Dell EMC and Intel Infrastructure Guide for Cloudera Data Platform Private Cloud

Abstract

This white paper provides infrastructure configuration and strategy guidance for customers planning new or upgraded data center deployments of Cloudera Data Platform Private Cloud on Intel architecture.

May 2020

Revisions

Date	Description
February 22, 2020	Initial draft
May 11, 2020	Initial release

Acknowledgments

This paper was produced by the Dell EMC Ready Solutions engineering team in collaboration with Cloudera and Intel. The authors are:

- Dell EMC – Mike Pittaro, Thaddeus Rogers
- Intel – Sandeep Togrikar, Bhasker Allene
- Cloudera – Ali Bajwa

Table of contents

Contents

Executive summary

This document provides advance infrastructure guidance and recommendations to customers who are installing or moving to Cloudera's forthcoming Cloudera Data Platform Private Cloud (CDP PVC) offering that delivers a cloud-like experience in customers' environments. CDP PVC is a new approach to data management and analytics that delivers powerful self-service analytics across hybrid and multi-cloud environments by leveraging disaggregated compute and storage models to provide simpler multi-tenancy and isolation, better infrastructure utilization, containerization and cloud native architecture. With Cloudera's evolution to cloud-native architecture, customers can generate the best results by ensuring that their best-in-class Dell infrastructure with Intel architecture is optimized for Cloudera deployments.

1 Introduction

Since 2014, Dell EMC, Intel, and Cloudera have collaborated to provide customers with guidance on optimal hardware to streamline the design, planning and configuration of their Cloudera deployments. This guide is based on our collective experience in deploying and running enterprise production environments. Cloudera extends the enterprise data cloud with its forthcoming Cloudera Data Platform (CDP) Private Cloud offering that delivers a cloud-like experience in customers' environments. CDP is a new approach to data management and analytics that leverages cloud-native architecture and containerization for powerful self-service analytics across hybrid and multi-cloud environments.

Platform upgrades are always an exciting event, since new features and capabilities become available. Successful upgrades require planning. For an enterprise data platform much of that planning involves the underlying infrastructure as much as the software. With Cloudera's evolution to cloud-native architecture, customers can generate the best results by ensuring that their best-in-class Dell EMC infrastructure - powered by Intel 2nd Gen Intel Xeon Scalable processors and Intel Optane persistent memory - is optimized for such deployments.

Dell EMC, Intel, and Cloudera have assembled this guide based on our collective experience in deploying and running enterprise production environments. Although the ideal upgrade approach always depends on individual customer environments, our experience shows that most installations and upgrades follow the same general patterns.

This document provides current best practices on migration and upgrade strategy. Customers can begin planning for their specific environments while maximizing their investment in current infrastructure. Cloudera Data Platform and the underlying infrastructure capabilities have evolved significantly since the earliest releases of Apache Hadoop. Appropriate infrastructure selection is important for long-term success.

This guide describes both general migration strategies, and specific recommendations for infrastructure. The Migration and upgrade scenarios section focuses on strategies for the migration process, covering new installations and upgrades to existing environments. The Infrastructure guidance section focuses on hardware specifics.

Contact your Dell EMC and Cloudera representatives for the latest information and assistance with specific deployment scenarios.

1.1 Audience

The intended audience for this document is data center managers and architects who are involved in infrastructure planning for Cloudera Data Platform Private Cloud. Data engineering and data science teams also benefit from some of the high-level guidance, but the focus of the document is primarily infrastructure rather than software details.

This document assumes some familiarity with Cloudera Data Platform capabilities and functions.

1.2 Cloudera Enterprise Data Cloud

Today, the world's most innovative enterprises unlock value from their data and compete in the data age with the most valuable and transformative business use cases:

- IoT-enabled predictive maintenance
- Genomics research
- Real-time compliance monitoring

These use cases require multiple analytics workloads and data science tools and machine learning algorithms to run against the same diverse datasets. However, many businesses still struggle with limited data visibility and insufficient insights, often caused by:

- Analytic workloads running independently—in silos—because even newer cloud data warehouses and data science tools were never designed to work together.
- Data being everywhere—in data centers, public clouds, and at the Edge—and companies have no practical way to run analytics or apply machine learning algorithms to all their data.
- An incoherent approach to data privacy or IP protection—often resulting in onerous controls that limit business productivity and increase costs.

Organizations need a new approach. An enterprise data cloud unlocks the power of your data to serve customers better, operate with greater efficiency, and strengthen security to protect your business:

- It is optimized for hybrid and multi-cloud environments, delivering the same data management capabilities across bare metal, private, and public clouds.
- It allows multiple analytic functions to work together on the same data at its source, eliminating costly and inefficient data silos.
- It maintains strict enterprise data security, governance, and control across all environments.
- It is one hundred percent open source, with open compute and open storage, ensuring zero vendor lock-in and maximum interoperability.

1.2.1 Cloudera Data Platform (CDP)

Cloudera Data Platform (CDP) combines the best of Hortonworks' and Cloudera's technologies to deliver the industry's first enterprise data cloud. CDP delivers powerful self-service analytics across hybrid and multi-cloud environments, along with sophisticated and granular security and governance policies that IT and data leaders demand.

Initially delivered as a public cloud service and followed up with Data Center, a comprehensive data management and analytics platform for on-premises IT environments, CDP includes:

- Data Warehouse and Machine Learning services and a Data Hub service for building custom business applications that are powered by our new Cloudera Runtime open-source distribution
- A unified control plane to manage infrastructure, data, and analytic workloads across hybrid and multi-cloud environments
- Consistent data security, governance, and control that safeguards data privacy, regulatory compliance, and prevents cybersecurity threats across environments
- One hundred percent open source, supporting your objectives to avoid vendor lock-in and accelerate enterprise innovation
- A clear path for extending your existing CDH and HDP investment to cloud-native architecture

