



HPE Reference Architecture for Cloudera Data Platform Private Cloud Base on HPE EPA Asymmetric Architecture for Big Data Analytics

Using HPE Apollo r2600 Gen10 for Compute and HPE Apollo 4200
Gen10 Storage

CONTENTS

Executive summary.....	3
Solution overview.....	4
HPE Elastic Platform for Big Data Analytics.....	4
Cloudera Data Platform (CDP) Private Cloud Base.....	5
Regular (Base) clusters.....	6
CDP Private Cloud Base tools.....	6
HPE ProLiant DL360 Gen10.....	7
HPE Apollo r2600/XL170r Gen10.....	8
HPE Apollo 4200 Gen10.....	8
Aruba CX 8325.....	9
Aruba 6300F.....	9
Solution Architecture.....	10
Solution components.....	12
Software.....	13
Best practices and configuration guidance for solution.....	14
HPE DL360 Gen10, HPE Apollo r2600/XL170r Gen10, and HPE Apollo 4200 Gen10 settings.....	14
Red Hat Enterprise Linux settings.....	14
Cloudera Data Platform (CDP) Private Cloud Base Cluster Configurations.....	15
Testing and Validation of Deployment.....	16
TeraSuite test.....	17
Summary.....	18
Appendix A: Bill of materials.....	18
Resources and additional links.....	21



EXECUTIVE SUMMARY

Big Data Analytics have moved far beyond just utilizing legacy Hadoop clusters for running batch analytics. There are now a plethora of workloads highly dependent upon the Data Pipeline and Data Analytics, including interactive analytics, streaming analytics, AI/ML, etc. It's no surprise to any enterprise that data continues to grow exponentially and that data growth is accelerating as more data comes in off the edge from remote locations. The challenge is the majority of enterprises are struggling to effectively manage all this data and to ultimately exploit that data to deliver better business outcomes, improved customer experiences, and to deliver new business insights. Enterprises have to do more than to just store this data. They need to effectively manage that data across a data pipeline can enable advanced analytics, to train models, and to serve data intensive workloads. Hewlett Packard Enterprise (HPE) and Cloudera allow you to derive new business insights from all your data by providing a platform to store, manage, and process dense data at scale. This Reference Architecture provides several performance-optimized configurations for deploying Cloudera Data Platform (CDP) on clusters using Hewlett Packard Enterprise infrastructure that provides a significant reduction in complexity and a recognized increase in value and performance.

The HPE Elastic Platform for Analytics (EPA) is designed as a highly modular, disaggregated architecture to address the need for scalable, dense storage, and to act as an elastic multi-tenant Big Data Platform for on-premises deployment. It has a wide selection of building blocks based on density, capacity, and performance to reduce the total cost of ownership and data center footprint while optimizing performance for Big Data Analytics workloads. It supports a wide range of workloads ranging from Extraction, Transform and Load (ETL) processing offloaded from traditional data warehouses (DW), SQL-based interactive analytics, and near real-time event processing of data streams to machine and deep learning applications with GPUs.

Customers can select modular building blocks of compute, storage, and networking from Hewlett Packard Enterprise's diverse product portfolio, and integrate these blocks with software that enables an on-demand and elastic infrastructure foundation for Big Data. These building blocks can be deployed with co-located compute and storage on the same node (Symmetric) or disaggregated (Asymmetric) over a high-bandwidth network as more blocks of compute and storage are added. Hewlett Packard Enterprise supports two different deployment models under this platform:

- **HPE Workload and Density Optimized system (Asymmetric Architecture)** – HPE's recommended direction is to deploy Asymmetric or **Workload and Density Optimized (WDO)** architectures. These harness the power of a faster ethernet network that enables a building block approach to independently scale, compute, and storage, and lets you consolidate your data and workloads growing at different rates. In this testing the HPE WDO system is based on the HPE Apollo 4200 storage block and the HPE Apollo r2600/XL170r compute block. HPE ProLiant DL3XX is also often used as a compute node in these big data analytics deployments.
- **HPE Balanced and Density Optimized (BDO) system (Symmetric Architecture)** – Supports Hadoop deployments that scale, compute, and storage together, with some flexibility in the choice of memory, processor, and storage capacity. This frequently utilizes the HPE ProLiant DL380 server platform, with density optimized variants using HPE Apollo 4200 and HPE Apollo 4510 servers.

This Reference Architecture has been created to assist in the rapid design and deployment of Cloudera Data Platform Private Cloud Base v7.1.7 on HPE Elastic Platform for Big Data Analytics (EPA) with Asymmetric architecture for various sizes of HDFS clusters. This paper highlights recognizable benefits and provides guidance on building Cloudera Data Platform for HDFS clusters that meets business needs.

Document purpose: This Reference Architecture provides comprehensive architectural guidelines and implementation of Big Data Analytics solutions on HPE EPA asymmetric architecture using Cloudera Data Platform Private Cloud Base 7.1.7. In addition to outlining the key solution components, this document also provides guidelines for optimizing infrastructure configuration as well as performance analysis for distributed training and streaming analytics.

Target audience: This document is intended for subject matter experts, domain experts, data engineers, IT managers, pre-sales engineers, services consultants, partner engineers, and customers who are interested in implementing Big Data Analytics workloads for batch and real-time processing capabilities in their existing or new deployments.

This document describes the high-level design, performance results, and best practices for deploying Cloudera Data Platform Private Cloud Base on bare-metal HPE infrastructure.

The purpose of this document is to describe a Reference Architecture, highlighting recognizable benefits to technical audiences.

This Reference Architecture describes solution testing performed in December 2022.



SOLUTION OVERVIEW

Modern big data analytics environments must operate at large scale with high performance and are usually implemented as scalable clusters with tightly coupled compute and storage. As these environments grow, operational and infrastructure complexity can limit the agility and flexibility that are required to support changing workloads and storage demands.

One technique used to simplify these environments is to separate the compute and storage functions. This approach allows independent scaling and management of compute and storage but must be carefully designed to avoid performance bottlenecks.

This architecture describes an implementation of Cloudera Data Platform (CDP) Private Cloud Base 7.1.7 on HPE infrastructure with independent compute and storage.

The Reference Architecture configurations are based on the Cloudera Data Platform Private Cloud Base version 7.1.7 and HPE EPA WDO systems which include the HPE ProLiant DL360 Gen10, and HPE Apollo r2600/XL170r Gen10 and HPE Apollo 4200 Gen10 servers. Our solution, HPE EPA combining with Cloudera Data Platform addresses the challenges and helps the organizations implement Big Data Analytics solution on-prem with Hadoop Distributed File System (HDFS).

This document describes the high-level design, performance results, and best practices for deploying Cloudera Data Platform Private Cloud Base on bare-metal infrastructure with HPE Apollo 4200 as storage and HPE Apollo r2600/XL170r servers as compute.

