Technical Report

Oracle Database 12c Performance: Protocol Comparison Using Clustered Data ONTAP

Rodrigo Nascimento, NetApp
February 2016 | TR-4490

Abstract

This technical report compares protocol performance with Oracle 12c Real Application Clusters (RAC) by using an OLTP workload with the NetApp® clustered Data ONTAP® 8.3.2 storage operating system. We evaluated the FC, software iSCSI, kernel-based NFS (kNFS), and Oracle Direct NFS (DNFS) protocols. This report demonstrates that Data ONTAP can provide a high-performance environment for Oracle Database regardless of the choice of protocol.

NetApp All Flash FAS (AFF) systems combine the high performance of flash media with the industry-leading NetApp Data ONTAP platform to provide performance acceleration, operational agility, best-in-class data protection, and business continuity for database deployments.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Overview

This document describes the performance of an Oracle Database 12c Enterprise Edition database workload with the NetApp AFF8080 storage system and clustered Data ONTAP 8.3.2. The purpose of this document is to demonstrate that all of the protocol options provided by clustered Data ONTAP deliver very high IOPS at extremely low latencies.

Protocol-independence delivers several advantages for businesses. It provides customers with the all-flash performance that they want while leveraging their current infrastructure, whether that infrastructure is block-based or file-based. When a new infrastructure is being deployed, a customer can focus on topics such as cost, manageability, and flexibility without worrying about sacrificing performance. And customers who are looking toward cloud deployments almost always need IP connectivity. The ability of IP-based protocols, such as DNFS, to deliver maximum performance allows cloud integration while still leveraging the full capabilities of an all-flash array.

All tests were performed in an environment with standard enterprise features enabled. This setup included a full high-availability configuration and the use of RAID data protection. Inline data compression was also enabled, because customers require efficiency features that recover the maximum usable capacity from their all-flash storage systems. Tests that disable such enterprise features should not be considered valid.

1.1 Data ONTAP FlashEssentials Empowers All-Flash FAS Performance

NetApp Data ONTAP FlashEssentials is the power behind the performance and efficiency of All Flash FAS. Data ONTAP is a well-known operating system, but it is not widely known is that Data ONTAP, with its WAFL® (Write Anywhere File Layout) file system, is natively optimized for flash media.

Data ONTAP includes the following key features to optimize solid-state drive (SSD) performance and endurance:

- NetApp storage efficiency technologies deliver space savings of up to tenfold or more. Features include inline compression, deduplication, and thin provisioning. Savings can be further increased by using NetApp Snapshot® and NetApp FlexClone® technologies.
- Multiple writes are coalesced and written as a unit. The resulting reduction in storage overhead during write workloads improves performance and flash media longevity.
- AFF systems include a flash-optimized I/O path to maximize performance in a pure flash environment.
- With advanced drive partitioning, SSDs can be shared among controllers, increasing usable capacity and allowing more flexibility in configuration.
- AFF controllers can be used within a larger clustered Data ONTAP cluster, enabling nondisruptive workload migration between flash and hybrid tiers.
- Quality-of-service capability safeguards service level objectives in multiworkload and multitenant environments.

1.2 NetApp Clustered Data ONTAP 8.3.2

An essential feature for Oracle databases that are deployed on shared enterprise storage is the ability to deliver consistent and dependable high performance. High performance must be coupled with nondisruptive operations, high availability, scalability, and storage efficiency to provide a complete set of enterprise-grade data management features. Customers can depend on clustered Data ONTAP 8.3.2 and AFF to provide these essential features.

Built on the clustered Data ONTAP unified scale-out architecture, AFF consistently meets or exceeds the high performance demands of Oracle databases. It also provides rich data management capabilities, such as integrated data protection and nondisruptive upgrades and data migration. With these features, customers can eliminate performance silos and seamlessly integrate AFF into a shared infrastructure.

Clustered Data ONTAP delivers enhanced inline compression that significantly reduces the amount of flash storage required and carries near-zero effects on system performance. It also provides industry-leading ecosystem integration with database applications that makes administration of databases and storage systems far more efficient when compared with other flash storage solutions on the market.

