



Technical Report

Oracle Databases on ONTAP Select

Tudor Pascu and Joe Scott, NetApp
June 2018 | TR-4690

Important

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

TABLE OF CONTENTS

1	Introduction	4
2	ONTAP Platforms	4
2.1	ONTAP with AFF and FAS Controllers, Including NetApp Private Storage	4
2.2	ONTAP Cloud	5
2.3	ONTAP Select	5
3	Planning and Licensing	5
3.1	ONTAP Select Licensing	6
3.2	Hypervisor Licensing	7
3.3	Oracle Licensing	7
4	Oracle Database Performance Optimization	7
4.1	Oracle Buffer Cache	7
4.2	Oracle In-Memory	7
5	ONTAP Select Deployment Options	7
5.1	High Availability	7
5.2	VM Type	8
5.3	Drive Count	8
5.4	Storage Virtual Machines	8
5.5	Compression, Compaction, and Deduplication	8
5.6	Thin Provisioning	9
5.7	Space Management	9
6	Oracle Data Protection	9
6.1	Snapshot-Based Backups	9
6.2	Consistency Group Backup	9
6.3	Oracle Snapshot-Optimized Backup	10
6.4	Oracle Hot Backup	10
6.5	Recovery Based on Snapshot	10
6.6	ONTAP Select and Third-Party Snapshots	11
6.7	Oracle Recovery Manager with S3	11
7	General Oracle Configuration	11
7.1	filesystemio_options	11
7.2	db_file_multiblock_read_count	12
8	General NFS Configuration	13

8.1	NFS Versions.....	13
8.2	TCP Slot Tables.....	13
8.3	Direct NFS	13
8.4	Direct NFS and Host File System Access.....	14
8.5	ADR and NFS.....	14
9	Oracle on ONTAP Select Solution Designs	14
9.1	Production Databases.....	14
9.2	Development and Testing.....	15
9.3	Disaster Recovery with SnapMirror.....	16
9.4	Disaster Recovery with Oracle DataGuard	17
10	Performance Testing.....	17
10.1	Configuration.....	17
10.2	Load Generation	18
10.3	Protocols.....	18
10.4	Metrics	18
10.5	Latency Cutoff Tests	19
10.6	Protocol Performance Tests.....	19
	Conclusion	20
	Where to Find Additional Information	21

1 Introduction

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

ONTAP Select can run on existing virtualized infrastructure, allowing internally or externally facing service providers to provide each tenant with their own data management environment. New tenants can be quickly deployed by adding additional ONTAP Select instances.

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 Select and the virtualized environments that support ONTAP Select 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](#)
- [TR-4613: ONTAP Select on KVM](#)
- [TR-4517: ONTAP Select Product Architecture and Best Practices](#)

2 ONTAP Platforms

NetApp 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
- NetApp 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, you can build data protection and disaster recovery strategies 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 database as a service (DBaaS) that involves 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 Web Services (AWS), Microsoft Azure, and IBM Cloud. This capability is delivered by AFF and FAS systems, including NetApp 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 your 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 you can resolve by deploying ONTAP Select on solid-state drives (SSDs).

3 Planning and Licensing

The most important aspect of NetApp ONTAP Select planning is that you must define the requirements of the database in terms of performance, capacity, and availability. The physical hosts on which ONTAP Select are installed play the most important part in determining whether the performance of the ONTAP Select instance can meet your database performance requirements.

Considerations include the following:

- How many random IOPS are required?
- What type of storage is used (SSD or HDD)?
- Is the storage local (direct-attached storage [DAS]) or shared?
- For local storage, how many drives are backing the datastore or storage pool where ONTAP Select is installed?
- What type (10Gb, 1Gb) and how many network ports are available to the ONTAP Select VM?
- What are the license requirements?
- What are the backup and recovery requirements?

3.1 ONTAP Select Licensing

ONTAP Select is licensed by capacity. NetApp offers two ONTAP Select license types: Premium and Standard. The Premium license is required for SSD storage. The Premium license allows for creating two types of ONTAP Select VMs: a small instance (four cores with 16GB of RAM) and a medium instance (eight cores with 64GB of RAM).