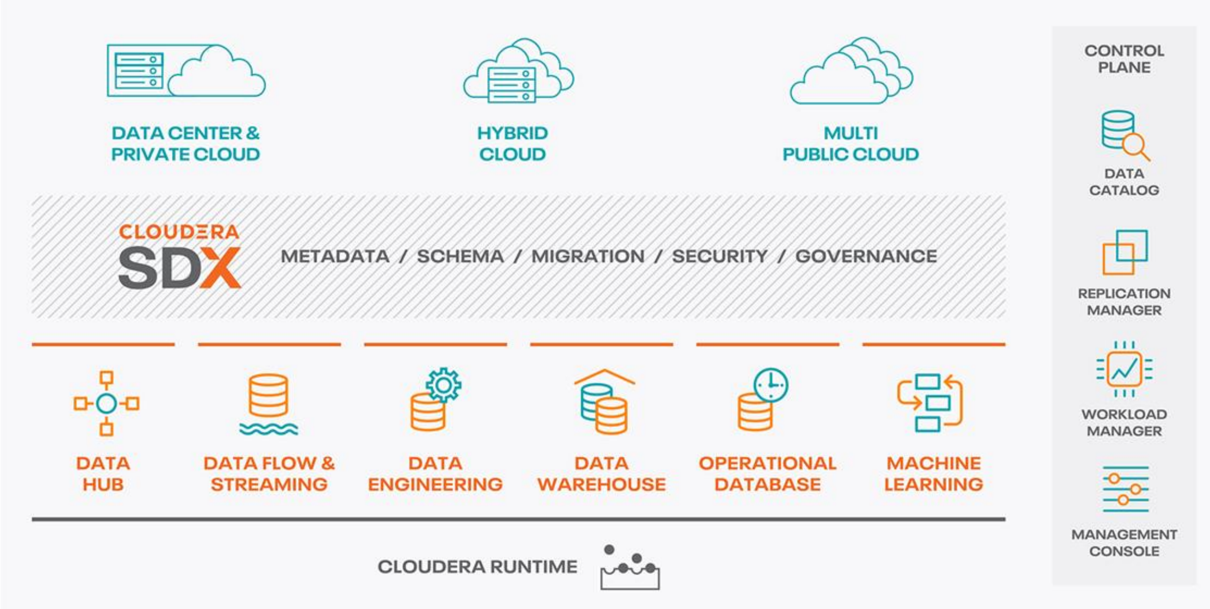

Figure 1 Cloudera Data Platform

1.2.2 Cloudera Data Platform Private Cloud

A fundamental challenge for legacy monolithic environments is the inability of enterprises to rapidly onboard new workloads and teams to meet business needs faster. These challenges are key to the success or failure of data analytics and AI and ML initiatives. CDP Private Cloud overcomes these challenges that are associated with monolithic environments by enabling isolation of workloads and multi-tenancy. CDP Private Cloud (PVC) simplifies deployment, scaling, and management of workload environments.

In the past, enterprises had to pre-plan many aspects of their deployments including hardware configuration, software versioning, and resource allocation across tenants and workloads. For example, the environment that is required to improve call center efficiency differs from the environment that is required to save lives by early detection of diseases. CDP PVC enables enterprises to make real-time decisions for all deployment aspects that are tailored to workloads like data warehouse, machine learning, and data engineering. CDP PVC delivers isolated compute environments, real-time compute provisioning with auto-scaling, multiple software versions running concurrently, and sharing of data, metadata, and security across various compute environments. This flexibility and improved resource utilization can help organizations ensure application performance and reduce costs by maximizing the use of resources.

CDP Private Cloud enables Data Warehouse, Data Engineering, and Machine Learning experiences coupled with a single pane of glass for management. CDP Data Center serves as a foundation in providing stateful elements like storage, schema, security, and governance.

There are two aspects of deploying Cloudera Data Platform Private Cloud – the data lake cluster (running CDP-DC) and the compute experiences cluster (running on OpenShift). These are two separate clusters that are independent tracks from a deployment planning perspective.

Figure 2 illustrates the primary components in a complete installation.

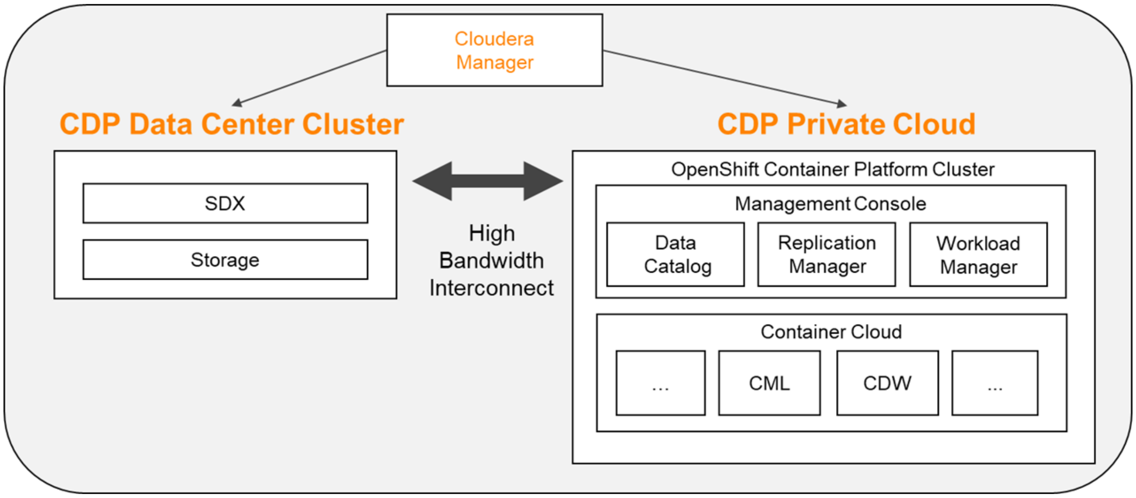

Figure 2 CDP Private Cloud components

A full installation consists of two cooperating clusters – CDP Data Center, and a private cloud platform running OpenShift. The primary functions are:

1. **Cloudera Manager** provides management, monitoring, and deployment for both clusters.
2. The **CDP Data Center** cluster provides the core Cloudera **SDX** services and **HDFS/Ozone storage.**
3. The **OpenShift Container Platform** provides the core capabilities necessary for the containerized cloud environment, including Kubernetes and its control plane.
4. The **Management Console** is the control plane for the container cloud, and provides the **Data Catalog**, **Replication Manager**, and **Workload Manager** services.
5. The **Container Cloud** runs the workloads, including **CDW and CML.** More workloads will be added over time.
6. A **High Bandwidth Interconnect** connects the two clusters to provide access from the container platform to storage on the CDP Data Center cluster.

2 Migration and upgrade scenarios

There are two aspects of deploying Cloudera Data Platform Private Cloud – the compute experiences cluster and the data lake cluster. This document discusses deploying the CDP-Private Cloud compute experiences scenario first, since it requires a fresh install. Then it discusses deploying the data lake, which includes possible upgrade scenarios.