HPE Elastic Platform for Big Data Analytics

Historically, Hadoop and HDFS have been deployed with storage and compute co-located and the main workload was a batch analytics process. The challenge with this design model was that as the cluster scaled out, compute and storage grew together, and typically one of the resources was over-provisioned to satisfy the needs of the other, which increased cost unnecessarily. With the rapid advancement of the next generation analytics applications and workload patterns (for example, streaming analytics, interactive analytics, end-to-end data pipelines, ML, AI, etc.), the architecture requires a disaggregated, scalable, and flexible design running on a high-performing platform.

Figure 1 illustrates an example of the data flow from the edge to the core to cloud, routed through a data pipeline that provides an infrastructure for data to not only flow bi-directionally but also allow for the implementation of analytic processes in real-time, near real-time, at rest, and AI modeling.

- Edge Acquisition and Edge Analytics are where data are acquired, processed, queued at the edge, and replicated in real-time to the core.
- Core/Cloud Analytics is deployed at the core data center or the cloud and integrated with an edge, it is where data is streamed, enriched, and analyzed in real-time, as well as being retained indefinitely for future analytics operations.
- AI/ML training is deployed at the core data center or in the cloud and is where the processed data is available for data scientists and data technicians, to understand the data, build machine learning models, train, and validate the models.
- Data Lake and Archives are deployed at the core data center or in the cloud to provide optimized storage; it is where all the ETL is performed and data is prepared for machine learning and deep learning in particular.



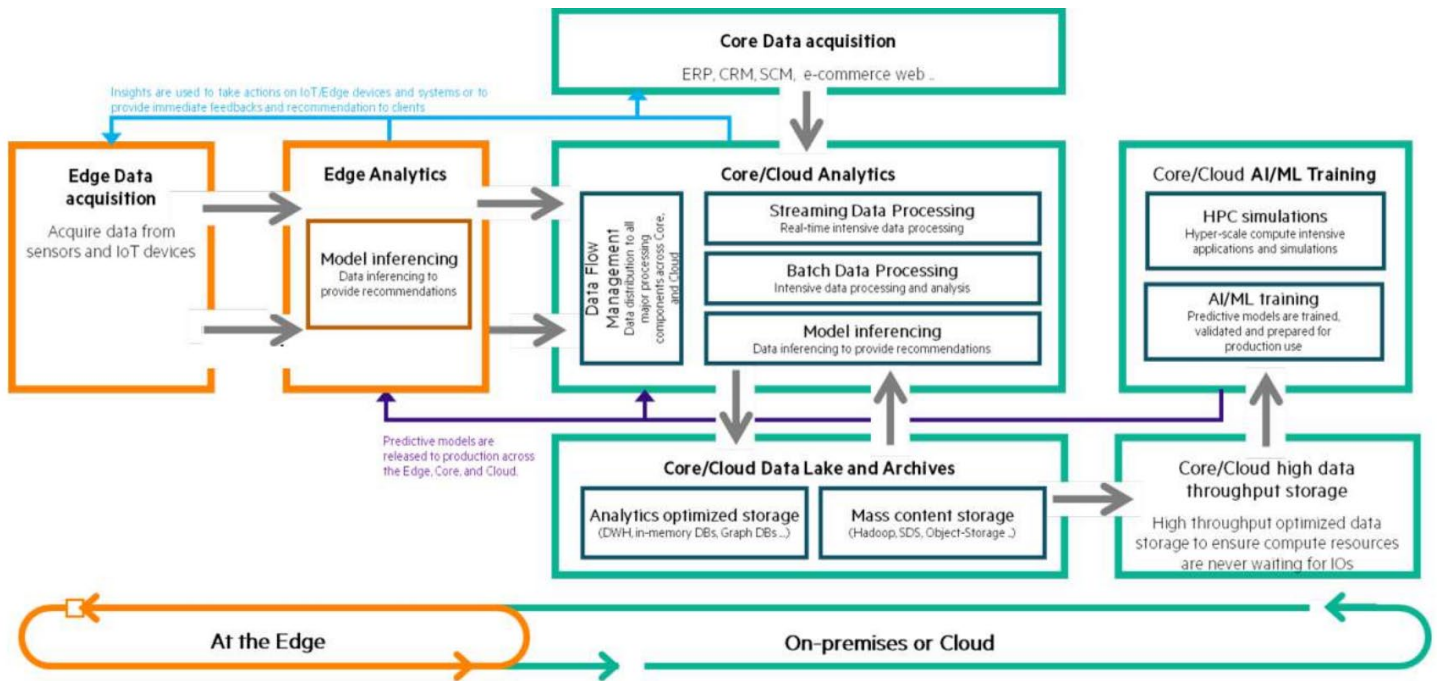


FIGURE 1. HPE Elastic Platform for Big Data Analytics (EPA) end-to-end data pipeline

For more information, refer the HPE Reference Configuration for Elastic Platform for Analytics (EPA) document located at, <https://www.hpe.com/psnow/doc/4AA6-8931ENW>.

Cloudera Data Platform (CDP) Private Cloud Base

Cloudera Data Platform (CDP) Private Cloud Base is an on-premises version of Cloudera Data Platform.

This new offering combines the best of Cloudera Enterprise Data Hub and Hortonworks Data Platform Enterprise, as well as new features and enhancements throughout the stack. This unified distribution is a scalable and adaptable platform through which we can securely run a wide range of workloads.

CDP Private Cloud Base supports a variety of hybrid solutions, including workloads created with CDP Private Cloud Data Services, in which compute tasks are separated from data storage and data can be accessed from remote clusters. By managing storage, table schema, authentication, authorization, and governance, this hybrid approach serves as a foundation for containerized applications.

The CDP Private Cloud Base shown in Figure 2 is made up of many different components, including Apache HDFS, Apache Hive 3, Apache HBase, and Apache Impala, as well as many others for specialized workloads. You can use any combination of these services to build clusters that meet your specific business needs and workloads. For common workloads, several pre-configured service packages are also available.





FIGURE 2. Architecture of Cloudera Data Platform (CDP) Private Cloud Base

These include the following:

Regular (Base) clusters

Develop and serve predictive models using the data engineering process. HDFS, Yet Another Resource Negotiator (YARN), YARN Queue Manager, Ranger, Atlas, Hive, Hive on Tez, Spark, Oozie, Hue, and Data Analytics Studio Data Mart were among the services provided. Interactively browse, query, and explore your data.

Data engineering

Process develop and serve predictive models.

Services included: HDFS, YARN, YARN Queue Manager, Ranger, Atlas, Hive, Hive on Tez, Spark, Oozie, Hue, and Data Analytics Studio

Data mart

Browse, query, and explore your data in an interactive way.

Services included: HDFS, Ranger, Atlas, Hive, and Hue.

Operational database

Real-time insights for modern data-driven business.

Services included: HDFS, Ranger, Atlas, and HBase.

Custom services

When you install a CDP Private Cloud Base cluster, you only need to install one parcel called Cloudera Runtime, which contains all of the components. See Cloudera Runtime Component Versions for a complete list of the included components.

CDP Private Cloud Base includes powerful tools to help manage, govern, and secure your cluster in addition to the Cloudera Runtime components.

