ORACLE

translab

Oracle Gold Partner

Oracle Private Cloud Appliance for Scalable Big Data Solution

Implementing Big Data Solution on PCA

January, 2021, Version 2.0

Public

Purpose statement

This document introduces Oracle Private Cloud Appliance as a scalable infrastructure for creating BigData solutions. We also present an overview of Cloudera Hadoop (CDP) benchmarking results performed on Oracle Private Cloud Appliance. It is intended solely to help you assess the business benefits of using PCA for Hadoop workloads and to plan your I.T. projects.

Disclaimer

This document in any form, software or printed matter, contains proprietary information that is the exclusive property of Oracle. Your access to and use of this confidential material is subject to the terms and conditions of your Oracle software license and service agreement, which has been executed and with which you agree to comply. This document and information contained herein may not be disclosed, copied, reproduced or distributed to anyone outside Oracle without prior written consent of Oracle. This document is not part of your license agreement nor can it be incorporated into any contractual agreement with Oracle or its subsidiaries or affiliates.

This document is for informational purposes only and is intended solely to assist you in planning for the implementation and upgrade of the product features described. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described in this document remains at the sole discretion of Oracle. Due to the nature of the product architecture, it may not be possible to safely include all features described in this document without risking significant destabilization of the code.

Disclaimer

The revenue recognition disclaimer on this page is required for any white paper that addresses future functionality or for products that are not yet generally available (GA). If you are unsure whether your statement of direction needs the disclaimer, read the revenue recognition policy. If you have further questions about your content and the disclaimer requirements, e-mail REVREC_US@oracle.com. If you do not need the disclaimer, you may delete it and the page that it appears. First, display hidden characters by clicking on the Paragraph icon on the Home toolbar. It is a small, square icon that appears to the left of the Quick Style Gallery. Then, highlight all the text on this page and press the Delete key. Notice that there is a section break displayed as a double-dotted line at the bottom of this page. Continue to press Delete until the page disappears and your cursor is on the Table of Contents page. Be sure not to remove the section break, or the formatting of the title page will be incorrect. Delete this note before publishing.

Table of contents

Purpose statement **2**
Disclaimer **2**
Disclaimer **2**
Executive Summary **5**
Oracle Private Cloud Appliance 5
Introduction to Big Data **6**
Cloudera **6**
Hadoop **7**
Apache Hadoop: Assumptions and Goals 7
Hadoop Components **7**
HDFS 7
YARN 8
MapReduce 8
Oracle Big Data Solutions **9**
Oracle Big Data Appliance (*No longer available for sale*) 9
Oracle Big Data Appliance Components 9
Hardware Components 9
Oracle Private Cloud Appliance 10
Oracle Private Cloud Appliance X8 Components 10
Hardware Components 10
Software Components 11
Support 12
Oracle for Hadoop Workloads **13**
Architecture 13
Ultra-Converged Infrastructure 13
Cloud-Native 14
100Gb Ethernet Network 15
DevOps tools 15
Fault Reduction 16
Other benefits provided by Oracle PCA 16
Hadoop Benchmarks on Oracle PCA **17**
TestDFSIO 17
TeraSort 17
KAFKA Perf Test 17
PCA Hadoop Architecture 18
PCA Configuration 18

Cloudera Components Distribution 18
Hadoop Configuration 19
Benchmark Results Summary **19**
Private Cloud Appliance 19
Improvement over BDA 20
Understanding the Results 20
Conclusions **21**
Final Thoughts 21
Authors **22**
About Translab Technologies **22**

Executive Summary

With multiple big data frameworks available on the market, choosing the right one is a challenge. A classic approach of comparing the pros and cons of each platform is unlikely to help, as businesses should consider each framework from the perspective of their particular needs. Oracle PCA is a multi-purpose Engineered System with High Density compute and ZFS Storage with 100Gbps backend Network. This appliance is already used extensively to as the Application/Middleware tier in business critical deployments. With direct connectivity to Oracle's flagship, Exadata database machine, we can create a robust unified data management solution.

The key objective of this document is to introduce Oracle Private Cloud Appliance (PCA) as the underlying infrastructure for building data-lakes using Cloudera-CDP to create scalable and easy to manage BigData solutions. It can be used for Hadoop/Spark workloads and streaming data processing. Oracle PCA is designed to run diverse workloads – from Hadoop-only workloads (Yarn, Spark, Hive etc.) to interactive SQL queries across Apache Kafka, Hadoop and NoSQL databases.

The document highlights the benchmark results obtained with Cloudera Data Platform (CDP) reference architecture implemented on Oracle PCA. The Benchmarking of different Hadoop applications was performed in the virtualized environment created on the Oracle Private Cloud Appliance (PCA). We used six VM's from six physical PCA compute-nodes to do the benchmarking. Storage was provisioned from the shared ZFS storage as block volumes to a Virtual Machine (VM).

Oracle Private Cloud Appliance

Traditional DYI infrastructure in the current IT landscape lack agility and introduces increased complexity because an efficient implementation requires tailor-made solutions and has dependency on distinct hardware/software components coming from multiple vendors. Organizations are looking at the cloud model to provide agility, speed and faster deployments.

Oracle Private Cloud appliance, an ultra-converged Engineered System, with Enterprise Manager providing a centralized management console, offers a compelling and elastic infrastructure solution to bring cloud-like flexibility to on-premises deployments. It is an appliance that comes pre-configured/pre-cabled from the factory. Oracle PCA uses Oracle Enterprise Manager to operate and manage the entire system as a single appliance; and not as individual servers, network and storage components. Provisioning new applications on Oracle PCA can be done in fraction of seconds with no compromises on security (PCIDSS compliant), performance and availability. It provides an open environment for innovation while maintaining tight integration and enterprise-level support. Oracle PCA is multitenant; it can be configured as a single cluster or as a set of clusters. This provides the flexibility customers need when deploying development, test and production clusters.