For both aspects, this section is primarily concerned with the general strategy and flow of the process, including both infrastructure and software. This section describes the infrastructure that is required in general terms; specific infrastructure recommendations are described in the Infrastructure guidance section.

2.1 Compute Experiences Deployment for CDP Private Cloud

Deploying the compute experiences requires a fresh deployment of CDP Private Cloud regardless of whether it is a new installation or an installation with an existing deployment of CDH or HDP.

2.1.1 New, Dedicated OpenShift Cluster

The OpenShift cluster infrastructure configuration generally follows the guidelines in the current Dell EMC Ready Stack for Red Hat OpenShift Container Platform Design Guide. In general, the worker nodes in the container platform are memory- and processor-optimized, and only include enough storage for runtime operations. The minimum number of nodes in the container platform is seven. There are three master nodes, three worker nodes, and one container services administration node. The container platform can be scaled by additional worker nodes after initial deployment. More storage for analytic workloads can also be added to the platform using the Kubernetes Container Storage Interface (CSI). This storage may be useful for additional analytic components that are used along with CDP Private Cloud.

The high-bandwidth interconnect uses standard Ethernet and is sized to provide adequate capacity for the data traffic between the CDP DC cluster and the container platform. One or more 100 GbE links are typically used.

Specific recommendations on infrastructure configurations are provided in Infrastructure guidance.

2.1.2 Existing, Shared OpenShift Cluster

This option will not be supported in the initial release of CDP Private Cloud but is planned for future releases.

2.2 Data Lake Deployment for CDP Private Cloud

The data lake consists of storage and SDX components for securing, governing, and providing context for the data. There are two possible deployment options: a fresh install of CDP-DC, or an in-place upgrade from CDH or HDP.

Although this document is focused on infrastructure, it is difficult to cleanly separate the infrastructure and software paths. To help you decide whether to follow the upgrade in place or new install paths, this section provides a summary of the major steps that are involved in the upgrade to CDP 7.x.

2.2.1 Fresh software install

The CDP DC Cluster infrastructure configuration generally follows the guidelines in the Cloudera Data Platform - Data Center Reference Architecture in addition to Dell EMC Ready Solution for Cloudera Hadoop 6.1 Architecture Guide and the Dell EMC Ready Solution for Hortonworks Hadoop 3.0 Architecture Guide. The main difference between those guides and the CDP DC recommendations here is this cluster infrastructure is optimized for storage performance, with less memory and processor resources. This is because the bulk of the data processing runs on the container infrastructure instead of the CDP DC cluster. The minimum number of nodes in the CDP DC cluster is seven. There are three infrastructure nodes and four worker nodes. This cluster can be scaled after initial deployment by adding additional worker nodes.

2.2.2 Data Lake in-place upgrade

A significant advantage of the in-place upgrade approach is reuse of the current CDH or HDP cluster infrastructure. Before upgrading a cluster, the infrastructure it is running on should be evaluated to determine whether it meets the needs of a full CDP DC environment.

1. Upgrades generally do not require additional infrastructure capacity. However, if a cluster is running at full capacity, additional nodes may be needed to provide headroom for test and verification activity. Some housekeeping activity to archive old or unused data or jobs is also worthwhile.
2. After CDP Private Cloud is deployed, most of the CDP DC compute experiences workloads will migrate to the new infrastructure. The CDP DC infrastructure becomes a storage-centric cluster. This configuration means that smaller memory, and lighter processor configurations, may be adequate for the CDP DC cluster as workloads migrate.
3. CDP Private Cloud relies heavily on the network infrastructure for data access. We recommend a minimum of two bonded 10 GbE Ethernet connections, with 25 GbE or faster preferred. Clusters with less network capacity should probably not be upgraded. Although faster network cards can be installed, the return on labor effort may not be worthwhile.
4. If the cluster is due for an infrastructure refresh, it is probably best to follow the CDP Private Cloud migration upgrade path.

2.2.2.1 Upgrade prerequisites

CDP DC 7.1 is expected to support upgrades from the following earlier releases:

- CDH 5.x, version 5.13 or later
- HDP 2.6.5
- CDP DC 7.0

If you are running versions of CDH older than 5.13 or HDP older than 2.6.5, then an upgrade to the later release is required before upgrading to CDP DC 7.1. Support for upgrades from HDP 3.x and CDH 6.x is planned for future releases of CDP DC 7.x.

2.2.2.2 Upgrading from CDH 5.13-5.16 to CDP DC 7.1

CDH upgrades involve the following steps:

1. Upgrade the OS and RDBMS servers.
2. Uninstall deprecated components (such as Flume.).

3. Upgrade Cloudera Manager to 7.1.
4. Upgrade the CDH cluster to CDP DC 7.1.
5. Perform post upgrade steps, including activation of new features or any required refactoring.

Note: The upgrade to CDP DC 7.1 includes a new version of Hive, and conversion to Atlas, Ranger, Tez, and the Capacity Scheduler.

2.2.2.3 Upgrading from HDP 2.6.5 to CDP DC 7.1

HDP upgrades involve the following steps:

1. Upgrade the OS and RDBMS servers.
6. Uninstall deprecated components (such as Flume).
7. Upgrade Ambari.
8. Upgrade the HDP cluster to CDP DC 7.1.
9. Transfer cluster management from Ambari to Cloudera Manager 7.1.
10. Perform post upgrade steps, including activation of new features or any required refactoring.

Note: The HDP upgrade to CDP DC 7.1 involves conversion to Cloudera Manager from Ambari.

---

3 Infrastructure guidance

This chapter provides specific infrastructure configuration guidance for deployment of Cloudera Data Center Private Cloud, based on the Migration and upgrade scenarios previously described.

The configurations are based on Dell EMC PowerEdge servers and Dell EMC PowerSwitch networking. The inclusion of Intel next gen technology optimized for this solution allows customers realize even more from their investments. Intel Xeon delivers powerful performance for even the most compute hungry workloads. The addition of Intel® Optane™ persistent memory helps businesses extract more actionable insights from data – from cloud and databases and in-memory analytics. Intel® Optane™ persistent memory is an innovative memory technology that delivers a unique combination of affordable large capacity and support for data persistence. This technology coupled with Dell EMC infrastructure and Cloudera's solution can help businesses get faster insights from their data-intensive applications as well as deliver the benefits of consistently improved service scalability.

