

Cloudera Data Platform (CDP) 7.1.6 with PowerFlex and PowerScale

August 2021

H18845.1

White Paper

Abstract

This white paper highlights the design considerations of CDP 7.1.6 on Dell EMC PowerFlex and PowerScale.

Dell Technologies Solutions

Copyright

The information in this publication is provided as is. Dell Inc. makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

Use, copying, and distribution of any software described in this publication requires an applicable software license.

Copyright © 2021 Dell Inc. or its subsidiaries. All Rights Reserved. Dell Technologies, Dell, EMC, Dell EMC and other trademarks are trademarks of Dell Inc. or its subsidiaries. Intel, the Intel logo, the Intel Inside logo and Xeon are trademarks of Intel Corporation in the U.S. and/or other countries. Other trademarks may be trademarks of their respective owners. Published in the USA 08/21 White Paper H18845.1.

Dell Inc. believes the information in this document is accurate as of its publication date. The information is subject to change without notice.

Contents

Revision history4

Executive Summary5

Product overview6

Solution overview.....9

CDP deployment14

Testing and Validation16

Conclusion.....19

Appendix A Solution components20

Appendix B References22

Revision history

Table 1. Revision history

| Date | Document revision | Description of changes |
|-------------|-------------------|--|
| August 2021 | 1.1 | Additions: <ul style="list-style-type: none">• Revision table Updates: <ul style="list-style-type: none">• Corrected PowerFlex terminology in VMware considerations• Footers display correct information• Document ID updated to H18845.1• Feedback email link updated with new document ID |
| July 2021 | 1.0 | Initial release |

Executive Summary

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 workload 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 7.1.6 on Dell EMC infrastructure with independent compute and storage. Dell EMC PowerFlex is used to provide the compute environment and high-performance scalable local storage. Dell EMC PowerScale provides scalable network attached storage for large datasets. A high-performance Dell EMC PowerSwitch network fabric connect these systems.

Using PowerFlex for the compute infrastructure provides operational simplicity and flexibility for the CDP environment. The PowerFlex storage layer provides scalable, high-performance access for runtime data needed by CDP.

PowerScale is used as the primary storage for the CDP environment. The PowerScale HDFS protocol adapter is a Cloudera Certified storage system that provides, scalable, high-performance storage with the durability and storage efficiency of OneFS. The Hadoop Name Node functionality that is integrated into PowerScale simplifies the deployment and scaling of CDP since dedicated Name Node servers are not required.

Unlike tightly coupled compute and storage environments, this architecture also provides the flexibility to add additional workloads outside the CDP environment. The PowerFlex capabilities support the deployment of additional virtual machines on the compute cluster with local storage. The multiprotocol capabilities of PowerScale support access to the CDP data over protocols like NFS.

Note: This document may contain language from third-party content that is not under Dell's control and is not consistent with Dell's current guidelines for Dell's own content. When such third-party content is updated by the relevant third parties, this document will be revised accordingly.

Audience

This white paper is intended for Cloudera administrators, system engineers, IT program managers, IT architects, and IT managers. Readers are expected to have an understanding and working knowledge of:

- Dell EMC PowerFlex family
- PowerScale
- Cloudera CDP deployment
- Kerberos

Terminology

Table 2 provides definitions for some of the terms that are used frequently in this document.

Table 2. Terminology

| Term | Definition |
|------|----------------------------------|
| HCI | Hyperconverged Infrastructure |
| SDS | Storage Data Server |
| SDC | Storage Data Client |
| VM | Virtual Machine |
| NAS | Network-attached Storage |
| HDFS | Hadoop Distributed File system |
| CDP | Cloudera Data Platform |
| KDC | Kerberos Key Distribution Center |

We value your feedback

Dell Technologies and the authors of this document welcome your feedback on the solution and the solution documentation. Contact the Dell Technologies Solutions team by [email](#) or provide your comments by completing our [documentation survey](#).

Author: Swathi Nagaram, Michael Pittaro, Kirankumar Bhusanurmam

Contributors: Nicholas Ruggiero, Dale McDonald

Note: For links to additional documentation for this solution, see [Dell Technologies Solution Info Hub for PowerFlex](#).

Product overview

This solution broadly consists of PowerFlex, PowerScale, and Cloudera CDP Base.