CDP Private Cloud Base tools

Cloudera Manager

Cloudera Manager is used by CDP Private Cloud Base to manage one or more clusters and their configurations, as well as to monitor cluster performance. Cloudera Manager is also used for managing installations, upgrades, maintenance workflows, encryption, access controls, and data replication. Cloudera Manager can also be used to manage Cloudera Enterprise CDH clusters. Cloudera Manager can be used to set up a virtual private cluster, which allows you to separate compute resources from data storage and share data storage among compute resources.



Apache Atlas

Apache Atlas, which is used to provide data governance, is also included in the CDP Private Cloud Base. Apache Atlas acts as a common metadata store, allowing metadata to be exchanged both inside and outside of the Hadoop stack. Because Atlas and Apache Ranger are tightly integrated, you can define, administer, and manage security and compliance policies consistently across all components of the CDP stack. Atlas replaces the Cloudera Navigator Metadata Server for Cloudera Enterprise customers. It offers the following capabilities:

- Flexible metadata models
- Entity search using model attributes, classifications (tags), and free text
- Lineage across entities is based on processes applied to the entities

Apache Ranger

Apache Ranger manages your CDP Private Cloud Base cluster auditing, authentication, and authorization.

Apache Ranger offers a centralized framework for collecting access audit history and reporting data, including parameter filtering. Through this centralized reporting capability, Ranger enhances audit information obtained from CDP components and provides insights.

Apache Ranger also manages access control via the user interface, ensuring that policies are applied consistently across CDP Private Cloud Base components. Security administrators can define database, table, column, and file security policies, as well as manage permissions for specific LDAP-based groups or individual users. Dynamic conditions such as time or geolocation can also be used to supplement an existing policy rule. Using a service-based definition, the Ranger authorization model can be easily extended to any data source.

For customers familiar with Cloudera Enterprise, Apache Ranger replaces Sentry and Navigator Audit Server and also provides the following capabilities:

- Better fine-grained access controls:
 - Dynamic Row Filtering
 - Dynamic Column Masking
 - Attribute-based Access Control
 - SparkSQL fine-grained access control
- Rich policy features
 - Allow/Deny constructs, Custom policy conditions/context enrichers, time bound policies, Atlas integration (for tag-based policies)
- Extensive Access Auditing with rich event metadata

HPE ProLiant DL360 Gen10 server

The HPE ProLiant DL360 Gen10 server delivers security, agility and flexibility without compromise. It supports the Intel® Xeon® Scalable processor with up to a 60% performance gain and 27% increase in cores, along with 2933 MT/s HPE DDR4 SmartMemory supporting up to 3.0 TB with an increase in performance of up to 82%. With the added performance that Intel® Optane™ persistent memory 100 series for HPE, HPE NVDIMMs and 10 NVMe bring, the HPE ProLiant DL360 Gen10 means business. Deploy, update, monitor and maintain with ease by automating essential server life cycle management tasks with HPE OneView and HPE Integrated Lights Out 5 (iLO 5). While HPE ProLiant Gen10 servers were used for this testing, HPE also supports the use of ProLiant DL3XX Gen10 Plus and Gen11 servers with similar configurations for compute nodes.



FIGURE 3. HPE ProLiant DL360 Gen10



HPE Apollo r2600/XL170r Gen10

The HPE Apollo r2600 Gen10 server, like its predecessor, is designed for Big Data, High-Performance Computing, and a host of other critical computing tasks that can take advantage of this density-optimized system. There are two HPE ProLiant blade servers supported on this system, the half-height, half-width HPE ProLiant XL170r Gen10 server and the full-height, half-width HPE ProLiant XL190r Gen10 server. Utilizing a standard form factor, this system will easily integrate with traditional data centers driving higher levels of compute density, efficiency, and system scaling.

The HPE R2600 Gen10 chassis supports the server nodes with power, cooling, and storage, while the server nodes provide the compute and I/O. Two power supply units up to 2200W provide power for up to four nodes. Four single rotor fans provide cooling but can be upgraded with four additional fans for redundancy. An optional Rack Consolidation Management module (RCM) on the back of the system enables iLO Aggregation and enables the chassis to be daisy-chained together and connect to a top of chassis management switch. Up to 20 r2600 Gen10 chassis can be installed in a standard 42U enclosure for up to 80 individual servers if outfitted with the HPE Apollo R2600 Gen10 chassis. HPE ProLiant Gen9 XL170r and HPE ProLiant Gen9 XL190r server nodes are not compatible with the HPE Apollo R2600 Gen10 Chassis, and HPE ProLiant Gen10 server nodes are not compatible with the Gen9 chassis.



FIGURE 4. HPE Apollo R2600 XL170r Gen10

HPE Apollo 4200 Gen10

The HPE Apollo 4200 Gen10 server offers an architecture optimized for Big Data Analytics, Software-Defined Storage, backup and archive, and other data storage - intensive workloads. Its unique, easily serviceable 2U design saves data center space with up to 28 LFF or 54 SFF hot-plug drives. It delivers accelerated performance with a superior bandwidth and balanced architecture, Intel® Xeon® Processors, and NVMe connected SSDs. The focus on security extends from FIPS 140-2 Level 1 validated storage controllers down to the system silicon level, taking full advantage of HPE innovations in firmware protection, malware detection, and recovery. With HPE GreenLake Flex Capacity and HPE Financial Services, you can combine the economic agility benefits of consumption-based IT with the performance and security of on-premises. While Apollo 4200 Gen10 servers were used for this testing, HPE also supports the use of Apollo 4200 Gen10+ and Alletra 4120 Gen11 servers with similar configurations for storage nodes.



FIGURE 5. HPE Apollo 4200 Gen10



Aruba CX 8325

The Aruba CX 8325 Switch Series offers a flexible and innovative approach to addressing the application, security, and scalability demands of the mobile, cloud and IoT era. These switches serve the needs of the next generation core and aggregation layer, as well as emerging data center requirements at the Top of Rack (ToR) and End of Row (EoR). They provide over 6.4Tbps of capacity, with line-rate Gigabit Ethernet interfaces including 1Gbps, 10Gbps, 25Gbps, 40Gbps, and 100Gbps. The Aruba CX 8325 series includes industry-leading line rate ports 1/10/25GbE (SFP/SFP+/SFP28) and 40/100GbE (QSFP+/QSFP28) with connectivity in a compact 1U form factor. These switches offer a fantastic investment for customers wanting to migrate from older 1GbE/10GbE to faster 25GbE, or 10GbE/40GbE to 100GbE ports.



FIGURE 6. Aruba CX 8325 Switch

Aruba 6300F

The Aruba CX 6300 Switch Series is a modern, flexible, and intelligent family of AOS-CX stackable switches ideal for access, aggregation, and data center top-of-rack (TOR) deployments. With a cloud-centric design that combines a fully programmable OS with the Aruba Network Analytics Engine, the Aruba CX 6300 extends industry-leading monitoring and troubleshooting capabilities to the access layer. Support of Aruba NetEdit and the Aruba CX Mobile App verify that configurations are flawless and easy to deploy.

