



# **Performance overview and validation in AWS - Cloud Volume 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.html> on October 20, 2023. Always check docs.netapp.com for the latest.

# Table of Contents

- Performance overview and validation in AWS ..... 1
  - Kafka in AWS cloud with NetApp Cloud Volumes ONTAP (high-availability pair and single node) ..... 1
- Methodology of testing ..... 12
- Observation ..... 12

# Performance overview and validation in AWS

[Previous: Why NetApp NFS for Kafka workloads?](#)

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 cloud with NetApp Cloud Volumes ONTAP (high-availability pair and single node)

A Kafka cluster with NetApp Cloud Volumes ONTAP (HA pair) 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 NAS.

Platform component	Environment configuration
Kafka 3.2.3	<ul style="list-style-type: none"><li>• 3 x zookeepers – t2.small</li><li>• 3 x broker servers – i3en.2xlarge</li><li>• 1 x Grafana – c5n.2xlarge</li><li>• 4 x producer/consumer — c5n.2xlarge *</li></ul>
Operating system on all nodes	RHEL8.6
NetApp Cloud Volumes ONTAP instance	HA pair instance – m5dn.12xLarge x 2node Single Node Instance - m5dn.12xLarge x 1 node

### NetApp cluster volume ONTAP setup

1. For the Cloud Volumes ONTAP HA pair, we created two aggregates with three volumes on each aggregate on each storage controller. For the single Cloud Volumes ONTAP node, we create six volumes in an aggregate.

aggr3

EBS Allocated Capacity:5.05 TB

EBS Used Capacity:298.21 GB

Volumes:3

kafka\_aggr3\_vol1 (1 TB)

kafka\_aggr3\_vol2 (1 TB)

kafka\_aggr3\_vol3 (1 TB)

AWS Disks:8

State:online

Underlying AWS Tier:Provisioned IOPS SSD (io1)

AWS Disk Size:2 TB

Underlying AWS Capacity:12 TB

Encryption Type:

Home Node:kafka\_nfs\_cvo\_ha1-01

Provisioned IOPS:80000

Close

aggr22

EBS Allocated Capacity:6.73 TB

EBS Used Capacity:280.95 GB

Volumes:3

kafka\_aggr22\_vol1 (1 TB)

kafka\_aggr22\_vol2 (1 TB)

kafka\_aggr22\_vol3 (1 TB)

AWS Disks:8

State:online

Underlying AWS Tier:Provisioned IOPS SSD (io1)

AWS Disk Size:2 TB

Underlying AWS Capacity:16 TB

Encryption Type:

Home Node:kafka\_nfs\_cvo\_ha1-02

Provisioned IOPS:20000

Close

aggr2

EBS Allocated Capacity:5.32 TB

EBS Used Capacity:209.90 GB

Volumes:6

kafka\_aggr2\_vol2 (1 TB)

kafka\_aggr2\_vol3 (1 TB)

kafka\_aggr2\_vol4 (1 TB)

AWS Disks:4

State:online

Underlying AWS Tier:Provisioned IOPS SSD (io1)

AWS Disk Size:2 TB

Underlying AWS Capacity:6 TB

Encryption Type:

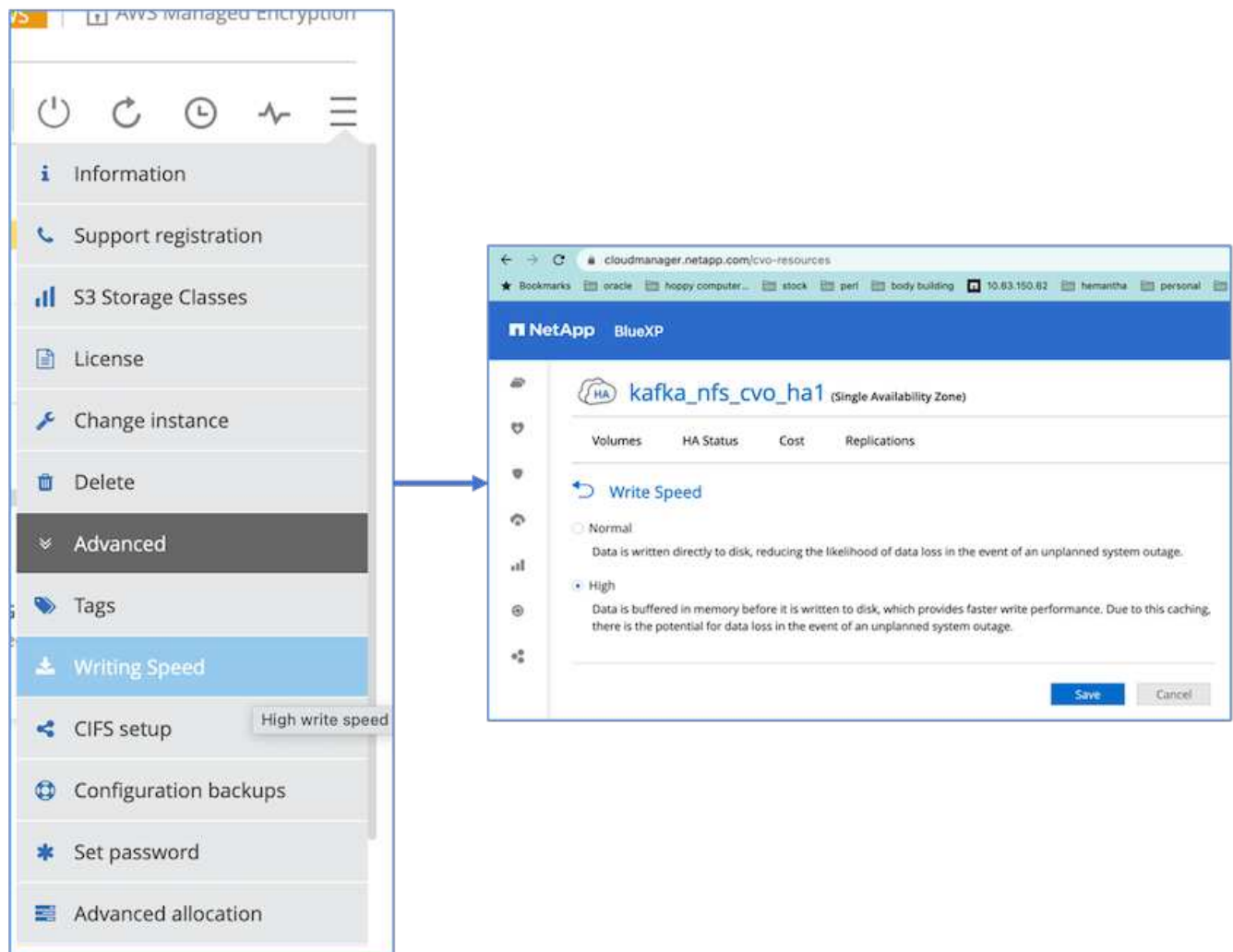
Home Node:kafka\_nfs\_cvo\_sn-01

Provisioned IOPS:80000

Close

2. To achieve better network performance, we enabled high speed networking for both the HA pair and the single node.

2

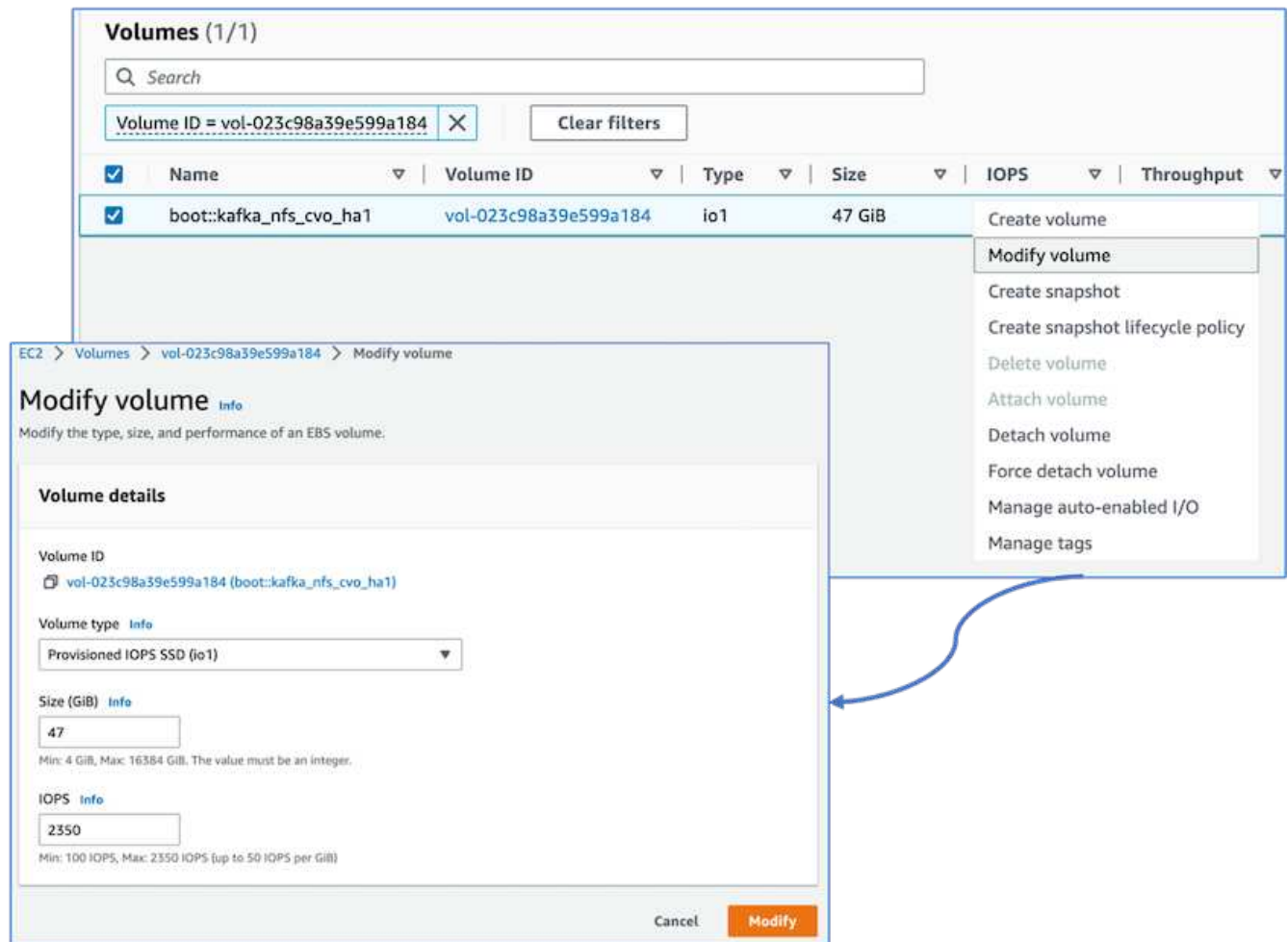


3. We noticed that the ONTAP NVRAM had more IOPS so we changed the IOPS to 2350 for the Cloud Volumes ONTAP root volume. The root volume disk in Cloud Volumes ONTAP was 47GB in size. The following ONTAP command is for the HA pair, and the same step is applicable for the single node.

```

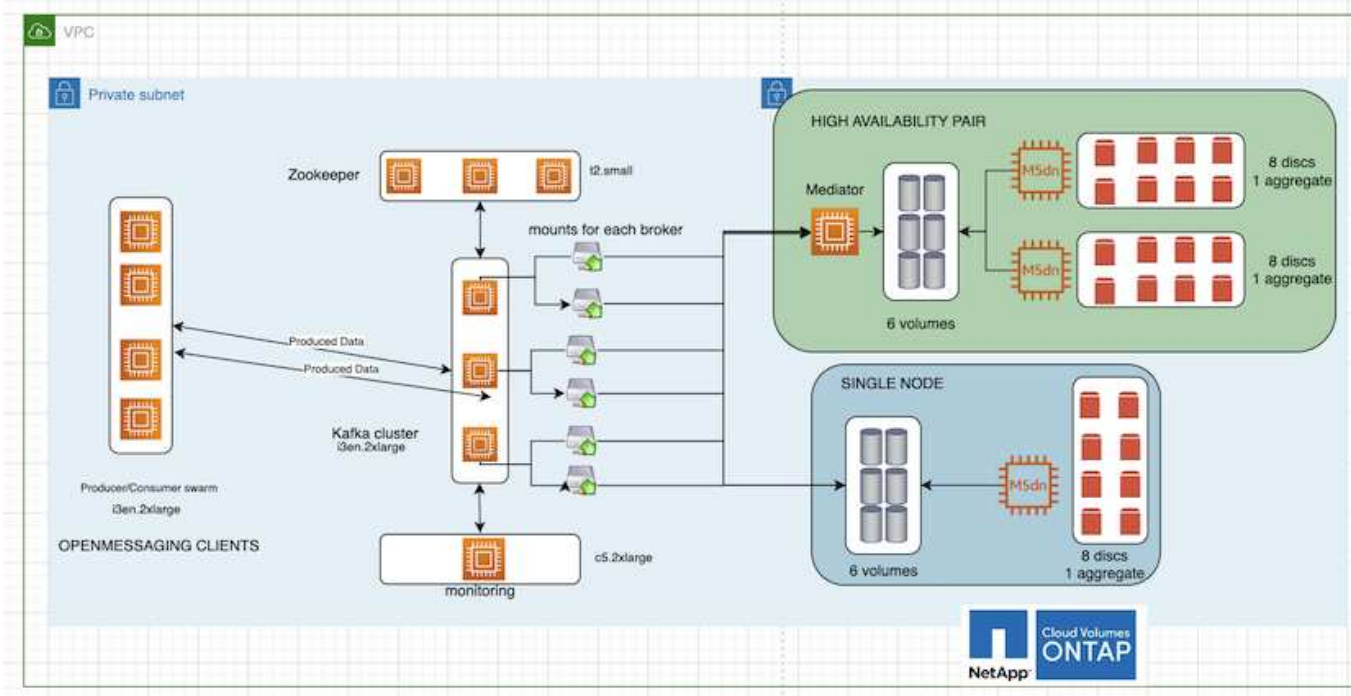
statistics start -object vnvram -instance vnvram -counter
backing_store_iops -sample-id sample_555
kafka_nfs_cvo_ha1:*> statistics show -sample-id sample_555
Object: vnvram
Instance: vnvram
Start-time: 1/18/2023 18:03:11
End-time: 1/18/2023 18:03:13
Elapsed-time: 2s
Scope: kafka_nfs_cvo_ha1-01
  Counter                                                    Value
  -----
  backing_store_iops                                         1479
Object: vnvram
Instance: vnvram
Start-time: 1/18/2023 18:03:11
End-time: 1/18/2023 18:03:13
Elapsed-time: 2s
Scope: kafka_nfs_cvo_ha1-02
  Counter                                                    Value
  -----
  backing_store_iops                                         1210
2 entries were displayed.
kafka_nfs_cvo_ha1:*>

```



The following figure depicts the architecture of an NAS-based Kafka cluster.

- **Compute.** We used a three-node Kafka cluster with a three-node zookeeper ensemble running on dedicated servers. Each broker had two NFS mount points to a single volume on the Cloud Volumes ONTAP instance through a dedicated LIF.
- **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 HA-pair Cloud volumes ONTAP instance with one 6TB GP3 AWS-EBS volume mounted on the instance. The volume was then exported to the Kafka broker with an NFS mount.