PowerFlex family

PowerFlex is a software-defined storage platform that is designed to significantly reduce operational and infrastructure complexity. This platform empowers organizations to move faster by delivering flexibility, elasticity, and simplicity with predictable performance and resiliency at scale. The PowerFlex family provides a foundation that combines compute and high-performance storage resources in a managed unified fabric. PowerFlex comes in flexible deployment options - rack, appliance, or ready nodes - that enable two-layer (compute and server SAN), single-layer (HCI), and storage-only architectures. PowerFlex is ideal for high-performance applications and databases, building an agile private cloud, or consolidating resources in heterogeneous environments.

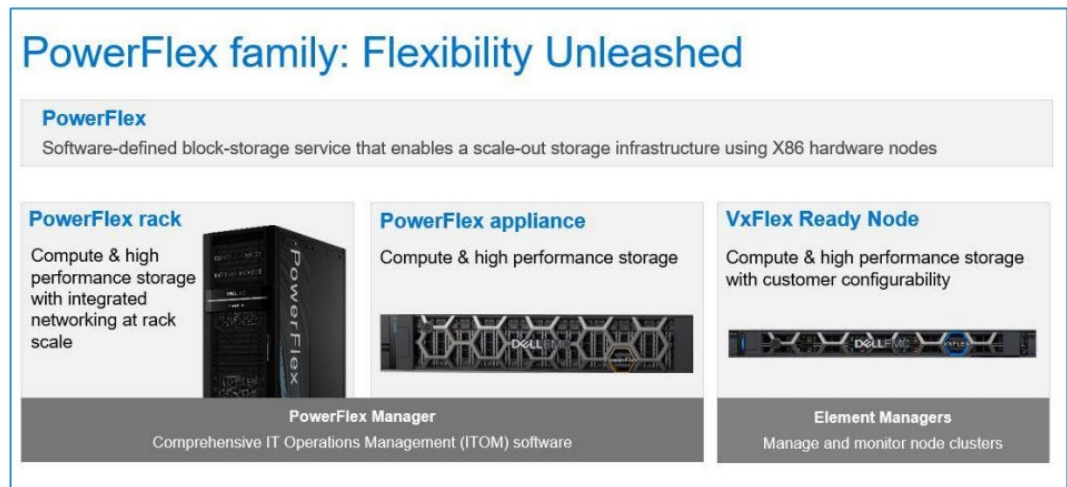


Figure 1. PowerFlex family

PowerFlex software components

Software is the key differentiation and the “secret sauce” in the PowerFlex offering. PowerFlex software components not only provide software-defined storage services. They also simplify infrastructure management and orchestration with comprehensive ITOM and LCM capabilities. These capabilities span compute and storage infrastructure, from BIOS and Firmware to nodes, software, and networking.

The core foundational component in the PowerFlex family that enables Software Defined Storage (SDS) services is called “PowerFlex.” This name represents the core value that it enables for the platform. Also, PowerFlex Manager is a comprehensive IT Operational Management (ITOM) and Life Cycle Management (LCM) tool that drastically simplifies management and ongoing operation.

PowerFlex

PowerFlex (previously VxFlex OS) is the software foundation of PowerFlex software-defined storage. It is a scale-out block storage service that is designed to deliver flexibility, elasticity, and simplicity with predictable high performance and resiliency at scale.

PowerFlex Manager

PowerFlex Manager is the software component in the PowerFlex family that enables ITOM automation and life cycle management capabilities for PowerFlex systems.

PowerScale

The Dell EMC PowerScale scale-out network-attached storage (NAS) platform provides Hadoop clients with direct access to big data through a Hadoop File System (HDFS) interface. Powered by the distributed Dell EMC PowerScale OneFS operating system, a PowerScale cluster delivers a scalable pool of storage with a global namespace. The distributed OneFS operating system combines the memory, I/O, CPUs, and disks of the nodes into a cohesive storage unit. This unit presents a global namespace as a single file system.

Hadoop compute clients access the data that is stored in a PowerScale cluster by using the HDFS protocol. Every node in the cluster can act as a NameNode and a DataNode. Each node boosts performance and expands cluster capacity. For Hadoop analytics, the PowerScale scale-out distributed architecture minimizes bottlenecks, rapidly serves big

data, and optimizes performance for analytics jobs. The NameNode daemon is a distributed process that runs on all the nodes in the cluster. A compute client can connect to any node in the cluster to access NameNode services. The nodes work together as peers in a shared-nothing hardware architecture with no single point of failure.

