Technical Report

Performance Characterization of NetApp Cloud Volumes ONTAP for Azure with Application Workloads

NetApp Data Fabric Group, NetApp
March 2019 | TR-4671

Abstract

This technical report examines the performance and fit of application workloads to NetApp® Cloud Volumes ONTAP® instances that are running in Microsoft Azure.

NetApp®

TABLE OF CONTENTS

LIST OF TABLES

1 Introduction

To help NetApp customers select the most appropriate solutions for their IT infrastructure, NetApp provides up-to-date documentation describing its products. This technical report describes the results of performance tests for the Cloud Volumes ONTAP data management software running on Microsoft Azure virtual machines. NetApp partners, customers, and employees can use this information to make informed decisions about which workloads are appropriate for Cloud Volumes ONTAP.

The test configuration environments described in this report consist of the following components:

- Cloud Volumes ONTAP single-node clusters of the following Microsoft Azure virtual machines:
  - Standard_DS14_v2 (Premium license)
  - Standard_DS4_v2 (Standard license)
  - Standard_DS3_v2 (Explore license)
- A Microsoft Windows Server 2012 R2 host running Iometer. For each test, the virtual machine used for the Iometer machine matched the Cloud Volumes ONTAP system virtual machine.
- The iSCSI protocol for block workloads.

2 Differences Among Virtual Machines

The Azure documents General purpose virtual machine sizes and Memory optimized virtual machine sizes describe the capabilities of the different Windows virtual machines in terms of network, CPU, and memory.

The specific combination of these components is what determines the overall performance of an instance. For Cloud Volumes ONTAP, the CPU capability contributes to general performance, whereas memory contributes heavily to read performance. The network capability acts as a throughput limiting factor that is independent of read and write performance levels.

Table 1) Capabilities of supported Azure Microsoft virtual machines.

VM Type	CPU	RAM	Expected Network Bandwidth (Mbps)
Standard_DS3_v2	4	14GB	3000
Standard_DS4_v2	8	28GB	6000
Standard_DS14_v2	16	112GB	12000

3 Test Configurations

All test configurations used the iSCSI protocol for block I/O connectivity. The tests focused on the following:

- Iometer was used to generate I/O workloads. We used different numbers of LUNs and different numbers of outstanding I/O operations to see the effect of concurrency level on IOPS, throughput, and latency.
- Normal writing speed was used for single-node systems. Under normal write speed, data is written directly to disk, minimizing the likelihood of data loss if an unplanned system outage occurs. Conversely, when high write speed is used, data is buffered in memory before it is written to disk. This configuration provides faster write performance but increases the potential for data loss if an unplanned system outage occurs. High write speed is recommended if fast write performance is required and you can tolerate the risk of data loss due to an unplanned system outage. For example, data loss might be handled by the application.
- We tested different Microsoft Azure virtual machines. The virtual machines tested had different hardware characteristics, and, because they belonged to different licensing packages, different

costs were incurred by Microsoft Azure and NetApp. A virtual machine from each licensing package was picked for the tests.

- All the tests were performed with Premium_LRS volumes. Cloud Volumes ONTAP supports an extra volume type: Standard_LRS.
- Cloud Volumes ONTAP was tested on three different Microsoft Azure virtual machines:
  - Standard_DS3_v2, included in the Explore license, which allows up to 2TB of storage
  - Standard_DS4_v2, included in the Standard license, which allows up to 10TB of storage
  - Standard_DS14_v2**,** included in the Premium license, which allows up to 252TB of storage

4 Summary of Test Results

Each tested configuration consisted of a unique workload that is representative of the workloads used in widely deployed POSIX applications:

- **OLTP workload.** 8KB block size, 100% random access I/O, and a mixture of 70% reads. This workload simulates database applications (SAP, Oracle, SQL) and OLTP servers.
- **Streaming reads.** 64KB block size, 100% reads, and 100% sequential access I/O. This workload simulates applications such as media servers (for example, video on demand) and Virtual Tape Libraries (VTL).
- **Streaming writes.** 64KB block size, 100% writes, and 100% sequential access I/O. This workload simulates applications such as media capture, virtual type libraries, medical imaging, archiving, backup, and video surveillance.
- **Analytics.** 16KB block size, 50% reads, and 100% random access I/O. Analytics workloads are unplanned by nature. They involve both read and write operations and require high throughput and low latency.

The test workloads differed significantly. Collectively, however, they represent workloads that might be considered for a Cloud Volumes ONTAP deployment. The results of each test are discussed separately in this section because the workloads were so different.

4.1 OLTP Workload

The 8KB block size workload simulates an OLTP transactional database. Transactional workloads tend to be read heavy as data about an item is retrieved, but they involve a smaller number of writes as transactions are committed. Transactional workloads are highly sensitive to write latency, especially for writes to the transaction log. Typically, it is optimal for log writes not to exceed 15ms of latency. Lower latency is always better.