Clustered Data ONTAP is the foundation for the Data Fabric enabled by NetApp. Data can be seamlessly moved and managed across internal and external clouds, unifying the on-premises IT environment with the cloud. Data replication, management, and protection and other key features of Data ONTAP are the same across environments. Customers can build their own cloud environment, enable multi-cloud environments with NetApp Private Storage, or run completely within public clouds by using a Cloud ONTAP® software subscription. Data ONTAP binds it all together.

NetApp is a global supplier of enterprise scale-out storage and data management fabric, and clustered Data ONTAP has been an industry-leading operating system since 2012. Onsite ready and cloud connected, clustered Data ONTAP is a complete, future-proof solution in a rapidly changing technology environment.

1.3 Oracle

Oracle Database 12c provides industry-leading performance, scalability, security, and reliability on clustered or single servers with a wide range of options that meet the business needs for critical enterprise applications. Oracle RAC brings an innovative approach to the challenges of rapidly increasing data volume and the demands for high performance. Oracle RAC uses a scale-out model in which active-active clusters use multiple servers to deliver high performance, scalability, and availability.

NetApp solutions for Oracle databases deliver superior scalability, continuous data access, and automated data management for an immediate response to business opportunities. NetApp has worked with Oracle for years to develop innovative, integrated solutions that reduce IT and business costs and complexity. NetApp leads the way with data storage, offering compelling solutions for Oracle databases.

2 Results

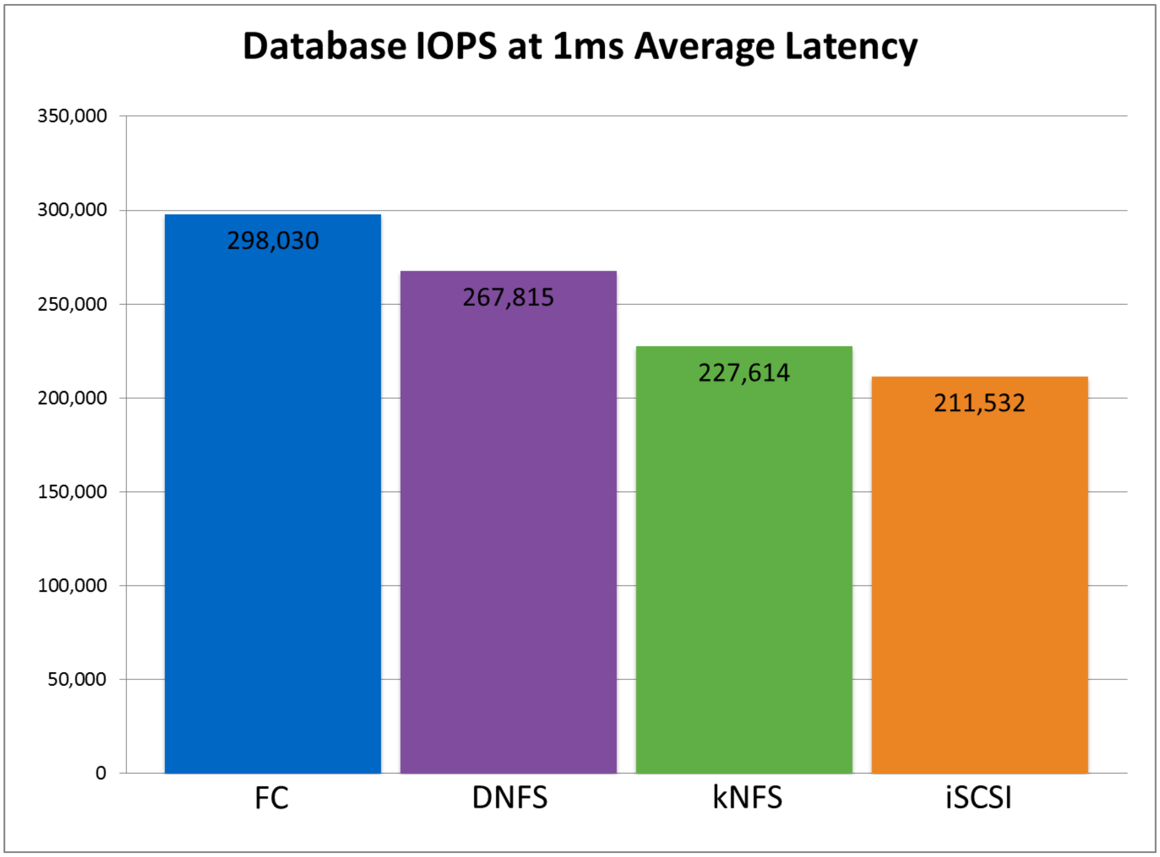
NetApp measured the Oracle server read latency and the IOPS of AFF8080 storage controllers that run clustered Data ONTAP 8.3.2 with an Oracle OLTP workload across the multiple protocols supported by Data ONTAP. The protocols that we used in our testing were FC with 8Gb connections, software iSCSI, kNFS, and Oracle DNFS with10GbE connections.