Introduction to Big Data

The current usage, as defined by industry analyst Doug Laney of Gartner, states that "Big Data" is data with large **V**ariety that is accumulating in huge **V**olumes and at high **V**elocity. The three Vs has since been expanded to also include **V**ariability and **V**eracity**.**

Volume: As cost of storage devices has dramatically reduced, organizations are not only collecting huge amounts of data from a wide variety of sources but also storing them for a longer period.

Velocity: Faster network connectivity has given rise to data streaming in from sources that were not available few years back (like, social media feeds, IoT sensors, mobile apps) in real-time at high speed.

Variety: As the complexity and type of application has grown, data now encompasses everything: traditional structured data, unstructured documents, emails, videos, audios, etc.

Variability: Data flows are unpredictable and changes depending on seasonality and event-based triggers.

Veracity: Veracity of data refers to the quality of the data and its different sources. As data is acquired from multiple sources with different formats, it must be matched, cleansed, and transformed for it to be useful.

Cloudera

Cloudera developed CDH (Cloudera's Distribution encompassing Apache Hadoop). CDH included various Big Data tools like Apache Hadoop, Apache Spark, Apache Flume, Apache Impala, Apache Kudu and Apache HBase needed by data scientists. In 2011, Hortonworks was founded by a group of engineers from Yahoo! Hortonworks released HDP (Hortonworks Data Platform), a competitor to CDH.

In 2019, Cloudera and Hortonworks merged, and the two platforms (CDH and HDP) were combined into a single platform: CDP (Cloudera Data Platform). CDP Private Cloud Base is the on-premises version of Cloudera Data Platform. It supports a variety of hybrid solutions where compute tasks are separated from data storage and where data can be accessed from remote clusters, including workloads created using CDP Private Cloud Experiences. CDP Private Cloud Base comprises a variety of components such as Apache HDFS, Apache Hive 3, Apache HBase, and Apache Impala, along with many other components for specialized workloads (see CDP Private Cloud Base (cloudera.com).

Some of the key components of CDP Private Cloud are:

- **Management Console** for managing environments, data lakes, environment resources, and users across all CDP services.
- **Environment**, which is a logical representation of the association between user account and underlying private cloud infrastructure.
- **Data Lake** for hosting the data with built-in security and governance mechanism.
- **Data Warehouse** to build a unified single source of truth with the organisation.
- **Machine Learning** for data scientists to develop, test, train, and deploy predictive applications.

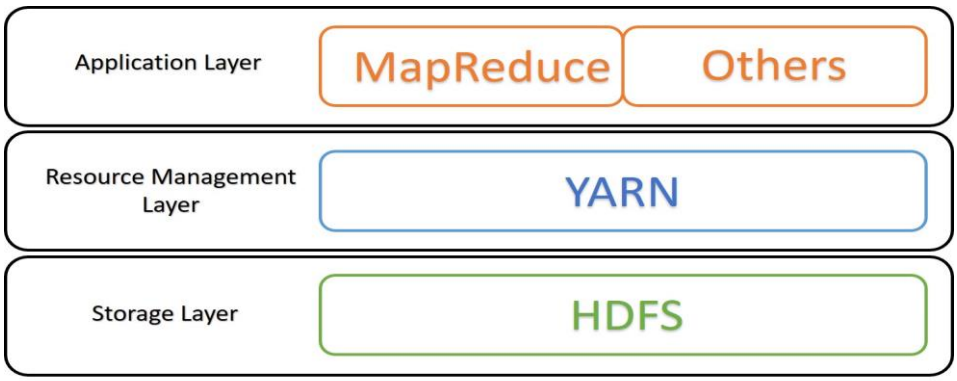
Hadoop

Apache Hadoop: Assumptions and Goals

Apache Hadoop Project started by making certain assumptions that were based on the prevalent technology. These assumptions came to define not only the architecture, but also its pros and cons.

- The first assumption was to make Hadoop run on commodity x86 architecture. By eschewing custom hardware, the project aimed to create an architecture that was both low-cost and scalable.
- The second assumption came from the first. As x86 hardware was neither very reliable nor very powerful in 2000s, in any scalable x86-based architecture, chances of hardware failure were non-trivial. This meant Hadoop had to detect faults and recover from them quickly and automatically.
- The limitations of networking and storage in 2000s meant that **"Moving Computation was Cheaper than Moving Data"**. It was more efficient to move the computational workload close to the data it operated on than send the data to the compute node that requested it. This is especially true when the system was expected to handle Petabytes and Exabyte of data. This assumption helped minimise network congestion and increase the overall bandwidth of the system.
- The final assumption was to have multiple copies of each file for data protection and computational performance. The number of copies of a file can be changed based on the need (**Tuneable Replication**).

Hadoop Components



Hadoop Architecture

HDFS

HDFS is short for Hadoop Distributed File System. It is the storage layer of Hadoop. It provides fault-tolerance and runs on commodity x86 hardware. To ensure fault-tolerance, data is stored as three separate copies across multiple nodes and server racks. This ensure, that if a node or, for that matter, an entire rack fails, the data is not lost.

It takes large datasets and divides it into smaller units called blocks. The blocks are distributed across the entire cluster. HDFS works on a master-slave architecture. The master node is called **NameNode**. It manages multiple slave nodes known as **DataNodes,** within the cluster.

NameNode maintains the file metadata (the file name, file permissions, IDs, locations, and the number of replicas) for the data blocks and forms the most-critical part of Hadoop. If a NameNode fails, HDFS cluster would lose the data in the DataNodes. Hence, an additional NameNode is generally available to mitigate this. This additional node can be in active-passive (Secondary NameNode) or active-active (Standby NameNode).

YARN

YARN stands for Yet Another Resource Negotiator. It is the default cluster management resource tool for Hadoop. YARN is composed of four components: Resource Manager**,** Node Manager, Application Master and Container**.**

Resource Manager runs on the NameNode and plays the most important role in resource allocation. It is tasked with the responsibility of allocating resources to the different applications that are running. It also accepts job submissions.