## OpenMessage Benchmarking configurations

1. For better NFS performance, we need more network connections between the NFS server and the NFS client, which can be created using `nconnect`. Mount the NFS volumes on the broker nodes with the `nconnect` option by running the following command:



```
[root@ip-172-30-0-121 ~]# cat /etc/fstab
UUID=eaalf38e-de0f-4ed5-a5b5-2fa9db43bb38/xfsdefaults00
/dev/nvme1n1 /mnt/data-1 xfs defaults,noatime,nodiscard 0 0
/dev/nvme2n1 /mnt/data-2 xfs defaults,noatime,nodiscard 0 0
172.30.0.233:/kafka_aggr3_vol1 /kafka_aggr3_vol1 nfs
defaults,nconnect=16 0 0
172.30.0.233:/kafka_aggr3_vol2 /kafka_aggr3_vol2 nfs
defaults,nconnect=16 0 0
172.30.0.233:/kafka_aggr3_vol3 /kafka_aggr3_vol3 nfs
defaults,nconnect=16 0 0
172.30.0.242:/kafka_aggr22_vol1 /kafka_aggr22_vol1 nfs
defaults,nconnect=16 0 0
172.30.0.242:/kafka_aggr22_vol2 /kafka_aggr22_vol2 nfs
defaults,nconnect=16 0 0
172.30.0.242:/kafka_aggr22_vol3 /kafka_aggr22_vol3 nfs
defaults,nconnect=16 0 0
[root@ip-172-30-0-121 ~]# mount -a
[root@ip-172-30-0-121 ~]# df -h
```

Filesystem	Size	Used	Avail	Use%	Mounted on
devtmpfs	31G	0	31G	0%	/dev
tmpfs	31G	249M	31G	1%	/run
tmpfs	31G	0	31G	0%	/sys/fs/cgroup
/dev/nvme0n1p2	10G	2.8G	7.2G	28%	/
/dev/nvme1n1	2.3T	248G	2.1T	11%	/mnt/data-1
/dev/nvme2n1	2.3T	245G	2.1T	11%	/mnt/data-2
172.30.0.233:/kafka_aggr3_vol1	1.0T	12G	1013G	2%	/kafka_aggr3_vol1
172.30.0.233:/kafka_aggr3_vol2	1.0T	5.5G	1019G	1%	/kafka_aggr3_vol2
172.30.0.233:/kafka_aggr3_vol3	1.0T	8.9G	1016G	1%	/kafka_aggr3_vol3
172.30.0.242:/kafka_aggr22_vol1	1.0T	7.3G	1017G	1%	/kafka_aggr22_vol1
172.30.0.242:/kafka_aggr22_vol2	1.0T	6.9G	1018G	1%	/kafka_aggr22_vol2
172.30.0.242:/kafka_aggr22_vol3	1.0T	5.9G	1019G	1%	/kafka_aggr22_vol3
tmpfs	6.2G	0	6.2G	0%	/run/user/1000

```
[root@ip-172-30-0-121 ~]#
```

2. Check the network connections in Cloud Volumes ONTAP. The following ONTAP command is used from the single Cloud Volumes ONTAP node. The same step is applicable to the Cloud Volumes ONTAP HA pair.

```
Last login time: 1/20/2023 00:16:29
kafka_nfs_cvo_sn::> network connections active show -service nfs*
-fields remote-host
```

node	cid	vserver	remote-host
------	-----	---------	-------------

[illegible]

```
kafka_nfs_cvo_sn-01 2315762677 svm_kafka_nfs_cvo_sn 172.30.0.223
kafka_nfs_cvo_sn-01 2315762678 svm_kafka_nfs_cvo_sn 172.30.0.223
kafka_nfs_cvo_sn-01 2315762679 svm_kafka_nfs_cvo_sn 172.30.0.223
48 entries were displayed.
```

```
kafka_nfs_cvo_sn::>
```