Two types of NFS were also tested. The standard NFS client provided by the operating system kernel is known as kNFS. Oracle introduced the DNFS client with 11gR1 to overcome performance and availability limitations sometimes encountered with kernel implementations. DNFS is the preferred, default option in recent Oracle releases.

2.1 IOPS at Latency Threshold

Figure 1 illustrates the performance comparison of multiple protocols at 1ms average latency for Oracle database I/O as seen by the host. Although all-flash arrays are capable of delivering much lower latency, the 1ms threshold is commonly used by customers when evaluating all-flash array performance. Traditional storage arrays that are configured with hard disk drives (HDDs) are unable to deliver I/O at 1ms latencies unless the data is resident in cache.

Figure 1) Oracle database IOPS at 1ms average latencies across all protocols served from AFF8080.

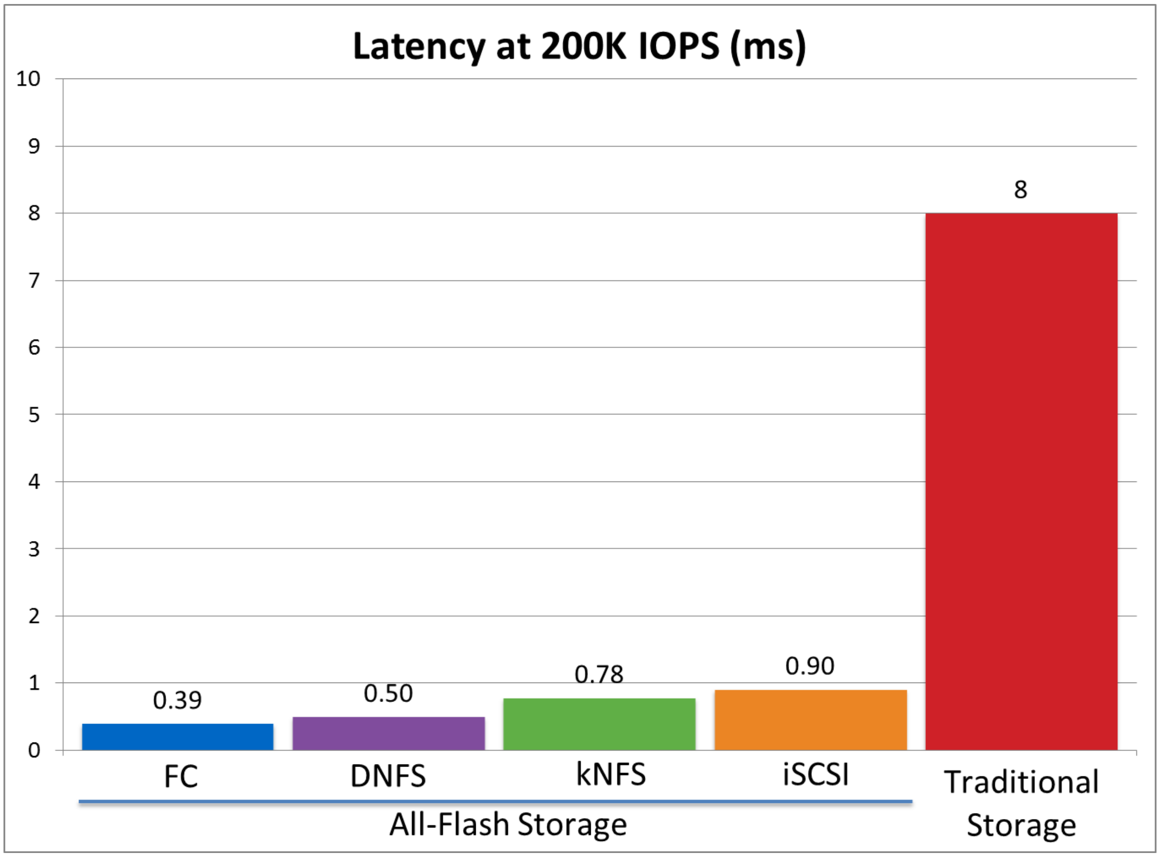


In Figure 1, the difference between FC and DNFS is about 10%. However, this number should not be interpreted to mean that placing a given workload on DNFS generally results in a 10% decrease in performance. Not all database performance problems result from storage I/O latency. Furthermore, the difference in performance shown in Figure 1 is at a very high I/O level. Lower IOPS levels show a smaller difference between protocols.

2.2 All-Flash Protocols and Traditional Media

Many customers who select an all-flash array are planning to migrate from a traditional storage array containing HDDs and no flash-based acceleration. Figure 2 shows the difference in protocol performance compared with traditional HDD-based storage.

Figure 2) Oracle average latencies across all protocols and also with non-flash storage media.



Because of the inherent latencies associated with spinning media, SAS-based HDD storage systems cannot achieve latencies better than approximately 8ms. Although differences exist between the protocols, any protocol with AFF offers an approximately tenfold improvement in performance.

Note: Although the performance of an Oracle database is essentially unaffected by the choice of protocol, some small differences do exist. All tests were conducted with an FC or IP network that was configured according to best practices in terms of host bus adapter (HBA) queue depth, TCP/IP kernel parameters, and so on. As operating systems and associated networking components evolve to better use the capabilities of modern all-flash storage arrays, the performance differences between protocols will continue to narrow.

2.3 Storage Latency and Wait Events