A PowerScale cluster is platform agnostic for compute. You can run most of the common Hadoop distributions with a PowerScale cluster. Clients running different Hadoop distributions or versions can simultaneously connect to the cluster.

Dell EMC PowerScale OneFS 8.2 and Cloudera Data Platform for Private Cloud Base 7.1.6 are Cloudera Quality Assurance Test Suite (QATS) certified. See [Certification Report](#) for the detailed information about the Certificate and QATS execution summary report.

Follow the [OneFS best practices guide](#) and [Dell EMC PowerScale Best Practice Guide for Hadoop Data Storage](#) for configuring and managing the HDFS service on a Dell EMC PowerScale cluster to optimize data storage for Hadoop analytics.

Dell Technologies recommends that you follow the [Dell EMC PowerScale network design best practice](#) for design considerations of the Dell EMC PowerScale external network to ensure maximum performance and an optimal user experience.

Cloudera Data Platform

Cloudera Data Platform (CDP) Private Cloud Base is the on-premises version of the Cloudera Data Platform. This new product combines the best of Cloudera Enterprise Data Hub and Hortonworks Data Platform Enterprise along with new features and enhancements across the stack. This unified distribution is a scalable and customizable platform where you can securely run many types of workloads.

CDP Private Cloud Base supports various hybrid solutions where compute tasks are separated from data storage and where data can be accessed from remote clusters. This hybrid approach provides a foundation for containerized applications by managing storage, table schema, authentication, authorization, and governance.

CDP Private Cloud Base consists of various components including Apache HDFS, Apache Hive 3, Apache HBase, Apache Impala, and many other components for specialized workloads. You can select any combination of these services to create clusters that address your business requirements and workloads. Several preconfigured packages of services are also available for common workloads. These packages include:

Table 3. CDP Private Cloud Base packages

| Workload | Use cases | Services |
|------------------|---|--|
| Data Engineering | Ingest, transform, and analyze data. | HDFS, YARN, YARN Queue Manager, Ranger, Atlas, Hive metastore, Hive on Tez, Spark, Oozie, Hue, Data Analytics Studio |
| Data Mart | Browse, query, and explore your data in an interactive way. | HDFS, YARN, YARN Queue Manager, Ranger, Atlas, Hive metastore, Impala, Hue |

| Workload | Use cases | Services |
|----------------------|--|----------------------------|
| Operational Database | Low latency writes, reads, and persistent access to data for Online Transactional Processing (OLTP) use cases. | HDFS, Ranger, Atlas, HBase |

When installing a CDP Data Center cluster, you install a single parcel, called Cloudera Runtime, that contains all the components. For a complete list of the included components and versions, see [Cloudera Runtime and Component Versions](#).

In addition to the Cloudera Runtime components, the CDP Data Center includes powerful tools to manage, govern, and secure your cluster.

For Cloudera Runtime components documentation, see [CDP - Data Center Component Documentation](#).

Cloudera Manager

Cloudera Manager is an end-to-end application for managing CDP clusters. Cloudera Manager provides granular visibility into, and control over, every part of the CDP cluster. These capabilities empower operators to improve performance, enhance quality of service, increase compliance, and reduce administrative costs.

Kerberos

Kerberos is a network authentication protocol that is designed to provide strong authentication for client/server applications by using secret-key cryptography in most distributed systems, including CDP. Kerberos provides secure and reliable authentication to multiple applications. PowerScale systems support the Kerberos authentication feature using Kerberos Key Distribution Center (KDC) services.

Solution overview

This section provides an overview of the logical architecture of CDP 7.1.6 deployment on PowerFlex and PowerScale.

Logical architecture

[Figure 2](#) shows the Cloudera CDP logical architecture as deployed upon PowerFlex and PowerScale.