While Resource Manager is the most important YARN component running on NameNode, Node Manager is its counterpart in DataNodes. Each DataNode has its own instance of Node Manager running in on it. It enrolls itself with the Resource Manager and primarily tasked with managing the application containers entrusted to it by the resource manager.

When a job or application is submitted, it is managed by a unique Application Master. Application Master coordinates the application's operation and monitors its health.

The final component is Container, which represents a collection of physical resources such as RAM, CPU cores, and disks on a single DataNode.

MapReduce

MapReduce forms the logic that processes data spread across the Hadoop cluster. Whenever possible, MapReduce runs locally on the DataNode where the data is located to reduce bandwidth usage and improve cluster efficiency.

The data passes through three stages in the MapReduce algorithm: map, shift and sort, and reduce. In the mapping stage, data is ingested and split into data blocks called input splits. These splits are stored as unique key-value pairs.

Shuffle is the process where the outputs from mapping processes are copied to the reducer nodes. The output is then sorted into a single input file for the reduce phase. The shuffle and sort stages run in parallel.

In the final reduce stage, the input file aggregates the values based on the corresponding mapped keys. The output results in a new key-value pair. This is stored in HDFS.

Oracle Big Data Solutions

Most are keenly aware that Big Data is at the heart of nearly every digital transformation taking place today. Business analysts and data scientists are developing a host of new analytical techniques and models to uncover the value provided by this data. Big Data solutions are helping to increase brand loyalty, manage personalized value chains, uncover truths, predict product and consumer trends, reveal product reliability, and discover real accountability.

There is good reason for you to look to Oracle as the foundation for your Big Data capabilities. Since its inception, 35 years ago, Oracle has invested deeply across nearly every element of information management – from software to hardware and to the innovative integration of both on premises and Cloud-based solutions.

Oracle Big Data Appliance (*No longer available for sale*)

Oracle announced Big Data Appliance (BDA) in 2011 at Oracle OpenWorld event. It is an engineered system for running different Big Data workloads, including Hadoop.

BDA has a distributed architecture following the design principles of Hadoop. Each node has both compute and storage. It is certified to run Cloudera's distribution of Hadoop (both Cloudera 5.x and Cloudera 6.x). It includes hardware (servers, InfiniBand-based network fabric, rack, PDUs) and software (operating system, Java, Cloudera) in a single SKU for easy ordering and implementation.

BDA offers flexibility with respect to configuration and scalability, allowing users to start with 6 nodes and increase one node at a time. The system comes with 40Gbps InfiniBand based switching fabric for high-speed and low latency. Oracle Enterprise Manager is the single point of management for the entire Big Data Appliance.

Oracle Big Data Appliance Components

Hardware Components

Component	Full Rack	Starter Rack
Nodes	18 x Compute / Storage Nodes*	6 x Compute / Storage Nodes*
Leaf Switch	2 x 32 Port QDR InfiniBand Leaf Switch • 32 x InfiniBand 40Gb ports • 8 x 10Gb Ethernet ports	2 x 32 Port QDR InfiniBand Leaf Switch • 32 x InfiniBand 40Gb ports • 8 x 10Gb Ethernet ports
Spine Switch	1 x 36 Port QDR InfiniBand Spine Switch • 36 x InfiniBand 40Gb Ports	1 x 36 Port QDR InfiniBand Spine Switch • 36 x InfiniBand 40Gb Ports
Additional Hardware Components	• Ethernet Administration Switch • 2 x Redundant Power Distributions Units (PDUs) • 42U rack packaging	• Ethernet Administration Switch • 2 x Redundant Power Distributions Units (PDUs) • 42U rack packaging

Oracle Private Cloud Appliance

Oracle Private Cloud Appliance (PCA) is also an engineered system which provides much more flexibility. It is described an "**On-Premises Cloud Native Ultra-Converged Infrastructure"** for consolidating business-critical workloads. PCA offers seamless workload migration to Oracle Cloud Infrastructure to provide customers with a hybrid cloud solution.

PCA's full hardware and software stack is managed by Oracle Enterprise Manager and supported by Oracle. It offers full complement of software to provide public cloud-like experience, including multi-tenancy, metering, billing, etc. It supports both VMs and containers for full cloud-native experience. It supports both Infrastructure as a Services (IaaS) and Platform as a Service (PaaS). Trusted Partitioning is supported across the system to reduce database licensing cost. Zero downtime VM upgrades ensure that application services are available 24x7. For DevOps, Oracle PCA offers application portability with Oracle Linux Cloud Native Environment with built-in Kubernetes Engine. Oracle has also included automation tools to make Day 2 operations manageable.

Like public cloud, Oracle Private Cloud Appliance is also highly scalable. It can scale up to 1200 cores and 2.3PB of usable storage in a single rack. Moreover, storage can be increased by connecting up to 23 additional disk shelves. It supports up to 8 fully isolated tenant groups to maximize system utilization and security.

Oracle Private Cloud Appliance X8 Components

Hardware Components

Component	Full Rack	Base Rack
Nodes	• 2 x Management Nodes[1] • Up to 25 Compute Nodes[2]	• 2 x Management Nodes[1] • 2 x Compute Nodes[2]
Leaf Switch	2 x 36 Port 100Gb Ethernet Switch[3]	2 x 36 Port 100Gb Ethernet Switch[3]
Spine Switch	2 x 36 Port 100Gb Ethernet Switch[3]	2 x 36 Port 100Gb Ethernet Switch[3]
Storage	Oracle ZFS Storage ZS7-2[4]	Oracle ZFS Storage ZS7-2[4]
Additional Hardware Components	• 48-port 1Gb Ethernet Management Switch • 2 x Redundant Power Distributions Units (PDUs) • 42U rack packaging • Optional Storage Trays	• 48-port 1Gb Ethernet Management Switch • 2 x Redundant Power Distributions Units (PDUs) • 42U rack packaging • Optional Storage Trays

[1] Oracle Server X8-2 Management Nodes

The Oracle Server X8-2 Management Nodes are used to manage the PCA. These 2 servers operate in high-availability mode to ensure that there is no management downtime. Each node has 2x Intel Xeon-5218 2.3GHz 16-core processors (Total 32 physical cores), 384GB RAM, 2x 1.2TB HDDs in RAID1, 1x Dual-port 100Gbit Ethernet HCA (CX5), 1x Gb management port (BASE-T) and 2x 10Gb/25Gb onboard Ethernet ports.

