WHITE PAPER

ORACLE FLASHSTACK WITH FLASHBLADE

DEPLOY A MODERN DATA WAREHOUSE

PURESTORAGE®

TABLE OF CONTENTS

INTRODUCTION

To stay relevant in today's competitive, digitally disruptive market, and to stay ahead of your competition, you have to do more than just store, extract, and analyze your data — you have to draw the true business value out of it. Fail to evolve, and your organization may be left behind as companies ramp up and speed up their competitive, decision-making environments. This means deploying cost-effective, energy-efficient solutions that allow you to quickly mine and analyze your data for valuable information, patterns, and trends, which in turn can enable you to make faster ad-hoc decisions, reduce risk, and drive innovation.

FlashStack™ is a flexible, all-flash converged infrastructure solution from Pure Storage® and Cisco® that brings the all-flash revolution to your data center, faster. It combines the latest in compute, network, storage hardware, and virtualization software into a single, integrated architecture that speeds time to deployment, lowers overall IT costs, and reduces deployment risk. Highly efficient components reduce the costs associated with power, cooling, and data center space. Powered by 100% flash storage, FlashStack provides the performance and reliability that business-critical applications demand.

This document describes a FlashStack reference architecture for deploying a highly available Oracle DB environment on Pure Storage FlashBlade using Cisco UCS compute servers, Cisco Nexus Switches, and Oracle Linux. Pure Storage has validated the reference architecture with a data warehouse workload in its lab. This document presents the hardware and software configuration of the components involved and the results of various tests, and offers implementation and best practices guidance.

FlashStack is designed to increase IT responsiveness to business demands while reducing the overall cost of computing. FlashStack components are integrated and standardized to help you achieve timely, repeatable, consistent deployments.

The combination of a Cisco UCS platform, Pure Storage FlashBlade, and Oracle Real Application Cluster (RAC) architecture can accelerate your IT transformation by enabling faster deployments, greater flexibility and efficiency, and lower risk.

GOALS AND OBJECTIVES

Oracle® data warehouse deployments are extremely complicated in nature and customers face enormous challenges in maintaining these landscapes in terms of time, effort, and cost. Oracle databases are often the mission critical components of a customer's business. Ensuring availability and lowering TCO is always a top priority.

A storage platform based on Oracle data warehouse and analytics solutions supported by an all-flash storage solution, such as Pure Storage FlashBlade and FlashStack, can help you solve the challenges of data warehousing, management, and analysis — no matter where your data is stored.

The goal of this document is to showcase the scalability, performance, manageability, and simplicity of the FlashStack Converged Infrastructure solution for deploying a modern data warehouse with Oracle databases.

The following are the specific objectives of this reference architecture document.

1. Build, validate, and predict the performance of server, network, and storage platforms on a per workload basis
2. Demonstrate the seamless scalability of performance and capacity to meet the growth needs of Oracle Database
3. Confirm the high availability of database instances, without performance compromise, through software and hardware upgrades
4. Ensure customer success with prebuilt, pre-tested drivers for Oracle database software

We will demonstrate the scalability and performance capabilities of FlashBlade by performing high-end data ingest using a TPC-H like benchmark and by executing business-related queries with a high degree of complexity.

AUDIENCE

The target audience for this document includes but is not limited to storage administrators, data center architects, database administrators, field consultants, IT Managers, and Oracle solution architects who want to implement a modern data warehouse on Linux in a FlashStack Converged Infrastructure solution. A working knowledge of Oracle Database, Linux, server, storage, and networks is assumed but is not a prerequisite to read this document.

REFERENCE ARCHITECTURE DESIGN PRINCIPLES

This document describes, at a high level, the components of an enterprise reference architecture to deploy a data warehouse using Oracle and Pure Storage FlashBlade and FlashStack for the purposes of information access and analysis. The rest of the paper is dedicated to defining key technical challenges and how Oracle Database 12c functionality is used to address them.

The guiding principles for implementing this reference architecture are:

- **Repeatability:** Create a scalable building block that can be easily replicated at any customer site. Publish the versions of various firmware under test and weed out any issues in the lab before customers deploy this solution.
- **Availability:** Create a design that is resilient and not prone to failure of a single component. For example, we include best practices to enforce multiple paths to storage, multiple NICs for connectivity, and high availability (HA) clustering with the use of Oracle RAC.
- **Efficiency:** Take advantage of the inline data reduction, higher bandwidth, and low latency of the Pure Storage FlashBlade used in the FlashStack solution.

- **Simplicity:** Avoid unnecessary and/or complex tweaks to make the results look better than a normal out-of-box environment.
- **Scalability:** Demonstrate the linear scaling of Oracle databases within the FlashStack architecture, even as we maintain best-in-class flash storage performance.

FLASHSTACK INTRODUCTION

FlashStack™ is a converged infrastructure solution with unmatched power, scale, and flexibility. 100% all-flash and built with best-of-breed components from Cisco and Pure Storage, FlashStack's key differentiators are precisely those that drive enterprise digital transformation: open standards and stateless operation, ample performance and easy scalability, and maximum density and minimal cabling. And while a FlashStack solution provides industry-leading simplicity, flexibility, and efficiency, it is also more economical than legacy disk-based converged infrastructure solutions.

> FlashStack is available from accredited FlashStack Partners, who help provide an excellent converged infrastructure ownership experience. FlashStack Partners have the knowledge and experience necessary to help streamline the sizing, procurement, and delivery of your entire system.

KEY BENEFITS OF THE FLASHSTACK SOLUTION

- **Consistent Performance and Scalability**
  - Consistent sub-millisecond latency with 100% flash storage
  - Consolidate 100s of enterprise-class applications in a single rack
  - Scale easily, without disruption
  - Repeatable growth through multiple FlashStack deployments
- **Operational Simplicity**
  - Fully tested, validated, and documented for rapid deployment
  - Reduced management complexity
  - No storage tuning or tiers necessary
- **Lowest TCO**
  - Dramatic savings in power, cooling, and space with all-flash
  - Data reduction through compression
- **Enterprise Grade Resiliency**
  - Highly available architecture and redundant components
  - Non-disruptive operations
  - Upgrade and expand without downtime or performance loss

FIGURE 1. FlashStack with FlashBlade

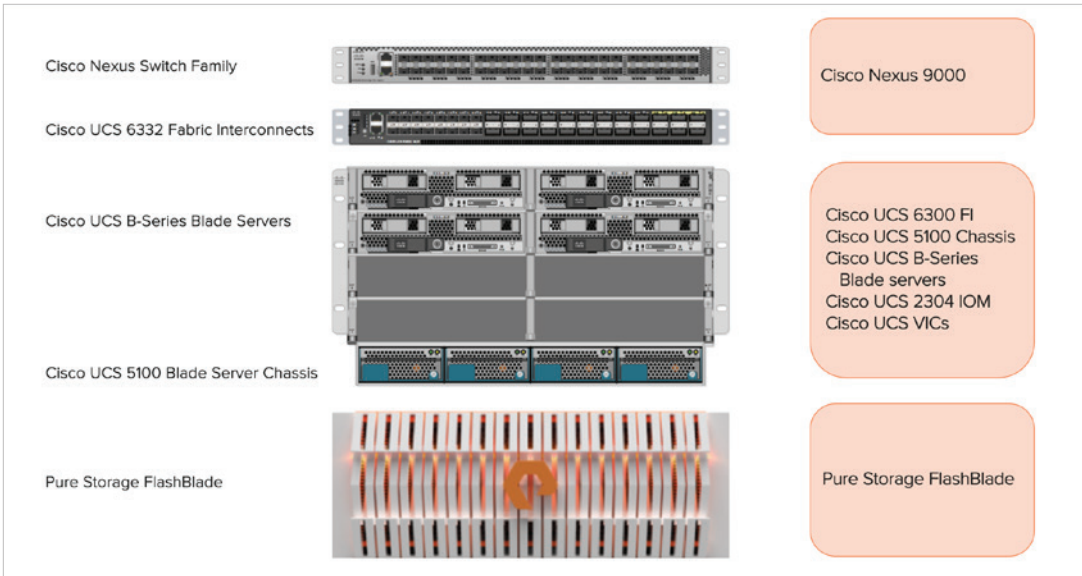


FIGURE 2. FlashStack system components

The above components are connected and configured according to best practices of both Cisco and Pure Storage and provide the ideal platform for running enterprise workloads with confidence. FlashStack can scale up for greater performance and capacity by adding compute, network, or storage resources individually as needed, or it can scale out for environments that require multiple consistent deployments.

COMPONENTS OF THE REFERENCE ARCHITECTURE

The IT industry has been transforming rapidly to converged infrastructure, which enables faster provisioning, scalability, lower data center costs, simpler management infrastructure, and future-proofing with technology advancement. The FlashStack solution provides best-of-breed technology to reap the benefits that converged infrastructure brings to the table. This section details the various infrastructure components that make up FlashStack.