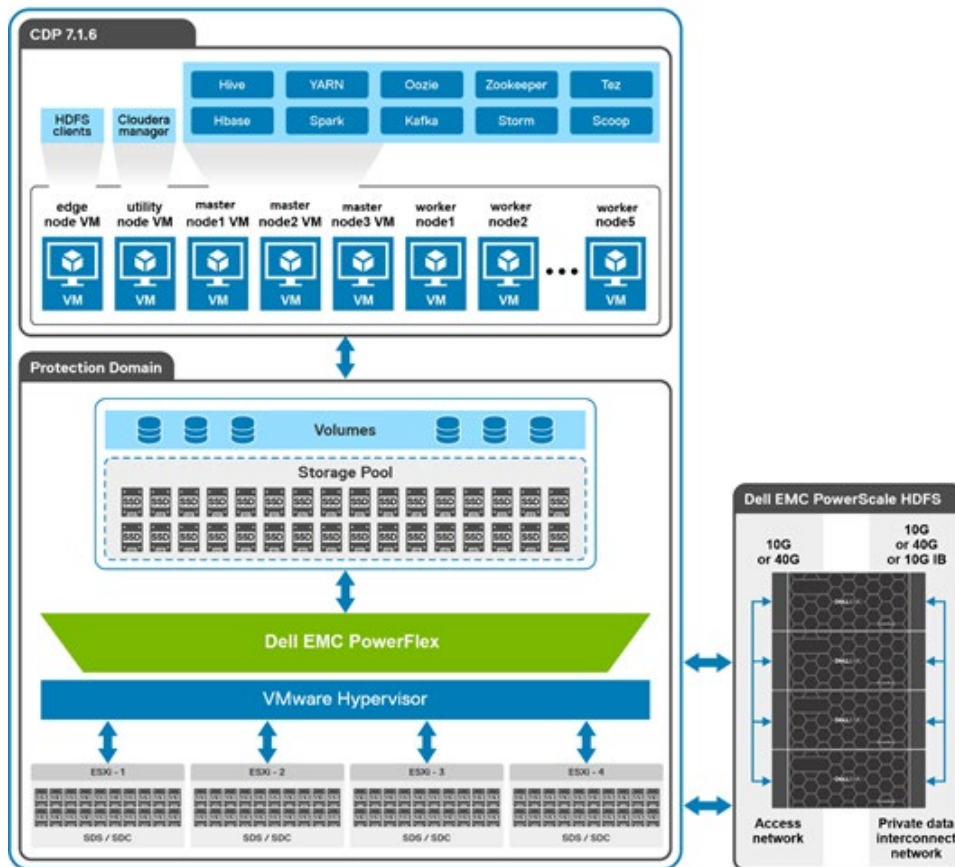


Figure 2. Logical architecture of CDP deployment on PowerFlex and PowerScale

The PowerFlex software-defined storage was configured in a VMware vSphere environment and deployed with a single-layer Hyperconverged Infrastructure (HCI) architecture option. The HCI configuration has both SDC and SDS software components that are configured on each node of the cluster. SDC provides the compute capabilities while SDS provides storage for CDP runtime components. The PowerScale scale-out network-attached storage (NAS) platform provides Hadoop clients with direct access to big data through a Hadoop File System (HDFS) interface. Powered by the distributed Dell EMC PowerScale OneFS operating system, a PowerScale cluster delivers a scalable pool of storage with a global namespace.

Hadoop compute clients access the data that is stored in a PowerScale cluster by using the HDFS protocol. Every node in the PowerScale cluster can act as a NameNode and a DataNode.

The first consideration for the implementation is the number of CDP clusters to be deployed. A common use case here is for Dev, UAT, and Prod Clusters. Each CDP cluster requires its own Access Zone on PowerScale, and its own pool of Linux VMs.

Networking design

The following figure shows the logical layout of PowerFlex rack access and aggregation with management aggregation architecture. From the different topologies below, Dell Technologies went with the Hyperconverged topology, mainly because of the small scale for the deployment.

In general, this decision should be made with the architecture endgame in mind. The amount of local storage that the VMs require is typically 10% to 20% of the usable HDFS storage. If the final solution needs 1 PB of HDFS storage, then the local storage would be at most 200 TB. The storage must be distributed across the PowerFlex nodes. If the final PowerFlex number was 20 nodes, then 10 TB per node seems easy enough for an HCI topology.

Note: There is an additional 1 Gb link from the PowerFlex controller nodes to the out-of-band management switch.