[2] Oracle Server X8-2 Compute Nodes

The total quantity of Oracle Server X8-2 Computes nodes depends on the PDU capacity, PCA can scale up to 25 nodes. Each node has 2x Intel Xeon-8260, 2.4GHz 24-cores, 165 watts processors (Total 48 physical cores) and is available in three memory configurations (384GB, 768GB and 1.5TB RAM). There are 2x 1.2TB HDDs in RAID1 for system software. For connectivity, there is 1x Dual-port 100Gbit Ethernet NIC for data, a separate 1x GbE port (BASE-T) for management and 2x 10Gb/25Gb onboard Ethernet ports (currently unused). All compute nodes run Oracle VM Server for x86 as the hypervisor.

[4] Ethernet Switches

There are two pairs of 100Gbps Ethernet switches in the PCA system. These offer 100GbE network connectivity within the rack and can be configured to support to support 10/25/40Gbps speed if required.
The Leaf switch pairs use QSFP28 ports to provide communication network between the internal PCA hardware. This network connects the compute nodes, system disk, management servers
The Spine switch pair also uses QSFP28 ports and provides communication between the PCA and other systems (like, other Oracle Engineered Systems, external storage or customer's data center network.)

[4] Oracle ZFS Storage ZS7-2

PCA comes with an enterprise-grade Oracle ZFS Storage ZS7-2 that combines extreme performance and extreme capacity. Oracle ZFS Storage Appliance is designed to extract maximum storage performance from standard enterprise grade hardware while providing robust data protection, management simplicity, and compelling economics.

Software Components

- **Oracle VM Server (OVM)**
  The foundation of PCA is built on Oracle VM, a highly scalable, low overhead, hypervisor. It supports VMs with up to 128 vCPUs and running a wide variety of guest OS, like Linux, Oracle Solaris, and Microsoft Windows. Oracle VM (OVM) supports Oracle VM Templates which enable rapid deployment of pre-configured VMs and entire Oracle application stacks.
  Oracle VM also support "Trusted Partition". Oracle allows OVM Server to limit the number of Oracle Processor licenses required by allowing only a subset of total physical cores to license. In a Trusted Partition, two virtual cores are counted as a physical core. Unlike other hypervisor, only the maximum number of vCPUs running at any point of time must be licensed in OVM.
  By default, if an Oracle software is certified for use with Oracle VM or Oracle Linux, it is certified for use with Oracle Private Cloud Appliance. This includes the Oracle Linux CNE, Oracle Database, Oracle Fusion Middleware, Oracle Applications, and Oracle Real Application Clusters.

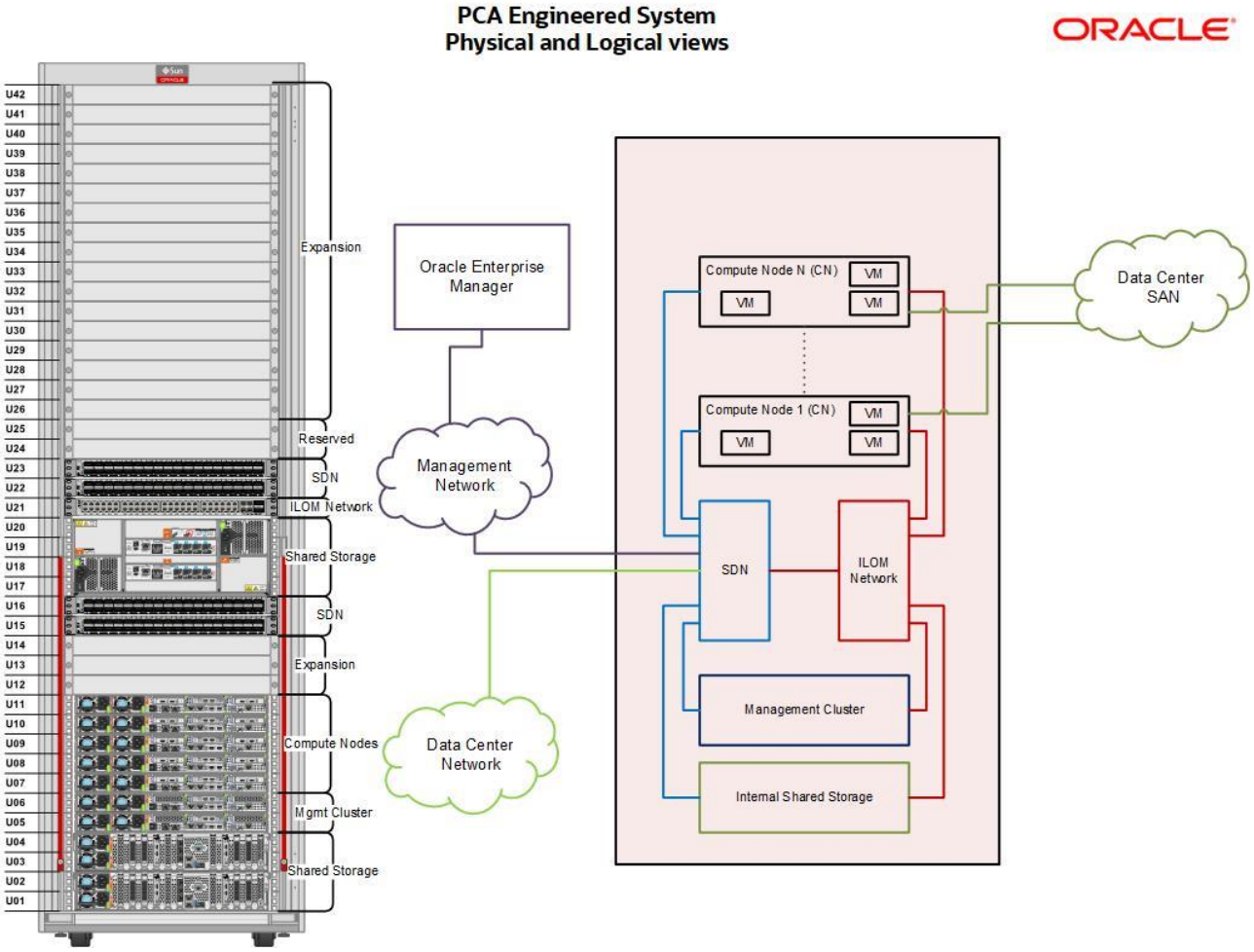
- **Oracle Private Cloud Appliance controller software**
  PCA controller software provide admins with a single pane of glass for management and monitoring of the hardware, upgrading software, deploying virtual resources (servers, networks, and storage). It provides telemetry data for monitoring resource utilization in real-time. It runs on two dedicated management nodes in high availability mode with automatic failover.

