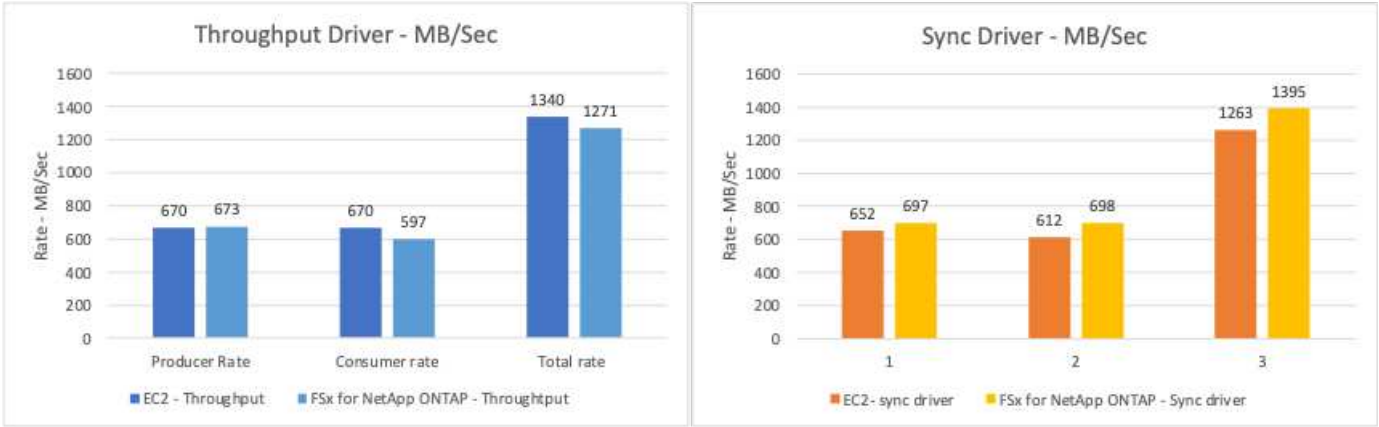


The below chart shows the 2GB/Sec FSx for NetApp ONTAP and 4GB/Sec performance for Kafka replication factor 3. The replication factor 3 does the read and write operation three times on the FSx for NetApp ONTAP storage. The total rate for throughput driver is 881 MB/Sec, which does read and write Kafka operation approximately 2.64 GB/Sec on the 2GB/Sec FSx for NetApp ONTAP filesystem and total rate for throughput driver is 1328 MB/Sec that does read and write kafka operation approximately 3.98 GB/Sec. Ther Kafka

performance is linear and scalable based on the FSx for NetApp ONTAP throughput.



The below chart shows the performance between EC2 instance vs FSx for NetApp ONTAP (Kafka Replication Factor : 3)



Next: Performance overview and validation with AFF on-premises.

Copyright information

Copyright © 2023 NetApp, Inc. All Rights Reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP “AS IS” AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

LIMITED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (b)(3) of the Rights in Technical Data -Noncommercial Items at DFARS 252.227-7013 (FEB 2014) and FAR 52.227-19 (DEC 2007).

Data contained herein pertains to a commercial product and/or commercial service (as defined in FAR 2.101) and is proprietary to NetApp, Inc. All NetApp technical data and computer software provided under this Agreement is commercial in nature and developed solely at private expense. The U.S. Government has a non-exclusive, non-transferrable, nonsublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b) (FEB 2014).

Trademark information

NETAPP, the NETAPP logo, and the marks listed at <http://www.netapp.com/TM> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.