Pure Storage FlashBlade

FlashBlade™ is a new, innovative scale-out storage system designed to accelerate modern analytics applications while providing best-of-breed performance in all dimensions of concurrency – including IOPS, throughput, latency, and capacity. FlashBlade is as simple as it is powerful, offering elastic scale-out storage services at every layer alongside DirectFlash technology for global flash management.

PURPOSE-BUILT FOR MODERN ANALYTICS

FlashBlade is the industry's first cloud-era flash purpose-built for modern analytics, delivering unprecedented performance for big data applications. Its massively distributed architecture enables consistent performance for all analytics applications using NFS, S3/Object, SMB, and HTTP protocols.

FAST

- Elastic performance that grows with data, up to 17 GB/s
- Always-fast, from small to large files
- Massively parallel architecture from software to flash

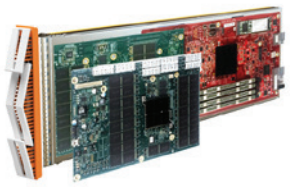
BIG

- Petabytes of capacity
- Elastic concurrency, up to 10s of thousands of clients
- 10s of billions of objects and files

SIMPLE

- Evergreen – don't rebuy TBs you already own
- "Tuned for Everything" design, no manual optimizations required
- Scale-out everything instantly by simply adding blades

THE FLASHBLADE DIFFERENCE



BLADE

SCALE-OUT DIRECTFLASH + COMPUTE

Ultra-low latency, 8, 17, and 52TB capacity options that can be hot-plugged into the system for expansion and performance

PURITY//FB

SCALE-OUT STORAGE SOFTWARE

The heart of FlashBlade, implementing its scale-out storage capabilities, services, and management

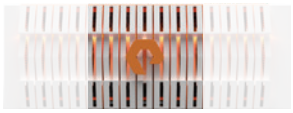
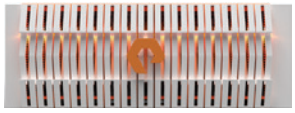
FABRIC

SOFTWARE- DEFINED NETWORKING

Includes a built in 40Gb Ethernet fabric, providing a total network bandwidth of 320Gb/s for the chassis

POWER, DENSITY, EFFICIENCY

FlashBlade delivers industry-leading throughput, IOPS, latency, and capacity – in 20x less space and 10x less power and cooling.

SPECIFICATIONS		8 TB BLADE	17 TB BLADE	52 TB BLADE
7 BLADES		98 TBs Usable	190 TBs Usable	591 TBs Usable
15 BLADES		267 TBs Usable	525 TBs Usable	1607 TBs Usable

* Usable capacity assumes 3:1 data reduction rate. Actual data reduction may vary based on use case.

PERFORMANCE

17 GB/s bandwidth with 15 blades

Up to 1M IOPS

CONNECTIVITY

8x 40Gb/s or 32x 10Gb/s Ethernet ports / chassis

PHYSICAL

4U

1,800 Watts (nominal at full configuration)

Purity for FlashBlade (Purity//FB 2)

FlashBlade is built on the scale-out metadata architecture of Purity for FlashBlade, capable of handling 10s of billions of files and objects while delivering maximum performance, effortless scale, and global flash management. The distributed transaction database built into the core of Purity means storage services at every layer are elastic: simply adding blades grows system capacity and performance, instantly.

Purity//FB 2 is the industry's most advanced software for scale-out storage, delivering amazing linear performance at unprecedented scale. It supports S3-compliant object store, offering ultra-fast performance at object store scale. It also offers a wave of new enterprise features, like snapshots, SMB, LDAP, network lock management (NLM), and IPv6, to extend FlashBlade into new use cases.

Pure1 – Intelligent, Self-Driving Storage

Pure1®, our cloud-based management, analytics, and support platform, expands the self-managing, plug-n-play design of our products with the machine learning predictive analytics and continuous scanning of Pure1 Meta™ to enable an effortless, worry-free storage experience.

PURE1 MANAGE

In the Cloud IT operating model, installing and deploying management software is an oxymoron: you simply login. Pure1 Manage is SaaS-based, allowing you to manage your array from any browser or from the Pure1 Mobile App – with nothing extra to purchase, deploy, or maintain. Just login, and from a single dashboard you can manage all your arrays, with full visibility on the health and performance of your storage.

PURE1 ANALYZE

For the first time in the industry, Pure1 Analyze delivers true performance forecasting – giving customers complete visibility into the performance and capacity needs of their arrays – now and in the future. Performance forecasting enables intelligent consolidation and unprecedented workload optimization.

PURE1 SUPPORT

Our philosophy is always to do what is right for the customer. No excuses, we focus on delivering a solution no matter the situation. We combine an ultra-proactive support team with the predictive intelligence of Pure1 Meta to deliver unrivaled support that's a key component in our proven 99.9999% availability. You'll be astounded when we fix issues you did not even know existed.

Evergreen Storage

The Evergreen™ Storage ownership model operates like SaaS and the cloud. Deploy storage once and benefit from a subscription to continuous innovation as you expand and improve performance, capacity, density, and/or features for 10 years or more – all without downtime, performance impact, or data migrations. Evergreen Storage provides expandability and upgradability for generations via its modular, stateless architecture, while FlashBlade's blade-based design delivers the linear scale of DirectFlash technology and compute simply by adding blades.

Cisco Unified Computing System

The Cisco Unified Computing System™ (Cisco UCS™) is the next-generation data center platform that unites compute, network, storage access, and virtualization into an organized structure aimed to reduce total cost of ownership and introduce vastly improved infrastructure deployment mechanisms at scale. UCS incorporates a unified network fabric with scalable, modular and powerful x86-architecture servers. With an innovative and proven design, Cisco UCS delivers an architecture that increases cost efficiency, agility, and flexibility beyond what traditional blade and rack-mount servers provide. Cisco makes organizations more effective by addressing the real problems that IT managers and executives face and solves them on a systemic level.

GREATER TIME-ON-TASK EFFICIENCY

Automated configuration can change an IT organization's approach from reactive to pro-active. The result is more time for innovation, less time spent on maintenance, and faster response times. These efficiencies allow IT staff more time to address strategic business initiatives. They also enable better quality of life for IT staff, which means higher morale and better staff retention—both critical elements for long-term efficiency.

Cisco UCS Manager is an embedded, model-based management system that allows IT administrators to set a vast range of server configuration policies, from firmware and BIOS settings to network and storage connectivity. Individual

servers can be deployed in less time and with fewer steps than in traditional environments. Automation frees staff from tedious, repetitive, time-consuming chores that are often the source of errors that cause downtime, making the entire data center more cost effective.

EASIER SCALING

Automation means rapid deployment, reduced opportunity cost, and better capital resource utilization. With Cisco UCS, rack-mount and blade servers can move from the loading dock and into production in a "plug-and-play" operation. Automatically configure blade servers using predefined policies simply by inserting the devices into an open blade chassis slot. Integrate rack-mount servers by connecting them to top-of-rack Cisco Nexus® fabric extenders. Since policies make configuration automated and repeatable, configuring 100 new servers is as straightforward as configuring one server – delivering agile, cost-effective scaling.

VIRTUAL BLADE CHASSIS

With a separate network and separate management for each chassis, traditional blade systems are functionally an accidental architecture based on an approach that compresses all the components of a rack into each and every chassis. Such traditional blade systems are managed with multiple management tools that are combined to give the illusion of convergence for what is ultimately a more labor-intensive, error-prone, and costly delivery methodology. Rack-mount servers are not integrated and must be managed separately or through additional tool sets, adding complexity, overhead, and the burden of more time.

Architecturally, Cisco UCS blade and rack-mount servers are joined into a single virtual blade chassis that is centrally managed yet physically distributed across multiple blade chassis, rack-mount servers, and even racks and rows. This capability is delivered through Cisco fabric interconnects that provide redundant connectivity, a common management and networking interface, and enhanced flexibility. This larger virtual chassis, with a single redundant point of management, results in lower infrastructure cost per server, with fewer management touch points, and lower administration, capital, and operational costs.