- **Oracle Enterprise Manager**
  Oracle Enterprise Manager is bundled with PCA. It allows rapid deployment and management an Oracle PCA. Once a PCA is power on and connected to network, Enterprise Manager automatically powers on and starts managing the internal hardware and can manage multiple PCA racks. IT department can set up internal Infrastructure as a Service (IaaS) and Database as a Service (DBaaS) for various LOBs to use. It has a self-service portal for users and developers to access to IaaS and DBaaS. Admins can maintain full control while offering agile cloud services. Metering and chargeback mechanisms are built in.

- **Oracle Linux**
  Oracle Linux is Oracle's own distribution of Linux for enterprise and cloud-native workloads. It is part of the PCA software stack.
  Built in with Oracle Linux is KSplice. With Kpslice, critical components of Oracle Linux installation can be updated without the need for system reboot. With zero-downtime updates, admins can ensure that their Linux environment is secured with latest patch without any disruption of business service.
  Other important component of Oracle Linux is Oracle Linux Cloud Native Environment (OLCNE). Oracle is a platinum member of both the Cloud Native Computing Foundation (CNCF) and the Linux Foundation. Oracle uses the Cloud Native Framework with container runtime environment for DevOps and Kubernetes for container orchestration. This simplifies and automates the process of deployment, scaling and management of Kubernetes clusters. In turn, this allows organizations adopt DevOps and agile methodology in their organization and modernize their application for cloud. DevOps projects can rapidly scale up and down compute and storage capacity as projects requirement changes.

Support

Oracle provides a single point of contact for all PCA components, hardware, and software. Oracle includes 24/7 remote monitoring for faults with notification within 5 minutes of any fault, restoration of service or escalation to development within 15 minutes and joint debugging with development team within 30 minutes. Oracle will also provide remote deployment of patch once every quarter.



Oracle for Hadoop Workloads

Oracle Private Cloud Appliance and Oracle Big Data Appliance are two engineered systems from designed for different business objectives. While BDA is purpose-built for running Big Data workloads like Hadoop, Oracle PCA is a highly flexible system to deliver private cloud. It is designed to run multiple types of application/middleware/database stacks, including Big Data workloads and analytics. PCA allows more modularity in scaling compute and storage. One can be increased without increasing the other. This leads to an optimum configuration of the Hadoop cluster.

Architecture

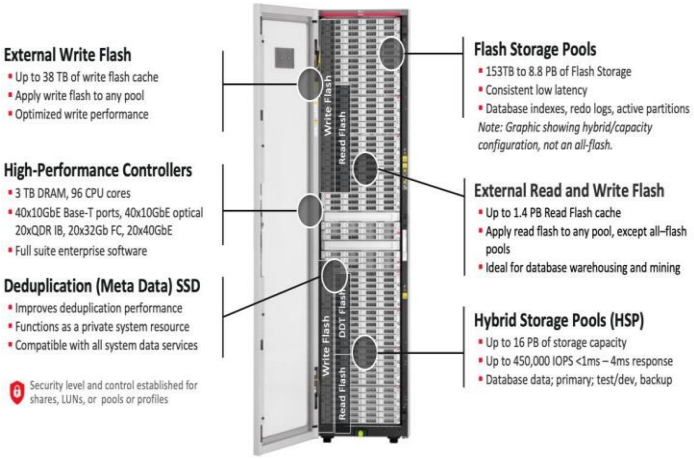
Ultra-Converged Infrastructure

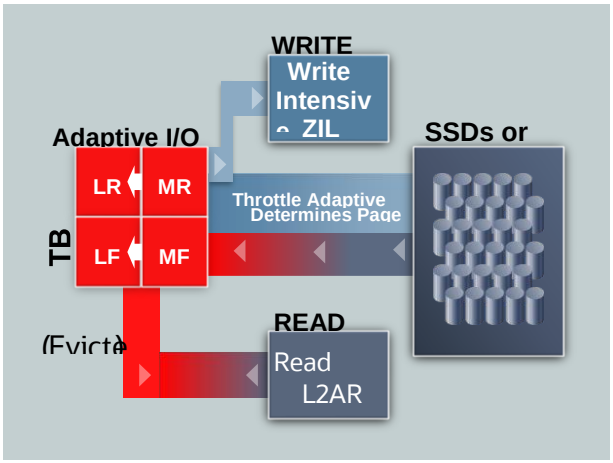
PCA is built on ultra-converged infrastructure. This means it has separate compute and storage devices which are connected over non-blocking 100Gbps Ethernet fabric provided by 4 100G switches. Each compute node has 200Gb and internal-storage-controllers has 160Gb of bandwidth. Since RAID/Mirroring is enabled by default, creating additional copies of data in ZS7-2 storage does not require Ethernet connectivity, this also reduces network bandwidth overhead. Since, Hadoop workload is more bandwidth dependent that IOPS dependent, PCA provides a greater advantage.