A powerful Aruba Gen7 ASIC architecture delivers fast, non-blocking performance, meaning your network is ready for tomorrow's unpredictable demands. Aruba Virtual Stacking Framework (VSF) allows for stacking of up to 10 switches, providing scale and simplified management. This flexible series has built-in high-speed uplinks and supports high density IEEE 802.3bt high power PoE with HPE Smart Rate multi-gigabit Ethernet for high-speed APs and IoT devices.



FIGURE 7. Aruba 6300F Switch



Solution architecture

The HPE EPA building blocks are the foundation of this Reference Architecture. The components can be combined in various ways to solve and address unique requirements. This section explains the detailed storage/compute, and control blocks based on the HPE ProLiant DL360 Gen10 and HPE ProLiant XL170r Gen10 servers required to build and implement a Cloudera Data Platform (CDP) Big Data Analytics solution on the HPE EPA WDO Platform. The blocks defined in this section may be modified (for example, processor model, memory, etc.) to address or meet the customer requirements accordingly.

The HPE EPA WDO solution infrastructure blueprints are composed of five blocks: control blocks, storage, compute blocks, network blocks, and rack blocks.

Following is the infrastructure of building blocks required to implement the solution with HDFS:

In this solution architecture, four HPE ProLiant DL360 servers, twelve Apollo r2600 compute servers, four Apollo 4200 storage servers and two Ethernet switches were used to build Cloudera Data Platform (CDP) Private Cloud Base Cluster. Deployment of the cluster should be planned and this reference link [CDP Private Cloud Base Installation Guide | CDP Private Cloud \(cloudera.com\)](#) provides guidance to build the cluster.

Figure 8 depicts the single-rack configuration for HPE EPA for WDO solution with Cloudera Data Platform (CDP) Private Cloud Base.

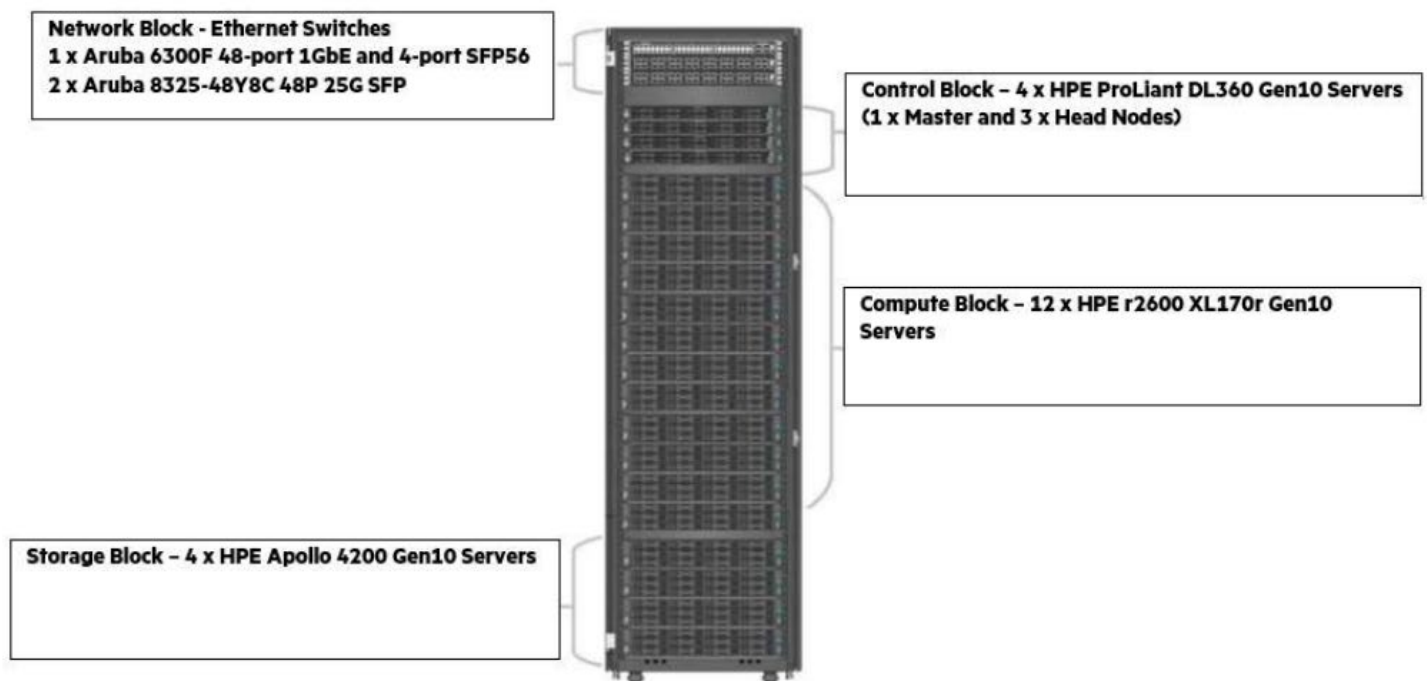


FIGURE 8. Basic conceptual diagram of single rack for HPE EPA with Cloudera Data Platform (CDP) Private Cloud Base with HDFS

Initially all the servers were deployed with Red Hat Enterprise Linux 7.9 Operating system and prerequisites were configured to meet the requirements. It is important to prepare each node before deploying the Cloudera Data Platform (CDP) Private Cloud Base cluster.



TABLE 1. Nodes for each Block

Block	Model
Control Block	HPE ProLiant DL360 Gen10
Storage Block	HPE Apollo 4200 Gen10
Compute Block	HPE Apollo r2600/XL170r Gen10
Network Block	Aruba 8325-48Y8C 48p 25G SFP+/28 8p 25G QSFP Aruba 6300F 48-port 1GbE and 4-port SFP56
Rack Block	1200mm or 1075mm

In this solution, the Cloudera CDP Cluster was designed with four Master nodes, twelve compute nodes, four storage nodes and two network switches. Table 2, Table 3, and Table 4 show the services configured on the servers and roles.

TABLE 2. HPE Cloudera Data Platform 7.1.7 Services and Roles

HPE ProLiant DL360 Gen10 Servers	Role	Services
Server No.1	Management Node	Cloudera Manager
Server No.2	Utility Node	Hive Metastore Server, Hue Load Balancer, Hue Server, Hue Kerberos Ticket Renewer, Impala Catalog Server, Cloudera Management Service Alert Publisher, Cloudera Management Service Event Server, Cloudera Management Service Host Monitor, Cloudera Management Service Service Monitor, ZooKeeper Server
Server No.3	Master node1	Atlas Server, Hbase RegionServer, HDFS NameNode, Hive Gateway, Impala Daemon, Solr Server, Tez Gateway, YARN JobHistory Server, YARN ResourceManager, ZooKeeper Server
Server No.4	Master Node2	HDFS SecondaryNameNode, Hive on Tez HiveServer2, YARN ResourceManager, ZooKeeper Server