The ONTAP Select capacity license defines the amount of storage that the ONTAP Select node can use to store end user data. A number of important clarifications are in order:

- When ONTAP Select is consuming storage from a datastore or storage pool, the storage space has already gone through numerous virtualization layers and logical representations. The storage overhead required by these layers is not visible to ONTAP Select, and it does not count against the capacity license. When you use ONTAP Select software RAID, the physical disks (SSDs and/or HDDs) are managed by ONTAP Select. Depending on the level of protection desired, a certain amount of storage is used for local data protection. This storage overhead, though managed by ONTAP Select, does not count against the capacity license. For the purpose of this definition, the storage space available to store end-user data is defined as raw space from the ONTAP Select point of view.
- ONTAP Select consumes space in addition to the storage licensed for end user data. This space is used to store ONTAP Select internal data structures and does not require a capacity license. This space is consumed during initial ONTAP Select cluster configuration, and it does not change through the life of the ONTAP Select cluster. The actual amount of storage varies depending on the ONTAP Select cluster size. This size does not exceed 266.5GB.
- For ONTAP Select high-availability (HA) pairs, each node maintains a synchronous copy of its peer's data. This secondary data copy contains user data and thus counts against the node level capacity license. This HA model is not configurable, and it is a fundamental aspect of how ONTAP Select provides HA functionality.
- You can use the ONTAP Select "raw" space (see the first point) to create the traditional ONTAP data management structures (aggregates and volumes). The level of utilization of these volumes as well as the storage efficiency ratios dictate the total amount of end-user data stored within the ONTAP Select node. Depending on your configuration, it is possible to store more data than the ONTAP Select capacity license. In other words, no action is required if or when the effective capacity of the ONTAP Select node exceeds the capacity license of that node.
- You can attach additional "raw" space (see the first point) to an ONTAP Select node nondisruptively. If the new "total raw" storage does not exceed the existing licensed capacity, no further action is required. If the new "total raw" capacity exceeds the existing node level capacity license, then that node is in violation of its license. A new license should be applied within 30 days to avoid license enforcement actions. Capacity licenses can be increased in 1TB increments and can be applied nondisruptively.

The ONTAP Select capacity license is node-locked, and each ONTAP Select node requires a separate license. At the time of writing this document, there was no option for an ONTAP Select cluster-level or site-level capacity license.

3.2 Hypervisor Licensing

ONTAP Select support VMware ESX and Kernel-Based Virtual Machine (KVM) hypervisors (RedHat and CentOS). See the NetApp Interoperability Matrix (IMT) for a list of supported hypervisor versions. Although licensing costs are one aspect of choosing a hypervisor, other aspects, such as operational experience with that hypervisor, should also be considered.

3.3 Oracle Licensing

You must also consider Oracle licensing costs when sizing a solution. The use of ONTAP Select does not directly affect Oracle licensing, but, because ONTAP Select runs in a virtualized hypervisor environment, customers might also wish to run their databases under virtualization. The Oracle license policy for virtualized instances depends on multiple factors. Contact your Oracle sales representative, and also review third-party documents such as VMware's [review](#) of their policies.

4 Oracle Database Performance Optimization

4.1 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.

4.2 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 ONTAP Select Deployment Options

NetApp ONTAP Select is more than just a storage system. It is a virtualized, consumption-based, software-defined storage system.

The following sections describe some of the important options for planning an Oracle on ONTAP Select 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 Select and ONTAP Deploy can be found on the NetApp Support site.

5.1 High Availability

ONTAP Select can be configured as a single-node cluster or as a multi-node cluster (2, 4, 6, and 8 node clusters are current possible configurations). In multi-node configurations, nodes form HA pairs within the cluster. Each node in an HA pair maintains a synchronous copy of its data on its HA partner. This configuration provides the data availability required by modern enterprises using commodity hardware. Failover between nodes is seamless, and applications remain online with zero downtime through a failover event.