3. We use the following Kafka `server.properties` in all Kafka brokers for the Cloud Volumes ONTAP HA pair. The `log.dirs` property is different for each broker, and the remaining properties are common for brokers. For broker1, the `log.dirs` value is as follows:

```
[root@ip-172-30-0-121 ~]# cat /opt/kafka/config/server.properties
broker.id=0
advertised.listeners=PLAINTEXT://172.30.0.121:9092
#log.dirs=/mnt/data-1/d1,/mnt/data-1/d2,/mnt/data-1/d3,/mnt/data-2/d1,/mnt/data-2/d2,/mnt/data-2/d3
log.dirs=/kafka_aggr3_vol1/broker1,/kafka_aggr3_vol2/broker1,/kafka_aggr3_vol3/broker1,/kafka_aggr22_vol1/broker1,/kafka_aggr22_vol2/broker1,/kafka_aggr22_vol3/broker1
zookeeper.connect=172.30.0.12:2181,172.30.0.30:2181,172.30.0.178:2181
num.network.threads=64
num.io.threads=64
socket.send.buffer.bytes=102400
socket.receive.buffer.bytes=102400
socket.request.max.bytes=104857600
num.partitions=1
num.recovery.threads.per.data.dir=1
offsets.topic.replication.factor=1
transaction.state.log.replication.factor=1
transaction.state.log.min.isr=1
replica.fetch.max.bytes=524288000
background.threads=20
num.replica.alter.log.dirs.threads=40
num.replica.fetchers=20
[root@ip-172-30-0-121 ~]#
```

- For broker2, the `log.dirs` property value is as follows:

```
log.dirs=/kafka_aggr3_vol1/broker2,/kafka_aggr3_vol2/broker2,/kafka_aggr3_vol3/broker2,/kafka_aggr22_vol1/broker2,/kafka_aggr22_vol2/broker2,/kafka_aggr22_vol3/broker2
```

- For broker3, the `log.dirs` property value is as follows:

```
log.dirs=/kafka_aggr3_vol1/broker3,/kafka_aggr3_vol2/broker3,/kafka_aggr3_vol3/broker3,/kafka_aggr22_vol1/broker3,/kafka_aggr22_vol2/broker3,/kafka_aggr22_vol3/broker3
```

4. For the single Cloud Volumes ONTAP node, The Kafka `servers.properties` is the same as for the Cloud Volumes ONTAP HA pair except for the `log.dirs` property.

- For broker1, the `log.dirs` value is as follows:

```
log.dirs=/kafka_aggr2_vol1/broker1,/kafka_aggr2_vol2/broker1,/kafka_aggr2_vol3/broker1,/kafka_aggr2_vol4/broker1,/kafka_aggr2_vol5/broker1,/kafka_aggr2_vol6/broker1
```

- For broker2, the `log.dirs` value is as follows:

```
log.dirs=/kafka_aggr2_vol1/broker2,/kafka_aggr2_vol2/broker2,/kafka_aggr2_vol3/broker2,/kafka_aggr2_vol4/broker2,/kafka_aggr2_vol5/broker2,/kafka_aggr2_vol6/broker2
```

- For broker3, the `log.dirs` property value is as follows:

```
log.dirs=/kafka_aggr2_vol1/broker3,/kafka_aggr2_vol2/broker3,/kafka_aggr2_vol3/broker3,/kafka_aggr2_vol4/broker3,/kafka_aggr2_vol5/broker3,/kafka_aggr2_vol6/broker3
```

5. The workload in the OMB is configured with the following properties: (`/opt/benchmark/workloads/1-topic-100-partitions-1kb.yaml`).

```
topics: 4
partitionsPerTopic: 100
messageSize: 32768
useRandomizedPayloads: true
randomBytesRatio: 0.5
randomizedPayloadPoolSize: 100
subscriptionsPerTopic: 1
consumerPerSubscription: 80
producersPerTopic: 40
producerRate: 1000000
consumerBacklogSizeGB: 0
testDurationMinutes: 5
```

The `messageSize` can vary for each use case. In our performance test, we used 3K.

We used two different drivers, Sync or Throughput, from OMB to generate the workload on the Kafka cluster.