TABLE 3. HPE Apollo r2600/XL170r Gen10 Servers

HPE Apollo r2600/XL170r Gen10 Servers	Role	Services
Server No.5	Compute Node / Worker Node	YARN NodeManager
Server No.6	Compute Node / Worker Node	YARN NodeManager
Server No.7	Compute Node / Worker Node	YARN NodeManager
Server No.8	Compute Node / Worker Node	HBase Master, Hive Gateway, Tez Gateway, YARN NodeManager
Server No.9	Compute Node / Worker Node	HBase RegionServer, Hive Gateway, Impala Daemon, Tez Gateway, YARN NodeManager
Server No.10	Compute Node / Worker Node	HBase RegionServer, Hive Gateway, Impala Daemon, Tez Gateway, YARN NodeManager
Server No.11	Compute Node / Worker Node	HBase RegionServer, Hive Gateway, Impala Daemon, Kafka Broker, Ranger Admin, Ranger Tagsync, Ranger Usersync, Tez Gateway, YARN NodeManager
Server No.12	Compute Node / Worker Node	HBase RegionServer, Hive Gateway, Impala Daemon, Kafka Broker, Tez Gateway, YARN NodeManager
Server No.13	Compute Node / Worker Node	HBase RegionServer, Hive Gateway, Impala Daemon, Impala StateStore, Kafka Broker, Tez Gateway, YARN NodeManager
Server No.14	Compute Node / Worker Node	YARN NodeManager
Server No.15	Compute Node / Worker Node	YARN NodeManager
Server No.16	Compute Node / Worker Node	YARN NodeManager



TABLE 4. HPE Apollo 4200 Gen10 Server

HPE Apollo 4200 Gen10 Servers	Role	Services
Server No.17	Storage Node / Worker Node	HDFS DataNode
Server No.18	Storage Node / Worker Node	HDFS DataNode, NiFi Node
Server No.19	Storage Node / Worker Node	HDFS DataNode, NiFi Registry
Server No.20	Storage Node / Worker Node	HDFS DataNode, HBase Thrift Server

SOLUTION COMPONENTS

The solution consists of four HPE Proliant DL360 Gen10, 12 HPE Apollo r2600/XL170r and four HPE Apollo servers Gen10 servers and the configuration details are as follows.

TABLE 5. Hardware

Component	Qty	Description
Cloudera Master/Data Flow Node – DL360 Gen10 8SFF model gnode32 2 x Intel Xeon-Gold 5218 (2.3GHz/16-core) processors 12x 32GB 3200MHz (384GB total) 6 x 1.2 TB 12G SAS 10K HDD (O/S and local data) 2 x HDD RAID-1, rest is JBOD PCI-E Slot 3ConnectX-6 100GE 1P NIC Embedded ALOM HPE Eth 10/25Gb 2p 640FLR-SFP28 Adptr - NIC Embedded Device HPE Smart Storage Battery Embedded Device Embedded Video Controller Embedded LOM HPE Ethernet 1Gb 4-port 331i Adapter - NIC Embedded RAID HPE Smart Array P408i-a SR Gen10	1	These servers were used for deploying HPE Cloudera CDP
Cloudera Master Node – DL360 Gen10 10SFF NVMe model gnode33 2 x Intel Xeon-Gold 5218 (2.3GHz/16-core) processors 24x 16GB 3200MHz (384GB Total) 1 x 800GB NVMe (O/S) PCI-E Slot 1HPE InfiniBand EDR/Ethernet 25Gb 2-port 841QSFP28 Adptr PCI-E Slot 2HPE InfiniBand EDR/Ethernet 25Gb 1-port 841QSFP28 Adapter Storage Backplane 1 HPE 10SFF NVMe/SAS 10/8 Bkpln Embedded Device HPE Smart Storage Battery Embedded Device Embedded Video Controller Embedded LOM HPE Ethernet 1Gb 4-port 331i Adapter Embedded RAID HPE Smart Array P408i-a SR Gen10	1	These servers were used for deploying HPE Cloudera CDP
Cloudera Master Nodes – DL360 Gen10 10SFF NVMe model gnode34-35 2 x 6252N 2.30GHz 24core 48 threads (gnode34) 2 x 6134 3.2GHz 8core 16 threads (gnode35) 12 x 32GB PC4-2933Y-R DIMM's (384GB) 1 x 800GB NVMe (O/S) PCI-E Slot 1HPE InfiniBand EDR/Ethernet 25Gb 2-port 841QSFP28 Adptr PCI-E Slot 2HPE InfiniBand EDR/Ethernet 25Gb 1-port 841QSFP28 Adapter Storage Backplane 1 HPE 10SFF NVMe/SAS 10/8 Bkpln Embedded Device HPE Smart Storage Battery Embedded Device Embedded Video Controller Embedded LOM HPE Ethernet 1Gb 4-port 331i Adapter Embedded RAID HPE Smart Array P408i-a SR Gen10	2	These servers were used for deploying HPE Cloudera CDP



Component	Qty	Description
Apollo r2600 Gen10 24SFF chassis - Compute gnode [07-09, 14] 4 x XL170r Gen10 Server nodes per chassis 2 x Intel Xeon-Gold 6134 (3.2GHz/16-core) processors 384GB RAM 1 x 800GB NVMe MU SFF SSD (O/S) 1 x 3.84TB NVMe RI SFF SSD (Local Data) PCI-E Slot 1 HPE Smart Array P408i-p SR Gen10 PCI-E Slot 2 HPE InfiniBand EDR/Ethernet 25Gb 1-port 841QSFP28 Adapter Storage Backplane 4 HPE 16SFF NVMe Backplane Embedded Device Embedded Video Controller Embedded mLOM HPE Eth 1Gb 2p 368FLR-MMT Adptr	4	These servers were used for HPE Cloudera CDP Compute
Apollo r2600 Gen10 24SFF chassis - Compute gnode15-17, 19, [11-13, 37] 2x Intel Xeon Gold 6226R 2.9 GHz 16core 32 threads 12x 32GB 3200MHz (384GB total) 1 x 800GB NVMe MU SFF SSD (O/S) 1 x 3.84TB NVMe RI SFF SSD (Local Data) Slot 1 HPE Ethernet 1Gb 4-port 366T Adapter Slot 2 HPE InfiniBand EDR/Ethernet 25Gb 1-port 840QSFP28 Adapter Embedded Device HPE Smart Storage Battery Embedded Device Embedded Video Controller Backplane 4HPE 16SFF NVMe Backplane	8	These servers were used for HPE Cloudera CDP Compute
Apollo 4200 Gen10 Data Nodes (Orange) gnode20-23 2x Intel Xeon Gold 6226R 2.9 GHz 16core 32 threads 12x 32GB 3200MHz (384GB total) 26 x 4TB 12G SAS 7.2k LFF HDD JBOD (HDFS Space) 1 x 800GB NVMe MU SFF SSD (O/S) 1 x 3.84TB NVMe RI SFF SSD (Local Data) Slot 1 HPE Ethernet 1Gb 4-port 366T Adapter Slot 2 HPE InfiniBand EDR/Ethernet 25Gb 1-port 840QSFP28 Adapter Embedded Device HPE Smart Storage Battery Embedded Device Embedded Video Controller Backplane 4HPE 16SFF NVMe Backplane	4	These servers were used for HPE Cloudera CDP Storage
Aruba 8325-48Y8C 48p 25G SFP+/28 8p 25G QSFP	2	Switch for 25GbE network connectivity
Aruba 6300F 48-port 1GbE and 4-port SFP56	1	For 1GbE network connectivity

TABLE 6. Software

Software

Component	Version	Description
Red Hat Enterprise Linux	7.9	Operating System
Open JDK	1.8.0_262	Open Java Development Kit
CDP Private Cloud	7.1.7	Cloudera Data Platform Private Cloud



BEST PRACTICES AND CONFIGURATION GUIDANCE FOR SOLUTION

Base cluster set-up requires system-level optimizations, detailed below while setting up the initial cluster to obtain a production-grade environment to scale to the high-throughput requirement for streaming analytics. These settings may require additional, specific optimizations as per the type of workloads.