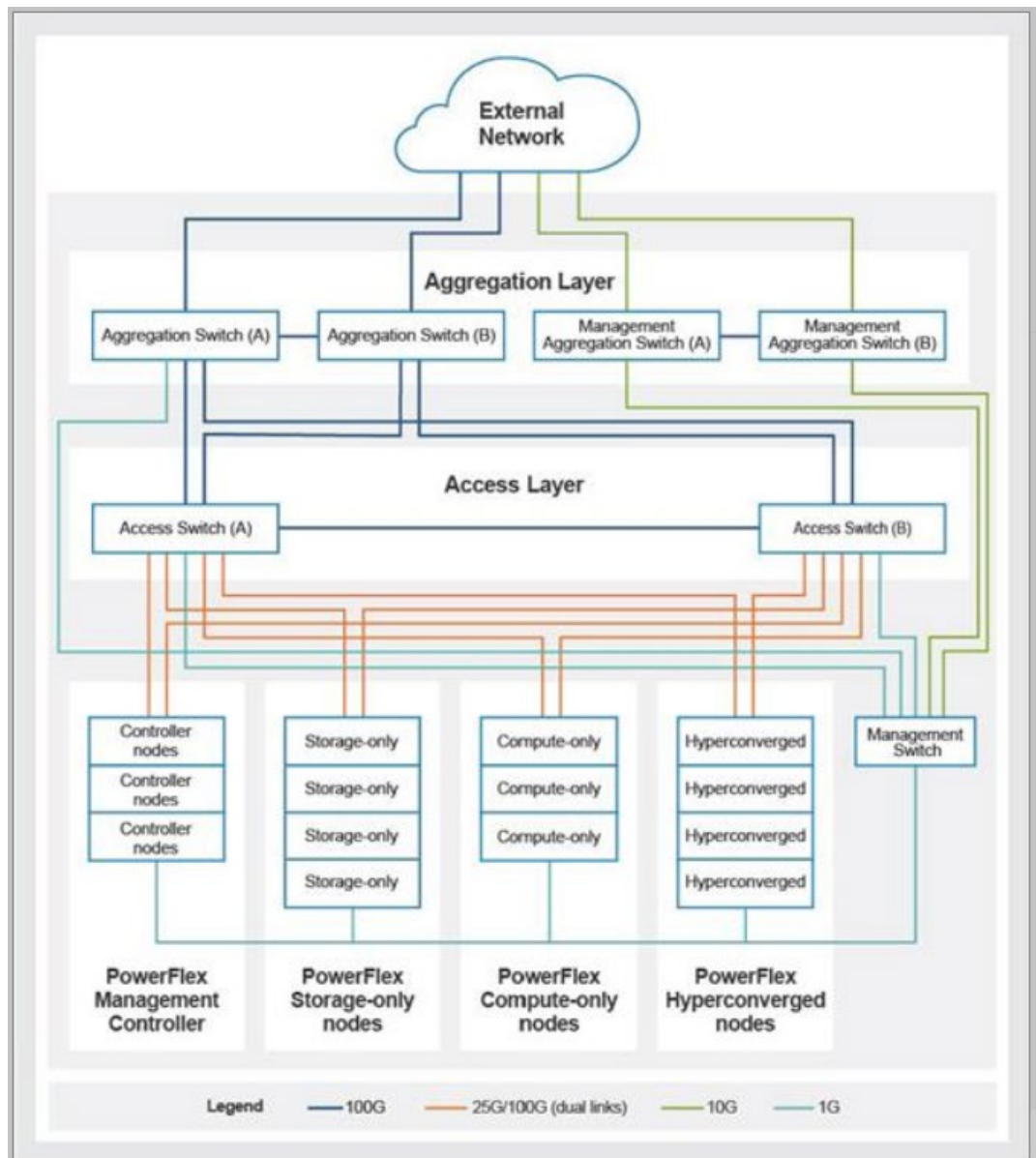


Figure 3. PowerFlex rack access and aggregation architecture

PowerFlex design

The Storage Data Server (SDS) aggregates the raw local storage in a node and serves it up as PowerFlex storage. A single protection domain is carved out of the SSD drives on the four PowerFlex hyperconverged nodes. A single storage pool is configured, and multiple volumes were carved out to meet the CDP requirements. These volumes are mapped to the ESXi cluster and added as a datastore. They are later mapped as disks drive to Hadoop virtual machines using VMware Paravirtual SCSI (PVSCSI) adapters.

Table 4. PowerFlex storage details for CDP virtual machines

| VM | Operating system | Data1 | Data2 |
|-----------------------------|------------------|-------|-------|
| Master [1-3], Utility, Edge | 500 GB | N/A | N/A |
| Worker [1-5] | 500 GB | 1 TB | 1 TB |

Note: Worker nodes require local storage for shuffle, Spark cache, and other functions. Dell Technologies recommends using the PowerFlex storage for this configuration. Leveraging two devices provides a good balance for simplicity, resiliency, and performance.

PowerScale design

SmartConnect is a module that specifies how the OneFS cluster handles connection requests from clients. For more information and best practices for SmartConnect, see the [Isilon External Network Connectivity Guide](#).

