Technical Report

Oracle Databases on ONTAP Cloud with Amazon Web Services

Jeffrey Steiner, NetApp
April 2018 | TR-4676

Important

Consult the Interoperability Matrix Tool (IMT) to determine whether the environment, configurations, and versions specified in this report support your environment.

TABLE OF CONTENTS

1 Introduction

NetApp® ONTAP® is powerful data management software with native capabilities that include integrated data protection and recovery, efficient cloning, and replication. ONTAP Cloud and ONTAP Select are software-defined storage options that extend the functionality of ONTAP beyond physical AFF and FAS hardware. Specifically, ONTAP Cloud takes control of hyperscaler storage volumes, while ONTAP Select runs in a more traditional virtualization infrastructure. ONTAP Cloud and ONTAP Select deliver the ONTAP feature set across more platforms to address the requirements of public and private cloud users.

From a technical standpoint, all of these options are still ONTAP; the difference between Cloud and Select lies in the business requirements. For example, some enterprises are looking to public cloud providers for a less expensive and more flexible disaster recovery capability. ONTAP Cloud delivers a simple and efficient replication capability to make on-premises data available within Amazon AWS or Microsoft.

As a further example, some service providers are building secure multitenant services that require integrated backup and restore capability. ONTAP Select can run on existing virtualized infrastructure, allowing the service provider to provide each tenant with their own data management environment. New tenants can be quickly deployed by adding additional ONTAP Select instances.

With respect to databases, there is a wide variety of user requirements, including database size, performance requirements, and data protection needs. Known deployments of NetApp storage include everything from a virtualized environment of approximately 6,000 databases running under VMware ESX to a single-instance data warehouse currently sized at 996TB and growing. As a result, there are few clear best practices for configuring an Oracle database on NetApp storage.

This document builds on the general best practices for managing Oracle databases on ONTAP storage described in TR-3633. The principles are unchanged overall, but there are a few considerations applicable to ONTAP Cloud and the hyperscaler environments that support ONTAP Cloud instances.

For more details, see the following resources:

- TR-3633: Oracle on NetApp ONTAP
- TR-4591: Database Data Protection
- TR-4592: Oracle on MetroCluster
- TR-4534: Migration of Oracle Databases to NetApp Storage Systems

2 ONTAP Platforms

ONTAP software is the foundation for advanced data protection and management. However, ONTAP refers only to software. There are several ONTAP hardware platforms from which to choose:

- ONTAP on AFF or FAS, including NetApp Private Storage (NPS)
- NetApp Private Storage for Cloud
- ONTAP Select
- ONTAP Cloud
- Cloud Volumes

The key concept is that ONTAP is ONTAP. Some hardware options offer better performance, others offer lower costs, and some run exclusively within hyperscaler clouds. The core functions of ONTAP are unchanged, with multiple replication options available to bind different ONTAP platforms into a single solution. As a result, data protection and disaster recovery strategies can be built on real-world needs, such as performance requirements, capex and opex considerations, and overall cloud strategy. The underlying storage technology runs anywhere in any environment.

2.1 ONTAP with AFF and FAS Controllers, Including NetApp Private Storage

For maximum performance and control of data, ONTAP on a physical AFF or FAS controller remains the leading solution. This approach is the standard option on which thousands of customers have relied for more than 25 years. ONTAP delivers solutions for any environment, ranging from three mission-critical databases to 60,000-database service provider deployments, instant restores of petabyte-scale databases, and DBaaS involving hundreds of clones of a single database.

NetApp introduced the NetApp Private Storage (NPS) option to address the needs of data-intensive workloads in the public cloud. Although many public cloud storage options exist, most of them have limitations in terms of performance, control, or scale.

ONTAP NPS delivers maximum storage performance, control, and flexibility to public cloud providers, including Amazon AWS, Microsoft Azure, and IBM Cloud. This capability is delivered by AFF and FAS systems, including MetroCluster™ options, in data centers connected directly to public clouds. The full power of the hyperscaler compute layer can be used without the limitations of hyperscaler storage. Furthermore, NPS enables cloud-independent and multicloud architectures because the data, such as application binaries, databases, database backups, and archives, all remains wholly within the NPS system. There is no need to expend time, bandwidth, or money moving data between cloud providers. Data remains under the ownership and control of the end user.

2.2 ONTAP Cloud

ONTAP Cloud runs in a hyperscaler cloud environment, bringing intelligence and Data Fabric connectivity to hyperscaler storage volumes. The overall best practices for running Oracle on ONTAP Cloud are the same as for Oracle on ONTAP generally. The primary considerations specific to Oracle on ONTAP Cloud surround performance and, to a lesser extent, cost.

ONTAP Cloud is partially limited by the performance of the underlying volumes managed by the cloud provider. The result is more manageable storage, and, in some cases, the caching capability of ONTAP Cloud offers a performance improvement. However, there are always some limitations in terms of input/output operations per second (IOPS) and latency due to the reliance on public cloud provider storage.

The prime use cases for ONTAP Cloud are currently disaster recovery (DR) and development and testing work, but some customers have used ONTAP Cloud for production activity as well. One particularly notable report was the use of Oracle’s In-Memory feature to mitigate storage performance limitations. This approach allows more data to be stored in RAM on the virtual machine (VM) hosting the database server, thus reducing performance demands on storage.

2.3 ONTAP Select

ONTAP Select is similar to ONTAP Cloud, except that it runs on a customer’s own virtualization infrastructure and delivers ONTAP intelligence and Data Fabric connectivity to the drives inside of white-box hardware. ONTAP Select also allows ONTAP and guest operating systems to share the same physical hardware as a highly converged infrastructure. The best practices for running Oracle on ONTAP Select are not affected. The primary consideration is performance, and ONTAP Select should not be underestimated.

An ONTAP Select environment does not match the peak performance of a high-end AFF system, but most databases do not require 500K IOPS. Typical databases require only around 5K to 10K IOPS, a target that can be met by ONTAP Select. Furthermore, most databases are limited more by storage latency than storage IOPS, a problem that can be addressed by deploying ONTAP Select on SSDs.

3 Planning and Licensing

The most important aspect of ONTAP Cloud planning is that you must first define the requirements of the database in terms of performance, capacity, and availability. Next, you must reduce to practice the required configuration given the capabilities and limitations of the virtual environment hosting ONTAP Cloud and the hosts utilizing ONTAP Cloud. Finally, you must assess the costs of the complete cloud environment.

Considerations include:

- How many random IOPS are required?
- How much bandwidth between the database server and ONTAP Cloud is required?
- What are the license requirements?
- What are the backup, recovery, and availability SLAs?
- How often is the cloud environment active?
- What are the backup and recovery requirements?

3.1 ONTAP Cloud Licensing

ONTAP Cloud is licensed in one of two ways: on demand through a hyperscaler cloud provider or as a bring-your-own-license (BYOL) model.

At the time of writing, on-demand license charges vary from $0.75 per hour to $2.90 per hour, depending on the features and capacities required. Detailed pricing information is available on the AWS marketplace. If the on-demand options are not suitable for your requirements, ONTAP Cloud licenses can be purchased directly from NetApp or a NetApp partner.

3.2 Oracle Licensing

You must also consider Oracle licensing costs when sizing a solution. Although AWS offers a lot of options for VM CPU, memory, storage, and bandwidth capability, they are not independently scalable. In some cases, you might require more bandwidth or RAM for a given VM than can be obtained without an undesirable concurrent increase in CPUs, which affects license costs. At the time of writing, Oracle's policy for cloud computing is as follows:

- AWS with hyperthreading: one Oracle CPU license per two vCPUs
- AWS without hyperthreading: one Oracle CPU license per vCPU

Furthermore, Oracle policy limits the use of Oracle Standard Edition to 8 AWS vCPUs. Use of Oracle Enterprise Edition is limited to 16 AWS vCPUs.

Note: This section addresses published Oracle licensing policy at the time of writing. This policy is subject to change by Oracle and furthermore can be modified by direct negotiation with Oracle. Many customers have enterprise licensing agreements that offer much more favorable terms.

4 AWS Storage Performance

The information in this document is based on the information currently published for Amazon AWS and ONTAP Cloud.

4.1 IOPS Targets

The most important requirement for a database is usually random, single-block IOPS. A typical enterprise resource planning (ERP) application that is driven by end users typically requires approximately 5K to 10K IOPS, while large data warehouses can drive IOPS to 500K and beyond.

ONTAP Cloud wraps itself around a collection of hyperscaler volumes to provide advanced capabilities such as backup and recovery, database cloning, and disaster recovery services based on Snapshot. ONTAP Cloud cannot improve performance much beyond the capabilities of the underlying hyperscaler volumes. In some cases, a database might have a highly cacheable workload that can be served from RAM on the ONTAP Cloud instance. In general, however, database performance of the underlying hyperscale volumes are neither improved nor reduced by the use of ONTAP Cloud. Exceptions include unusual workloads or when the overall configuration is not optimal.

4.2 ONTAP Cloud Checksums

ONTAP Cloud provides an additional layer of data integrity in the form of checksums. As data is written to disk, a checksum is calculated and stored alongside the data. When the data is later read, the checksum is recalculated to make sure that it matches. With this process, ONTAP can detect any corruption that might occur on the hyperscaler storage volume.