Not all database performance limitations are based on storage latency. In general, customers who have submillisecond latency see their primary performance limitation in the CPU of the database server or servers. Storage latency is not a significant factor.

As an example, Table 1 shows the five primary wait events on an Oracle database connected to traditional storage. This information is from an actual environment of a customer considering migrating to NetApp AFF. An Oracle wait event is a timed event that results in a delay to Oracle processing. It is essentially a list of the primary bottlenecks for a database.

Table 1) Wait events for traditional HDD storage.

Event	Number of Waits	Time (s)	Average Wait (ms)	% Database Time	Wait Class
Database file sequential read	479,476	2,722	6	71.02	User I/O
Database server CPU	–	591	–	15.42	–
Log file sync	183,807	395	2	10.31	Commit
Database file scattered read	12,445	88	7	2.29	User I/O
Direct path read temp	4,937	34	7	0.88	User I/O

The critical column is % Database Time, which shows the percentage of time that the database spent waiting for the events. The primary wait event is the database file sequential read, which counterintuitively represents random I/O, at 71% of all lost time. This result is a common pattern for traditional HDD-based storage or hybrid systems that lack sufficient caching capability. The next wait event is database server CPU time, responsible for 15% of delays.

In contrast, Table 2 shows the five primary wait events on an Oracle database server connected to AFF.

Table 2) Wait events for all-flash arrays.

Event	Number of Waits	Time (s)	Average Wait (ms)	% Database Time	Wait Class
Database server CPU	–	1,691,541	–	82.83	–
Database file sequential read	145,028,104	123,278	0.85	6.04	User I/O
Database file scattered read	15,299,292	71,477	4.67	3.50	User I/O
Database file parallel read	2,068,331	12,665	6.12	0.62	User I/O
SQL *Net more data from client	848,851	8,229	9.69	0.40	Network

The primary bottleneck is now the database server CPU. The storage latency for random I/O operations has dropped to less than 1ms, and, for multiple-block operations, it is slightly higher. Storage latency is only responsible for about 10% of all delays. Any performance difference that results from the specific protocol chosen would be nearly undetectable because storage latency is no longer a significant limiting factor.

Therefore, under these conditions, a customer can maximize the CPU utilization of the server. CPUs cost money, and Oracle licenses are almost always provided on a per-CPU basis. Therefore, an idle CPU is a wasted investment.