This guidance is based on experience with Cloudera Enterprise, Hortonworks Data Platform, and OpenShift container platforms. At this point in the development cycle, it is a best-effort guidance. It is intended to assist with infrastructure planning and is subject to change. Before making any purchasing decisions, contact your Dell EMC and Cloudera representatives for the latest information and assistance with specific deployment scenarios.

3.1 New installation recommendations

A fresh installation of CDP PVC involves completely new infrastructure. The recommendations provided here are for production deployments using current technology and best practices.

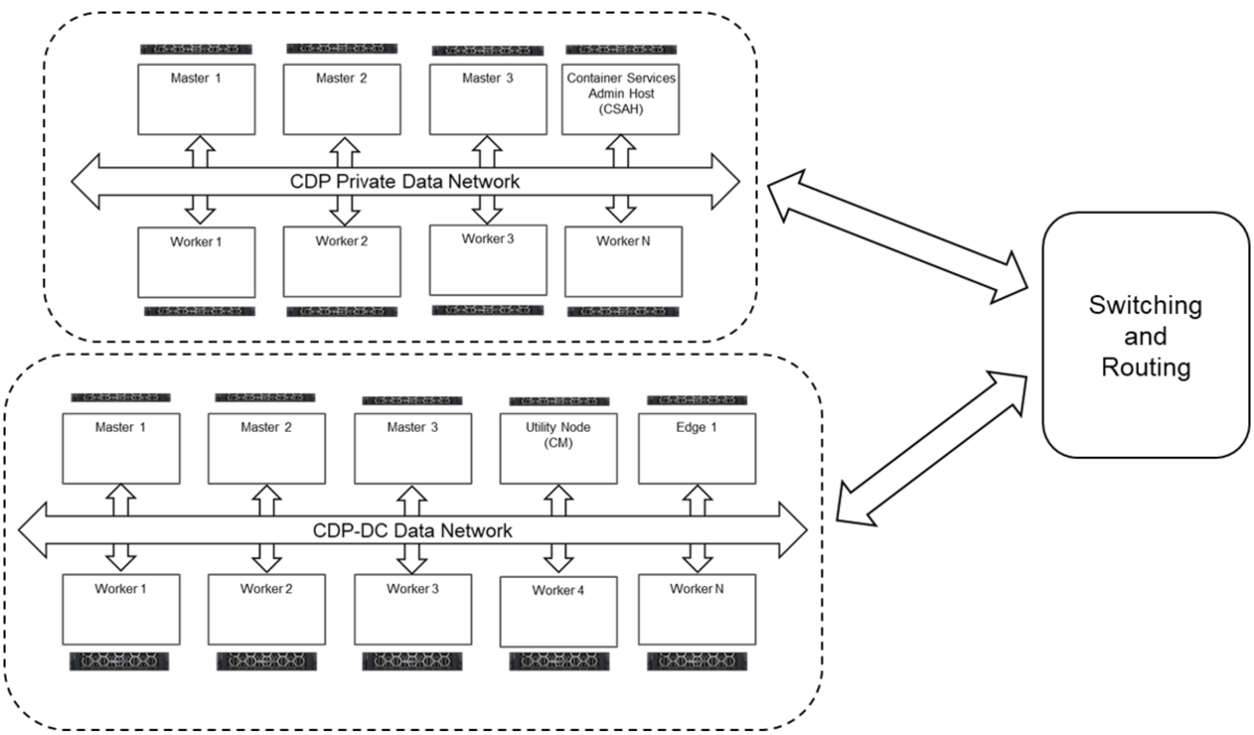

Figure 3 CDP Private Cloud Infrastructure

Figure 3 Illustrates the node level detail of a new deployment. The following sections provide details on the configurations of the individual nodes in the two clusters and the recommended network infrastructure. Table 1 provide a summary of the software services that each node provides.

Physical Node	Software Function
CDP DC Edge Node	Hadoop Clients Cloudera Manager
CDP DC Master Nodes 1-3	NameNode Resource Manager ZooKeeper
CDP DC Worker Nodes	DataNode NodeManager CDP DC (YARN) workloads
CDP PVC Master Nodes 1-3	OpenShift services Kubernetes Services
CDP PVC Worker Nodes	Kubernetes Operators CDP PVC workload pods

Table 1 CDP DC and CDP PVC Nodes and Roles

3.1.1 CDP Data Center infrastructure nodes

At least three infrastructure nodes are required. More nodes with the same configuration can be used for edge nodes. This configuration is sized for approximately 1 PB of cluster storage or 250 worker nodes.

Machine Function	CDP DC Infrastructure Nodes
Platform	Dell EMC PowerEdge R640
Chassis	2.5 in. chassis with up to 10 hard drives and 2 PCIe slots
Processor	Dual Intel Xeon Gold 6234 3.3 GHz (8 Core) 24.75 M cache
RAM	384 GB (12 x 32 GB 2933 MT/s)
Network Daughter Card	Mellanox ConnectX-4 Lx Dual Port 25 GbE DA/SFP rNDC
Boot Configuration	From PERC controller
Storage Controller	Dell EMC PERC H740P RAID Controller, 8 GB NV Cache, Minicard
Storage - HDD	8 x 1 TB 7.2 K RPM NLSAS 12 Gbps
Storage - SSD	2 x 480 GB SSD SAS 12 Gbps 512e 2.5in Hot-Plug

Table 2 CDP DC Infrastructure node configuration

3.1.2 CDP Data Center worker nodes

A minimum of three worker nodes is required, but we recommend at least five to support high availability and downtime for maintenance.

3.1.2.1 Recommended configuration

This configuration is the recommended general-purpose CDP Data Center worker node.

Machine Function	CDP DC Worker Node
Platform	Dell EMC PowerEdge R740xd server
Chassis	Chassis with up to 12 x 3.5 in. HDD, 4 x 3.5 in. HDDs on MP and 4 x 2.5 in. HDDs on Flex Bay

Processor	Dual Intel Xeon Gold 6246R 3.4 GHz, 16 Core, 35.75 M Cache
RAM	192 GB (12 x 16 GB 2933 MT/s)
Network Daughter Card	Mellanox ConnectX-4 Lx Dual Port 25 GbE DA/SFP rNDC
Boot Configuration	BOSS controller card + with 2 M.2 Sticks 240 GB
Storage Controller	Dell EMC PERC HBA330 Controller, 12 Gb Minicard
Storage – HDD	16 x 4 TB 7.2 K RPM SATA 6 Gbps 512n 3.5 in. hot-plug hard drive
Storage – SSD	4 x 480 GB SSD SAS mixed-use 12 Gbps

Table 3 CDP Data Center Worker Node - Recommended Configuration

This configuration assumes workloads such as CDW, CML, and Data Hub are also running on the CDP DC nodes. If most workloads will be running on the CDP PVC cluster, then the Storage-only configuration may be more appropriate.