- The yaml file used for Sync driver properties is as follows (/opt/benchmark/driver-kafka/kafka-sync.yaml):

```
name: Kafka
driverClass:
  io.openmessaging.benchmark.driver.kafka.KafkaBenchmarkDriver
# Kafka client-specific configuration
replicationFactor: 3
topicConfig: |
  min.insync.replicas=2
  flush.messages=1
  flush.ms=0
commonConfig: |

bootstrap.servers=172.30.0.121:9092,172.30.0.72:9092,172.30.0.223:9092
2
producerConfig: |
  acks=all
  linger.ms=1
  batch.size=1048576
consumerConfig: |
  auto.offset.reset=earliest
  enable.auto.commit=false
  max.partition.fetch.bytes=10485760
```

- The yaml file used for the Throughput driver properties is as follows (/opt/benchmark/driver-kafka/kafka-throughput.yaml):

```

name: Kafka
driverClass:
io.openmessaging.benchmark.driver.kafka.KafkaBenchmarkDriver
# Kafka client-specific configuration
replicationFactor: 3
topicConfig: |
  min.insync.replicas=2
commonConfig: |

bootstrap.servers=172.30.0.121:9092,172.30.0.72:9092,172.30.0.223:909
2
  default.api.timeout.ms=1200000
  request.timeout.ms=1200000
producerConfig: |
  acks=all
  linger.ms=1
  batch.size=1048576
consumerConfig: |
  auto.offset.reset=earliest
  enable.auto.commit=false
  max.partition.fetch.bytes=10485760

```

## 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 Cloud Volumes ONTAP HA pair:

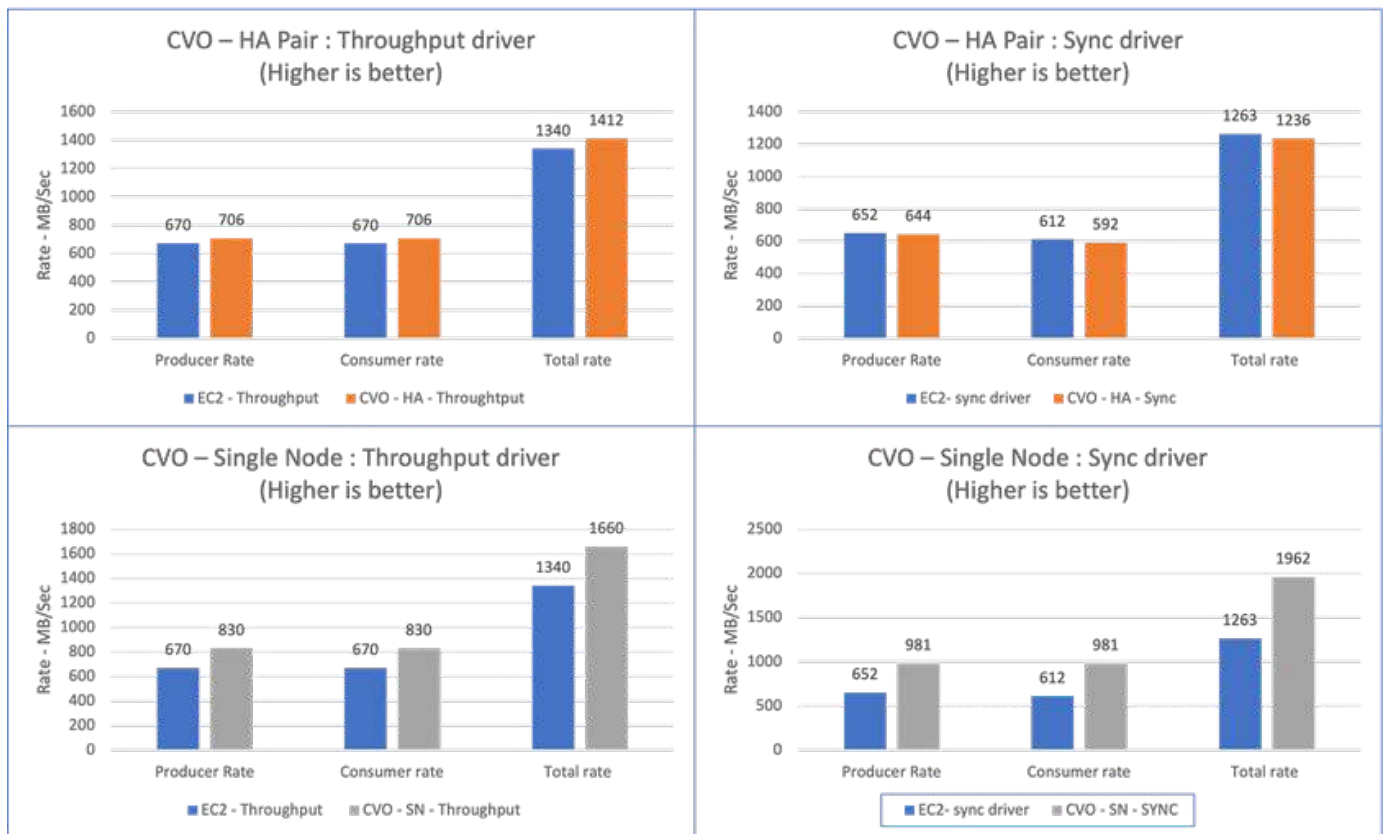
- Total throughput generated consistently by the Sync driver: ~1236 MBps.
- Total throughput generated for the Throughput driver: peak ~1412 MBps.