It is important to remember that checksums increase the overall IOPS demand. For example, a database performing 5000 highly random IOPS requires 5000 data IOPS from the ONTAP Cloud instance and up to 5000 additional IOPS to read the checksums. Both sources of storage I/O must be considered when provisioning the hyperscaler storage volumes. As a conservative measure, you should double the database IOPS estimates when deciding how many IOPS are required from the underlying EBS volumes.

4.3 EBS Volumes

The most common volume type used in AWS environments is the general purpose SSD type, called gp2. The performance capability of these volumes is based mostly on the size of the volume and is set at three IOPS per GB.

The optimal EBS volume size is 3.33TB. This results in a volume with 10K IOPS, which is the maximum available in a gp2 volume. Using larger volumes does not yield better performance and using smaller volumes results in proportionally lower performance volumes. Going directly to the 3.33TB volume size helps make sure that, as the environment grows, more EBS volumes are not required to improve performance. If an ONTAP Cloud aggregate is built with multiple smaller EBS volumes, there is a risk that the maximum of six is reached, leaving no options for adding IOPS capability to the aggregate.

As an example, consider a 1TB database with an IOPS requirement of 10K IOPS. This database could require up to 10K additional checksum IOPS for a total of 20K IOPS. You can determine sizing based on performance or capacity:

$$1\text{TB} \times 3\text{ IOPS per GB} = 3{,}000\text{ IOPS}$$

$$3.3\text{TB} \times 3\text{ IOPS per GB} = 10{,}000\text{ IOPS}$$

A single TB of storage has a baseline I/O capability of 3K IOPS, which is insufficient for the 20K IOPS requirement. Provisioning a pair of 3.3TB volumes of storage meets the 20K IOPS performance requirement but results in threefold overprovisioning.

If the performance requirements are met by overprovisioning ordinary gp2 volumes, the cost is $300 per month:

$$6.6\text{TB} \times \$0.10\text{ per GB per month} = \$660/\text{month}$$

AWS also offers an option to provision bytes and IOPS independently using a volume type called io1. The same 1TB database could be based on four 250GB io1 volumes that deliver 5K IOPS each. This approach precisely meets both the capacity and performance requirements, but the cost is higher than for overprovisioning gp2 volumes:

$$1\text{TB} \times \$0.10\text{ per GB per month} = \$100/\text{month}$$

$$20\text{K IOPS} \times \$0.065\text{ per IOPS per month} = \$1300/\text{month}$$

There is also a bursting capability that is applicable to volumes smaller than 1TB. The purpose of the burst capability is to support initial startup I/O such as booting a VM or starting a database. The bursting limit is 3,000 IOPS, and the duration of the burst period varies based on the size of the volume.

If the example 1TB database is based on six 250GB gp2 volumes, it would have a burst capability of 18K IOPS. However, after approximately 40 minutes, the burst credits would be exhausted, and performance is limited to the baseline 3K IOPS. If I/O remains below 3K IOPS, the burst credits are gradually replenished. If the performance target of 20K IOPS for the example 1TB database does not apply 100% of the time, then relying on burst credits might be a more affordable option.

Caution: Bursting behavior complicates performance testing with any AWS VM, including ONTAP Cloud, with volumes less than 1TB in size because performance varies significantly depending on the available burst credits.

4.4 Latency

Although hyperscaler storage options include all-flash media, the observed latencies are higher than those of a typical physical, on-premises all-flash storage array. The expected latency with AWS gp2 volumes is between 1ms and 3ms. In comparison, a physical storage array such as a NetApp AFF or FAS system immediately acknowledges writes after they are committed to nonvolatile NVRAM media. The write to the actual storage devices occurs much later. The result is that write latencies are typically between 100µs and 200µs.

These latency differences should be considered when planning a cloud-based strategy for any relational database. Testing realistic workloads, including write I/O, is a requirement.

4.5 Bandwidth

Although bandwidth limits exist for cloud storage, the limiting factor for relational database workloads is usually IOPS, not MBps. Still, we must consider the bandwidth limits of cloud VMs.

For example, assume a performance target of 5,000 IOPS for a database using an 8K block size. This target is approximately 50MBps of I/O, which is well within the bandwidth capabilities of most hyperscaler VMs. The more likely limitation would be from the IOPS capabilities of the underlying storage resources.

5 Oracle Database Performance Optimization

In many cases, reaching the required IOPS with acceptable latency is a challenge, especially when moving a database from an on-premises environment to a hyperscaler cloud. In some cases, the only option might be to pay for cloud storage volumes that offer sufficient performance. However, alternatives exist that allow a database to completely avoid I/O, which can deliver improved performance and/or reduce costs.

5.1 Oracle Buffer Cache

The Oracle buffer cache is RAM reserved by Oracle to cache recently used blocks. Many production databases and typical customer practices have their roots in very old versions of Oracle for which server RAM was at a premium. As a result, many storage systems are significantly overbuilt to meet IOPS requirements that could have been avoided entirely if more RAM on the host were made available to the database. Adding a small amount of RAM to the Oracle system global area (SGA) can frequently pay for itself through storage savings. This saving applies equally inside and outside of the cloud.

Options are somewhat limited for RAM in AWS environments. VMs are available with anything from 512MB of RAM up to 488GB, but in general the CPUs allocated to a VM increase with RAM. The best possible Oracle performance would probably be seen with a VM that includes the maximum 488GB of RAM, but those VMs also currently include 64 virtual CPUs, which is extremely expensive.

More options exist at a smaller scale. For example, an AWS m4.2xlarge includes eight vCPUs and 32GB of RAM, while an r4.2xlarge contains the same eight vCPUs and 61GB of RAM. That doubles the RAM with only a small increase in cost.

There is no good way to determine the precise amount of Oracle buffer cache required when moving a workload into the cloud because the RAM in a VM is only one of many variables. NetApp recommends remembering that two options exist when working with storage sizing:

- Configure the storage system to meet established I/O requirements.
- Configure the database server to reduce and avoid performing I/O in the first place.

5.2 Oracle Smart Flash Cache

Oracle Smart Flash Cache was introduced with Oracle 11gR2 as a less expensive way to improve Oracle performance without a requirement to invest in RAM or storage systems. It functions as a victim cache, in a manner similar to that of NetApp Flash Cache™ technology. In a typical single cache architecture, data blocks are constantly aged out of the main RAM-based cache as new blocks are read in. If those blocks are needed again, they must be read back into cache from the drives.

Oracle databases tend to have very concentrated workloads. Although a very small increase in RAM on a storage system can deliver a very large increase in performance, the cost of RAM in a storage system can be significant. NetApp and Oracle both introduced features branded as a "flash cache" for similar reasons. They function as extensions of the main cache to avoid storage I/O.

An increasing number of hyperscaler VM options include locally attached SSD storage. Local storage is not desirable for important data such as normal datafiles or redo logs, but it can be safely used as a Oracle Smart Flash Cache. This strategy can help overcome some of the IOPS limits on hyperscaler persistent storage volumes.

For example, consider the options listed in Table 2.

Table 1) Direct-attached storage options.

VM Type	Cores	RAM	Direct-Attached Storage	Cost
r4.2xlarge	8	61GB	None	$0.532/hour
i3.2xlarge	8	61GB	1 x 1,900 NVMe SSD	$0.624/hour

A 17% increase in the hourly cost of the VM delivers nearly 2TB of NVMe-based storage that can be used for both Oracle Smart Flash Cache and temporary tablespaces.

5.3 Oracle In-Memory

Oracle's In-Memory feature allows a database server to efficiently store a compressed columnar representation of columns or full tables in local RAM. This feature can also reduce the bandwidth and IOPS demands on a storage system and has been validated by customers. It is a licensed option, but the costs might be outweighed by the savings.

5.4 Oracle Temporary Tablespaces

Most databases are limited by random read latency, but, in a public cloud environment, the bandwidth demands of temporary tablespace activity can become a problem. As shown in section 5.2, "Oracle Smart Flash Cache," you can acquire a large quantity of locally attached SSD for a modest increase in cost. You can then use this space for both Oracle Smart Flash Cache and temporary tablespaces.

6 ONTAP Cloud Deployment Options

ONTAP Cloud is more than just a storage system. It is a virtualized, consumption-based, on-demand storage system. There are no capital acquisition costs, no time spent awaiting shipment, and no effort to cable a new system. Individual ONTAP Cloud instances can be stopped, started, created, destroyed, and reconfigured as required. This approach opens up many options for configuring a database environment that would be impossible with physical hardware.

The following sections describe some of the important options for planning an Oracle on ONTAP Cloud deployment. Understanding these options is critical to selecting the best overall configuration for a database project that meets SLAs for performance, data protection, and availability at the lowest possible cost.

A complete explanation ONTAP Cloud and ONTAP Cloud Manager can be found on the NetApp Support site.

6.1 Write Speed

ONTAP Cloud has two options for write speed: normal and high. In normal speed write mode, data is written directly to disk before being acknowledged to the operating system. In high-speed write mode, data is buffered in memory for up to 10 seconds before being written to disk. The result is much higher write performance, but at the risk of data loss if a crash of the ONTAP Cloud instance occurs while data is held in the buffer.

This result might be acceptable in some cases. For example, an Oracle database holding critical business information would normally run exclusively in normal write mode because loss of redo log data could mean permanent data loss. In contrast, a database copy used for development purposes would not normally require the same level of protection. In the event of a crash of the ONTAP Cloud instance, the state of the aggregate would essentially jump backward in time up to 10 seconds. The database would otherwise be crash consistent and should start up again as usual.