We recommend Intel Xeon Gold processors for best performance when using erasure coding with HDFS. The HDFS erasure coding feature uses the Intel Storage Acceleration Library, which uses the AES-NI, SSE, AVX, AVX2, and AVX 512 instruction sets that Xeon Gold processors support.

3.1.2.2 Storage-only configuration

This configuration is recommended for deployments where the CDP-DC cluster will be providing HDFS storage with minimal processing, and most of the processing will be performed on the CDP-PVC cluster.

Machine Function	Worker Nodes
Platform	Dell EMC PowerEdge R740XD
Chassis	Chassis with up to 12 x 3.5 in. HDD, 4 x 3.5 in. HDDs on MP and 4 x 2.5 in. HDDs on Flex Bay
Processor	Intel® Xeon® Gold 5217 Processor (8 core, 11.0 M Cache, 3.0 GHz)
RAM	92 GB (12 x 16 GB RDIMM, 2933MT/s)
Network Daughter Card	Mellanox ConnectX-4 Lx Dual Port 25 GbE DA/SFP rNDC
Boot Configuration	BOSS controller card + with 2 M.2 Sticks 240 GB
Storage Controller	Dell EMC PERC HBA330 Controller, 12 Gb Minicard
Storage – HDD	12 x 4 TB 7.2 K RPM SATA 6 Gbps 512n 3.5 in. hot-plug hard drive
Storage – SSD	2 x 480 GB SSD SAS mixed-use 12 Gbps

Table 4 CDP DC Worker Node – Storage-only configuration

3.1.3 CDP PVC master nodes

The Master nodes host OpenShift services. Four nodes are required - three master nodes and a Container Services Administration Host. This master node configuration is adequate for OpenShift container clusters up to 250 nodes and rarely needs to be customized.

Machine Function	Master Node
Platform	Dell EMC PowerEdge R640
Chassis	2.5 in. chassis with up to 10 hard drives, 8 NVMe drives, and 3 PCIe slots, 2 CPU only
Power Supply	Dual, hot-plug, redundant power supply (1+1), 750 W

Processor	Dual Intel Xeon Gold 6226 2.7 G, 12C/24T, 10.4 GT/s, 19.25M cache)
RAM	192 GB (12 x 16 GB 2933MT/s)
Network Daughter Card	Mellanox ConnectX-4 Lx Dual Port 25 GbE SFP 28 rNDC
Network - additional	Mellanox ConnectX-4 Lx dual port 25 GbE SFP 28 NIC, low profile
Storage Controller	Dell EMC PERC HBA330 Controller, 12 Gb Minicard
Storage – OS and Data	1 x Dell 3.2 TB, NVMe, Mixed Use Express Flash, 2.5 SFF Drive, U.2, P4610 with Carrier

Table 5 CDP PVC Master Node Recommended Configuration

3.1.4 CDP PVC worker nodes

A minimum of four worker nodes are required. For a new deployment, we recommend 10 to 20 worker nodes.

OpenShift supports heterogenous node configurations, and it is possible to create specialized node configurations. These recommendations assume usage of CDW and CML on the container cloud. Contact your Dell EMC representative for assistance in sizing and customizing any of these configurations.

3.1.4.1 Recommended configuration

This worker node configuration provides a good balance of compute and memory for typical CDP cloud workloads. The primary big data storage is HDFS on the Cloudera-DC cluster, so local storage is limited. There is enough local storage to support temporary files, disk caches, and storage for other applications.

Machine Function	Compute Nodes
Platform	Dell EMC PowerEdge R640
Chassis	2.5 in. chassis with up to 10 hard drives, 8 NVMe drives, and 3 PCIe slots, 2 CPU only
Processor	2x Intel Xeon Gold 6248R processor (24 cores, 3.0 GHz, 35.75 MB Cache)
RAM	384 GB (12 x 32 GB RDIMM, 2933MT/s, Dual Rank)
Network Daughter Card	Mellanox ConnectX-4 Lx Dual Port 25 GbE SFP 28 rNDC
Additional Network Card	Mellanox ConnectX-4 Lx dual port 25 GbE SFP28 NIC, low profile
Boot Configuration	UEFI BIOS boot mode with GPT partition
Storage Controller	HBA330 12 Gbps SAS HBA Controller (NON-RAID), Minicard
Storage – OS + Data	1 x 6.4 TB, NVMe, Mixed Use, 2.5in, U.2, Intel P4610

Table 6 CDP PVC Worker Node - Recommended Configuration

The minimum memory requirement for CDP PVC worker nodes is 192 GB. We recommend 384 GB because the cloud worker nodes run multiple containers and can take advantage of the larger memory available automatically.

3.1.4.2 Large memory configuration

This worker node is a large memory configuration (1.5 TB) with additional processor capacity. This configuration may be suitable for workloads that benefit from keeping large datasets in memory.

Machine Function	Compute Nodes
Platform	Dell EMC PowerEdge R640
Chassis	2.5 in. chassis with up to 10 hard drives, 8 NVMe drives, and 3 PCIe slots, 2 CPU only
Processor	2x Intel Xeon Gold 6258R processor (28 cores, 2.7 GHz, 38.75 MB Cache)
RAM	384 GB (12 x 32 GB RDIMM, 2933MT/s, Dual Rank)
Persistent Memory	1536 GB (12 x 128 GB Intel Optane DC persistent memory)
Network Daughter Card	Mellanox ConnectX-4 Lx Dual Port 25 GbE SFP 28 rNDC
Additional Network Card	Mellanox ConnectX-4 Lx dual port 25 GbE SFP28 NIC, low profile
Boot Configuration	UEFI BIOS boot mode with GPT partition
Storage Controller	HBA330 12 Gbps SAS HBA Controller (NON-RAID), Minicard
Storage - Data	1 x 6.4 TB, NVMe, Mixed Use, 2.5in, U.2, Intel P4610

Table 7 CDP PVC Worker Node - Large Memory Configuration

The Intel Optane DC persistent memory that is used in this configuration enables large memory capacity, with higher performance than traditional storage.