3 Configuration and Testing

NetApp studied the performance of an AFF8080 system to determine its sustained IOPS and Oracle server read latency across several protocols. The following sections describe the methodology and design considerations used to test the AFF8080 by running a standard Oracle workload.

3.1 Test Methodology

In this study, we used the SLOB2 load-generation tool to simulate an OLTP workload against the Oracle 12*c* RAC database test configuration. The workload generated a database select-update ratio of approximately 75:25 against the Oracle database in the test configuration.

Our tests measured and compared the performance across a variety of protocols of the Oracle 12c RAC database with an OLTP workload connected to NetApp AFF8080 storage controllers running NetApp clustered Data ONTAP 8.3.2.

All test configurations used one of the protocols tested, including FC with 8Gb connections, software iSCSI, kernel NFS, and Oracle DNFS with 10GbE connections, between the database servers and the AFF8080 storage controller. Additionally, the Oracle DNFS protocol was used from within the database server application. Oracle DNFS is an optimized NFS client that provides faster and more scalable access to NFS storage located on NAS devices. DNFS is accessible over TCP/IP.

The tested software iSCSI, kNFS, and Oracle DNFS configurations used an Intel 82599ES 10GbE controller in the database servers connected to the NetApp 10GbE unified target adapters installed in the AFF8080 controllers. The tested FC configuration used a QLogic QLE2562 in the database servers connected to onboard 8Gb FC ports in the AFF8080 controllers through a Brocade 6510 switch.

All tests used a five-node Oracle 12c RAC implementation configured on five physical servers accessing the AFF8080 controllers through the respective protocol.

We used NetApp clustered Data ONTAP 8.3.2, the latest release in the Data ONTAP family, to demonstrate the performance capabilities that this operating system offers to NetApp customers. Inline compression and deduplication are enabled by default on the AFF8080 running clustered Data ONTAP 8.3.2, and we ran our tests with these storage efficiency features enabled.

We also used jumbo frames for the tests that used 10GbE network connections.

3.2 Hardware and Software

For this study, we configured five Oracle Database 12*c* database servers on five Fujitsu Primergy RX300 S7 servers. We connected the five servers to a two-node AFF8080 system through a variety of protocols. We connected each node of the AFF8080 to a single DS2246 shelf and populated each shelf with 24 800GB SSDs.

Table 3 and Table 4 list the hardware and software components that we used for the Oracle performance test configuration.

Table 3) Oracle hardware and software components.

Hardware and Software Components	Details
Oracle Database 12*c* servers	Five Fujitsu Primergy RX300 S7
Server operating system	Red Hat Enterprise Linux 6.6
Oracle Database version	12*c* (RAC)
Processors per server	Two 6-core Xeon E5-2630 @ 2.40 GHz
Physical memory per server	48GB
FC network	8Gb FC with multipathing
FC HBA	QLogic QLE2562 dual-port PCIe
Converged network adapter per server for Oracle data traffic	QLogic QLE8152 dual-port 10GbE

Hardware and Software Components	Details
Oracle interconnections per server	One Intel 82599EB dual-port 10GbE
Dedicated public 1GbE ports for management	Two Intel 1350GbE ports
8Gb FC switch	Brocade 6510 24-port
10GbE IP switch	Cisco Nexus 5596