Cisco Unified Computing System Components

CISCO UCS 6300 SERIES FABRIC INTERCONNECTS

The Cisco UCS 6300 Series Fabric Interconnects provides line-rate, low-latency, lossless, 10 and 40-Gigabit Ethernet (varies by model), Fibre Channel over Ethernet (FCoE) and Fibre Channel (FC) connectivity. Cisco UCS 6300 Series Fabric Interconnects provide the management and communication backbone for Cisco UCS B-Series Blade servers, and Cisco UCS 5100 Series Blade Server Chassis.

https://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-6300-series-fabric-interconnects/index.html

CISCO NEXUS 9000 SERIES SWITCHES

The Cisco Nexus 9000 Series offers modular 9500 switches and fixed 9300 and 9200 switches with 1/10/25/40/50/100 Gigabit Ethernet switch configurations. 9200 switches are optimized for high-performance and density in NX-OS mode operations.

https://www.cisco.com/c/en/us/products/switches/nexus-9000-series-switches/index.html

CISCO UCS 5100 SERIES BLADE SERVER CHASSIS

The Cisco UCS 5108 Blade Server Chassis is six rack units (6RU) high, can mount in an industry standard 19-inch rack, and uses standard front-to-back cooling. A chassis can accommodate up to eight half-width, or four full-width, Cisco UCS B-Series Blade Server form factors within the same chassis.

https://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-5100-series-blade-server-chassis/index.html

CISCO UCS B-SERIES BLADE SERVERS

Cisco UCS B-Series Blade servers are based on Intel Xeon processor E7 and E5 product families and work with virtualized and non-virtualized applications to increase:

- Performance
- Energy efficiency
- Flexibility
- Administrator productivity

https://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-b-series-blade-servers/index.html

CISCO UCS 2300 SERIES FABRIC EXTENDER

The Cisco UCS 2300 series Fabric Extender brings the unified fabric into a blade server enclosure, providing multiple 10 and 40 Gigabit Ethernet connections between blade servers and the fabric interconnect, simplifying diagnostics, cabling, and management.

https://www.cisco.com/c/en/us/products/collateral/servers-unified-computing/ucs-6300-series-fabric-interconnects/datasheet-c78-675243.html

CISCO UCS MANAGER

Cisco UCS Manager provides unified, centralized, embedded management of all Cisco UCS software and hardware components across multiple chassis and thousands of virtual machines. Administrators use the software to manage the entire Cisco UCS as a single logical entity through an intuitive GUI, a command-line interface (CLI), or an XML API.

Cisco UCS management software provides a model-based foundation for streamlining the day-to-day processes of updating, monitoring, and managing computing resources, local storage, storage connections, and network connections. By enabling better automation of processes, Cisco UCS Manager allows IT organizations to achieve greater agility and scale in their infrastructure operations while reducing complexity and risk.

Cisco UCS Manager provides an easier, faster, more flexible, and unified solution for managing firmware across the entire hardware stack than traditional approaches to server firmware provisioning. Using service profiles, administrators can associate any compatible firmware with any components of the hardware stack. The firmware's auto-installation capability simplifies the upgrade process by automatically sequencing and applying upgrades to individual system elements.

Some of the key elements managed by Cisco UCS Manager include:

- Cisco UCS Integrated Management Controller (IMC) firmware
- RAID controller firmware and settings
- BIOS firmware and settings, including server universal user ID (UUID) and boot order
- Converged Network Adapter (CAN) firmware and settings, including MAC addresses and worldwide names (WWNs) and SAN boot settings
- Virtual port groups used by virtual machines, using Cisco Data Center VM-FEX technology
- Interconnect configuration, including uplink and downlink definitions, MAC address and WWN pinning, VLANs, VSANs, quality of service (QoS), bandwidth allocations, Cisco Data Center VM-FEX settings, and Ether Channels to upstream LAN switches

Cisco UCS Manager provides end-to-end management of all the devices in the Cisco UCS domain it manages. Devices that are uplinked from the fabric interconnect must be managed by their respective management applications. Cisco UCS Manager is provided at no additional charge with every Cisco UCS platform.

For more information on Cisco UCS Manager, visit:
https://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-manager/index.html

CISCO UCS SERVICE PROFILES

Service Profiles are essential to the automation functions in Cisco UCS Manager. They provision and manage Cisco UCS systems and their I/O properties within a Cisco UCS domain. Cisco UCS resources are abstract in the sense that their identify, I/O configuration, MAC addresses and worldwide names (WWNs), firmware versions, BIOS boot order, and networking attributes (including quality of service (QoS) settings, pin groups, and threshold policies) are all programmable using a just-in-time deployment model. The manager stores this identity, connectivity, and configuration information in service profiles that reside on the Cisco UCS 6300 Series Fabric Interconnects. A service profile can be applied to any blade server to provision it with the characteristics required to support a specific software stack. A service profile allows server and network definitions to move within the management domain, enabling flexibility in the use of system resources. Service profile templates allow different classes of resources to be defined and applied to number of resources, each with its own unique identities assigned from predetermined pools.

Oracle Linux 7.3

Oracle Linux, formerly known as Oracle Enterprise Linux, is a Linux distribution based on Red Hat Enterprise Linux (RHEL), repackaged and freely distributed by Oracle, available under GNU General Public License (GPL) since late 2006. Oracle Linux can be downloaded through Oracle's E-Delivery service or from a variety of mirror sites, and can be deployed and distributed freely. Commercial technical support is available through Oracle's Oracle Linux Support program, which supports Oracle Linux, and existing RHEL or CentOS installation.

Oracle Corporation distributes Oracle Linux with two alternative kernels:

- **Red Hat Compatible Kernel (RHCK)** – identical to the kernel shipped in Red Hat Enterprise Linux
- **Unbreakable Enterprise Kernel (UEK)** – based on newer mainline Linux kernel versions, with Oracle's own enhancements for OLTP, Infiniband, and SSD disk access, NUMA-optimizations, Reliable Datagram Sockets (RDS), async I/O, OCFS2, and networking.

Oracle Linux Support Program provides support for KVM components as part of the Oracle Linux 5, Oracle Linux 6, Oracle Linux 7, RHEL5, RHEL6, and RHEL7. This does not include Oracle Product support on KVM offerings

Oracle Database Management System

Oracle revolutionized the field of enterprise database management systems with the most extensive self-management capabilities in the industry, ranging from zero-overhead instrumentation to integrated self-healing and business-driven management. Oracle Database 12c, the next generation of the world's most popular database, makes DBA lives easier by providing various features like change and configuration management, patching, provisioning, testing, performance management, and automatic tuning. Oracle Database high-availability (HA) technologies, collectively referred to as Oracle Maximum Availability Architecture (MAA), provide complete resiliency against all types of outages – from component failures to natural disasters. Industry-leading Oracle HA technology such as Oracle Real Application Clusters (Oracle RAC) provides the highest levels of server HA while Oracle Active Data Guard protects data and applications against sitewide outages.

The FlashStack solution for Oracle includes the following Oracle 12c components and/or features:

- Oracle Database 12c Release 2 (12.2.0.1) Enterprise Edition
- Oracle Grid Infrastruture 12c (12.2.0.1)
- Oracle Flex ASM & ASM Cluster File System (ACFS)
- Oracle Direct NFS Client

Oracle and Cisco UCS – Benefits

AVAILABILITY

- UCS failover capabilities protect against common hardware failures, which is a paramount requirement for mission-critical Oracle databases

SCALABILITY/FLEXIBILITY

- Capacity on demand and dynamic resource allocation
- Stateless blades enable rapid provisioning of nodes

MANAGEABILITY

- Single management interface to manage the hardware

PERFORMANCE

- Improved interconnect and Cache Fusion performance
- Fast, low latency, and lossless 10 Gb Ethernet enables large RAC clusters

Oracle dNFS

Oracle dNFS (direct Network File System) is the NFS client functionality directly integrated in the Oracle RDBMS server. dNFS makes the task of configuring Oracle database on NAS storage much simpler compared to Standard NFS (aka Kernel NFS). Direct NFS Client on Oracle 11g, 12c, or higher supports NFSv3, NFSv4, and NFSv4.1 protocols to access the NFS server.

> The key benefits of Direct NFS Client include simplicity, ease of administration, load balancing, high availability and cost effectiveness. Oracle has optimized the I/O code path by avoiding kernel overhead and, as such, it can improve I/O performance.

Direct NFS Client is capable of performing concurrent direct I/O by bypassing Operating System level caches. It also performs asynchronous I/O, which allows processing to continue while the I/O request is submitted and processed. These two key performance and scalability features provide unparalleled performance when compared to Kernel NFS clients. Another key feature of Direct NFS Client is high availability. Direct NFS Client delivers optimized performance by automatically load balancing requests across all specified paths (up to 4 parallel network paths). If one network path fails, then Direct NFS Client will reissue I/O commands over any remaining paths, ensuring fault tolerance and high availability.

One of the primary challenges of Kernel NFS administration is inconsistency with configurations across different platforms. Direct NFS Client eliminates this problem by providing a standard NFS client implementation across all platforms supported by Oracle Database. This also makes NFS a viable option on platforms like Windows, which doesn't natively support NFS.

As NFS is a shared file system, it supports Real Application Cluster (RAC) databases as well as single-instance databases. Direct NFS Client recognizes when an instance is part of an RAC and automatically optimizes the mount points for RAC, relieving the administrator of manually configuring the NFS parameters.