5.2 VM Type

ONTAP Select supports two VM sizes: small and medium. The only difference between the VMs is the amount of memory and the number of CPU cores that are reserved. These parameters cannot be changed after deployment. A medium VM instance requires a Premium license.

5.3 Drive Count

An ONTAP Select Premium instance requires a minimum of four SSD drives. A larger number of local SSDs is required to achieve peak performance. The performance results shown in section 10, “Performance Testing” were generated using a medium ONTAP Select instance with a Premium license and 24 SSDs.

5.4 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 with 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 on FAS/AFF can be considered an aggregator of workloads, deployed in a central location at the heart of your datacenter. Multiple SVMs are the norm when using ONTAP on FAS or AFF. ONTAP Select supports multiple SVMs as well. However, ONTAP Select is a software-defined storage solution. Deploying ONTAP Select can be done programmatically in a manner of minutes for each instance. This feature is more congruent with an internal cloud architecture in which a virtual instance of ONTAP Select can be aligned with a specific use case or application and managed together as a converged or a hyper converged solution. For these types of deployments, a single SVM per ONTAP Select cluster can suffice.

There is, however, no technical reason multiple SVMs cannot be used. The ONTAP Select maximum cluster size is currently eight nodes, and such a cluster can contain a mix of ONTAP Select VM types and storage configurations. Separating workloads and authentication domains are examples of situations where multiple SVMs should be used.

5.5 Compression, Compaction, and Deduplication

You can use data efficiency features such as compression, compaction, and deduplication for a database on ONTAP Select. An ONTAP Select medium VM with local SSDs (a DAS configuration) is conceptually close to an AFF array because both can rely on ultra-low-latency storage that can deliver extreme IOPS levels. This allows the use of maximum efficiency settings with minimal concern for the performance effect on the underlying disks.

In contrast, an ONTAP Select small VM with SAS/SATA drives is conceptually closer to a FAS array because the underlying drives used for data storage have more limitations. Operations such as post-processing deduplication that rely on disk reads might affect performance.

For storage efficiency best practices depending on the ONTAP Select configuration, see the technical reports [TR-4517](#) and [TR-4613](#).

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 Select architecture.

5.6 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 Select environment.

5.7 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.

6 Oracle Data Protection

6.1 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

6.2 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.

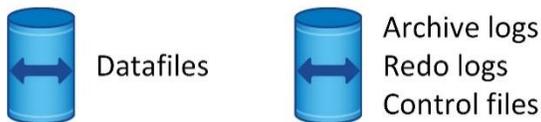
6.3 Oracle Snapshot-Optimized Backup

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 database administrators (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.

6.4 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.

6.5 Recovery Based on Snapshot

NetApp volume-based SnapRestore® (VBSR) technology enables 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.

6.6 ONTAP Select 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 Select and NetApp Oracle management products comply with these requirements.

6.7 Oracle Recovery Manager with Amazon S3

Although the majority of all backup and recovery scenarios should be possible by 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.

7 General Oracle Configuration

The following parameters generally apply to all configurations.

7.1 `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.

7.2 `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.

8 General NFS Configuration

8.1 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.

8.2 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 optimal 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 optimal 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.

8.3 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 on the storage system because I/O is performed in the most efficient way possible. In addition, Oracle DNFS provides multipathing and fault tolerance. For example, two 10Gb interfaces can be bound together to offer 20Gb of bandwidth. 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.

8.4 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.

8.5 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.

9 Oracle on ONTAP Select Solution Designs

The following sections outline some of the primary use cases for Oracle on ONTAP. The best approach is to first define the business requirements in terms of performance, data protection, recovery point objective (RPO), recovery time objective (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 Select environment to a particular Oracle database requirement.

9.1 Production Databases

Operating a production database on ONTAP Select 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 Availability

A production database generally requires ONTAP Select in HA mode to protect against VM failure.

Performance

ONTAP Select Premium with local SSD drives delivers the maximum performance available for production databases.