The workload tested consisted of 8KB block size, 70% reads, and 100% random access I/O.

Table 2 shows the results of 8KB blocks, 70% reads, and 100% random access I/O for the different virtual machines.

Table 2) OLTP workload IOPS and latency.

VM Type	IOPS for Single Node (Ops/s)	Latency for Single Node (msec)	IOPS for HA pair (Ops/s)	Latency for HA pair (msec)	License
Standard_DS14_v2	18120	9	17195	10	Premium
Standard_DS4_v2	18770	7	15995	11	Standard
Standard_DS3_v2	15755	8	Not applicable	Not applicable	Explore

4.2 Streaming Read Workload

We tested a workload with a large contiguous read request size that simulates applications such as media servers (for example, video on demand) and Virtual Tape Libraries (VTL). The workload consisted of 64KB block size, 100% reads, and 100% sequential access I/O.

The most significant measurement for this type of workload is throughput, as Table 3 shows.

Table 3) Streaming read workload.

VM Type	Throughput for Single Node (MBps)	Throughput for HA pair (MBps)	License
Standard_DS14_v2	427	380	Premium
Standard_DS4_v2	411	346	Standard
Standard_DS3_v2	340	Not applicable	Explore

Cloud Volumes ONTAP has robust performance for streaming reads. The differences between the virtual machines are caused by the networking capabilities of the instance and the number of CPUs available.

4.3 Streaming Write Workload

We tested a workload with a large contiguous write request size that simulates applications such as media capture, Virtual Tape Libraries (VTL), medical imaging, archiving, backup, video surveillance, and reference data. The workload consisted of 64KB block size, 100% writes, and 100% sequential access I/O.

The most significant measurement for this type of workload is throughput, as Table 4 shows.

Table 4) Streaming write workload throughput.

VM Type	Throughput for Single Node (MBps)	Throughput for HA pair (MBps)	License
Standard_DS14_v2	257	243	Premium
Standard_DS4_v2	258	205	Standard
Standard_DS3_v2	47	NA	Explore

Cloud Volumes ONTAP performed well for the streaming write workload.

4.4 Analytics Workloads

We tested a workload with analytics characteristics. Analytics contain a mixture of read and write, for which throughput and latency are the most important measurements. The workload we tested consisted of 16KB block size, 50% read, and 100% random access I/O.

Table 5) Analytics workload throughput.

VM Type	Throughput for Single Node (MBps)	Latency for Single Node (msec)	Throughput for HA pair (MBps)	Latency for HA pair (msec)	License
Standard_DS14_v2	145	14	152	12	Premium
Standard_DS4_v2	135	9	128	14	Standard
Standard_DS3_v2	90	20	NA	NA	Explore

The results show that Cloud Volumes ONTAP is a good fit for analytics workloads, because it can maintain high throughput under strict latency constraints.

4.5 Data on FabricPool

In the ONTAP 9.5 release, all inactive or cold data in the active file system can be tiered to the Azure Blob. We tested workloads on volumes that had the relevant tiering policy (auto, backup) to cool inactive data to the Azure Blob.

Writing to Azure Blob

We worked with Iometer and the iSCSI protocol to stream a write workload on a volume with tiering capacity set to 'auto' while the data was written to EBS. In the background, data that was considered cold was transferred to Azure Blob. We did not see any change in write performance.

Reading from Azure Blob

We worked with the FIO tool on Linux as a client and the NFS protocol, reading from a mounted volume, while all the inactive data was cooled to the Azure Blob. Running 64K BS, random read, average read size was 16K. We saw degradation of about 30%-50% in throughput and IOPS, comparing to reading all the data from EBS drives.

Conclusion

Cloud Volumes ONTAP was found to be a good fit for OLTP workloads, streaming read workloads, streaming write workloads, and analytics workloads. NetApp has a long history of providing high-performance and feature-rich storage systems. Cloud Volumes ONTAP extends this legacy to Azure. With Cloud Volumes ONTAP, NetApp continues to develop leading-edge storage solutions that provide the agility and mobility that current NetApp customers need and that future NetApp customers will want. Cloud Volumes ONTAP is part of a family of products that stretch from the private cloud to the hybrid cloud to the public cloud. The product family runs the NetApp ONTAP data management software. Understanding the performance characteristics of Cloud Volumes ONTAP is critical for setting our customers' expectations and enabling their continued success.

Where to Find Additional Information

To learn more about the information described in this document, refer to the following website:

- NetApp Documentation Center

Version History

Version	Date	Document Version History
Version 1.1	March 2019	Updated performance with Cloud Volumes ONTAP v9.5 Single and HA pair
Version 1.0	March 2018	Updated performance with Cloud Volumes ONTAP 9.3

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.

Copyright Information

Trademark Information