As the data is stored centrally on ZS7-2 which has dual-redundant controllers, the data needs be moved only once. While HDFS does keep multiple copies of data by design, the number of copies can be reduced from 3 to 2, thereby increasing storage efficiency. The same 100 TB of data can be stored in 250 TB of raw drive capacity (considering RAID overheads) with 4 copies of data.

The unique architecture of ZS7, based upon the Hybrid Storage Pool model and wide variety of advanced data services make the Oracle ZFS Storage Appliance an excellent choice for a wide variety of enterprise storage workloads that demand high performance.

The internal Oracle ZFS Storage ZS7-2 has 2x 7.68TB Readzilla SSDs for ReadCache, 20x 14TB serial-attached SCSI (SAS) HDDs (providing 100TB usable capacity) for data storage and 2x 200GB Write Flash Accelerators. It provides direct access to ZFS shares from virtual machines. It is geared for both, IOPS–intensive workloads like OLTP databases, and bandwidth-driven workloads such as data-warehousing and Hadoop. Management tools provide real-time analytics to visualize and drill down into specific workloads to understand data patterns and quickly root cause congestion scenarios. The storage capacity can be expanded up to 8PB or 16PB via external Oracle ZFS Storage Appliance (ZS7-HE).





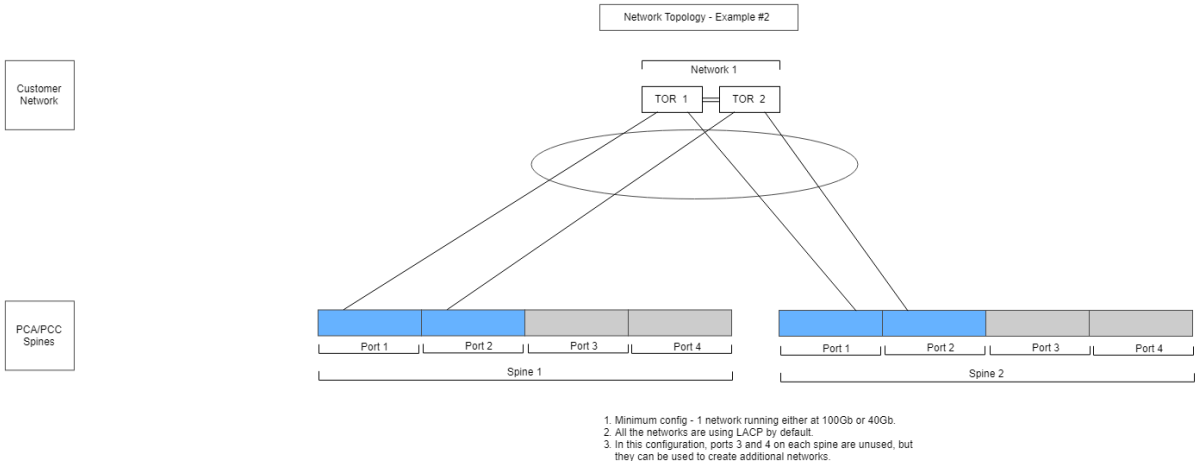
Oracle ZFS Storage Appliance for Backup Data Sheet

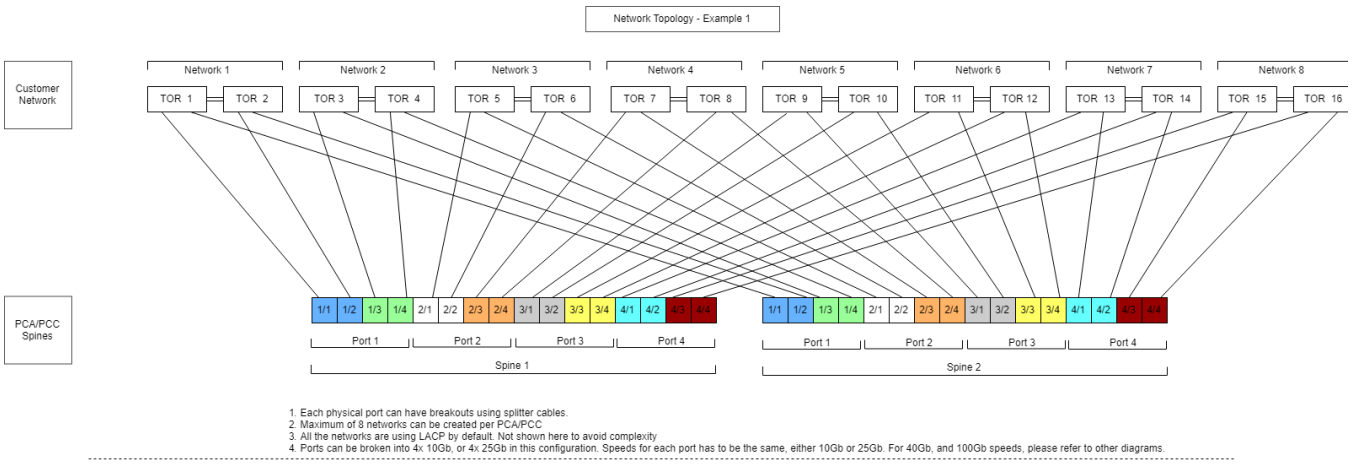
Cloud-Native

With the advancement in virtualization technologies, given sufficient resources, VMs are able to match bare-metal performance on many use cases. Oracle PCA enables VM support via the low-overhead OVM hypervisor and also supports containers which require significantly less resource overhead. Kubernetes has matured as orchestration tool for containers. PCA is built-on cloud-native principles and support containers and Kubernetes as shipped from factory. This makes PCA the tight choice for developers, operators, and businesses.

100Gb Ethernet Network

PCA uses 4x 36-port 100Gbps Ethernet switches for the network fabric. Each PCA rack provides 8x100Gb ports (4 from each Spine switch) to create custom networks. PCA's Ethernet switches implements Software Defined Networking (SDN) which permits creation of multiple isolated virtual networks on top of the same physical network. This allows network isolation for different tenants without investing in and managing multiple separate physical networks.