Direct NFS Client usage

To use Direct NFS Client, the NFS file systems must first be mounted and available over regular NFS mounts. The mount options used in mounting the file systems are not relevant, as Direct NFS Client manages the configuration after installation. Direct NFS Client searches for mount entries in the following order. It will use the first entry found if duplicate entries exist in the configuration files.

1. $ORACLE_HOME/dbs/oranfstab
2. /etc/oranfstab
3. /etc/mtab

The **oranfstab** configuration in **$ORACLE_HOME/dbs** is local to the database under **$ORACLE_HOME**, whereas the **oranfstab** in **/etc/oranfstab** applies to all Oracle databases on that server.

> Note: If a volume is not listed in **oranfstab**, Oracle will look through the OS mount tab (**/etc/mtab**) to find a match. If that fails, control is handed back to the database and file access is made through Kernel NFS.

For multiple interfaces that are on the same subnet, NIC bonding or static path[1] has to be setup on Linux to get increased bandwidth. For all other platforms you can include the attribute **dontroute** inside the **oranfstab** file.

DSS – Data Load & DSS queries

A TPC-H like benchmark was used to generate and test the data load, as well as DSS queries to validate the data warehouse type workload on FlashStack for Oracle. The TPC-H like benchmark consists of a suite of business-oriented ad-hoc queries and concurrent data modifications. The benchmark illustrates decision support systems that examine large volumes of data and execute queries with a high degree of complexity.

[1] See Oracle Support Document "How to Setup Direct NFS client multipaths in same subnet Doc ID # 822481.1"

SOLUTION DESIGN

This section describes the design considerations for the Oracle Database 12c RAC on FlashStack deployment with FlashBlade. In this solution design, we have used two chassis with 8 identical Intel CPU-based Cisco UCS B-series B200 M4 blade servers for hosting the 8-node Oracle RAC database.

FIGURE 3. Oracle Database 12c RAC on FlashStack deployment with FlashBlade

Physical Topology

Pure Storage FlashStack consists of a combined stack of hardware (storage, network, and compute) and software (Cisco UCS Manager, Oracle Database, Purity OS for FlashBlade, Oracle Linux).

- **Network:** Cisco Nexus 9372PX and Cisco UCS Fabric Interconnect 6332 16UP for external and internal connectivity of IP network
- **Storage:** Pure Storage FlashBlade with 14 x 17TB blades with 4 x 40Gb Ethernet connectivity
- **Compute:** Cisco UCS B200 M4 Blade Servers

The following figure shows the physical connectivity between the hardware components that makes up the FlashStack Converged Infrastructure with FlashBlade.

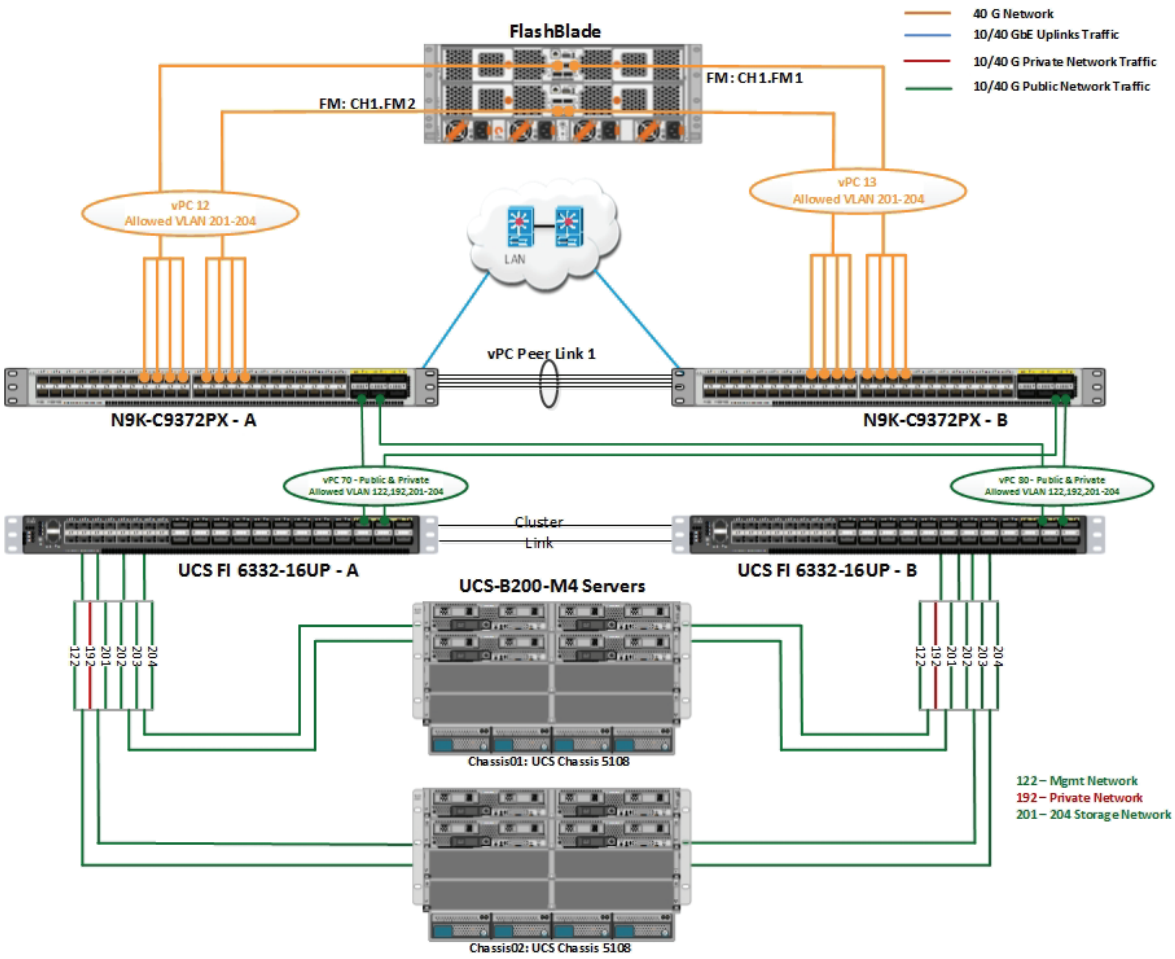


FIGURE 4. FlashStack Converged Infrastructure for Oracle physical connectivity

Server Configuration

Two chassis with 4 identical Intel CPU based Cisco UCS B-series B200 M4 blade servers were deployed for hosting the 8-node Oracle RAC database.

The server has Cisco UCS VIC 1340 cards and they were connected by four ports from each Cisco Fabric extender of the Cisco UCS chassis to the Cisco Fabric Interconnect, which were in turn connected to the Cisco Nexus 9K switch for upstream connectivity to access the Pure Storage FlashBlade NFS volumes. The server configuration is described below.

COMPONENT	DESCRIPTION
PROCESSOR	2 X INTEL XEON E5-2697 V3 2.6 GHZ (2 CPUS WITH 14 CORES EACH)
MEMORY	256GB @ 2.1GHZ (8 X 32GB)
HBA	4 X 10G PORTS ON CISCO UCS VIC 1340 (UCS-IOM-2208XP) 40GBPS
NIC	4 X 10G PORTS ON CISCO UCS VIC 1340
BIOS	TURBO BOOST, HYPER THREADING, VIRTUALIZATION TECHNOLOGY (VT), VT FOR DIRECTED IO, INTEL ATS SUPPORT WERE ENABLED
UCS FIRMWARE (ACROSS ALL COMPONENTS)	3.1 (2B)

TABLE 1. UCS Blade configuration

Cisco UCS Service Profile Configuration

In order to facilitate rapid deployment of UCS servers, a service profile template was created with the following attributes.

- Ethernet adapter policy was set to Linux
- Configured six vNIC Ethernet adapters with one vNIC for Management network, one for private Oracle RAC interconnect traffic and the remaining four for FlashBlade traffic
- All other settings were left at default
- Deployed 8 service profiles from the template and associated with the blade servers

For this FlashStack solution design, we have configured six VLANs as described below.

VLAN CONFIGURATION	VLAN
PUBLIC VLAN	122
PRIVATE VLAN (RAC INTERCONNECT)	192
STORAGE VLAN – 1	201
STORAGE VLAN – 2	202
STORAGE VLAN – 3	203
STORAGE VLAN – 4	204

TABLE 2. VLAN configuration

FlashBlade Configuration

The FlashStack design includes FlashBlade for increased bandwidth and throughput. The table below shows the components of the array.

There are no special configuration or value changes from any normal confirmation. There are no performance knobs to tune on FlashBlade. The hosts are redundantly connected to the controllers through 4 x 40Gb connections (2 x 40Gb per controller module) from the redundant Cisco Nexus 9K switches. FlashBlade was loaded with Purity//FB version 2.0.6, which includes a NLM (Network Lock Manager) feature for the NFS protocol to provide advisory file locking semantics.