This section provides best practices and a recommended configuration for HPE Cloudera Data Platform (CDP) Private Cloud Base cluster running on HPE DL360 Gen10, HPE Apollo 4200 Gen10, and HPE Apollo r2600 Gen10 Servers. The settings were performed on all the nodes in the cluster.

HPE DL360 Gen10, HPE Apollo r2600/XL170r Gen10, and HPE Apollo 4200 Gen10 settings

Hewlett Packard Enterprise recommends changing the default BIOS settings to the following using Workload Profile High-Performance Compute (HPC) on all servers hosting Hadoop to ensure the highest performance.

1. Power on the server. During the POST, Press F9 for the server to boot to BIOS (System Utilities).
2. Select System Configuration menu.
3. Select BIOS/Platform Configuration (RBSU) menu.
4. Select Workload Profile. Select the **Workload** Profile as **High_Performance_Compute** option from the drop down menu.
5. Select Power and Performance Options. Select the Power Regulator as **Static_High_Performance_Mode**.

Red Hat Enterprise Linux settings

We have performed the CPU, RAM, Security, and Network settings listed below for optimal performance on all the nodes in the Cloudera Data Platform (CDP) Private Cloud Base Cluster.

CPU

1. Enable Multi-Threading.
2. Set the BIOS settings for CPU and memory to Maximum Performance mode.

Memory

1. Minimize anonymous page faults. Minimize anonymous page faults by setting `vm.swappiness = 1`, which frees them from the page cache before swapping application pages (this reduces the OOM-killer invocation).
2. Edit the `/etc/sysctl.conf` file in your editor of choice and add the following line.

```
vm.swappiness=1 Then, run the following: # sysctl -p # sysctl -a | grep "vm.swappiness"
```

3. Disable transparent huge-page compaction.

```
echo "never" > /sys/kernel/mm/redhat_transparent_hugepage/enabled
```

4. Disable transparent huge-page defragmentation.

```
echo "never" > /sys/kernel/mm/redhat_transparent_hugepage/defrag
```

5. Add these commands to `/etc/rc.local` to ensure that transparent huge page compaction and defragmentation remain disabled across reboots.

Security

1. Disable SELinux - SELinux can be disabled on RHEL or CentOS by editing `/etc/selinux/config` and setting `SELINUX=disabled`. This change must be done as root (or with proper sudo access) and requires a reboot.
2. Disable Firewall - Disable the host-based firewalls on all the machines in the cluster.

```
systemctl stop firewalld.service
systemctl disable firewalld.Service
```

Network

1. Enable the NTP - Enable the NTP daemon by running the following below commands.

```
systemctl start ntpd.service
systemctl enable ntpd.service
```

2. Compute nodes network tuning:

- a. Add the following parameters in `/etc/sysctl.conf`. Disable TCP timestamps to improve CPU utilization (this is an optional parameter and will depend on your NIC vendor).



- b. Disable TCP timestamps to improve CPU utilization (this is an optional parameter and depends on your NIC vendor):
`net.ipv4.tcp_timestamps=0.`

3. Enable TCP sacks to improve throughput:

```
net.ipv4.tcp_sack=1
```

4. Increase the maximum length of processor input queues:

```
net.core.netdev_max_backlog=250000
```

5. Increase the TCP max and default buffer sizes using `setsockopt()`:

```
net.core.rmem_max=4194304
net.core.wmem_max=4194304
net.core.rmem_default=4194304
net.core.wmem_default=4194304
net.core.optmem_max=4194304
```

6. Increase memory thresholds to prevent packet dropping:

```
net.ipv4.tcp_rmem=4096 87380 4194304
net.ipv4.tcp_wmem=4096 65536 4194304
```

NOTE

If you want to run this from the command line, then quote the values being set. For example: `sysctl -w net.ipv4.tcp_rmem="4096 87380 4194304"`.

7. Set the socket buffer to be divided evenly between TCP window size and application buffer.

```
net.ipv4.tcp_adv_win_scale=1
```

8. Verify NIC advanced features - Verify which features are available with your NIC using `ethtool`.

```
$ sudo ethtool -k <ethx>
```

9. There are various offload capabilities that modern NICs (and especially high-performance NICs) have, and it is always recommended to enable them. A few features such as `tcp-segmentation-offload` (TSO), `scatter-gather` (SG), and `generic segmentation-offload` (GSO) are good features to enable (if not enabled by default).

- a. NIC ring buffer configurations
- b. Check existing ring buffer sizes. `ethtool -g <ethx>`

10. After checking the preset maximum values and the current hardware settings, the following command can be used to resize the ring buffers:

```
# ethtool -G <interface> rx <newsized>
```

Or

```
# ethtool -G <interface> tx <newsized>
```

Cloudera Data Platform (CDP) Private Cloud Base Cluster Configurations

Cloudera Data Platform Private Cloud Base includes many default parameter settings. The Figure 9 depicts tuning parameters set to achieve the best performance from the testing cluster.



```
mapreduce.map.memory.mb=4096
mapreduce.map.java.opts=-Xmx3072m
mapreduce.map.cpu.vcores=1
mapreduce.map.output.compress=true
mapreduce.map.output.compress.codec=org.apache.hadoop.io.compress.Lz4Codec
mapreduce.map.speculative=true
mapreduce.reduce.memory.mb=4096
mapreduce.reduce.java.opts=-Xmx3072m
mapreduce.reduce.cpu.vcores=1
mapreduce.reduce.speculative=true
mapreduce.task.io.sort.factor=64
mapreduce.task.io.sort.mb=768
mapred.map.tasks=1248
mapred.reduce.tasks=624
mapreduce.job.reduce.slowstart.completedmaps=0.85
mapreduce.reduce.shuffle.parallelcopies=12
```

FIGURE 9. Tuning parameters

TESTING AND VALIDATION OF DEPLOYMENT

This section provides details of the validation done using HPE EPA to accomplish the following test objectives and use cases:

Testing the Cluster:

1. Access cluster nodes via ssh – validate root level access to the cluster nodes.
2. Access Cloudera Manager – validate admin level access to Cloudera Manager.
3. In Cloudera Manager, ensure that required services are setup and running and that all health checks are Green.
4. Ensure that Cluster has been set up per proper Cloudera supported guidelines.
5. Run a PiEstimator job to manually verify that the CDP Private Cloud Base installation was successful.