6.2 High Availability

ONTAP Cloud in an Amazon AWS environment includes an option to run as an HA pair rather than just a single instance. In this configuration, each node maintains a copy of data on its locally attached Elastic Block Store (EBS) and a mirror on the EBS volumes of its HA partner. This configuration is protective against data loss caused by failure of an EBS volume. Furthermore, this type of configuration is typically used with each node in a different AWS availability zone to protect against an outage affecting an entire AWS data center.

6.3 Instance Type

The VM hosting ONTAP Cloud can be changed as requirements change. For example, a database used for basic functional testing can be hosted on a smaller, less costly VM. If more demanding load tests are required or the number of users accessing the database increases, ONTAP Cloud can be migrated to a larger and more powerful VM. The change is not disruptive when ONTAP Cloud is configured as an HA pair. Single-node configurations would experience an I/O interruption during the change. Depending on the OS configuration in use, this approach might require a restart of the Oracle database, or it might just cause a pause.

6.4 Purchasing Options

Hyperscaler cloud providers offer multiple purchasing options ranging from on-demand instances to reserved instances to dedicated instances. The costs can differ substantially. For example, a top-of-the-line ONTAP Cloud instance running on an AWS c4.8xlarge instance currently has an on-demand rate of $1.591/hour or an annualized rate of about $14,000, assuming that the VM runs 24 hours per day for a full year. If an ONTAP Cloud application or database VM is expected to run continuously, substantial discounts are available. Reserving an instance and paying in advance currently yield savings of about 40%, bringing the cost to approximately $8,300.

Current AWS pricing is available here.

6.5 Operating Schedules

Not all ONTAP Cloud instances need to run full time. For example, four production database servers continuously running in AWS with a central ONTAP Cloud also running continuously are substantially more expensive than an ONTAP Cloud instance that is only used for disaster recovery. In the case of DR, the database servers would not normally need to be running at all, and the ONTAP Cloud instance might only need to be activated for 30 minutes per day to update the mirrors. The ability to turn resources on and off on demand is a critical benefit of hyperscaler cloud services and should be leveraged as much as possible.

6.6 ONTAP Cloud VM Options

ONTAP Cloud is available on VMs of various sizes, ranging from 4 to 16 cores with capacities from 1TB to 368TB. The hourly cost varies correspondingly from $0.95/hour to $4.78/hour. Selecting the right option requires first defining all the expected workloads in terms of IOPS, bandwidth, and capacity and then ensuring ONTAP Cloud is sized appropriately.

For example, ONTAP Cloud running on an AWS m4.xlarge is the least expensive option, but this option is limited to approximately 94MBps total bandwidth available for serving data from NFS file systems or iSCSI LUNs.

The most powerful ONTAP Cloud instances runs on a c4.8xlarge instance, with a maximum backend storage bandwidth capability of 500MBps and 32K IOPS. The cost is currently $4.78/hour, which is nearly 5x the cost of the entry-level m4.xlarge option. If the purpose of the architecture is to provide an around-the-clock cloning environment, the total costs of running the ONTAP Cloud instance could be enough to warrant a closer look at the least costly option that still meets requirements. However, if ONTAP Cloud is only hosting a disaster recovery environment that wakes once per day to transfer changes, the cost might be considered negligible.

7 ONTAP Cloud Best Practices

A complete description of the configuration of ONTAP Cloud is beyond the scope of this document. The following sections cover the key considerations of an ONTAP Cloud deployment. TR-3633: Oracle Databases on ONTAP covers general best practices for Oracle on ONTAP and is applicable to ONTAP Cloud. Variations or additions to TR-3633 are discussed in the following sections.

7.1 Quality of Service

ONTAP contains advanced QoS capabilities to manage both IOPS and bandwidth, but it is unlikely that QoS would be of value in a Cloud ONTAP environment. The reason is the relatively low IOPS capabilities of the hyperscaler volumes. A typical ONTAP Cloud deployment might be capable of 10K IOPS. Subdividing 10K IOPS further by QoS would probably result in NFS or iSCSI clients with too little IOPS to accomplish much work.

In addition, the primary value of QoS applies to physical hardware systems with inherent physical limitations, capital costs, and long procurement lead times. The lack of flexibility makes QoS important for management of those limited resources. In contrast, you can create ONTAP Cloud instances on demand to address needs as they arise and change, meaning an ONTAP Cloud instance functions as a basic unit of QoS. Furthermore, using multiple ONTAP Cloud resources as a QoS method can save money because you can shut them down individually when they are not in use.

7.2 Storage Virtual Machines

Storage virtual machines (SVMs, previously called Vservers) are the basic functional unit of storage. It is useful to compare an SVM on ONTAP to a guest on a VMware ESX server. When first installed, ESX has no preconfigured capabilities, such as hosting a guest OS or supporting an end-user application. It is an empty container until a VM is defined. ONTAP is similar. When first installed, ONTAP has no data-serving capabilities, and an SVM must be defined. It is the SVM personality that defines the data services.

ONTAP Cloud is typically used with just a single SVM. Like QoS, the purpose of SVMs is to better manage resources in a defined physical system that cannot be easily changed. If you require multiple SVMs, it would usually be easier to create additional ONTAP Cloud instances.

There is, however, no technical reason why you cannot use multiple SVMs. For example, an application associated with a particular database might need to store certain data with enhanced security. In this case, you could use a second SVM with a different management team.

7.3 Compression, Compaction, and Deduplication

Data efficiency features such as compression, compaction, and deduplication should be used with caution in an ONTAP Cloud environment due to the comparatively limited IOPS capabilities of the backend hyperscaler storage volumes.

Efficiency features such as NetApp Snapshot™ copies and NetApp FlexClone® volumes focus more on logical data than physical storage. They are extremely useful and should be part of almost any ONTAP Cloud architecture.

7.4 Thin Provisioning

Thin provisioning refers to the configuration of more space on a storage system than is technically available. The most common type of thin provisioning is the creation of a LUN that appears to a SAN host as its full size, but it only starts occupying space on the storage system as data is written.

Compression, Snapshot copies, and FlexClone volumes are also forms of thin provisioning because they create more logical data than physical space, although there is always a chance that unexpected change rates consume all physical space.

Snapshot copies, FlexClone volumes, and thin-provisioned LUNs are all appropriate types of thin provisioning for an ONTAP Cloud environment.

7.5 Space Management

Some care must be taken with thin provisioning an Oracle environment because data change rates can increase unexpectedly. For example, space consumption due to Snapshot copies can grow rapidly if tables are reindexed, or a misplaced Oracle Recovery Manager (RMAN) backup can write a large amount of data in a very short time. Finally, it can be difficult to recover an Oracle database if a file system runs out of free space during datafile extension.

Fortunately, these risks can be addressed with careful configuration of `volume-autogrow` and `snapshot-autodelete` policies. As their names imply, these options enable a user to create policies that automatically clear space consumed by Snapshot copies or grow a volume to accommodate additional data. Many options are available, and needs vary by customer.

See the [ONTAP Logical Storage Management Guide] for your version of ONTAP for a complete discussion of these features.

7.6 Volume Reallocation

The most important type of I/O for a database is usually random read I/O. This fact is the primary reason why all-flash storage systems have come to dominate the database market: they deliver consistent low latency for random read operations.

Sequential I/O operations have historically been less of a problem because they are an inherently efficient type of I/O. Data is read in large blocks, and after a storage system detects a sequential I/O operation in progress, it performs readahead to load data from the drives before the database has even requested it.

When storage systems were based exclusively on spinning media, the relatively low IOPS capability of the drives led to some problems related to sequential I/O processing with databases. As ONTAP processed the overwrites of datafiles, the layout on disk could become suboptimal. This fact has no effect on random I/O performance, but sequential I/O operations such as full table scans or RMAN backups slowed because more I/O was required to assemble the large blocks being read.

Although almost all hyperscaler storage used by ONTAP Cloud is solid state, the relatively limited IOPS capabilities of volumes require periodic optimization to make sure that sequential I/O performance does not degrade over time. Scheduled reallocation scans of a volume might be required. For example, a standard Oracle standby database performs only random overwrites of the datafiles. The command `volume reallocation` can be used to measure and, if necessary, optimize the layout on disk. This approach might not be required in all cases. See section 11, “Oracle on ONTAP Cloud Solution Designs,” for specific notes about the use of reallocation scans in different scenarios.

Caution: The `read_realloc` option operates in a similar manner, but it was developed for storage systems with spinning drives and is not suitable for ONTAP Cloud.

8 Oracle Data Protection

8.1 Snapshot-Based Backups

A more complete explanation of Oracle data protection with ONTAP can be found in TR-4591: Database Data Protection: Backup, Recovery, Replication, and DR. For comprehensive assistance with design, contact NetApp or a NetApp partner.

The most important consideration for a file system layout is the plan for leveraging NetApp Snapshot technology. There are three primary approaches:

- Consistency group snapshots
- Snapshot-optimized backups
- Snapshot-protected hot backups

8.2 Consistency Group Backup

Placing an entire Oracle database in a single volume, including datafiles, archive logs, redo logs, and control files, is a valid backup, restore, and replication method.