Backup and Recovery

Many options exist for backup and recovery, including:

- A simple snapshot schedule on ONTAP Select
- 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.

See section 6, “Oracle Data Protection,” for additional details.

9.2 Development and Testing

ONTAP Select, including single-node deployments and small instances, is a good fit for development and testing work because it can be deployed on demand on almost any hardware in your datacenter. This includes servers with local storage or shared storage (ONTAP Select vNAS). There are no procurement delays for new equipment, nor is money wasted on unused infrastructure during quieter periods.

Data Availability

In contrast to production environments, development databases do not generally require high availability or even backups. A single instance of ONTAP Select offers both cost savings and lower performance demands on the virtualization infrastructure.

Data Seeding with ONTAP

Testing a database on ONTAP Select 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 on the ONTAP Select instance. A better option is 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 Select 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 Select 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 Select. The LUN is mapped to the 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 Select. 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 Select. 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 Select 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.

9.3 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 Select 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 Select just as with any other type of data. The LUN is mapped to the 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 Availability

A production database typically requires ONTAP Select in HA mode to protect against VM failure.

Performance

ONTAP Select Premium with local SSD drives delivers the maximum performance available for production databases. Although a DR copy of a database might not create significant performance demands outside of a disaster situation, the solution must still be sized to support the DR workload. Furthermore, many customers leverage the DR infrastructure for day-to-day development and testing work, which creates additional performance demands on storage.

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 the following:

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

See section 6, “Oracle Data Protection,” for additional details.

9.4 Disaster Recovery with Oracle DataGuard

Oracle DataGuard can also be used to maintain a disaster recovery copy on ONTAP Select. 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.

Data Availability

A production database generally requires ONTAP Select in HA mode to protect against VM failure.

Performance

ONTAP Select Premium with local SSD drives delivers the maximum performance for production databases. The write workload of a DataGuard standby database can be substantial, so the ONTAP Select instance hosting the copy should have performance capabilities equivalent to the source database.

Furthermore, many customers leverage DR infrastructure for day-to-day development and testing work, which creates additional performance demands on storage.

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 the following:

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

See section 6, “Oracle Data Protection,” for additional details.

10 Performance Testing

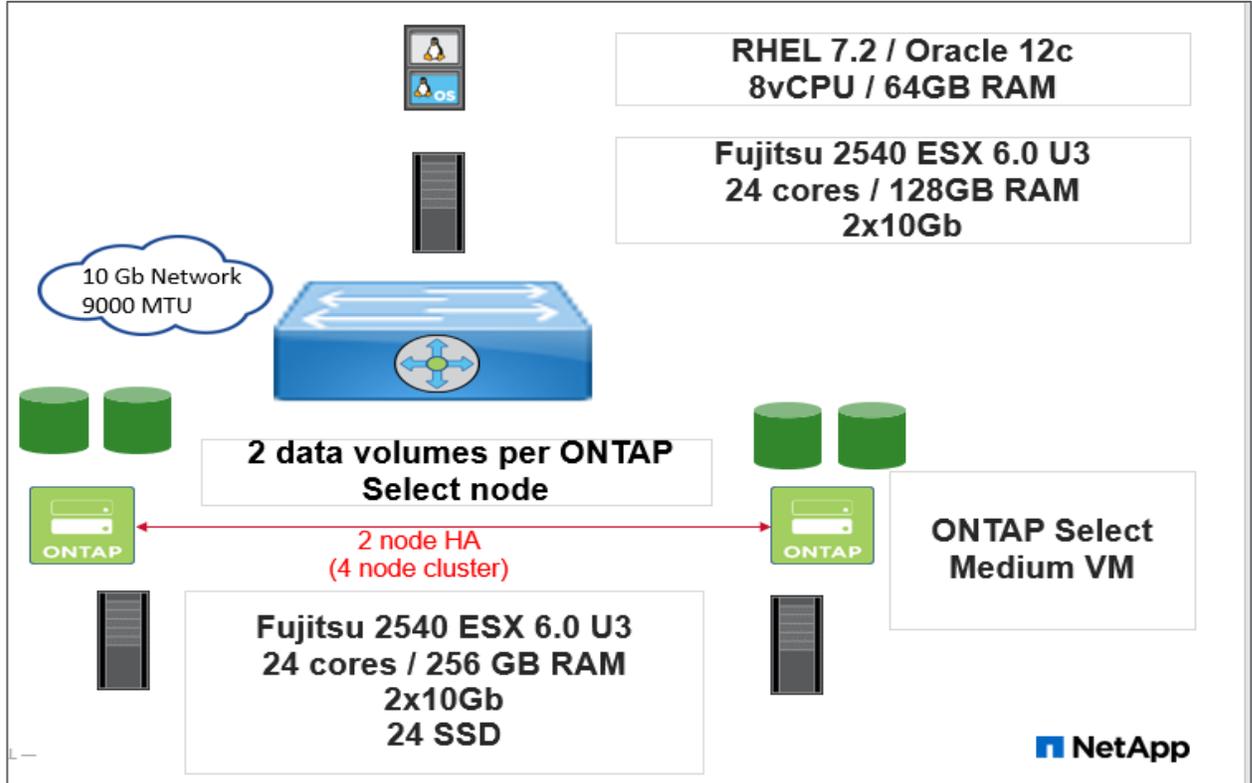
Confirming that an ONTAP Select solution meets the requirements of a production database requires careful planning and testing. White box or commodity servers differ in performance even if the number of disk drives and number of CPU cores appear to be the same. The choice of RAID controllers (if present), hypervisors, drive characteristics, and myriad other variables result in a much larger number of price and performance points when compared with an engineered storage system such as FAS or AFF.