NOTE

If you have a secure cluster, use the kinit command-line tool to authenticate Kerberos.

6. Log in to a host in the cluster.
7. Run the Hadoop PiEstimator example using the following command.

```
yarn jar /opt/cloudera/parcels/CDH/lib/hadoop-mapreduce/hadoop-mapreduce-examples.jar
pi 10 100
```

8. In Cloudera Manager, navigate to **Cluster > ClusterName > yarn Applications**.
9. Check the results of the job as shown in the Figure 10.



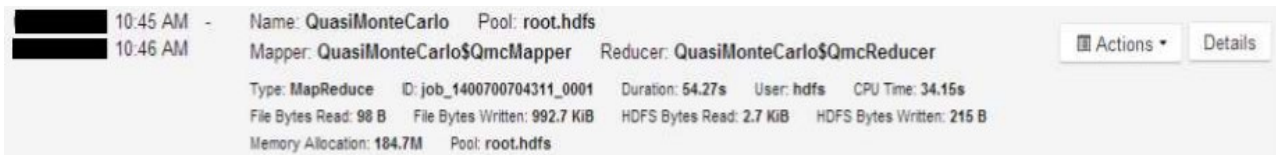


FIGURE 10. Example of the results

10. Validation of MapReduce operations using TeraSuite test:

- a. Run teragen with 1 mapper per Datanodes/Storage disk in the Cluster, with 1x replication.
- b. Run teragen with 1 mapper per datanode/Storage disk in the Cluster, with default (3x) replication.
- c. Run terasort with a varying number of reducers to find best throughput.

TeraSuite test

The solution was validated with the MapReduce benchmark tools Teragen and Terasort. The goal was to validate the functionality of different components and services rather than stress the cluster for maximum performance. This test was designed to exercise a default replication factor of 3. The solution was operationally validated with different workload sizes of 100GB, 500GB, and 1TB. The Cloudera Data Platform (CDP) Private Cloud Base Cluster performed optimally, indicating the resources such as I/O and network have been configured and tuned to optimally run and support different workload sizes.

The performance lab environment was configured with four (4) HPE ProLiant DL360 Gen10, 12 HPE Apollo r2600/XL170r Gen10 and four(4) Apollo 4200 Gen10 servers with the back-end network between nodes is with 100Gbps.

The main configuration applied is the size of each mapper and reducer. Using some simple YARN calculations, we used 1 vCore for each container, and since the entire cluster has 240 vCores, this would yield 240 containers. Allocated memory was four (4) GB per container, resulting in 960 GB of the total 1.4 TB used for each run.

Steps

1. The Figure 11 shows the **teragen** command to generate 1TB of data.

```
time hadoop jar /opt/cloudera/parcels/CDH/jars/hadoop-mapreduce-examples-3.1.1.7.1.7.1000-141.jar teragen \
-Dmapreduce.map.log.level=INFO -Dmapreduce.reduce.log.level=INFO \
-Dyarn.app.mapreduce.am.log.level=INFO -Dio.file.buffer.size=131072 \
-Ddfs.blocksize=256M -Ddfs.replication=1 -Dmapreduce.map.max.attempts=1 \
-Dmapreduce.map.memory.mb=4096 -Dmapreduce.map.java.opts=-Xmx3072m \
-Dmapreduce.map.cpu.vcores=1 -Dmapreduce.map.output.compress=true \
-Dmapreduce.map.output.compress.codec=org.apache.hadoop.io.compress.Lz4Codec \
-Dmapreduce.map.speculative=true -Dmapreduce.task.io.sort.factor=64 \
-Dmapreduce.task.io.sort.mb=768 -Dmapreduce.job.maps=1248 10000000000 \
-Dyarn.app.mapreduce.am.resource.mb=1024 -Dyarn.app.mapreduce.am.command.opts=768 \
/user/data/sandbox/poc/teragen
```

FIGURE 11. Teragen Command

2. The Figure 12 shows the **terasort** command to sort the generated 1TB data.



```
time hadoop jar /opt/cloudera/parcels/CDH/jars/hadoop-mapreduce-examples-3.1.1.7.1.7.1000-141.jar terasort \
-Dmapreduce.map.log.level=INFO -Dmapreduce.reduce.log.level=INFO \
-Dyarn.app.mapreduce.am.log.level=INFO -Dio.file.buffer.size=131072 \
-Ddfs.blocksize=256M -Ddfs.replication=1 -Dmapreduce.map.max.attempts=1 \
-Dmapreduce.map.memory.mb=4096 -Dmapreduce.map.java.opts=-Xmx3072m \
-Dmapreduce.map.cpu.vcores=1 -Dmapreduce.map.output.compress=true \
-Dmapreduce.map.output.compress.codec=org.apache.hadoop.io.compress.Lz4Codec \
-Dmapreduce.map.speculative=true -Dmapreduce.task.io.sort.factor=64 \
-Dmapreduce.task.io.sort.mb=768 -Dmapreduce.job.maps=1248 \
-Dmapreduce.reduce.memory.mb=4096 -Dmapreduce.reduce.java.opts=-Xmx3072m \
-Dmapreduce.reduce.cpu.vcores=1 -Dmapreduce.reduce.maxattempts=1 \
-Dmapreduce.reduce.speculative=true -Dyarn.app.mapreduce.am.resource.mb=1024 \
-Dmapreduce.job.reduces=624 -Dyarn.app.mapreduce.am.command.opts=-Xmx768m \
-Dmapreduce.job.reduce.slowstart.completedmaps=0.85 \
-Dmapreduce.reduce.shuffle.parallelcopies=12 -Dmapreduce.terasort.output.replication=1 \
/user/data/sandbox/poc/teragen /user/data/sandbox/poc/terasort
```

FIGURE 12. Terasort command

- The Figure 13 shows the **teravalidate** command for the generated data and has been sorted.

```
time hadoop jar /opt/cloudera/parcels/CDH/jars/hadoop-mapreduce-examples-3.1.1.7.1.7.1000-141.jar teravalidate \
-Dio.file.buffer.size=131072 -Ddfs.blocksize=256M \
-Dmapreduce.map.memory.mb=4096 -Dmapreduce.map.java.opts=-Xmx3072m \
-Dmapreduce.reduce.memory.mb=4096 -Dmapreduce.reduce.java.opts=-Xmx3072m \
-Dmapreduce.task.io.sort.factor=64 -Dmapreduce.task.io.sort.mb=768 \
-Dmapreduce.job.maps=1248 -Dmapreduce.job.reduces=624 \
-Dyarn.app.mapreduce.am.resource.mb=1024 -Dyarn.app.mapreduce.am.command.opts=768 \
/user/data/sandbox/poc/terasort /user/data/sandbox/poc/teravalidate
```

FIGURE 13. teravalidate command

With these test results, HPE validated the Cloudera Data Platform (CDP) Private Base Cluster HDFS and MapReduce layers with effective use of compute and storage resources. The goal of these tests was not to measure performance metrics. They were to confirm that all services deployed in the Cloudera Data Platform (CDP) Private Cloud Base cluster were working correctly.