A crash-consistent backup of a database requires the capture of the entire database structure, including datafiles, redo logs, and control files, at a single point in time. If the database is stored in a single NetApp FlexVol® flexible volume, then the process is simple; a Snapshot copy can be created at any time. If a database spans volumes, a consistency group (CG) Snapshot copy must be created. Several options exist for creating CG Snapshot copies, including NetApp SnapCenter® software, the NetApp Snap Creator® framework, NetApp SnapManager® for Oracle (SMO), NetApp SnapDrive® for UNIX, and user-maintained scripts.

This option is generally best when the recovery requirement is the point of the backup itself.

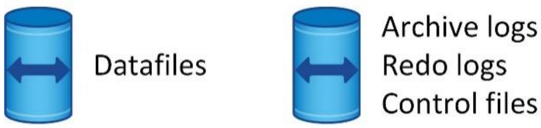
8.3 Oracle Snapshot-Optimized Backup

Snapshot-based backup and recovery become even simpler with Oracle 12c because there is no need to place a database in hot backup mode. The result is an ability to schedule snapshot-based backups directly on a storage system and still preserve the ability to perform complete or point-in-time recovery.

Two sets of data are required to recover from a snapshot-optimized backup:

- A snapshot or set of CG snapshots of the data files
- A synchronized set of archive logs, redo logs, and control files created after the data file snapshots

Figure 1) Two-volume model.



Although the hot backup recovery procedure is more familiar to DBAs, it has for a long time been possible to use snapshots that were not created while the database was in hot backup mode. Extra manual steps were required with Oracle 10g and 11g during recovery to make the database consistent. With Oracle 12c, `sqlplus` and `rman` contain the extra logic to replay archive logs on datafile backups that were not in hot backup mode.

For complete details, see the section "Recovery Using Storage Snapshot Optimization" available in various releases of the Oracle 12c documentation. Also, see Oracle document ID 604683.1 regarding Oracle third-party snapshot support.

8.4 Oracle Hot Backup

Two sets of data are required to recover from a hot backup:

- A snapshot of the datafiles in backup mode
- The archive logs created while the datafiles were in hot backup mode

If complete recovery, including all committed transactions, is required, a third item is required:

- A set of current redo logs

There are a number of ways to drive recovery of a hot backup. Many customers restore snapshots by using the ONTAP CLI and then using Oracle `rman` or `sqlplus` to complete the recovery. This approach is especially common with large production environments in which the probability and frequency of database restores are extremely low and any restore procedure is handled by a skilled DBA. For complete automation, solutions such as NetApp SnapCenter include an Oracle plug-in with both command line and graphical interfaces.

8.5 Recovery Based on Snapshot

NetApp volume-based SnapRestore® (VBSR) technology allows a volume to be almost instantly reverted to an earlier point in time. Because all the data on the volume is reverted, VBSR might not be appropriate for all use cases. For example, if an entire database, including datafiles, redo logs, and archive logs, is stored on a single volume and this volume is restored with VBSR, then data is lost because the newer archive log and redo data are discarded.

VBSR is not required for restore. Many databases can be restored by using file-based single-file SnapRestore (SFSR) or by simply copying files from the Snapshot copy back into the active file system.

VBSR is preferred when a database is very large or when it must be recovered as quickly as possible, and the use of VBSR requires isolation of the datafiles. In an NFS environment, the datafiles of a given database must be stored in dedicated volumes that are uncontaminated by any other type of file. In a SAN environment, datafiles must be stored in dedicated LUNs on dedicated FlexVol volumes. If you use a volume manager (including Oracle Automatic Storage Management [ASM]), the disk group must also be dedicated to datafiles.

Isolating datafiles in this manner allows them to be reverted to an earlier state without damaging other file systems.

8.6 ONTAP Cloud and Third-Party Snapshots

Oracle doc ID 604683.1 explains the requirements for third-party snapshot support and the multiple options available for backup and restore operations.

The third-party vendor must guarantee that the company's snapshots conform to the following requirements:

- Snapshots must integrate with Oracle's recommended restore and recovery operations.
- Snapshots must be database crash consistent at the point of the snapshot.
- Write ordering is preserved for each file in a snapshot.

ONTAP Cloud and NetApp Oracle management products comply with these requirements.

8.7 Oracle Recovery Manager with S3

Although the majority of all backup and recovery scenarios should be possible using the local snapshots, Oracle RMAN has integration with Amazon S3 as a low-cost streaming backup option. RMAN with S3 is not a substitute for local snapshots because performance is much slower during backup and recovery while it copies files. However, it is a good option for long-term backups located on different media in addition to the local snapshots.

9 General Oracle Configuration

The following parameters are generally applicable to all configurations.

9.1 filesystemio_options

The Oracle initialization parameter `filesystemio_options` controls the use of asynchronous and direct I/O. Contrary to common belief, asynchronous and direct I/O are not mutually exclusive. NetApp has observed that this parameter is frequently misconfigured in customer environments, and this misconfiguration is directly responsible for many performance problems.

Asynchronous I/O means that Oracle I/O operations can be parallelized. Before the availability of asynchronous I/O on various OSs, users configured numerous dbwriter processes and changed the server process configuration. With asynchronous I/O, the OS itself performs I/O on behalf of the database software in a highly efficient and parallel manner. This process does not place data at risk, and critical operations, such as Oracle redo logging, are still performed synchronously.

Direct I/O bypasses the OS buffer cache. I/O on a UNIX system ordinarily flows through the OS buffer cache. This approach is useful for applications that do not maintain an internal cache, but Oracle has its own buffer cache in the SGA. In almost all cases, it is better to enable direct I/O and allocate server RAM to the SGA rather than to rely on the OS buffer cache. The Oracle SGA uses the memory more efficiently. In addition, when I/O flows through the OS buffer, it is subject to additional processing, which increases latencies. The increased latencies are especially noticeable with heavy write I/O when low latency is a critical requirement.

The following options are available for `filesystemio_options`:

- **async.** Oracle submits I/O requests to the OS for processing. This process allows Oracle to perform other work rather than waiting for I/O completion and thus increases I/O parallelization.
- **directio.** Oracle performs I/O directly against physical files rather than routing I/O through the host OS cache.
- **none.** Oracle uses synchronous and buffered I/O. In this configuration, the choice between shared and dedicated server processes and the number of dbwriters are more important.
- **setall.** Oracle uses both asynchronous and direct I/O.

In almost all cases, the use of `setall` is optimal, but consider the following issues:

- Some customers have encountered asynchronous I/O problems in the past, especially with previous Red Hat Enterprise Linux 4 (RHEL4) releases. These problems are no longer reported, however, and asynchronous I/O is stable on all current OSs.
- If a database has been using buffered I/O, a switch to direct I/O might also warrant a change in the SGA size. Disabling buffered I/O eliminates the performance benefit that the host OS cache provides for the database. Adding RAM back to the SGA repairs this problem. The net result should be an improvement in I/O performance.
- Although it is almost always better to use RAM for the Oracle SGA than for OS buffer caching, it might be impossible to determine the best value. For example, it might be preferable to use buffered I/O with very small SGA sizes on a database server with many intermittently active Oracle instances. This arrangement allows the flexible use of the remaining free RAM on the OS by all running database instances. This situation is highly unusual, but it has been observed at some customer sites.

  Note: The `filesystemio_options` parameter has no effect in direct NFS (DNFS) and ASM environments. The use of DNFS or ASM automatically results in the use of both asynchronous and direct I/O.

NetApp recommends the following practice:

- Set `filesystemio_options` to `setall`, but be aware that under some circumstances the loss of the host buffer cache might require an increase in the Oracle SGA.

9.2 db_file_multiblock_read_count

The `db_file_multiblock_read_count` parameter controls the maximum number of Oracle database blocks that Oracle reads as a single operation during sequential I/O. This parameter does not, however, affect the number of blocks that Oracle reads during any and all read operations, nor does it affect random I/O. Only sequential I/O is affected.

Oracle recommends that the user leave this parameter unset. Doing so allows the database software to automatically set the optimum value. This approach generally means that this parameter is set to a value that yields an I/O size of 1MB. For example, a 1MB read of 8KB blocks would require 128 blocks to be read, and the default value for this parameter would therefore be 128.

Most database performance problems observed by NetApp at customer sites involve an incorrect setting for this parameter. There were valid reasons to change this value with Oracle versions 8 and 9. As a result, the parameter might be unknowingly present in `init.ora` files because the database was upgraded in place to Oracle 10 and later. A legacy setting of 8 or 16, compared to a default value of 128, significantly damages sequential I/O performance.

NetApp recommends the following:

- The `db_file_multiblock_read_count` parameter should not be present in the `init.ora` file. NetApp has never encountered a situation in which changing this parameter improved performance, but there are many cases in which it caused clear damage to sequential I/O throughput.

10 General NFS Configuration

10.1 NFS Versions

Oracle currently limits NFS support of the mounted file systems to NFS version 3. For this reason, NetApp does not support the use of file systems mounted through NFSv4, NFSv4.1, or pNFS with Oracle databases. This document will be updated if Oracle's support stance changes.

10.2 TCP Slot Tables

TCP slot tables are the NFS equivalent of the host bus adapter (HBA) queue depth. These tables control the number of NFS operations that can be outstanding at any one time. The default value is usually 16, which is far too low for optimum performance. The opposite problem occurs on newer Linux kernels, which can automatically increase the TCP slot table limit to a level that saturates the NFS server with requests.