10.1 Configuration

The test environment consisted of a four-node ONTAP Select cluster and used a medium ONTAP Select VM with a Premium license. Each node had exclusive access to a local VMware Virtual Machine File

System (VMFS) datastore composed of 24 SSDs. Only two nodes out of the four-node cluster were used, and these nodes were part of a single HA pair. The other HA pair in the four-node cluster was idle and not included in the test. All the ONTAP Select servers were the Fujitsu 2540 model. A separate Fujitsu 2540 server was used to generate the Oracle workload. Figure 2 shows the test architecture.

Figure 2) Test architecture.



10.2 Load Generation

Load was created using the SLOB2 (Silly Little Oracle Benchmark) utility, available [here](#). SLOB drives storage I/O from an Oracle database and allows the 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.

10.3 Protocols

Two protocols were tested on the same hardware setup: iSCSI and NFS. ONTAP Select does not support FC.

10.4 Metrics

Read and write IOPS were based on the Oracle Automatic Workload Repository (AWR) report ILoad 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 is based on the db file sequential read statistic, and write latency is based on db file parallel write. Many other options were 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.

10.5 Latency Cutoff Tests

Table 1 shows the maximum IOPS capabilities of the test configuration using a variety of read/write ratios at the 1ms mark. The latency cutoff is based on the database read latency because this is the primary storage limitation for Oracle databases. As expected, minor differences in performance exist, but NFS and iSCSI deliver equivalent performance overall.

Table 1) Performance at 1ms.

	NFS IOPS (Total)	Read Latency (ms)	iSCSI IOPS (Total)	Read Latency (ms)
50% reads	29,730.8	0.836	33,701.7	1.062
75% reads	38,385.8	1.017	39,430.3	1.025
80% reads	38,460.4	0.974	41,549.1	1.011
90% reads	50,276.4	1.028	46,226.1	0.987

10.6 Protocol Performance Tests

Figure 3 and Figure 4 show read and write latencies for NFS and iSCSI using a 75% read and 25% write workload over an increasing IOPS load.

Figure 3) IOPS/Latency with iSCSI.