3.1.5 Network recommendations

We recommend the use of Dell EMC PowerSwitch hardware for the cluster networking. Dell EMC PowerSwitch networking provides:

- Disaggregated-hardware and software switching solutions
- Support for Open Network Install Environment (ONIE), enabling zero-touch installation of alternate network operating systems
- Your choice of network operating system to help simplify data-center fabric orchestration and automation.
- A broad ecosystem of applications and tools, both open-source and Linux-based, providing more options to optimize and manage your network.
- Dell EMC high-capacity network fabrics are cost-effective and easy to deploy and provide a clear path to a software-defined data center. They offer high density for 25/40/50/100 GbE deployments in top-of-rack, middle-of-row, and end-of-row deployments.

The Dell EMC Networking OS10 Enterprise Edition is a network operating system supporting multiple architectures and environments, as shown in Figure 4.

OS10 allows multi-layered disaggregation of network functions. OS10 contributions to Open Source provide users with the freedom and flexibility to pick their own third-party networking, monitoring, management, and orchestration applications. OS10 Enterprise Edition bundles an industry-hardened networking stack featuring standard L2 and L3 protocols over established northbound interfaces such as CLI and SNMP. The Switch Abstraction Interface (SAI) and Control Plane Services (CPS) abstraction layers provide disaggregation at the Network Processing Unit (NPU), and for the software applications written on top of Linux kernel.

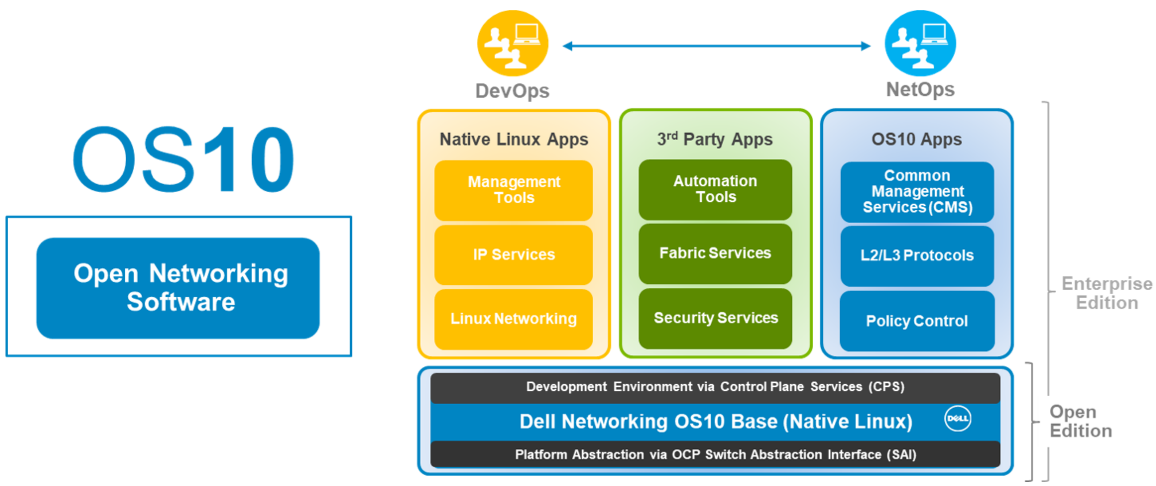

Figure 4 Dell EMC OS 10

3.1.5.1 Cluster networks

The cluster networks use a leaf-spine architecture. In the simplest deployment, each server in the cluster is attached to leaf switches in each rack and leaf switches have multiple connections to spine switches. Also, servers are connected by their iDRAC port to a 1 GbE management switch providing out of band access to the iDRAC interface.

The cluster network designs are based on the current [Dell EMC Ready Solution for Cloudera Hadoop 6.1 Architecture Guide,] [Dell EMC Ready Solution for Hortonworks Hadoop 3.0 Architecture Guide,] and [Dell EMC Ready Stack for Red Hat OpenShift Container Platform 4.2 Design Guide.]

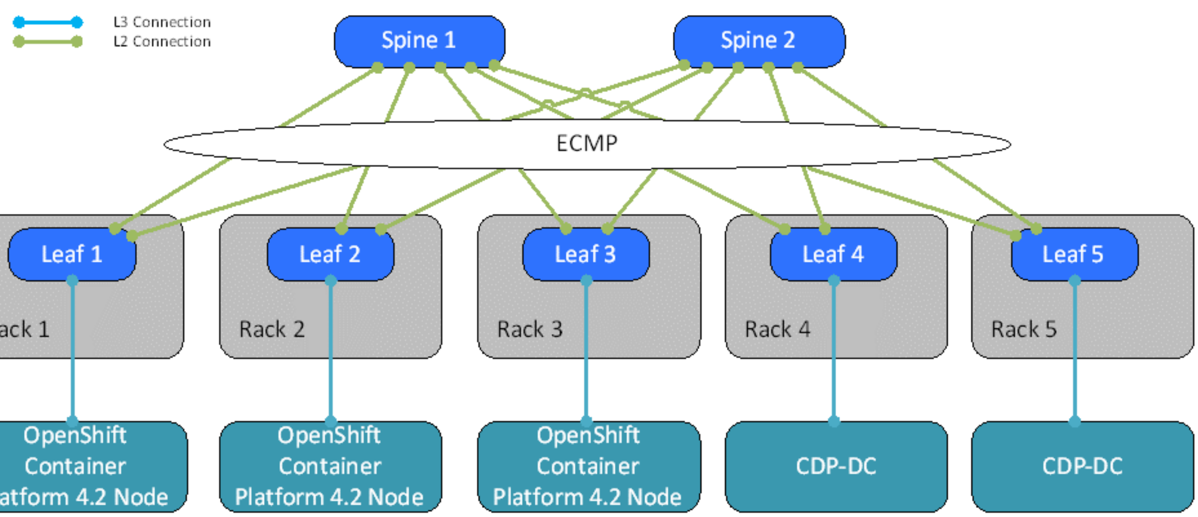

Figure 5 Leaf – Spine Cluster network

Table 8 lists the recommended Dell EMC PowerSwitch models.

Switch Model	Switch Configuration	Description
Dell EMC PowerSwitch S5248F-ON	48 x 25 GbE 6 x 100 GbE	Leaf Switch
Dell EMC PowerSwitch S5232F-ON	32 x 100 GbE	Spine Switch

Dell EMC PowerSwitch Z9264F-ON	64 x 100 GbE	Spine Switch (Large Cluster)
Dell EMC PowerSwitch S3148-ON	48 x 1 GbE	Management Switch

Table 8 Recommended Switches

Figure 6 illustrates the networks that are used for the CDP Data Center and CDP Private Cloud clusters, including the interconnect.