DevOps tools

PCA is built for DevOps and agile methodology. Various tools required for DevOps, like self-service portal, metering, automation, and billing, are included by default. Developers can easily provision resources needed without complicated approvals and workflows. Moreover, each business unit can be billed separately for cost control. Organisations benefit from quicker time to market and better total cost of operations. Kubernetes is support as shipped from factory, making deployment and management of containerised application easier.

Fault Reduction

Hadoop design makes assumption that all hardware will fail at some point. To mitigate this problem more nodes and more copies of data are needed within the cluster. When a particular node fails, the computation is just moved to another node where the secondary copy of data is available. As number of nodes increase, so does the probability of node failure. Hence, more redundancy and more inter-node traffic for data replication is required. This creates an overhead on the networking and reduces storage performance.

PCA is built to resist faults. It has no single point of failure in any of the components. Even the management node is fully redundant and highly available. If any single component fails, the cluster just routes around the failed component. This means fewer nodes are required to deliver the same capability/performance. KSplice is one of the modules of Oracle Linux which allows updating Oracle Linux without any need for system reboot. Zero-downtime updates ensure greater security and lower disruption.

Other benefits provided by Oracle PCA

At first glance, PCA's ultra-converged architecture looks different from traditional Hadoop distributed architecture. However, we believe that rather than being a drawback, it is a refreshing change needed for running new generation of Big Data workloads, including Hadoop.

First, PCA offers more comprehensive and modern tools for resource management than Hadoop. YARN is limited to scheduling of Hadoop jobs only. It cannot automate various administrative activities, like upgrading the software and applying new firmware patches. PCA can easily automate these activities using built-in automation tools.

Second, as more and more Big Data applications become cloud-native; an increasing number of DataOps teams are running Hadoop on hypervisor and cloud, instead of bare metal. PCA can also be a critical component of an organization's hybrid cloud

Third, HDFS supports only access control lists (ACLs) and a traditional file permissions model. PCA can integrate LDAP with HDFS to provide increased security.

Finally, Hadoop does not have encryption at rest or during transit. ZS7-2 Storage and the Ethernet switches support both encryption at rest and during transit, safeguarding the system and the data.

Hadoop Benchmarks on Oracle PCA

Several Benchmarks were performed on the Cloudera distribution with the following standard benchmarking tools:

- TestDFSIO
- Teragen, Terasort, and Teravalidate
- Kafka Perf Test

TestDFSIO

The TestDFSIO benchmark is a read and write test for HDFS. It is helpful for tasks such as stress testing HDFS, to discover performance bottlenecks in the network, to shake out the hardware, OS and Hadoop setup of the cluster machines (particularly the Name Node and the Data Nodes) and to get a first impression of how fast the cluster is in terms of I/O.

TeraSort

The TeraSort test sorts a large number of 100-byte records. It does considerable computation, networking, and storage I/O, and is often considered to be representative of real Hadoop workloads. A full TeraSort benchmark run consists of the following three steps:

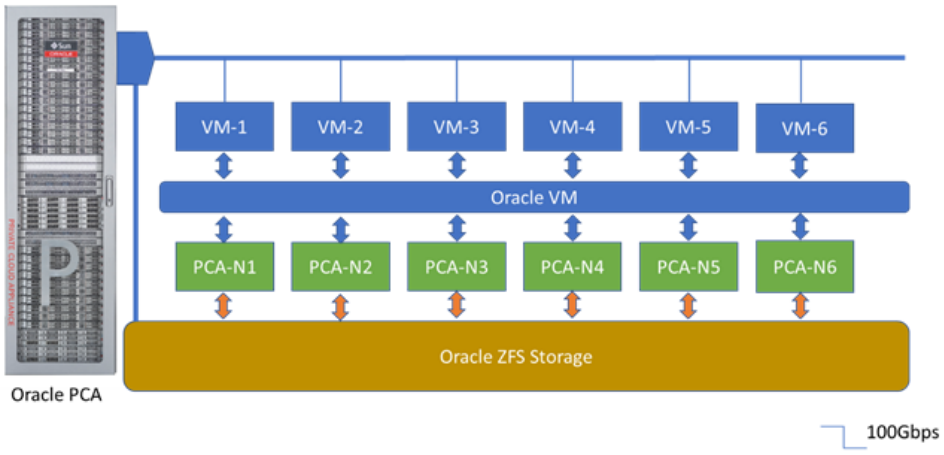
- Generating the input data via **TeraGen**
- Running the actual **TeraSort** on the input data
- Validating the sorted output data via **TeraValidate**
-

KAFKA Perf Test

Apache Kafka is a streaming message platform. It is designed to be high performance, highly available, and redundant. The kafka-perf-test tools is primarily used for:

- Measuring read and/or write throughput
- Stress testing the cluster based on specific parameters (such as message size)
- Load testing for the purpose of evaluating specific metrics or determining the impact of cluster configuration changes

PCA Hadoop Architecture

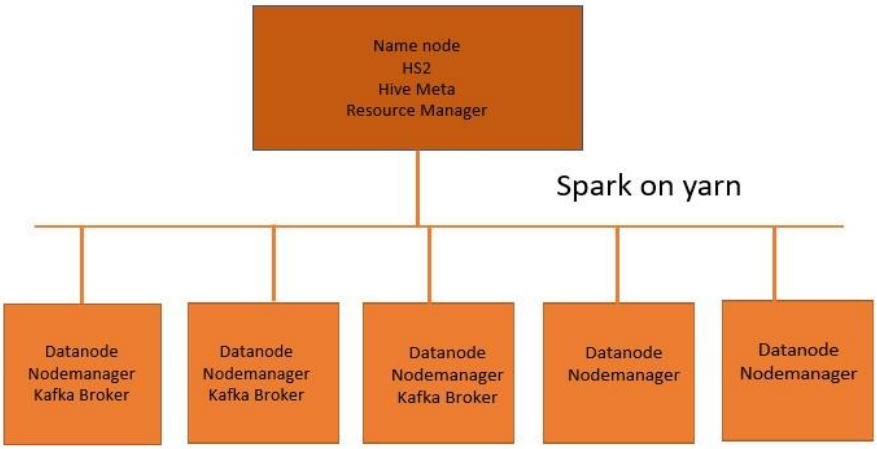


PCA Configuration

The Hadoop cluster was deployed on Oracle PCA with the following configuration:

- Nodes: CPU – 16 cores, RAM – 64 GB
- Storage: 6 *300GB volumes mapped directly from Oracle ZFS capacity pool
- Network: 100Gbps network connectivity between nodes; 160Gb (4x40Gb) bandwidth available for ZFS
- Hypervisor: Oracle OVM Server
- Operating System: Oracle Linux 7
- Hadoop: Cloudera7.1, Hadoop- 3.0, Spark 2.4, Kafka 2.2.1

Cloudera Components Distribution



Hadoop Configuration

Configuration item	Final tuning
dfs.namenode.handler.count	160
fs.trash.interval	default
io.file.buffer.size	default
yarn.nodemanager.resource.memory-mb (All nodes:- 50*5)	250GB
yarn.nodemanager.resource.cpu-vcores(Average)	50
yarn_scheduler_minimum_allocation_mb	1GB
yarn.scheduler.maximum-allocation-mb	32GB
mapreduce.map.memory.mb	2GB
mapreduce.reduce.memory.mb	4GB
mapreduce.map.java.opts	defaults
mapreduce.reduce.java.opts	defaukts
mapred_compress_map_output	TRUE
mapred_map_output_compression_codec	snappy
mapred_reduce_parallel_copies	32
mapreduce.task.io.sort.mb	512
mapreduce.map.sort.spill.percent	0.8
mapreduce.task.io.sort.factor	64
CDP version	7.1

Benchmark Results Summary

Private Cloud Appliance

The PCA benchmark was run on a six node PCA cluster. However, the resource was capped by YARN to 50 vCPUs and 250 GB. The following results were obtained:

Kafka Benchmark

Test Name	**Test Results**	**Throughput Average**
Single producer, no consumer, no replication	1 500 373 msg/sec	148 MB/sec
Single producer, no consumer,3x async replication	1 430 083 msg/sec	141MB/sec
Single producer, no consumer,3x sync replication	1 210 132 msg/sec	132 MB/sec
3producer, no consumer,3x async replication	3 118 082 msg	308 MB/sec
No producer, Single consumer	3 814 096 msg/sec	310 MB/sec
3 consumer, no replication	8 811 965 msg/sec	850 MB/sec

Benchmark	Size (in TB)	Time Taken (in sec)
TeraSort	1	760
DFSIO write	1	532
DFSIO read	1	388

Improvement over BDA

To have a proper comparison with BDA, we needed to normalise the data size and memory. Compute capacity is not required to be normalised as the difference is too small. *Ceteris paribus*, doubling RAM reduces time taken by 40% and doubling data size increases time taken by 1.5x.

Benchmark	Size (in TB)	Memory (in GB)	Improvement observed from BDA → PCA
TeraSort	1	250	60%
DFSIO write	1	250	45%
DFSIO read	1	250	41%

Understanding the Results

The principal factor driving the significant performance gains is the 100Gbps network infrastructure and the state-of-the-art ZFS storage appliance. Since Hadoop is throughput intensive, having 100Gbps network infrastructure leads to not only better storage performance but also compute performance.

Conclusions

The result shows that the Cloudera Hadoop can provide extremely compelling performance on the Oracle PCA Virtualization environment. Running more worker nodes on the Oracle PCA and using multi-pathing for storage can further improve the performance and scale.

Final Thoughts

Over the years Hadoop has gone through different versions and iterations. Many new concepts and modules in Hadoop have been developed in version 2.0 and 3.0.

PCA took these newer concepts and modules into its design consideration. It incorporates newer technologies, like 100Gbps Ethernet, containers and IT automation, for improved management. PCA enables rapid transition from idea to prototype to product. It gives organizations a better edge over competition. Self-service portal, metering and billing creates a private cloud for DevOps and DataOps.

PCA is equally equipped for Big Data, Cloud-native and traditional workloads, making it a Swiss Army knife of an organization's IT landscape. This make it a better investment for organizations.

Cloudera is modernizing their distribution of Hadoop to be containerized. This requires the underlying platform to also be container-ready and Kubernetes-ready. PCA has native support for containers and Kubernetes orchestration.

Finally, PCA architecture is a big improvement in performance and management. It simplifies Day 1 and Day 2 operations.

All these make PCA the right choice for Hadoop workloads

Authors

Gaganjit Singh Wazir, CEO at Translab Technologies
Neeraj G Dhopte, CTO at Translab Technologies
Shirshendu Bikash Mandal, VP, Technology Solutions at Translab Technologies

Lalit Bhola, Principal Product Manager for Oracle PCA/PCC at Oracle
Scott Ledbetter, Principal Solutions Engineer for Oracle PCA/PCC at Oracle

About Translab Technologies

Translab Technologies enables digital Transformation for enterprises worldwide by providing seamless customer experience, business agility and actionable insights. Utilizing disruptive technologies, like Big Data, Analytics, Internet of Things, Automation, Mobility and Cloud, we offer solutions that have applications across industry sectors. Translab Technologies is headquartered in Bangalore and provides support to its customers worldwide.

Translab has designed solutions for data-warehousing and Bigdata analytics for customers in BFSI, Insurance & healthcare to provide business insights on large volume of data that yield results for various use cases. Some of the key use cases are Customer360, Sentimental Behaviour Analytics, and Descriptive Analytics etc.

Translab Technologies has been Oracle Gold Partner with deep Expertise on Engineered systems, Private Cloud Appliance, databases, and Analytics Tools.

Translab has also partnered with Cloudera to address the bigdata challenges and has deep expertise to design and deploy Bigdata clusters using Cloudera technology.