For optimum performance and to prevent performance problems, adjust the kernel parameters that control the TCP slot tables.

Run the `sysctl -a | grep tcp.*.slot_table` command and observe the following parameters:

```
# sysctl -a | grep tcp.*.slot_table
sunrpc.tcp_max_slot_table_entries = 128
sunrpc.tcp_slot_table_entries = 128
```

All Linux systems should include `sunrpc.tcp_slot_table_entries`, but only some include `sunrpc.tcp_max_slot_table_entries`. They should both be set to 128.

10.3 Direct NFS

Oracle's DNFS client is designed to bypass the host NFS client and perform NFS file operations directly on an NFS server. Enabling it requires only changing the Oracle Disk Manager library. Instructions for this process are provided in the Oracle documentation.

Using DNFS results in a general improvement in I/O performance and decreases the load on the host and the storage system because I/O is performed in the most efficient way possible. In addition, Oracle DNFS provides multipathing and fault tolerance. A failure of one interface results in I/O being retried on the other interface. The overall operation is very similar to FC multipathing.

If you use DNFS, it is critical that you install all patches described in Oracle doc 1495104.1. If a patch cannot be installed, the environment must be evaluated to make sure that the bugs described in that document do not cause problems. In some cases, an inability to install the required patches prevents the use of DNFS.

Caution
• Before using DNFS, verify that the patches described in Oracle doc 1495104.1 are installed. • Starting with Oracle 12c, DNFS includes support for NFSv3, NFSv4, and NFSv4.1. NetApp support policies cover v3 and v4 for all clients, but at the time of writing, NFSv4.1 is not supported for use with Oracle DNFS. • Do not use DNFS with any type of round-robin name resolution, including DNS, DDNS, NIS, or any other method. This restriction includes the DNS load balancing feature available in ONTAP. When an Oracle database using DNFS resolves a host name to an IP address, it must not change on subsequent lookups. Having the IP address change can result in Oracle database crashes and possible data corruption.

10.4 Direct NFS and Host File System Access

Using DNFS can occasionally cause problems for applications or user activities that rely on the visible file systems mounted on the host because the DNFS client accesses the file system out of band from the host OS. The DNFS client can create, delete, and modify files without the knowledge of the OS.

When the mount options for single-instance databases are used, they enable caching of file and directory attributes, which also means that the contents of a directory are cached. Therefore, DNFS can create a file, and there is a short lag before the OS rereads the directory contents and the file becomes visible to the user. This delay is not generally a problem, but, on rare occasions, utilities such as SAP BR*Tools might have issues. If a problem happens, address it by changing the mount options to use the recommendations for Oracle Real Application Cluster (RAC). This change results in the disabling of all host caching.

Change mount options only when DNFS is used and a problem results from a lag in file visibility. If DNFS is not in use, using Oracle RAC mount options on a single-instance database results in degraded performance.

10.5 ADR and NFS

Some customers have reported performance problems resulting from an excessive amount of I/O on data in the `ADR` location. The problem does not generally occur until a lot of performance data has accumulated. The reason for the excessive I/O is unknown, but this problem appears to be a result of Oracle processes repeatedly scanning the target directory for changes.

Removal of the `noac` and/or `actimeo=0` mount options allows host OS caching to occur and reduces storage I/O levels.

NetApp recommends that you do not place `ADR` data on a file system with `noac` or `actimeo=0` because performance problems are likely. Separate `ADR` data into a different mount point if necessary.

11 Oracle on ONTAP Cloud Solution Designs

The following sections outline some of the primary use cases for Oracle on ONTAP Cloud and how they affect the overall ONTAP Cloud configuration. These are only examples; there is no single best option. The best approach is to first define the business requirements in terms of performance, data protection, RPO, RTO, and availability. The technical architecture should then be designed based on those defined requirements.

The following sections illustrate some of the options available to tailor an ONTAP Cloud environment to a particular Oracle database requirement.

11.1 Production Databases

Operating a production database on ONTAP Cloud generally requires the highest possible levels of data protection and availability. This fact means zero tolerance for data loss (RPO = 0) and minimal time required to recover after an outage. High availability is also a requirement, and performance demands would probably be much higher than for a database used for simple development purposes.

Data Durability and Availability

A production database is generally run with ONTAP Cloud in normal write mode to avoid the risk of data loss if the ONTAP Cloud VM crashes. In an AWS environment, ONTAP Cloud is also likely to be used in HA mode with the two nodes located on different availability zones.

Oracle DataGuard can be configured in multiple modes:

- Maximum protection mode guarantees that all commitments are recorded by both the primary and standby database before proceeding.
- Maximum availability mode normally operates with an RPO of 0, but, if the replication link is lost for a configurable amount of time, the primary database resumes operations until the link is repaired.
- Maximum performance mode is an asynchronous replication option. The standby database usually lags the primary database by a few minutes.

In some cases, a standby database might be preferable in AWS in place of ONTAP Cloud in HA mode. For example, a standby database managed by Oracle DataGuard has more flexible replication options and can be switched between synchronous and asynchronous modes. It can even be used to replicate between different cloud providers. The overall value of ONTAP Cloud is unaffected by the choice of replication technology.

Performance

Making sure that an ONTAP Cloud solution meets the requirements of a production database requires careful planning and testing. Although the hyperscaler volumes used by ONTAP Cloud are based on flash, they have more limited IOPS capabilities with higher latencies than are seen in on-premises all-flash arrays. For most database use cases, the IOPS capabilities of an ONTAP Cloud instance approximate the IOPS capabilities of the underlying hyperscaler volumes, and those values can be used for general planning. However, testing is still required, and you must follow all best practices carefully, especially the requirements for reallocation scans.

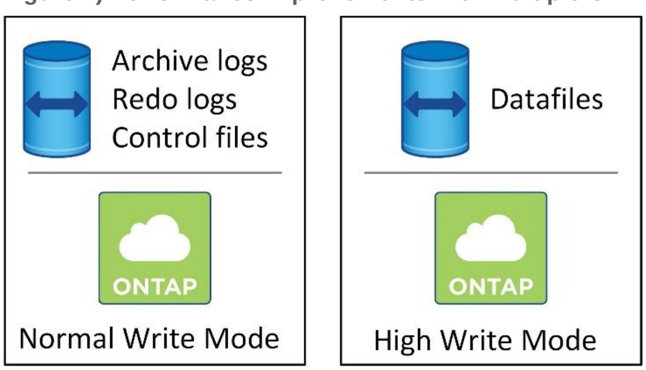
In addition, in an AWS environment, the EBS volumes should be at least 1TB in size to avoid the complications of EBS I/O bursting. Erratic performance can result depending on the available I/O burst credits when a particular query is executed.

Other measures to improve performance include the use of larger SGA sizes, Oracle Flash Cache, and local temporary files, as explained in section 5, “Oracle Database Performance Optimization.”

Multiple Instances

One option for improving performance is the use of multiple ONTAP Cloud instances, with one of them operating in high-speed write mode.

Figure 2) Performance improvements with multiple ONTAP Cloud instances.



When used properly, multiple instances do not affect recoverability. The approach is as follows:

- The datafiles are hosted on an ONTAP Cloud instance operating in high-speed write mode.
- The control files, archive logs, and redo logs are located on a second controller running with normal write speed.

There are two basic failure scenarios, not including disasters such as a widespread loss of EBS volumes:

- If the ONTAP Cloud instance hosting the log data crashes, no data is lost because all writes are written to persistent storage before being acknowledged.
- If the ONTAP Cloud instance hosting the datafiles crashes, up to 10 seconds of data could be lost.

The first scenario might cause a database crash, but no data is lost, and the Oracle database should start normally. The second scenario results in loss of the datafiles, but the log volume still contains synchronized copies of the redo log, archive log, and control files data. This data can be quickly replayed on the datafiles with the sqlplus command `recover automatic`.

One important requirement for this approach is that you must confine the datafiles to a single aggregate to make sure that they are all synchronized with one another in the event of an ONTAP Cloud crashes.

Backup and Recovery

Many options exist for backup and recovery, including:

- A simple Snapshot schedule on ONTAP Cloud
- The NetApp SnapCenter Oracle plug-in
- Oracle RMAN, including Oracle RMAN backups to an Amazon S3 destination

The majority of all backup and recovery scenarios should be possible using the local Snapshot copies. This option offers convenient, simple, and near-instantaneous restoration of a database, irrespective of size. It should not, however, be the only option for restoration because the blocks containing Snapshot data share volumes with the active file system data. If the EBS volumes are lost, the backups are also lost. This situation is a borderline disaster, but it is a possibility that you should consider.

See section 8, “Oracle Data Protection,” for additional details.

11.2 Development and Testing

Possibly the most important use case for ONTAP Cloud is for development and testing work because you can scale up and down as required. There are no procurement delays for new equipment, nor is money wasted on unused infrastructure during quieter periods.

Data Durability and Availability

In contrast to production environments, development databases do not generally require high availability or even backups. Single-instance ONTAP Cloud offers both cost savings and improved performance and would be a natural first choice for development work. Likewise, running in high-speed write mode would be a better option. Some data could be lost in the unlikely event of an ONTAP Cloud instance crashes, but, as long as the database exists in a single aggregate, it restarts as usual when you restart the instance. If particular tests required normal write speed mode to guarantee data integrity, the mode could be changed temporarily. This process is disruptive for a single ONTAP Cloud instance, but you can do it easily with Cloud Manager as needed.