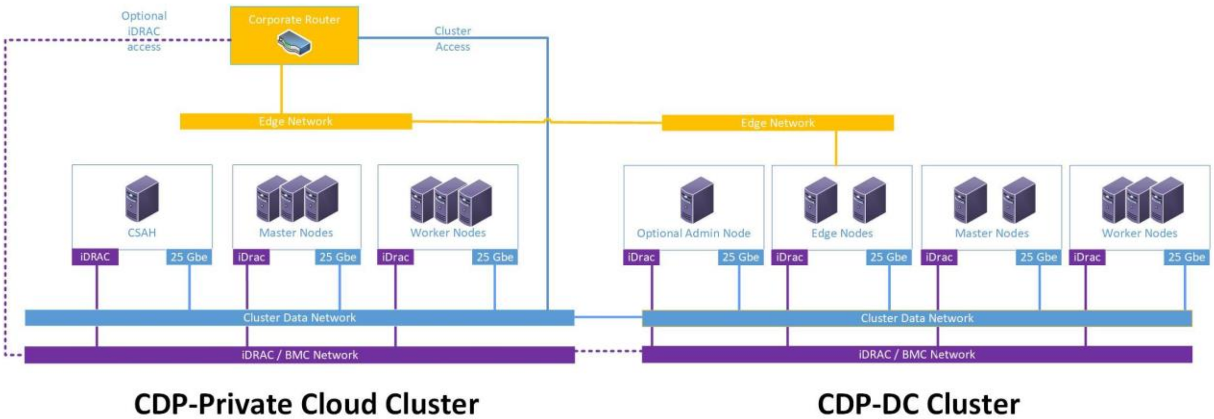

Figure 6 Cluster Networks

Table 9 describes the functions of the networks.

Network	Description	Available Services
CDP DC Cluster Data Network	The Data network carries the bulk of the traffic within the cluster. This network is aggregated within each pod, and pods are aggregated into the cluster switch.	The CDP DC services are available on this network. Note: The CDP DC services do not support multihoming and are only accessible on the Cluster Data Network.
CDP Private Cluster Data Network	The Data network carries the bulk of the traffic within the cluster. This network is aggregated within each pod, and pods are aggregated into the cluster switch.	The CDP PVC services are available on this network.
iDRAC/BMC	The BMC network connects the BMC or iDRAC ports and the out-of-band management ports of the switches. It is used for hardware provisioning and management. This network is aggregated into a management switch in each rack.	This network provides access to the BMC and iDRAC functionality on the servers. It also provides access to the management ports of the cluster switches.
Edge Network	The Edge network provides connectivity from one or more Edge Nodes to an existing premise network, either directly, or by the pod or cluster aggregation switches.	SSH access to one or more Edge Node is available on this network, and other application services may be configured and available.

Table 9 Cluster Networks

Connectivity between the clusters and existing network infrastructure can be adapted to specific installations. Common scenarios are:

1. The Cluster Data network is isolated from any existing network and access to the cluster is by the Edge network only.

2. The Cluster Data network is exposed to an existing network. In this scenario, the Edge network is either not used, or is used for application access or ingest processing.

3.1.5.2 iDRAC management network

Besides the Cluster Data network, a separate network is provided for cluster management – the iDRAC (or BMC) network.

The iDRAC management ports are all aggregated into a per rack Dell EMC Networking S3048-ON switch, with dedicated VLAN. This arrangement provides a dedicated iDRAC/BMC network, for hardware provisioning and management. Switch management ports are also connected to this network.

If out of band management is required, the management switches can be connected to the core or connected to a dedicated Management network.

3.1.5.3 Cluster interconnect sizing

The two clusters are assumed to reside in separate racks and are connected by an interconnect uplink.

Cloudera recommends a minimum bandwidth of 1 Gbps between each CDP DC and CDP PVC node. Based on this and the planned cluster sizes, the required bandwidth can be calculated. For example:

Parameter	Value
Network Adapters per node	1 x 25Gbe
Minimum bandwidth	1 Gbps
CDP DC nodes	10
CDP DC bandwidth required	250 Gbps
CDP PVC Nodes	20
CDP PVC jobs per node	10
CDP PVC bandwidth required	200 Gbps
Interconnect Bandwidth required	200 Gbps

The Interconnect link in this case should be 200 Gbps to provide the required sustained bandwidth. The interconnect can be scaled in 100Gbs increments for higher bandwidth needs as the size of the clusters increase. The interconnect can also be oversubscribed if necessary. The switches in Table 8 support multiple 100 GbE uplinks to support the interconnect.

3.2 Upgrade recommendations

We always recommend a new deployment of the CDP Private Cloud infrastructure. For CDP DC, existing infrastructure may be suitable when using the CDP Private Cloud in-place upgrade strategy.

3.2.1 CDP DC upgrade recommendations

Figure 7 shows the CDP DC cluster nodes for reference.

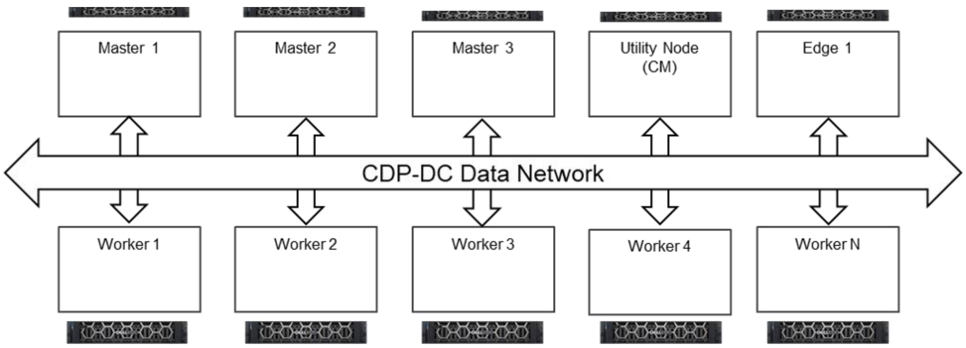

Figure 7 CDP DC cluster nodes

A significant advantage of upgrades is reuse of the current CDH or HDP cluster infrastructure. Since deployed clusters vary in configuration and technology, it is difficult to make specific upgrade recommendations. Before upgrading a cluster to CDP DC, the infrastructure it is running on should be evaluated to determine whether it meets the needs of a full CDP DC environment. For in-place upgrades, the primary requirement is that the cluster infrastructure should meet the same requirements as those specified for CDH or HDP in the Cloudera Enterprise Reference Architecture for Bare Metal Deployments.