COMPONENT	DESCRIPTION
FLASHBLADE	7 X 17TB BLADES TO 14 X 17TB BLADES
RAW CAPACITY RANGE	65 TB – 167 TB
CONNECTIVITY	4 X 40 GB/S REDUNDANT ETHERNET
PHYSICAL	4U
O.S VERSION	PURITY 2.0.6
PROTOCOL USED	NFS V3

TABLE 3. FlashBlade configuration

The FlashBlade network settings were configured with five subnets across five VLANs. The NFS filesystems are to be mounted on four of these subnets on to the target host.

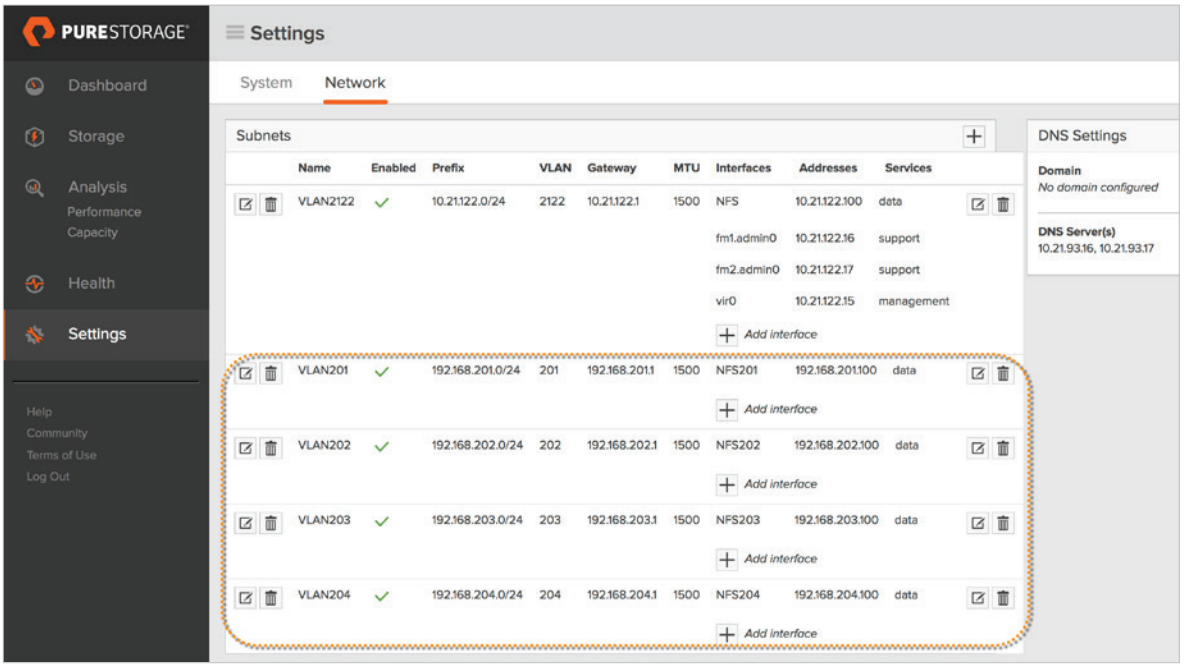


FIGURE 5. FlashBlade network settings

Oracle Configuration

Oracle 12c Release 2 Grid Infrastructure (GI) was installed on the first node as oracle user. The installation also configured and added the remaining 7 nodes as part of the GI setup. ASM was configured with NFS in the Flex ASM mode. Oracle Cluster Registry (OCR) files, voting disk files, and other clusterware files were placed on an NFS filesystem (named OCR) hosted on FlashBlade.

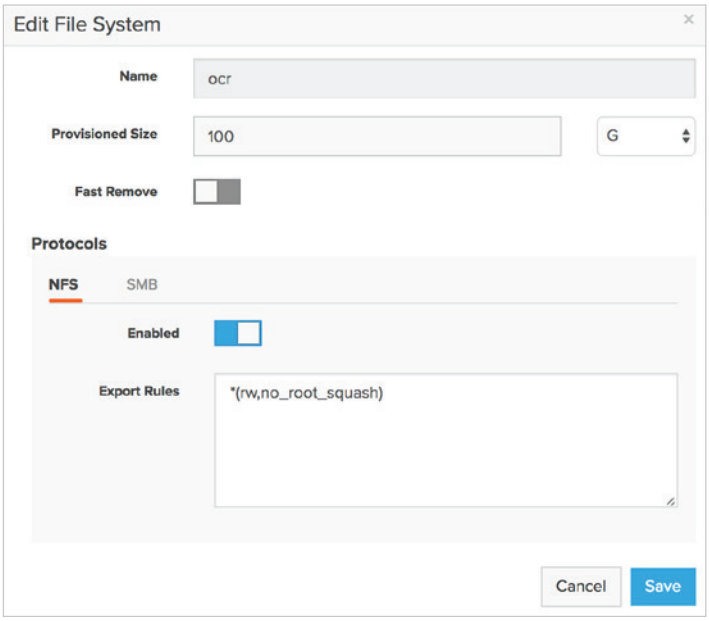


FIGURE 6. Editing an NFS file system

ASM was installed in Flex ASM mode with an ASM instance count of 3 out of the 8-node RAC.

```
2. oracle@dnfsnode1:~ (ssh)
[oracle@dnfsnode1 ~]$ srvctl config asm
ASM home: <CRS home>
Password file: +ocrvfdg/orapwASM
Backup of Password file:
ASM listener: LISTENER
ASM instance count: 3
Cluster ASM listener: ASMNET1LSNR_ASM
[oracle@dnfsnode1 ~]$ asmcmd
ASMCMD> showclustermode
ASM cluster : Flex mode enabled
ASMCMD> exit
[oracle@dnfsnode1 ~]$
```

The following table illustrates the software version and components of Oracle used in this reference architecture.

COMPONENT	DESCRIPTION
RDBMS /GI VERSION	ORACLE 12.2.0.1
SCALABILITY	8-NODE RAC
FILESYSTEM	NFS
ORACLE COMPONENTS	GRID INFRASTRUCTURE, DNFS, ACFS
ORACLE USER/GROUP	ORACLE, OINSTALL (DBA)

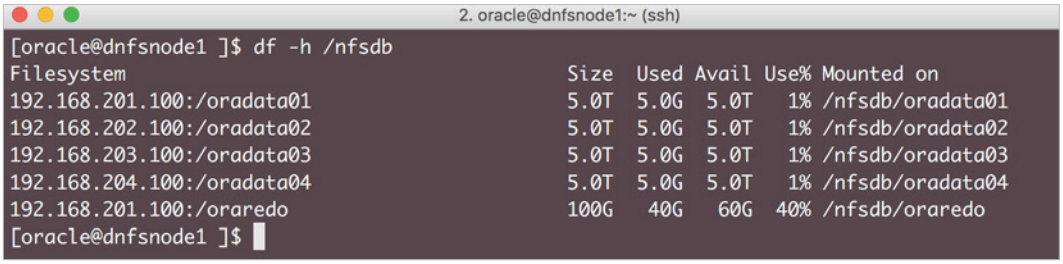
TABLE 4. Oracle Software configuration

The data warehouse database was created using Oracle Database Configuration Assistant (DBCA). The database-related files (data files, redo logs, control files) were placed on FlashBlade NFS volumes as listed below.

NFS VOLUMES	SIZE	DESCRIPTION
ORADATA01	5 TB	ORACLE DATABASE DATAFILES & CONTROLFILE
ORADATA02	5 TB	DATAFILES
ORADATA03	5 TB	DATAFILES
ORADATA04	5 TB	DATAFILES
ORAREDO	100 GB	REDO LOG FILES & CONTROLFILE

TABLE 5. Database-related files on FlashBlade volumes

These NFS filesystems were mounted on all 8 nodes with similar mount names on the storage VLANs (201 – 204). 5 x 1GB redo log files were created for each instance, totaling 40 redo log files.



The NFS filesystems were mounted with the following options.

```
192.168.201.100:/oraredo      /nfsdb/oraredo nfs rw,bg,nointr,hard,tcp,vers=3,actimeo=0
192.168.201.100:/oradata01    /nfsdb/oradata01 nfs rw,bg,nointr,hard,tcp,vers=3,actimeo=0
192.168.202.100:/oradata02    /nfsdb/oradata02 nfs rw,bg,nointr,hard,tcp,vers=3,actimeo=0
192.168.203.100:/oradata03    /nfsdb/oradata03 nfs rw,bg,nointr,hard,tcp,vers=3,actimeo=0
192.168.204.100:/oradata04    /nfsdb/oradata04 nfs rw,bg,nointr,hard,tcp,vers=3,actimeo=0
```

Oracle dNFS Configuration

Oracle dNFS was enabled at the RDBMS level on all the database nodes, and the **oranfstab** was updated to reflect the same across all nodes.