Each SmartConnect zone represents a specific pool of IP addresses. You associate a SmartConnect zone with an access zone. OneFS then allows only clients that connect through the IP addresses in the SmartConnect zone to reach the HDFS data in the access zone. A root HDFS directory is specified for each access zone. This configuration isolates data within access zones and allows you to restrict client access to the data.

CDP virtualization design

The CDP cluster environment consists of multiple software services running on multiple virtual machines. The implementation divides the server VMs into several roles, and each VM has a configuration that is optimized for its role in the cluster.

The VM configurations are divided into two broad classes:

- **Worker VMs** - Worker VMs handle the bulk of the Hadoop processing.
- **Master VMs** - Master VMs support services that are needed for the cluster operation.

Table 5. VM configurations

| VM | vCPU | Memory |
|--------------|------|--------|
| Master [1-3] | 12 | 64 GB |
| Utility | 4 | 16 GB |
| Edge | 2 | 16 GB |
| Worker [1-5] | 16 | 175 GB |

Note: Dell Technologies does not recommend overcommitting the memory.

Table 6. VM definitions

| VM | Count | Definition |
|---------|-------|---|
| Master | 3 | Runs all the daemons that are required to manage the cluster storage and compute services |
| Worker | 5 | Runs all the services that are required to store blocks of data on the local hard drives and run processing tasks against that data |
| Edge | 1 | Runs Cloudera Manager and the Cloudera Management services |
| Utility | 1 | Contains all client-facing configurations and services, including gateway configurations |

VMware considerations

The choice of which flavor of VM should be mapped to which physical host is a dynamic decision. The HDFS data can be accessed from any physical node, and the runtime storage on PowerFlex can also be accessed from any physical node. This design provides flexibility in VM placement since vMotion can be used to migrate VMs at any time to alleviate hotspots and or perform node maintenance. The design also provides flexibility in VM sizing, allowing the VM configurations to be changed based on workload requirements.

For the architecture testing, Dell Technologies chose a simple deployment scenario using four physical PowerFlex nodes and ten virtual machines with fixed mapping of VMs to physical nodes. See [Table 7, VM services](#).

Table 7. VM services

| VM | CDP service | PowerFlex physical node number |
|------------|--|--------------------------------|
| Master VM1 | FailoverController, YARN ResourceManager, ZooKeeper, JobHistory Server, Spark History Server, Kudu master | Node-1 |
| Master VM2 | FailoverController, YARN ResourceManager, ZooKeeper, JobHistory Server, Spark History Server, Kudu master | Node-3 |
| Master VM3 | ZooKeeper, Kudu master (All require an odd number of masters for high availability.) | Node-4 |
| Utility VM | Cloudera Manager, Cloudera Management Service, Hive Metastore, Impala Catalog Server, Impala StateStore, Oozie, ZooKeeper (Requires a dedicated disk), Apache Atlas, Apache Ranger | Node-3 |
| Edge VM | Hue, HiveServer2, Gateway configuration | Node-1 |

| VM | CDP service | PowerFlex physical node number |
|-----------------|--|---|
| Worker VM [1-5] | NodeManager, Impalad, Kudu tablet server | Node-1, Node-2 (two workers), Node3, Node-4 |

Note: Unlike standard CDP installations, the Master VMs do not run Cloudera NameNode processes, and the Worker VMs do not run DataNode processes. The PowerScale HDFS protocol provides these services.

CDP deployment

The deployment process for this architecture is similar to the deployment of CDP with PowerScale. See the [PowerScale CDP private base Install guide](#). The main differences are the use of PowerFlex and virtualization for the compute platform. This section summarizes the deployment steps and references additional documents for the installation details. The main steps in the deployment are:

1. [PowerFlex setup](#)
2. [PowerScale setup](#)
3. [Cloudera Manager setup](#)
4. [Kerberos configuration](#)
5. [CDP Private Cloud Base deployment](#)

PowerFlex setup

The PowerScale infrastructure should be installed following the normal Dell EMC recommendations. For assistance, contact your Dell EMC or partner representative.

Once the infrastructure is ready, deploy the PowerFlex cluster. For more details, see the [PowerFlex deployment guide](#). Then configure the network. See [Networking design](#) for more details.

Once the cluster is up and running, create 10 Linux virtual machines that are based on the recommendations in CDP virtualization design.