For a single Cloud Volumes ONTAP node:

- Total throughput generated consistently by the Sync driver: ~ 1962MBps.
- Total throughput generated by the Throughput driver: peak ~1660MBps

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.



Be sure to check the storage throughput when performing throughput or sync driver benchmarking.

## Performance



Hour

Day

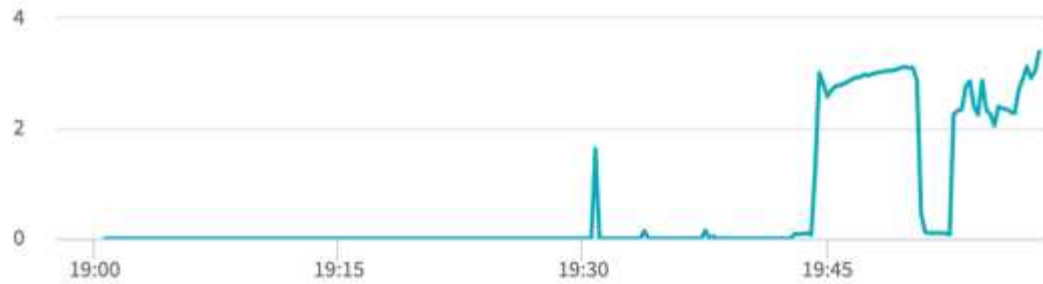
Week

Month

Year

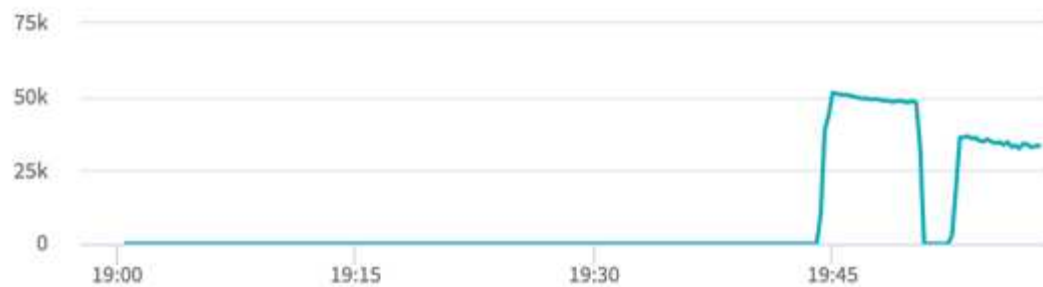
### Latency

2.99 ms



### IOPS

32.16 k



### Throughput

1,906.55 MB/s



Next: Performance overview and validation in AWS FSxN NetApp ONTAP.



## 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.