> Note: Oracle dNFS is by default enabled on Oracle 12c. To disable dNFS, the RDBMS should be rebuilt with the **dnfs_off** option. Please check the Best Practices section for enabling/disabling Oracle dNFS.

Ways to check if Oracle dNFS is enabled at the database level and working as expected:

- Check the **alert.log** on the database server. It should show the following entry.

```
Oracle instance running with ODM: Oracle Direct NFS ODM Library Version 4.0
```

- Run a SQL query against **v$dnfs_servers** that should show the details of the dNFS mounts.

```
SQL> select svrname, dirname, mntport, nfsport, wtmax, rtmax from v$dnfs_servers;

SVRNAME         DIRNAME       MNTPORT     NFSPORT       WTMAX       RTMAX
--------------- ---------- ---------- ---------- ---------- ----------
192.168.201.100 /oradata01        2049        2049      524288      524288
192.168.201.100 /oraredo          2049        2049      524288      524288
192.168.204.100 /oradata04        2049        2049      524288      524288
192.168.202.100 /oradata02        2049        2049      524288      524288
192.168.203.100 /oradata03        2049        2049      524288      524288

5 rows selected.
```

SCALABILITY TEST AND RESULTS

Pure Storage FlashBlade is purpose-built for modern analytics. FlashBlade can handle high volumes of data, leveraging massively parallel architecture, and delivers a powerful data warehouse system that is fast, big, and simple. In this FlashStack solution, we will test the scalability and high throughput performance of the FlashBlade for Oracle data warehouse use case.

Database Workload Configuration

We used a TPC-H like benchmark toolkit that is slated to mimic decision support systems. The benchmark consists of set of business-oriented ad-hoc queries that examine large volumes of data with a high degree of complexity, and give answers to critical business questions. The toolkit also provides a data population option to generate large volumes of data.

For this solution, we created a DSS (Decision Support System) database to demonstrate the high-bandwidth, high throughput data warehouse type workload laid out on FlashBlade. The DSS database was approximately 15TB in size. The data warehouse workloads are generally sequential in nature, read intensive, and exercise large IO size.

Data generation from the TPC-H like benchmark was used to generate the datasets that are loaded into the Oracle database.

COMPONENT	DESCRIPTION		
SCALE FACTOR	**10,000**		
FILESYSTEM – RAW DATAFILES	**ORACLE ASM CLUSTERED FILE SYSTEM (ACFS)**		
FILESYSTEM – DATABASE	**NFS**		
DATASET SIZE	**10 TB**		
RECORDS	**TABLE**	**RECORDS**	**TABLE SIZE**
	REGION	5	64K
	NATION	25	64K
	SUPPLIER	100M	15G
	CUSTOMER	1.5B	264G
	PART	2B	260G
	PARTSUPP	8B	1.3T
	ORDERS	15B	1.8T
	LINEITEM	60B	7.4T

TABLE 6. Database workload configuration

Tests

To illustrate FlashBlade's high throughput and scalability performance, the following tests were performed.

- **Data Ingest Test** – Load 10TB of data onto the DSS database, which tests the loading portion of the E-T-L operation. This test should showcase the speed and scale of data to be handled in environments like a data warehouse.
- **Blade Scalability Test** – Perform DSS queries across 7, 11, and 14 blades. This test should show not just the scalability aspect of FlashBlade, but the scale-out functionality that makes it very simple to add and gain performance.

Data Ingest Test Results

One of the objectives of the data load test is to attain the highest achievable data load based on the write bandwidth of FlashBlade.

- Highest achievable data load (ingest) rate

To simulate the real-world scenario for data loading where the data would predominantly be read from a storage system that is different from the target system, we setup another Pure Storage FlashArray system (FA //M50) to hold the data. This avoids reading and writing to the same, which would skew the metrics.

To improve the overall ingest time and to reduce read time of the source data, Oracle ASM Clustered File System (ACFS) was configured using 8 equally sized volumes from the //M50, which was mounted automatically on all the RAC nodes. FlashBlade was configured with 7 blades.

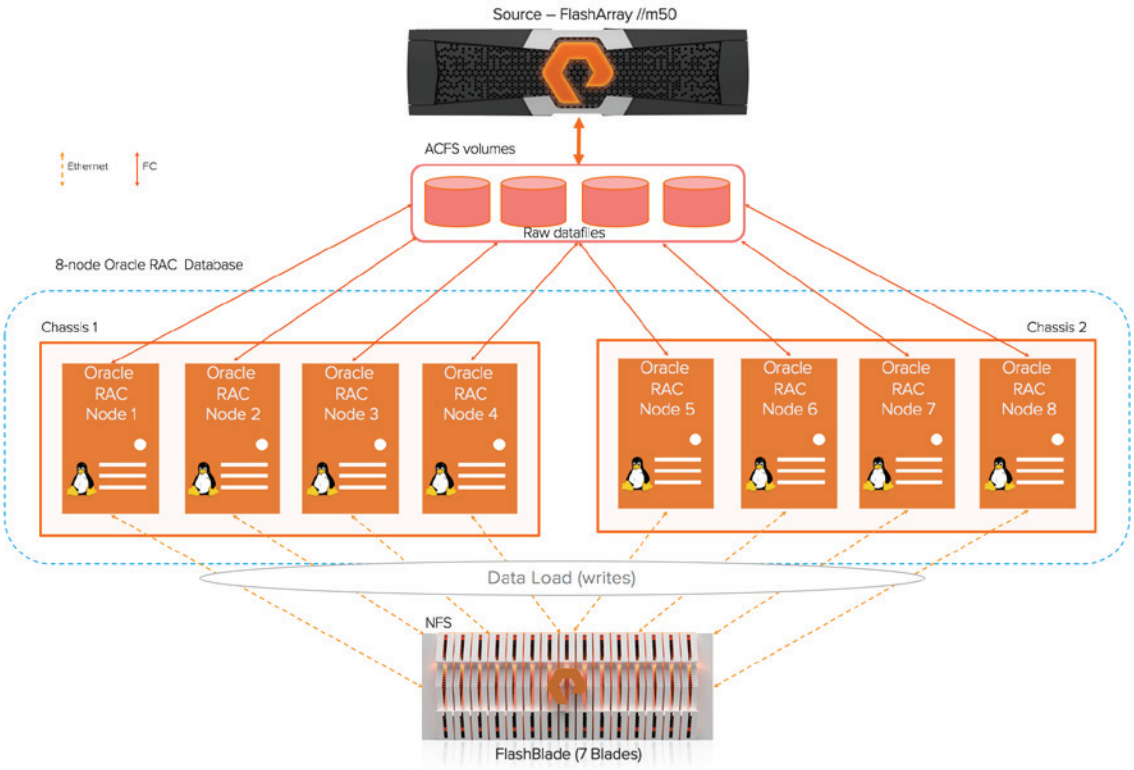


FIGURE 7. System connectivity

The following process was followed to generate the data files and load them onto Oracle.

1. ACFS was setup on Pure FlashArray //M50 using 8 volumes and mounted across all 8 nodes
2. Data was generated with the scale factor of 10,000 on ACFS
3. External tables were setup inside Oracle database pointing to the raw datafiles on ACFS
4. A permanent table structure was created on the database where the data will be loaded
5. Records from the external tables were selected using parallel clause and inserted into the permanent tables with parallel clause

The following table illustrates the time taken to load the data onto Oracle. The tables were loaded serially but each individual insert as select was run using parallel clause.

TABLE	RECORDS	SIZE	LOAD TIME (M:S)
REGION	5	64 KB	0:00.05
NATION	25	64 KB	0:00.06
SUPPLIER	100M	15 GB	0:10.16
CUSTOMER	1.5B	264 GB	1:54.18
PART	2B	260 GB	1:43.86
PARTSUPP	8B	1.3 TB	7:58.76
ORDERS	15B	1.8 TB	12:46.80
LINEITEM	60B	7.4 TB	56:14.05
TOTAL LOAD TIME			1H 20M 48S
TOTAL SIZE (TB)		10.85 TB	
OVERALL WRITE RATE			2.29 GB/S
INGEST RATE			8.05 TB/HR

TABLE 7. Time taken to load data onto Oracle

```
(Throughput)/(hour)=(Total Size (TB))/(Elapsed time (min) )*60=>10.85/80.8*60= 8.05 TB/hr
```

The data load test achieved **8.22 TB/hour ingest rate** on FlashBlade with 7 blades. The ingest rate will certainly go up with more blades in the FlashBlade system, as the maximum write rate a 15-blade system can achieve is around 3.75 GB/s, which equates to 13 TB/hr.

Blade Scalability Test Results

As part of the blade scalability tests, a set of business-oriented queries that examines large volumes of data with a high degree of complexity were identified (see Appendix A for the set of queries tested) and performed by varying the blade count in the Pure Storage FlashBlade. We also tested a full table scan query against the biggest table (**LINEITEM**) which had 60 billion records across 7, 11, and 14 blades on FlashBlade.