Data Seeding with ONTAP

Testing a database on ONTAP Cloud typically requires copying a database from an external source. You can perform this process with typical copy methods such as FTP, but it takes time to copy the data from the source and then restore it onto the ONTAP Cloud instance. A better option would be to leverage NetApp SnapMirror® replication.

Seeding with SnapMirror requires a source storage system running ONTAP using either NFS, iSCSI, or FC. The simplest option is to use a source database on which all datafiles are confined to a single FlexVol volume and then replicate to ONTAP Cloud with SnapMirror. This process provides the baseline database. In addition, a FlexVol volume containing archive logs, at least one redo log, and one control file is required to bring the database replica to a fully consistent state. These two datasets can be used in one of two ways. First, you can break the mirror during testing and then resynchronize it when testing is complete. Second, you can clone the replicated Snapshot copies in ONTAP Cloud using a variety of options.

SnapCenter is the most comprehensive option because it provides complete GUI-based automation of the entire database. SnapCenter can not only automate the entire cloning process, but also protect the source database and manage replication. Simpler dev/test requirements can be met with OS-based scripts.

Any protocol can be used, including FC. You can replicate FC LUN data and any other type of data to ONTAP Cloud. The LUN is mapped to the cloud hosts as an iSCSI LUN and otherwise operates as usual.

Data Seeding with RMAN

If the source database is not on ONTAP but an ONTAP system is available, then you can use Oracle RMAN to maintain an up-to-date image copy of the database, which you can then replicate to ONTAP Cloud. One benefit to this approach is that the image is usable by itself without any requirement for log replay.

The setup process starts when you restore a backup of the source database datafiles on ONTAP Cloud. After you restore the datafiles, you can establish a method for log shipping. The goal is to create a constant feed of archive logs generated by the primary database and then replay them on the restored database to keep them both close to the same state. You can do this replaying with rsync or FTP or as a standard Oracle standby database. You can also configure a simple log-shipping migration out of band with respect to the original database, which makes it particularly useful for mission-critical databases. No configuration changes are required for the source database, and the restoration and initial configuration of the migration environment have no effect on production operations.

You can then use the database in ONTAP Cloud with a variety of options. For example, you can create a baseline Snapshot copy as a backout path. You can then open the database for testing, and when testing is complete, you can restore the database to the base state using the previously created Snapshot copy. Replication from the source can then be resumed. For more dynamic requirements, you can clone a standby database using SnapCenter.

Backup and Recovery

Backup and recovery of a test database are not generally required, but Snapshot copies of a database are simple to create. Developers can easily drive the process themselves using the ONTAP CLI to conveniently and flexibly create backups. For example, before starting a series of repeated tests, a developer can shut down a database and create baseline Snapshot copies using a command as simple as `snapshot create -volume mydatabase -snapshot baseline`. The developer can then perform the required tests, shut down the database, and type `snapshot restore -volume mydatabase -snapshot baseline` to reset the database to its original state.

Other options exist to manage backup and recovery of a test database, but in a testing environment it is often easier to give the developers a login to the ONTAP CLI and issue commands directly.

11.3 Disaster Recovery with SnapMirror

Disaster recovery of a production database using SnapMirror requires ONTAP for the source database. The simplest option is the use of a source database for which all datafiles are confined to a single FlexVol volume and then replicated to ONTAP Cloud with SnapMirror. This process provides the baseline database. In addition, a FlexVol volume containing archive logs, at least one redo log, and one control file is required to bring the database replica to a fully consistent state.

Any protocol can be used, including FC. You can replicate FC LUN data to ONTAP Cloud just as with any other type of data. The LUN is mapped to the cloud hosts as an iSCSI LUN and otherwise operates as usual.

In the event of a disaster, the basic procedure is nothing more than breaking the mirrors and replaying the available logs. The process is simple enough to perform manually, but if you are using multiple databases and applications, OnCommand® Workflow Automation might be a better option.

Data Durability and Availability

A production database, including its disaster recovery copies, is generally run in normal write mode to avoid the risk of data loss if the ONTAP Cloud VM crashes. In an AWS environment, ONTAP Cloud is also likely to be used in HA mode with the two nodes located on different availability zones.

Performance

Making sure that an ONTAP Cloud solution meets the requirements of a production database requires careful planning and testing. This requirement includes creating a disaster recovery database that might be activated in the event of a disaster. As explained in section 11.1, “Production Databases,” the IOPS capabilities of an ONTAP Cloud instance approximate the IOPS capabilities of the underlying hyperscaler volumes, and those values can be used for general planning. Testing is still required, and all best practices must be followed carefully, especially the requirements for reallocation scans.

Cost Savings

You can activate ONTAP Cloud on demand for certain tasks. For example, some databases might require a disaster recovery RPO of only 12 hours. Therefore, the ONTAP Cloud instance would need to run briefly only twice per day as it receives SnapMirror updates from the source.

Multiple Instances

Using multiple ONTAP Cloud instances with slightly different configurations might help reduce costs and improve performance with a disaster recovery database.

For example, a database using ONTAP Cloud can be split across two ONTAP Cloud instances. The datafiles are hosted by a powerful, high-end ONTAP Cloud environment, while the archive logs are hosted by a smaller system. The ONTAP Cloud instance hosting datafiles needs to run for only a short time once per day to receive the latest copy of the datafiles, while the smaller archive log ONTAP Cloud instance runs constantly with an update interval of 15 minutes or less.

Furthermore, the ONTAP Cloud instance hosting datafiles can be run in high-speed write mode as a single instance, while the ONTAP Cloud instance hosting log data runs in a normal write speed mode in an HA configuration. If the ONTAP Cloud instance hosting datafiles unexpectedly crashes, the log data would still be current and able to provide an RPO = 0.

This approach can deliver substantial cost savings while also making sure that the ONTAP Cloud environment has enough performance capacity to support the active DR workload.

Backup and Recovery

One of the primary benefits of standardizing on ONTAP is the ability to leverage the same backup and recovery procedures across the enterprise, including in the cloud.

Many options exist for backup and recovery, including:

- A simple Snapshot schedule on ONTAP Cloud
- The NetApp SnapCenter Oracle plug-in
- Oracle RMAN, including Oracle RMAN backups to an Amazon S3 destination

See section 8, “Oracle Data Protection,” for additional details.

11.4 Disaster Recovery with Oracle DataGuard

Oracle DataGuard can also be used to maintain a disaster recovery copy in the cloud. DataGuard replication is protocol agnostic, meaning that the protocols on the source and destination do not need to match. Furthermore, Oracle RAC databases can be replicated to a single-instance database.

Availability

A production database, including its disaster recovery copies, is generally run in normal write mode to avoid the risk of data loss if the ONTAP Cloud VM crashes. In an AWS environment, ONTAP Cloud is also likely to be used in HA mode, with the two nodes located on different availability zones.

Performance

As explained in section 11.1, "Production Databases," the IOPS capabilities of an ONTAP Cloud instance approximate the IOPS capabilities of the underlying hyperscaler volumes, and those values can be used for general planning. However, testing is still required, and you must follow all best practices carefully, especially the requirements for reallocation scans.

Cost Savings

You can activate ONTAP Cloud on demand for certain tasks. For example, some databases might require a disaster recovery RPO on the order of only 12 hours. You can configure DataGuard in maximum performance mode.

Multiple Instances

Using multiple ONTAP Cloud instances with slightly different configurations might help reduce costs and improve performance with a disaster recovery database.

You can run the ONTAP Cloud instance hosting datafiles in high-speed write mode as a single instance, while the ONTAP Cloud instance hosting log data can be run in normal write speed mode in an HA configuration. If the ONTAP Cloud instance hosting datafiles unexpectedly crashes, the log data is still current and able to provide an RPO = 0.

This approach can deliver substantial cost savings while also making sure that the ONTAP Cloud environment has enough performance capacity to support the active DR workload.

Backup and Recovery

One of the primary benefits of standardizing on ONTAP is the ability to leverage the same backup and recovery procedures across the enterprise, including the cloud.

Many options exist for backup and recovery, including:

- A simple Snapshot schedule on ONTAP Cloud
- NetApp SnapCenter Oracle plug-in
- Oracle RMAN, including Oracle RMAN backups to an Amazon S3 destination

See section 8, "Oracle Data Protection," for additional details.

12 Performance Testing

This section presents examples of Oracle database performance on ONTAP Cloud and also shows how some configuration operations affect performance. This is not intended to be an exhaustive analysis. There are many ways to configure an Oracle database and workloads can vary dramatically.

12.1 Configuration

Performance testing was done using the SLOB2 (Silly Little Oracle Benchmark) utility, available here. SLOB drives storage I/O from an actual Oracle database and allows easy comparison of different configurations. IOPS and latency numbers are then extracted directly from the Oracle database to see what performance looks like from the database point of view.

ONTAP Cloud was configured by using an AWS c4.8xlarge instance. Storage was based six 4TB EBS volumes, which means each EBS volume was capable of 10K IOPS, the maximum currently available from an EBS gp2 volume.

The database used was Oracle 12cR2, version 12.2.0.1 running on a c5.4xlarge instance with a database size of approximately 1TB.