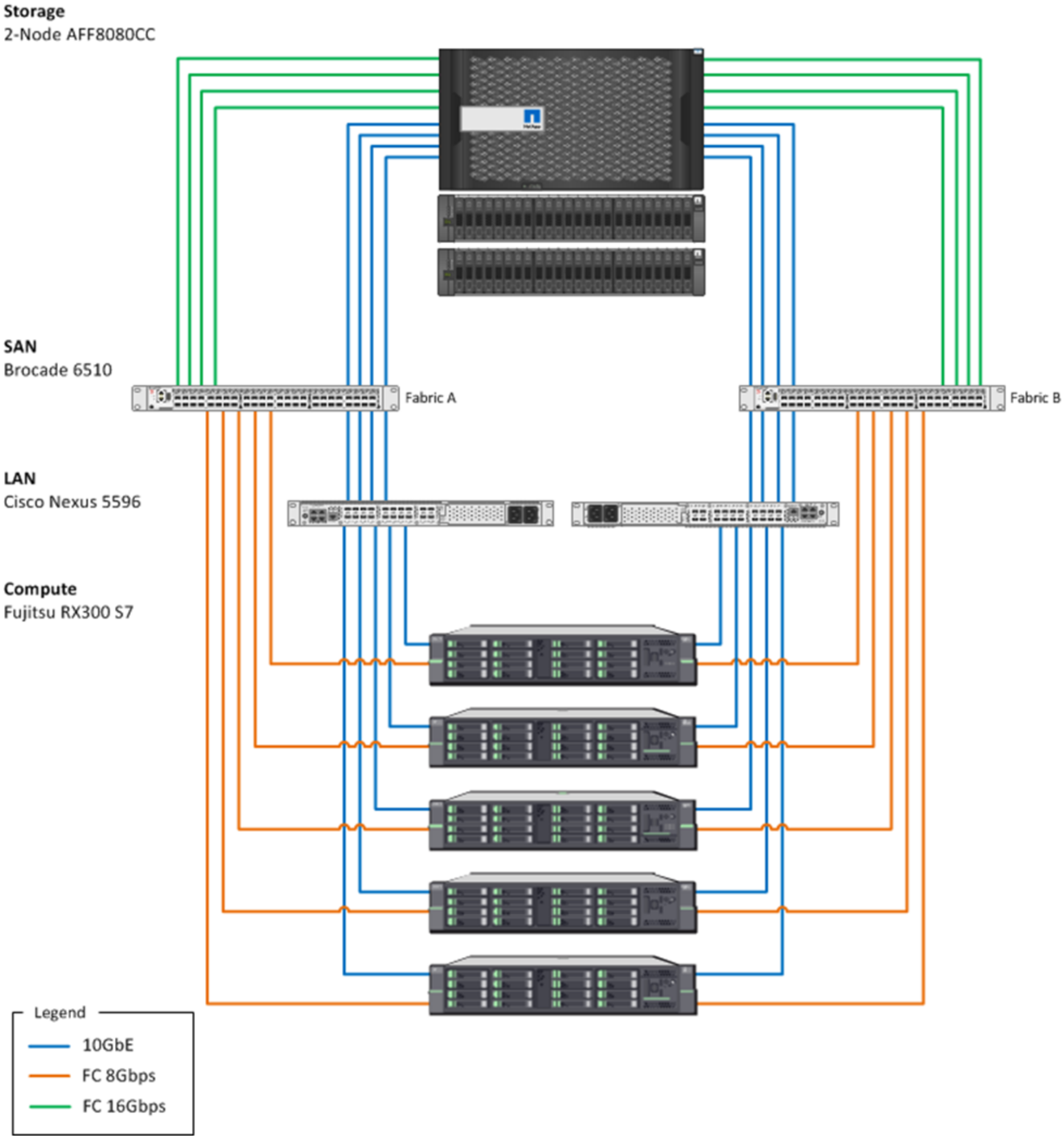
Table 4) NetApp storage system hardware and software.

Hardware and Software Components	Details
Storage system	NetApp AFF8080 configured as a high-availability (HA) active-active pair
Clustered Data ONTAP version	8.3.2
Total number of drives	48
Drive size	800GB
Drive type	SSD
FC target ports	Eight 16Gb ports
Ethernet ports	Eight 10GbE ports
Storage virtual machines (SVMs)	One across both node aggregates
Ethernet logical interfaces (LIFs)	Four 1Gb management LIFs (2 per node connected to separate private VLANs)
FC LIFs	Eight 16Gb data LIFs
iSCSI LIFs	Eight 10Gb data LIFs
NFS LIFs	Eight 10Gb data LIFs

3.3 Network Design

This section provides the network connectivity details for the tested configurations. The SAN diagram in Figure 3 shows that the FC SAN was deployed with a Brocade 6510 8Gb FC switch. Each storage node had four ports connected to the FC switch. Each server had two ports connected to the switch. The multiple ports used in the FC SAN configurations provided HA (through multipathing) and increased bandwidth. At no point in the testing did the network connectivity create a bottleneck.

Figure 3) Network design.