The following graph illustrates the first blade scalability test results. The total response time of five business-oriented queries that were run on the 8-node DSS database were plotted against 7, 11, and 14 blades.

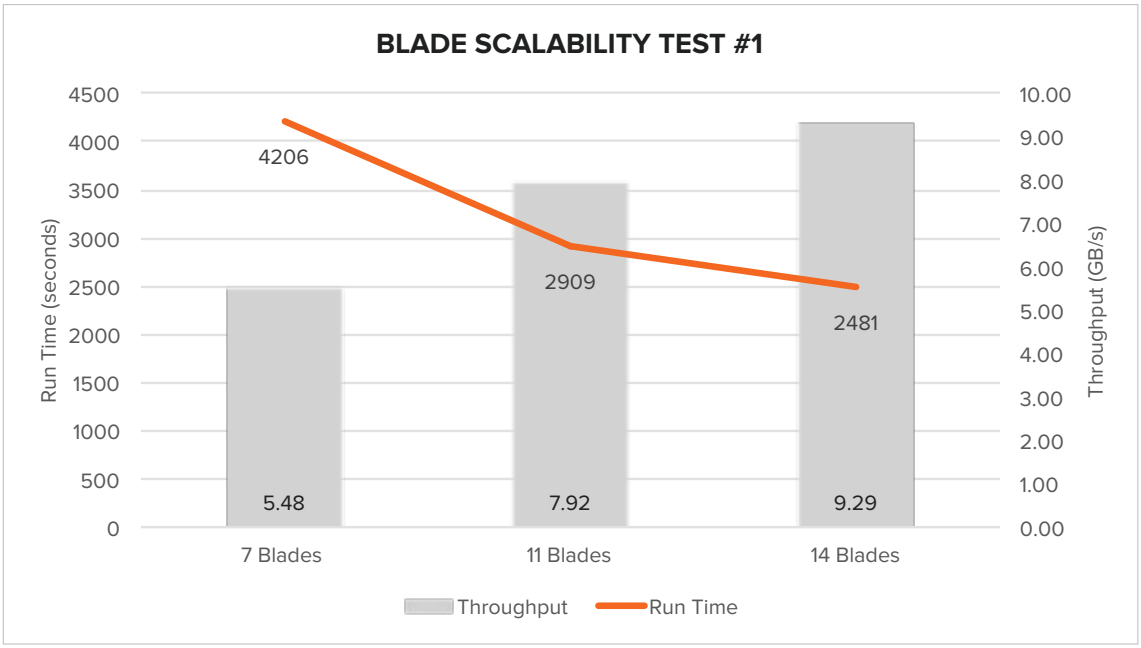


FIGURE 8. Blade Scalability Test #1

> Note: Throughput was calculated by dividing total physical reads performed by five business queries, divided by total run time. As the runtime is the end to end time reported by Oracle in executing the queries and not just limited to the I/O time waited by Oracle, the throughput doesn't reflect the maximum throughput FlashBlade can achieve; rather, it reflects only that for the queries.

The results show the scalability aspect of FlashBlade, as throughput went from 5.48 GB/s with 7 blades, to 7.92 GB/s with 11 blades, and 9.29 GB/s with 14 blades at 92% and 85% scalability.

The second blade scalability test was performed by running a full table scan against the biggest table (**LINEITEM**) that was 7.5TB in size with parallel clause, and captured the run time while varying blade counts on FlashBlade. In this case, most of the time was spent on "direct path reads", which is performing reads directly into the user space and bypassing Oracle's buffer cache.

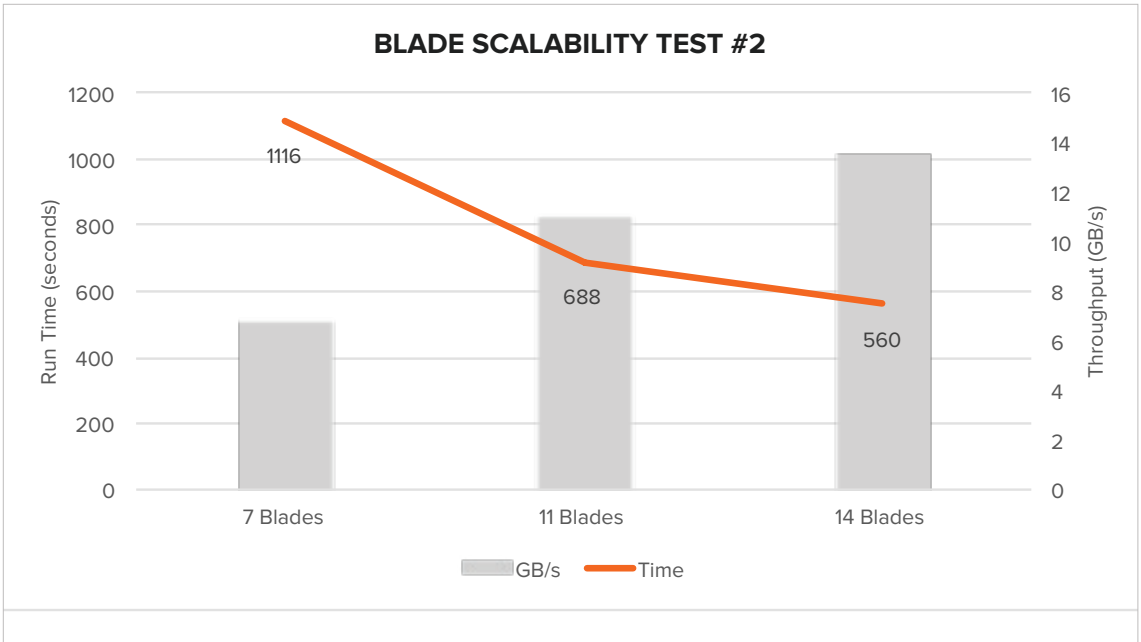


FIGURE 9. Blade Scalability Test #2

The results clearly show a linear scalability from 7 blades to 11 blades and 14 blades. Even though the throughput numbers are tad lower than what the FlashBlade team has seen in their internal tests, this still shows linear scalability. As the response time includes Oracle's end-to-end time, the throughput numbers don't reflect the maximum achievable throughput on FlashBlade.

Oracle also equally distributed the parallel queries across the 8 RAC nodes, which helped in extracting high performance across all the infrastructure components through which I/O passes between hosts and storage (I/O modules, Fabric Interconnects, 9K switches).

This test also validates FlashBlade's simplicity approach to scale out and gain performance.

NFS FILESYSTEMS FOR ORACLE

Networked filesystems provide the flexibility to mount the same filesystem across multiple hosts, which meets the shared storage requirement for Oracle RAC. NFS filesystems can certainly be used for single instances as well. Make sure the NFS protocol is selected when creating the filesystem on FlashBlade.

USE DNFS OVER KERNEL NFS

To scale up bandwidth on FlashBlade, enable numerous connections from the client rather than a single connection. Oracle's Direct NFS creates a separate connection to the storage system for every server process, as opposed to a single connection per mount point via Kernel NFS.

ENABLE PARALLELISM

To increase read and write bandwidth on FlashBlade, use client level parallelization techniques like parallel queries and multiple RMAN channels based on the CPU availability of your host in conjunction with dNFS. This increases the number of connections to FlashBlade, especially with dNFS.

USE MULTIPLE NETWORK INTERFACES

To enhance network bandwidth, be sure to have multiple network interfaces on the client. These multiple interfaces can be configured on a single subnet or on multiple subnets.

SINGLE SUBNET

IO performance on a dNFS environment with multiple interfaces and a single subnet is limited to the speed of the first interface that is picked by the OS. This is because the OS returns the first path when multiple paths are available on the subnet and thus traffic is always routed through the first path.

For example, in the setup below, FlashBlade has a single IP address on which the NFS filesystem will be mounted from the client, and the client has two interfaces.

```
Client interface      NFS server
ens7f0 10.21.108.193   10.21.108.10
ens7f1 10.21.108.194   10.21.108.10

[oracle@rlxora-b01-11-06 ~]$ route -n
Kernel IP routing table
Destination     Gateway         Genmask         Flags Metric Ref    Use Iface
0.0.0.0         10.21.108.1     0.0.0.0         UG    0      0        0 ens7f0
10.21.108.0     0.0.0.0         255.255.255.0   U     0      0        0 ens7f0 <-- first route
10.21.108.0     0.0.0.0         255.255.255.0   U     0      0        0 ens7f1 <-- route ignored
```

As per the routing table, traffic can go through the first interface (**ens7f0**) as the destination, and the mask fits for both routes. OS will invariably choose the first route.

In this case, to enhance bandwidth, it is recommended to use NIC bonding at the client level. Oracle support document (Doc ID 833481.1) provides other means to enable multiple paths in a single subnet using static routing, but it doesn't address the issue of availability when a NIC fails or a network cable is pulled, as the routing table will not be updated.