A series of SLOB2 tests using a 75%/25% read/write ratio were conducted with increasing load.

12.2 Metrics

Read and write IOPS were based on the Oracle AWR report Load Average. The SLOB2 database was configured to force random I/O, and all read and write operations were presumed to be single block. Read latency was based on the db file sequential read statistic, and write latency was based on the db file parallel write statistic. Many other options are available, but these two are used most often by customers. Redo logging is reported as total throughput, but latency is not included because the latency for a redo log write varies greatly depending on the size of the write operation. An absolute latency number for a redo logging operation is not meaningful by itself.

12.3 Protocols

The first two graphs show iSCSI and NFS performance using ONTAP Cloud configured in HA mode. You must be careful when comparing the two graphs because the axes scales differ.

Figure 3) iSCSI performance.

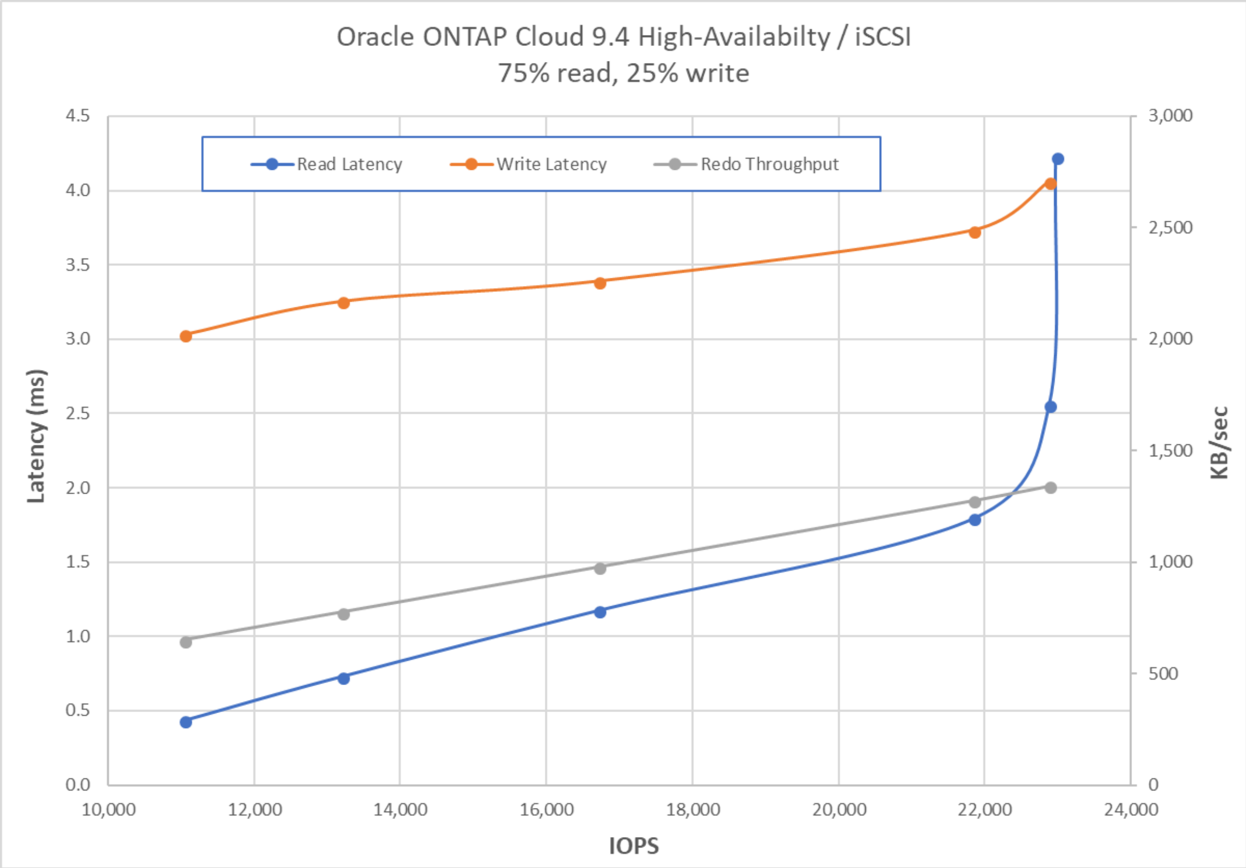
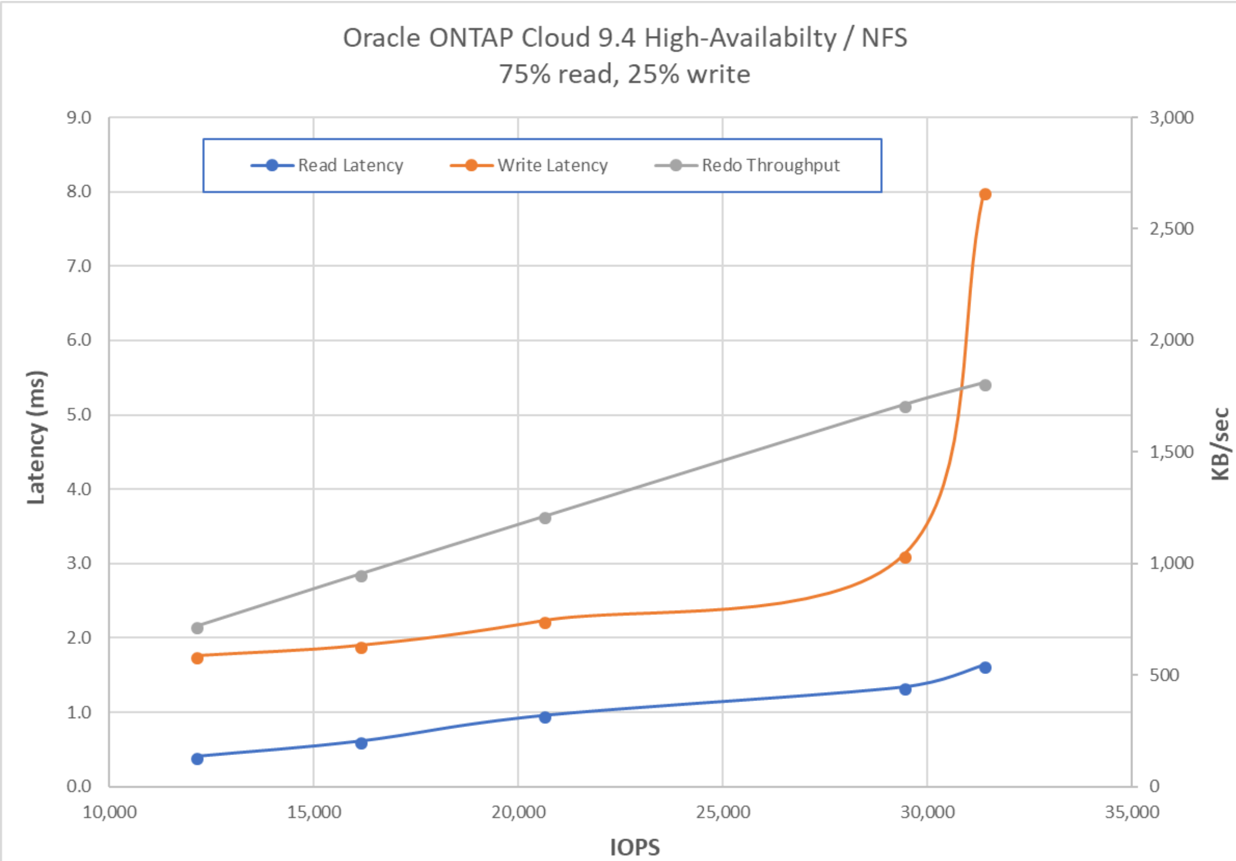


Figure 4) NFS performance.



Overall performance is comparable. Both protocols showed read latency around 1ms for most of the performance curve, and the corresponding redo logging throughput is similar. The write latency was somewhat higher with iSCSI, but this is based on the datafile writes. This type of write is generally a background operation, which means higher or lower latency is not usually a problem as long as the database is able to free buffers fast enough to accommodate new reads.

The only notable difference was the saturation point. The iSCSI test reaches a saturation point at approximately 22K IOPS, whereas the NFS test showed saturation around 29K IOPS. This may seem like a significant difference, but the latencies climb rapidly as the saturation point nears. Most databases would not be operating at this level, but NFS delivers the highest IOPS at the lowest latency if performance is extremely critical.

The choice of protocol should generally be based on management needs. For example, if NFS is used on a physical AFF system on the primary site and the data is replicated to ONTAP Cloud, then NFS is the natural choice. Likewise, if current customer practices use the ASM standard, there is no reason not to choose ASM on iSCSI within ONTAP Cloud.