3.4 Database Layout and Storage Provisioning Design

Table 5 summarizes the layout for the Oracle database. We used one Oracle RAC database to host the simulated OLTP environment. Each storage system node housed a single aggregate containing 24 800GB SSDs that were subdivided into NetApp RAID DP® groups plus one spare drive. We configured the two data aggregates into a single SVM and created a single database using Oracle RAC.

Table 5) Database configuration.

Storage	Aggregate Name	Volume Name	LUN Size (GB)	Vol Size (GB)	Description
Per node	–	–	–	–	Used advanced drive partitioning
	aggr0	root	–	55	Total aggregate size = 55GB
	aggr	–	–	–	21 data + 2 parity RAID DP + 1 spare Total aggregate size = 12.7TB
Oracle RAC configuration	–	db1_vol1	200	220	Data files
	–	db1_vol2	200	220	
	–	db1_vol3	200	220	
	–	db1_vol4	200	220	
	–	db1_vol5	200	220	
	–	db1_vol6	200	220	
	–	db1_vol9	200	220	
	–	db1_vol10	200	220	
	–	db1_vol11	200	220	
	–	db1_vol12	200	220	
	–	db1_vol17	40	100	Redo log
	–	db1_vol18	40	100	

For SAN protocol testing, we used one initiator group (igroup) per server to contain the FC initiators. We then created disk groups with an allocation unit size of 64MB by using the Oracle Automatic Storage Management (ASM) volume manager. Those ASM disk groups provided the storage that was required to create the tablespaces. The FC SAN was configured on the Brocade switch. Clustered Data ONTAP provided Asymmetric Logical Unit Assignment communication to the initiators so that optimal paths were used for host I/O access according to the multipath I/O load-balancing policies on the host.

We deployed zoning in our configuration to balance the FC connections, using eight paths per LUN. We used two HBA ports per server and four LIFs per node. One server port was zoned for two LIFs per node, and the other port was zoned for the other two LIFs per node.

For our IP testing, we used one iSCSI igroup per server to contain iSCSI Qualified Name initiators. All LUNs were unmapped from all FC igroups and then mapped to the iSCSI igroup.

Each storage node had four 10GbE ports. For this testing, one LIF per port was created in two different network segments. Each server had eight iSCSI sessions (four sessions using NIC1 and four sessions using NIC2), and each session represented one path to the LUNs.

3.5 Workload Design

We used SLOB2 to generate our OLTP workload. Each database server applied the workload to the Oracle database, log, and temp files. We configured the workload to be 75% selects and 25% updates with a block size of 8KB.

To collect our performance results, we tested the environment by increasing the number of Oracle users in SLOB2 from a minimum of 32 users up to a maximum of 440 users. At each load point, we verified that the storage system and the Oracle servers could maintain steady-state behavior without failure. We also confirmed that there were no bottlenecks across servers or networking systems.

Note: We took care in these test steps to simulate real database and customer workloads, but we acknowledge that workloads vary across databases. In addition, these test results were obtained in a closed lab environment with no competing workloads on the same infrastructure. In a typical shared-storage infrastructure, other workloads share resources. Your results might vary from those found in this report.

4 Conclusion

The NetApp AFF8080 solution provides very high IOPS at extremely low latencies across all protocols when serving an Oracle Database 12*c* OLTP workload. Our testing showed that the AFF8080 cluster was able to achieve between 210,000 and 300,000 IOPS at a maximum of 1ms read latency across all of the protocols in our study. Although differences exist between the protocols, any protocol choice with AFF offers an approximately tenfold increase in performance.

We also demonstrated in our testing that all of the protocols provide outstanding performance, although FC is the best option. This feature enables a customer to focus on topics such as cost, manageability, and flexibility without concern for sacrificing performance.

Although there are some differences, in the context of real database workloads, all protocols offer comparable performance. When compared with traditional storage that primarily leverages spinning media, all of these protocols deliver latency improvements that eliminate the storage system as a bottleneck. As a result, customers can choose the right protocol based on their business requirements, not on the technical limitations of a protocol.

References

The following references were used in this technical report:

The Silly Little Oracle Benchmark v2.2 (SLOB2)
http://kevinclosson.net/2012/02/06/introducing-slob-the-silly-little-oracle-benchmark/

Version History

Version	Date	Document Version History
Version 1.0	February 2016	Initial release

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.