MULTIPLE SUBNETS

"Direct NFS client best practices recommend always to use multipaths in separate subnets."

> Oracle recommends using a separate subnet for each , and it supports up to four subnets. With multiple subnets, there is no need to bond the network interfaces to aggregate bandwidth across the available interfaces. The routing will be automatic in the case of multiple subnets.

```
Client interface      NFS server
ens7f0 10.21.108.193   10.21.108.10
ens7f1 10.21.107.194   10.21.107.10

[oracle@rlxora-b01-11-06 ~]$ route -n
Kernel IP routing table
Destination     Gateway         Genmask         Flags Metric Ref    Use Iface
0.0.0.0         10.21.108.1     0.0.0.0         UG    0      0        0 ens7f0
10.21.108.0     *               255.255.255.0   U     0      0        0 ens7f0
10.21.107.0     *               255.255.255.0   U     0      0        0 ens7f1
```

In this case, these are two dynamic routes and, based on the traffic, the route is selected automatically.

As such, if you decide to use multiple subnets, they should be configured on both the client and the FlashBlade sides. Multiple subnets can be configured in the FlashBlade GUI under **Network Settings**.

Be sure to update the **oranfstab** with the subnets and the mount point details. In an RAC environment, all RAC nodes should have the appropriate **oranfstab** file configured.

NFS VOLUMES AND MOUNT POINT REQUIREMENTS

It is not required to have as many NFS filesystems/volumes as subnets for dNFS to be effective. Also, it is not required to mount a single volume on to all subnets for dNFS to be effective either. Oracle dNFS reads the **oranfstab** and, based on the storage paths and mount details, it will create multiple paths when the database files are accessed.

For example, with two subnets and two mounts, dNFS would create four paths to the storage system for every server process.

LINUX MOUNT OPTIONS

For mounting the NFS filesystem on Linux, use the following mount options. Do not specify the **rsize**, **wsize** options as the system can get the default offered by FlashBlade, which is 524288.

rw, bg, nointr, hard, tcp, vers=3, nolock, noac, actimeo=0

ENABLING & DISABLING DIRECT NFS CLIENT

To enable the Direct NFS Client, the standard Oracle Disk Manager (ODM) library that supports Direct NFS Client should be used. It can be enabled as follows (beginning with Oracle 11.2).

```
cd $ORACLE_HOME/rdbms/lib
make -f ins_rdbms.mk dnfs_on
```

To disable the Direct NFS Client, perform the following:

```
cd $ORACLE_HOME/rdbms/lib
make -f ins_rdbms.mk dnfs_off
```

VERIFYING THE USE OF DIRECT NFS CLIENT

- If dNFS is enabled, the **alert.log** will show the following entry when the database is started:

```
Oracle instance running with ODM: Oracle Direct NFS ODM Library Version 4.0
```

- Check the dNFS server information from **v$dnfs_servers** view.

```
SQL> select svrname, dirname, mntport, nfsport, wtmax, rtmax from v$dnfs_servers;
SVRNAME               DIRNAME                MNTPORT    NFSPORT      WTMAX      RTMAX
-------------------- -------------------- ---------- ---------- ---------- ----------
FlashBlade           /rman01                    2049       2049     524288     524288
```

> Note: Even though dNFS is enabled, Oracle only mounts the volume/filesystem and opens the files when they are accessed. If no datafiles are accessed, then the above view will return no rows.

CONCLUSION

Oracle data warehouse deployments are not easy. Ensuring performance, availability, and low TCO are always top priorities for IT managers and DBAs. In this paper, we have demonstrated how to address those challenges using the right tools from Oracle running on the right infrastructure from Pure Storage. This solution enables your user community to perform ad-hoc analysis across all of the data in your Oracle data warehouse quickly, cost-effectively, and with a solution that scales out virtually on the fly.

Furthermore, through the use of an OLAP-like benchmark with varying users and read/update workload characteristics, we have also demonstrated:

- Excellent throughput performance, achieved across data warehouse workloads through high-end data ingest and business-related queries
- Data load and high-end ingest capability by loading data warehouse datasets with a higher degree of scaling factor (**8.22 TB/hour ingest rate** on FlashBlade with 7 blades)
- The scalability of FlashBlade, in which throughput went from 5.48 GB/s with 7 Blades to 7.92 GB/s with 11 blades and 9.29 GB/s with 14 blades, at 92% and 85% scalability

Overall, the test results prove that a data warehouse solution on Oracle RAC is an ideal use case for FlashStack Converged Infrastructure with FlashBlade. Excellent bandwidth, high throughput, and scalability were achieved – showcasing the versatility of FlashBlade within the FlashStack CI solution. Finally, factors like service profiles via UCS, Oracle dNFS, ease of storage management, and lower costs for power, cooling, and rack space make the TCO for large scale deployments even more attractive.

REFERENCES

The following documents and links were referred to in preparing this document.

1. **Cisco UCS Manager GUI Configuration Guide**
   https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/sw/gui/config/guide/2-2/b_UCSM_GUI_Configuration_Guide_2_2.html

2. **Oracle 12c Release 2 Documentation**
   http://docs.oracle.com/database/122/index.htm

3. **Various Knowledge base documents at Oracle Support** (My Oracle Support credentials required to access)
   https://support.oracle.com

4. **Pure Storage Support pages**
   http://support.purestorage.com/

APPENDIX A: DSS QUERIES

Following are the five business-oriented queries that were run as part of the blade scalability tests with FlashBlade on the 8-node Oracle RAC database. These were generated using the query generator tool from the benchmark kit.

QUERY 4

```
select /* qry 4.0 */ o_orderpriority, count(*)  order_count
  from orders
 where o_orderdate between '01-JUL-93' and '01-NOV-93'
   and exists ( select *
                  from lineitem
                 where l_orderkey = o_orderkey
                   and l_commitdate < l_receiptdate)
 group by o_orderpriority
 order by o_orderpriority;
```

QUERY 5

```
select /* qry 5.0  */ n_name,
        sum(l_extendedprice * (1 - l_discount))  revenue
  from customer,
       orders,
       lineitem,
       supplier,
       nation,
       region
 where c_custkey = o_custkey
   and l_orderkey = o_orderkey
   and l_suppkey = s_suppkey
   and c_nationkey = s_nationkey
   and s_nationkey = n_nationkey
   and n_regionkey = r_regionkey
   and r_name = 'ASIA'
   and o_orderdate between '01-JAN-94' and '01-JAN-95'
 group by n_name
order by revenue desc;
```

QUERY 11

```
select /* qry 11.0 */ * From (
select ps_partkey, sum(ps_supplycost * ps_availqty) as value
  from partsupp,
       supplier,
       nation
 where ps_suppkey = s_suppkey
   and s_nationkey = n_nationkey
   and n_name = 'GERMANY'
 group by ps_partkey
having sum(ps_supplycost * ps_availqty) > (
        select sum(ps_supplycost * ps_availqty) * 0.0000000100
          from partsupp,
               supplier,
               nation
         where ps_suppkey = s_suppkey
           and s_nationkey = n_nationkey
           and n_name = 'GERMANY')
 order by value desc)
where rownum <=100;
```

QUERY 13

```
select /* qry 13.0 */ c_count, count(*) as custdist
  from ( select c_custkey, count(o_orderkey) c_count
           from customer
                left outer join orders
             on c_custkey = o_custkey
            and o_comment not like '%special%requests%'
          group by c_custkey
        ) c_orders
 group by c_count
 order by custdist desc, c_count desc;
```

QUERY 16

```
select /* qry 16.0 */ * From (
       select p_brand, p_type, p_size,
              count(distinct ps_suppkey) as supplier_cnt
         from partsupp,
              part
        where p_partkey = ps_partkey
          and p_brand <> 'Brand#45'
          and p_type not like 'MEDIUM POLISHED%'
          and p_size in (49, 14, 23, 45, 19, 3, 36, 9)
          and ps_suppkey not in (
              select s_suppkey
                from supplier
               where s_comment like '%Customer%Complaints%')
        group by p_brand, p_type, p_size
        order by supplier_cnt desc, p_brand,
              p_type, p_size)
 where rownum <= 100;
```

ABOUT THE AUTHOR

Somu Rajarathinam is the Pure Storage Oracle Solutions Architect responsible for defining database solutions based on the company's products, performing benchmarks, and developing reference architectures for Oracle databases on Pure.

Somu has over 20 years of Oracle database experience, including as a member of Oracle Corporation's Systems Performance and Oracle Applications Performance Groups. His career has also included assignments with Logitech, Inspirage, and Autodesk, ranging from providing database and performance solutions to managing infrastructure, to delivering database and application support, both in-house and in the cloud.

Web: www.somu.us
Twitter: @purelydb

ps_wp39p_oracle-flashstack-with-flashblade _ltr_01