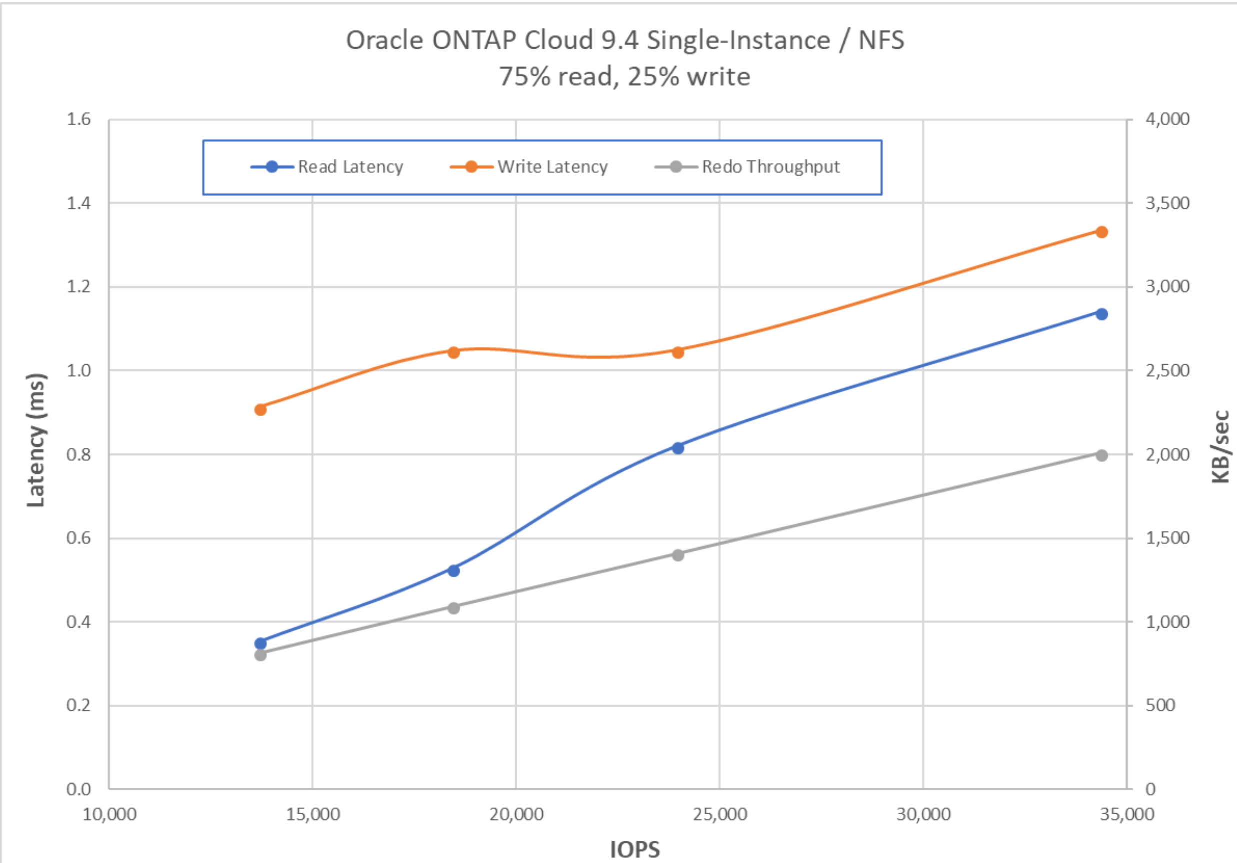
12.4 High Availability and Single-Instance

Figure 5 shows a repeat of the preceding NFS test, but with single-instance ONTAP Cloud configuration rather than a multi-instance HA version. Normal write speed mode was used, which removes the synchronous replication overhead required in an HA environment.

The results were as expected:

- The saturation point occurred around 35K IOPS as opposed to 29K IOPS, an improvement of approximately 20%.
- We observed redo logging throughput increases consistent with the additional overall IOPS being serviced.
- Read latency is comparable because all reads in an HA environment are serviced locally. However, HA mode does show slightly higher latency due to the general overhead of maintaining an HA environment.
- Write latency is much better in a single-instance configuration because the additional synchronous replication to the alternate set of EBS volumes is not required.

Figure 5) Single instance ONTAP Cloud.



12.5 Normal and High Speed Write Mode

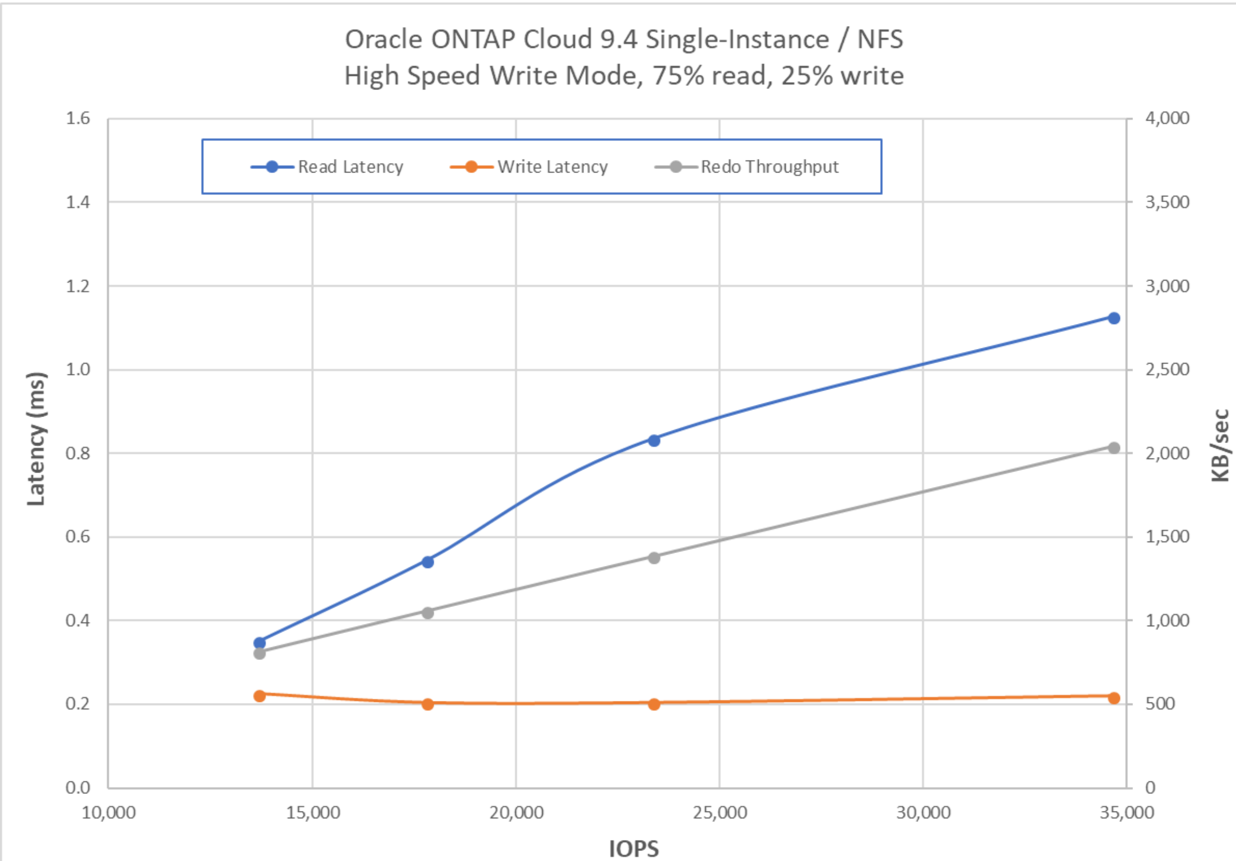
Figure 6 shows a repeat of the preceding NFS single-instance test, but with high-speed write mode enabled. This means that writes are cached and acknowledged before they are written to persistent media. A crash of the ONTAP Cloud instance would result in the loss of some data, but not overall data consistency. The state of the data would effectively jump backward in time to the most recent ONTAP checkpoint.

As discussed in section 6.1, this option is most appropriate for two scenarios. Basic development and testing generally requires a functional database. The loss of a few changes would be unlikely to cause problems. Second, a production environment could be configured with the critical log files on an ONTAP Cloud instance using normal write speed mode, but with the datafiles on high write-speed mode. A crash of the ONTAP Cloud instance hosting the datafiles would require additional log replay to make it usable, but this should be measured in seconds.

The results of the testing are as expected:

- The saturation point still occurs around 35K IOPS as ONTAP Cloud reaches the maximum IOPS capable of being serviced by the underlying volumes.
- Read latency is nearly identical.
- Write latency is much better in high write-speed mode. For example, at the 25K IOPS point the write latency is around 0.6ms with normal write-speed mode, but 0.2ms with high write-speed mode.

Figure 6) High speed write mode enabled.



Conclusion

ONTAP Cloud is an ideal solution for delivering complicated data management needs to the public cloud, including a full production database, disaster recovery, or simple cloning environment. ONTAP Cloud also interconnects with other ONTAP endpoints to deliver a true hybrid cloud capability that includes seamlessly moving and replicating data among different cloud providers and on-premises ONTAP systems.

The right database solution depends on both the technical details of the implementation and the business requirements driving the project. NetApp and partner professional services experts are available for assistance in complex projects.

Where to Find Additional Information

To learn more about the information described in this document, refer to the following documents and/or websites:

- TR-3633: Oracle on NetApp ONTAP
  http://www.netapp.com/us/media/tr-3633.pdf
- TR-4591: Database Data Protection
  http://www.netapp.com/us/media/tr-4591.pdf
- TR-4592: Oracle on MetroCluster
  http://www.netapp.com/us/media/tr-4592.pdf
- TR-4534: Migration of Oracle Databases to NetApp Storage Systems
  http://www.netapp.com/us/media/tr-4534.pdf

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.

TR-4676-0418

NetApp