SUMMARY

Hewlett Packard Enterprise and Cloudera allow one to derive new business insights from Big Data by providing a platform to store, manage, and process data at scale. However, designing and ordering Hadoop clusters can be both complex and time-consuming. This Reference Architecture provides several configurations for deploying clusters of varying sizes with the Cloudera Data Platform (CDP) Private Cloud Base 7.1.7 on HPE infrastructure and management software. These configurations leverage Hewlett Packard Enterprise to balance building blocks of servers, storage, and networking, along with integrated management software and bundled support. In addition, this Reference Architecture has been created to assist in the rapid design and deployment of Cloudera Data Platform (CDP) Private Cloud Base software on HPE infrastructure for clusters of various sizes.

APPENDIX A: BILL OF MATERIALS

This section provides the details of the Bill of Materials used for the solution built for this Reference Architecture. Based on the user needs the solution is flexible to scale out by adding symmetric nodes of HPE Apollo r2600/XL170r Gen10 and HPE Apollo 4200 Gen10 Servers.



NOTE

Part numbers are at time of testing and publication and subject to change. The bill of materials does not include complete support options or other rack and power requirements. If you have questions regarding ordering, please consult with your Hewlett Packard Enterprise Reseller or Hewlett Packard Enterprise Sales Representative for more details. hpe.com/us/en/services/consulting.html

TABLE A1. Bill of materials

Part number	Quantity	Description
Rack		
P9K10A	1	HPE 42U 600mmx1200mm G2 Kitted Advanced Shock Rack with Side Panels and Baying
P9K10A 001	1	HPE Factory Express Base Racking Service
BW932A	1	HPE 600mm Rack Stabilizer Kit
BW932A B01	1	HPE 600mm Rack include with Complete System Stabilizer Kit
Master Nodes		
P56958-B21	4	HPE ProLiant DL360 Gen10 5218 2.3GHz 16-core 1P 32GB-R MR416i-a NC 8SFF BC 800W PS Server
629135-B21	4	HPE Ethernet 1Gb 4-port 331i Adapter
872726-B21	4	HPE InfiniBand EDR/Ethernet 100Gb 2-port 841QSFP28 Adapter
Storage Nodes		
P28700-B21	4	HPE Apollo 4200 Gen10 Plus 24LFF TAA-compliant Configure-to-order System
P42912-B21	8	Intel Xeon-Gold 5318Y 2.1GHz 24-core 165W Processor Kit for HPE Apollo 4200 Gen10 Plus
P06031-K21	48	HPE 16GB (1x16GB) Dual Rank x8 DDR4-3200 CAS-22-22-22 Registered Smart Memory Kit
P28702-B21	4	HPE Apollo 4200 Gen10 Plus 4LFF SAS/SATA LP Drive Cage 3 Upgrade Kit
P38441-001	384	HPE 4TB SATA 6G Business Critical 7.2K LFF LP 1-year Warranty Multi Vendor HDD
P28711-B21	4	HPE Apollo 4200 Gen10 Plus PCIe Primary Riser Kit
629135-B21	4	HPE Ethernet 1Gb 4-port 331i Adapter
872726-B21	4	HPE InfiniBand EDR/Ethernet 100Gb 2-port 841QSFP28 Adapter
804338-B21	4	HPE Smart Array P816i-a SR Gen10 (16 Internal Lanes/4GB Cache/SmartCache) 12G SAS Modular Controller
P38997-B21	4	HPE 1600W Flex Slot Platinum Hot Plug Low Halogen Power Supply Kit
Compute Nodes		
867158-B21	3	HPE Apollo r2600 Gen10 24 SFF Premium Configure-to-order Chassis
867055-B21	12	HPE ProLiant XL170r Gen10 1U Node Configure-to-order Server
830272-B21	12	HPE 1600W Flex Slot Platinum Hot Plug Low Halogen Power Supply Kit
P42920-B21	24	Intel Xeon-Gold 6326 2.9GHz 16-core 185W Processor Kit for HPE Apollo r2600
P06033-K21	144	HPE 32GB (1x32GB) Dual Rank x4 DDR4-2933 CAS-21-21-21 Registered Smart Memory Kit
P47837-K21	12	HPE 800GB NVMe Gen4 Mainstream Performance Mixed Use SFF BC U.3 Static Multi Vendor SSD
P47846-K21	12	HPE 3.84TB NVMe Gen4 Mainstream Performance Read Intensive SFF BC U.3 Static Multi Vendor SSD
830272-B21	12	HPE 1600W Flex Slot Platinum Hot Plug Low Halogen Power Supply Kit
629135-B21	12	HPE Ethernet 1Gb 4-port 331i Adapter
872726-B21	12	HPE InfiniBand EDR/Ethernet 100Gb 2-port 841QSFP28 Adapter



Part number	Quantity	Description
Network		
JL624A	1	Aruba 8325-48Y8C 48p 25G SFP+/28 8p 25G QSFP
ROZ26A	20	HPE Aruba 100G QSFP28 to QSFP28 5M DAC Cab
JL448A	1	HPE Aruba X2C2 RJ45 to DB9 Console Cable
JL667A	1	Aruba 6300F 48-port 1GbE and 4-port SFP56



RESOURCES AND ADDITIONAL LINKS

HPE Servers for Big Data Analytics and Hadoop, [hpe.com/info/Hadoop](https://www.hpe.com/info/Hadoop)

HPE Apollo 4000 Systems, <https://www.hpe.com/us/en/storage/apollo-4000.html>

HPE ProLiant 360 Gen10, https://www.hpe.com/psnow/doc/a00008159enw?jumpid=in_lit-psnow-red

Cloudera Data Platform: <https://www.cloudera.com/products/cloudera-data-platform.html>

HPE Reference Architectures, [hpe.com/info/ra](https://www.hpe.com/info/ra)

HPE Servers, [hpe.com/servers](https://www.hpe.com/servers)

HPE Storage, [hpe.com/storage](https://www.hpe.com/storage)

HPE Networking, [hpe.com/networking](https://www.hpe.com/networking)

HPE Technology Consulting Services, [hpe.com/us/en/services/consulting.html](https://www.hpe.com/us/en/services/consulting.html)

To help us improve our documents, please provide feedback at [hpe.com/contact/feedback](https://www.hpe.com/contact/feedback).

LEARN MORE AT

HPE Apollo 4200 Gen10 Plus Data Storage Server Product Tour, https://www.youtube.com/watch?v=5cpHX_W4fZY

HPE Apollo 4200 Gen10 Plus System QuickSpecs, <https://www.hpe.com/psnow/doc/a50002573enw>

HPE Apollo 4200 Gen10 Plus Server QuickSpecs, <https://www.hpe.com/psnow/doc/a00056091enw>