PowerScale setup

The PowerScale infrastructure should be installed following the normal Dell EMC recommendations. For assistance, contact your Dell EMC or partner representative.

After the infrastructure is deployed, OneFS should be configured for use with CDP. Follow the instructions in the *Installing OneFS with Cloudera Manager* section of the [PowerScale CDP Private Cloud Base Install guide](#). In particular, an access zone must be created for the CDP Base cluster to use.

Cloudera Manager setup

Install Cloudera Manager on the Utility VM. Follow the Cloudera installation process at [Cloudera Manager Download Information and Installing Cloudera Manager](#).

Install the PowerScale Custom Service Descriptor

After Cloudera Manager is installed, then download and install the PowerScale Custom Service Descriptor (CSD) into Cloudera Manager. See the [PowerScale CDP Private Base Install guide](#) section 4.4.1, *Download PowerScale CSD into Cloudera Manager host*.

Kerberos configuration

Kerberos must be installed and configured on both PowerScale and Cloudera Manager. This process is different from the standard Cloudera process for enabling Kerberos since Kerberos must be enabled before deploying CDP Base.

First, enable Kerberos on the PowerScale cluster by following the instructions in the [PowerScale CDP private base Install guide](#) section 3.2.3, *Enable Kerberos on the HDFS Zone and view HDFS settings*.

Then enable Kerberos on Cloudera Manager. See [Install Cloudera Manager and CDP](#) for more details about enabling Kerberos on CDP 7.1.6 and the [PowerScale CDP private base Install guide](#) section 4.5, *Kerberizing Cloudera Manager* for more information.

CDP Private Cloud Base deployment

CDP Private Cloud Base can now be deployed from Cloudera Manager into the remaining VMs. Follow the guidelines in VM service for the service assignments.

Note: Do not deploy the Cloudera HDFS service. Instead, use the Dell EMC PowerScale service.

See the [PowerScale CDP private base Install guide](#) for the installation details. Additional information can be found in the [CDP Private Cloud Base Installation Guide](#).

Validating the deployment

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.

1. Log in to a host in the cluster.

Run the Hadoop PiEstimator example using the following command.

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

2. In Cloudera Manager, go to **Cluster > Cluster Name > YARN Applications**.
3. Check the results of the job. It appears similar to the following example:

| | | | | | |
|--|---|------------------------------------|--------------------------------------|---------|---------|
| 05/22 10:45 AM | - | Name: QuasiMonteCarlo | Pool: root.hdfs | | |
| 05/22 10:46 AM | | Mapper: QuasiMonteCarlo\$QmcMapper | Reducer: QuasiMonteCarlo\$QmcReducer | Actions | Details |
| Type: MapReduce ID: job_1400700704311_0001 Duration: 54.27s User: hdfs CPU Time: 34.15s | | | | | |
| File Bytes Read: 98 B File Bytes Written: 992.7 KIB HDFS Bytes Read: 2.7 KIB HDFS Bytes Written: 215 B | | | | | |
| Memory Allocation: 184.7M Pool: root.hdfs | | | | | |

Figure 4. Job results

Testing and Validation

This section describes functional validation with multiple workloads to ensure validity and lack of bottlenecks of the solution. Performance numbers are beyond the scope of this white paper.

Overview

The objective was to run a series of industry-standard workloads to determine the CDP Worker Nodes configuration that provides optimal compute, network, and storage resources utilization.

Benchmarks tools

A suite of three industry-standard workload tools was used to validate the configurations.

Testing with DFSIO

The TestDFSIO workload tool is a read and write test for HDFS. It is a utility that comes with CDP. TestDFSIO was run with many files to create multiple execution threads. This benchmark utility is like a “fire hose” test for the environment and shows that an optimal network architecture is in place.

This test is designed to test the aggregate bandwidth between CDP clients and PowerScale nodes.

Dell Technologies ran the read and write commands for each test case. The sum of read/write execution times determined the ranking; lower being the better.

Example DFSIO commands:

```
yarn jar /opt/cloudera/parcels/CDH/jars/hadoop-mapreduce-client-jobclient-3.1.1.7.1.6.0-297-tests.jar TestDFSIO -write -nrFiles 1000 -size 1GB
```

```
yarn jar /opt/cloudera/parcels/CDH/jars/hadoop-mapreduce-client-jobclient-3.1.1.7.1.6.0-297-tests.jar TestDFSIO -read -nrFiles 1000 -size 1GB
```



Figure 5. Aggregate network bandwidth between compute nodes and PowerScale

Testing with Terasuite

The Terasuite workload tool combines testing of HDFS and MapReduce layers of a Hadoop cluster. The goal is to generate, sort, and validate 1 TB of data (or any other amount of data) as fast as possible.