Beyond those recommendations, our general guidance is:

1. After CDP Private Cloud is deployed, most of the CDP DC compute workloads will begin migrating to the new infrastructure. The CDP DC infrastructure becomes a storage-centric cluster. This configuration means that smaller memory and lighter processor configurations may be adequate for the CDP DC cluster.
11. CDP Private Cloud relies heavily on the network infrastructure for data access. We recommend a minimum of two bonded 10 GbE Ethernet connections, with 25 GbE or faster preferred. Clusters with 1 GbE networks should probably not be upgraded. Although faster network cards can be installed, the infrastructure and labor costs should be considered.
12. If the cluster is due for an infrastructure refresh, it is probably best to follow the CDP Private Cloud migration upgrade path.

3.3 Additional infrastructure topics

3.3.1 Repurposing infrastructure

Repurposing infrastructure often comes up in upgrade conversations, so this document provides some suggestions that are based on experience.

The recommended CDP DC workers are storage heavy configurations, while the CDP PVC worker nodes are memory and compute heavy. Older CDH, HDP, or CDP DC worker nodes generally cannot be reused in the CDP PVC cluster for the cloud workloads. Those nodes are potentially useful in another environment for storage heavy applications. In some instances, it may be possible to add them to the OpenShift cluster for non-CDP PVC workloads.

After a cluster refresh, we often see that customers repurpose older CDH or HDP nodes for development, test, or disaster recovery clusters. In many instances, those nodes are added to existing clusters that are not performance critical.

3.3.2 Heterogenous nodes

The core OpenShift platform supports heterogenous node types, including compute, memory, and accelerator optimized configurations. The necessary support for seamless use of heterogenous nodes is not currently available in either OpenShift 4.2 or CDP PVC. They are not recommended for initial deployments. Software support in this area is evolving rapidly, so if these configurations are of interest, contact your Dell EMC or Cloudera representative for the latest status.

3.3.3 Using an existing OpenShift cluster

Depending on the OpenShift node configurations, this reuse may be possible. Contact your Dell EMC or Cloudera representative for the latest status.

3.3.4 Datacenter and physical topics

The recommended configurations use 19” rackmount servers in 1U and 2U form factors. Depending on available power and cooling capacity, approximately 30 CDP PVC nodes or 12 CDP DC nodes fit in a standard 42U rack. Figure 8 shows typical rack installations.

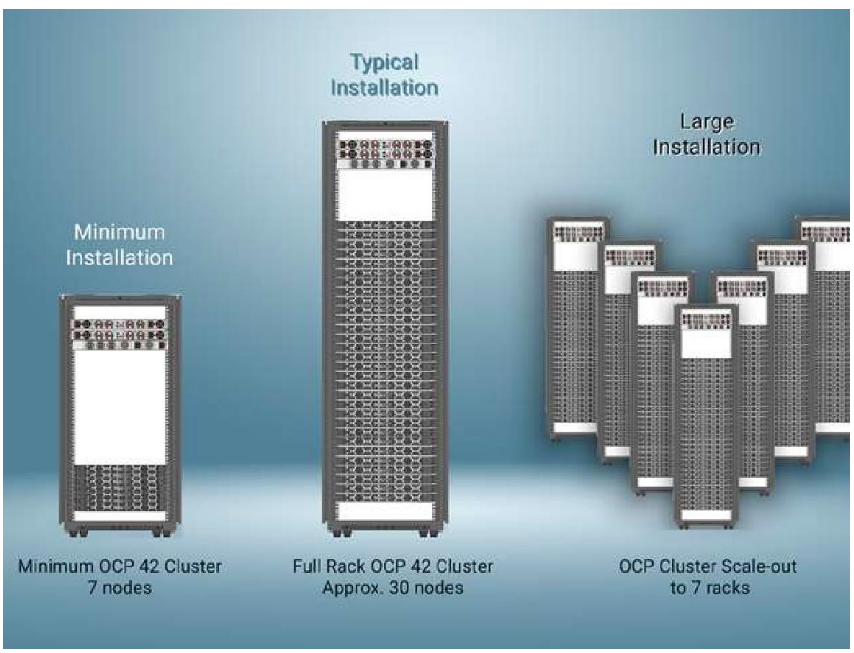

Figure 8 Sample rack installations

Rack enclosures and power distribution units are site-specific. It is essential to review the physical dimensions and power requirements during a site survey. Accurate power and cooling information for specific server configurations can be calculated using the Dell EMC Enterprise Infrastructure Planning Tool.

We recommend the use of dual redundant power supplies and dual power distribution units (PDU) for each rack. Figure 9 shows power distribution in a typical rack.

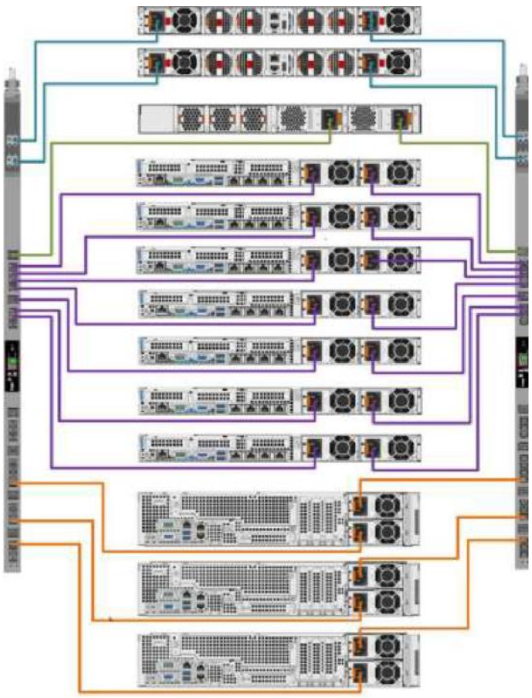

Figure 9 Power distribution example

3.3.5 Ozone object storage planning

Ozone is a scalable, redundant, and distributed object store optimized for big data workloads. Ozone can scale to billions of objects of varying sizes and can function effectively in containerized environments such as Kubernetes and YARN.

Note: Ozone is available for technical preview and is still under development. It should not be used in production systems yet.

From an infrastructure planning perspective, the initial release of Ozone is expected to run on the CDP DC cluster and use storage resources from that infrastructure. For customers interested in Ozone, contact your Cloudera representative for the latest status.