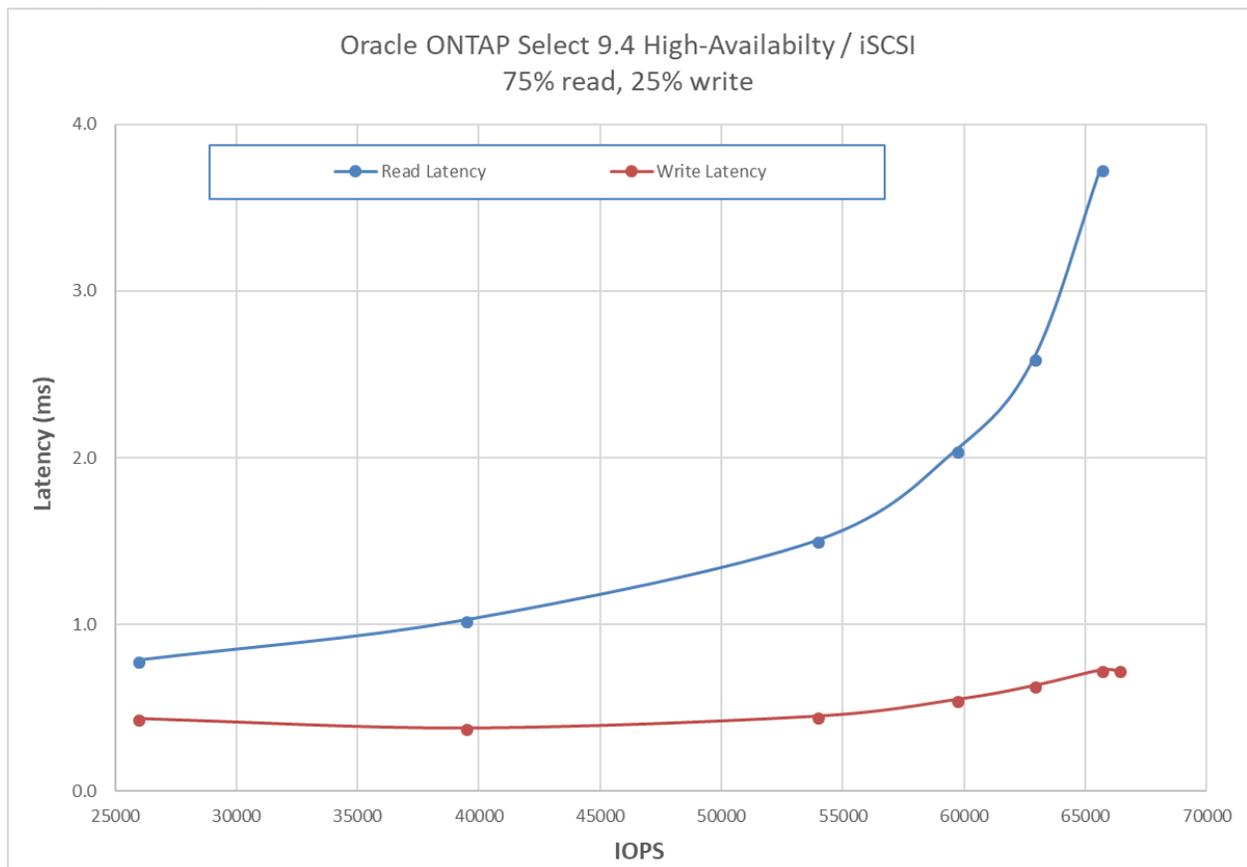
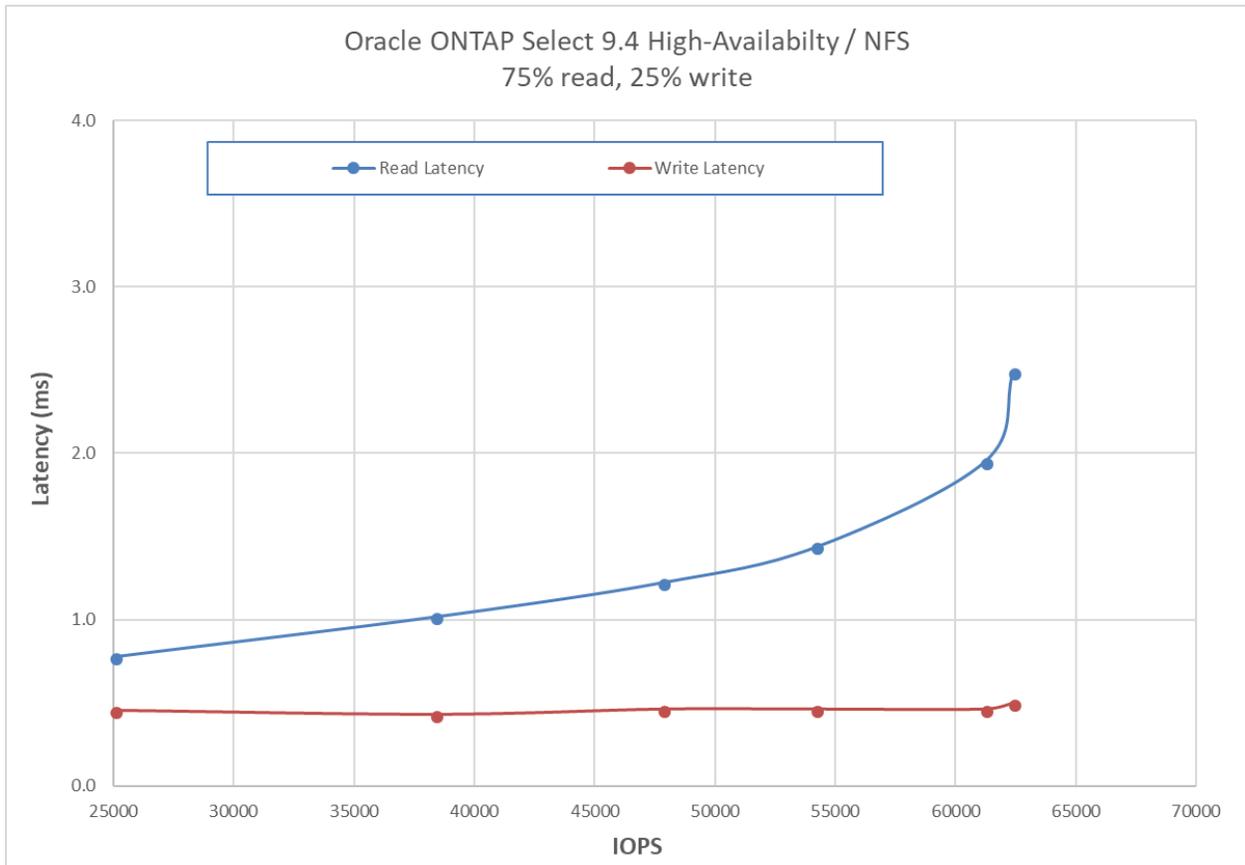


Figure 4) IOPS/Latency with NFS.



Overall performance is comparable. Both protocols showed read latency under 2ms and write latency of about 0.5ms for most of the performance curve. The saturation point for both protocols is approximately 62K IOPS.

These are the expected results for ONTAP on any platform. The differences in protocol overhead are negligible for database I/O patterns. 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 Select, then NFS is the natural choice. Likewise, if current customer practices use the ASM standard, then there is no reason not to choose ASM on iSCSI within ONTAP Select.

Conclusion

ONTAP Select is an optimal solution for delivering complicated data management using software-defined storage, including a full production database, disaster recovery, or simple cloning environment. ONTAP Select 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 with 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>
- TR-4613: ONTAP Select on KVM
<https://www.netapp.com/us/media/tr-4613.pdf>
- TR-4517: ONTAP Select Product Architecture and Best Practices
<https://www.netapp.com/us/media/tr-4517.pdf>
- TR-4700: SnapCenter Plug-In for Oracle Database
<https://www.netapp.com/us/media/tr-4700.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.

Copyright Information

Copyright © 2018 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.

RESTRICTED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.277-7103 (October 1988) and FAR 52-227-19 (June 1987).

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.

TR-4676-0418