This test is designed to exercise the compute and local storage configurations with concurrent HDFS access.

Dell Technologies ran the Terasuite commands for each of the test cases. The sum of Teragen, Terasort, Teravalidate execution times determined the ranking, lower being the better.

Example Teragen, Terasort, Teravalidate commands

Example commands for Teragen, Terasort, and Teravalidate include:

```
time yarn jar /opt/cloudera/parcels/CDH/lib/hadoop-mapreduce/hadoop-mapreduce-examples.jar -Dmapred.map.tasks=320
teragen 10000000000 /benchmarks/tera32-32-run3
```

```
time yarn jar /opt/cloudera/parcels/CDH/lib/hadoop-mapreduce/hadoop-mapreduce-examples.jar terasort
/benchmarks/tera32-32-run3 /benchmarks/terasort32-32-run3
```

```
time yarn jar /opt/cloudera/parcels/CDH/lib/hadoop-mapreduce/hadoop-mapreduce-examples.jar teravalidate
/benchmarks/terasort32-32-run1 /benchmarks/teraval32-32-run3
```



Figure 6. The Terasort job and the aggregate network bandwidth



Figure 7. Host information during the Terasort job

Testing with YCSB

YCSB was used to characterize the performance of the HBase NoSql database. The YCSB simulated workloads used were:

- Workload A: Mixed workload: 50/50% mix of reads and inserts
- Workload B: Read mostly workload: 95/5% mix of reads and updates
- Workload D: Reads and Inserts: 95/5 % mix of reads and inserts



Figure 8. YCSB testing and the HBASE metrics from Cloudera Manager

One billion records were loaded and then workloads A, B, D were run. The YSCB results were ranked using an average of the overall operations per second. A higher rate of operations is considered better.

Test Scenarios

Three different scenarios were evaluated, based on the PowerFlex compute configuration and using five virtualized CDP Worker Nodes.

Table 8. Test scenarios

| Test Scenarios | Physical cores | Worker VM cores | YARN container cores |
|----------------|----------------|-----------------|----------------------|
| Scenario 1 | 16 | 16 | 16 |
| Scenario 2 | 16 | 16 | 32 |
| Scenario 3 | 16 | 32 | 32 |

Note: YARN container cores were changed by modifying the YARN configuration parameters `yarn.nodemanager.resource.cpu-vcores` and `yarn.scheduler.maximum-allocation-vcores`.

Scenario 1 used an exact mapping of worker VM cores and YARN containers cores to available physical cores. Scenario 2 used a 2:1 mapping between YARN container cores and physical cores, allowing YARN to manage task parallelism. Scenario 3 used a 2:1 mapping for both VM cores and YARN cores to physical cores, allowing VMware to do the hyperthreading.

The DFSIO, Terasuite, and YCSB tests were run for each scenario, and the results were evaluated.

Validation Results

All the test suites ran successfully, with good utilization of available compute and storage resources.

The relative performance for the different scenarios is shown in [Table 9](#), with 1 being the best, and 3 being the least performant.

Table 9. Scenario ranking

| Benchmark tool | TestDFSIO ranking | Terasuite ranking | YSCB ranking |
|----------------|-------------------|-------------------|--------------|
| Scenario 1 | 3 | 2 | 2 |
| Scenario 2 | 2 | 3 | 3 |
| Scenario 3 | 1 | 1 | 1 |

The difference in performance for the three scenarios were relatively small. Behavior may be different in a large environment. The flexibility of virtualized CDP nodes with decoupled storage made it simple to run these different scenarios. This capability is important when optimizing the configuration for specific workloads.

Conclusion

By following the guidance in this document, you can deploy Cloudera Private Cloud Base with a modern decoupled compute and storage architecture.

This configuration provides independent scaling and management of the storage and compute resources. The PowerFlex infrastructure provides flexibility and operational simplicity for the compute aspects of CDP and supports the optimum utilization of available resources. The PowerScale infrastructure provides reliable and efficient HDFS storage with the full operational management and administration capabilities of OneFS. It simplifies CDP deployment by eliminating the need for NameNode services on the CDP servers.

Appendix A Solution components

This appendix provides the hardware and software components that are used in the Cloudera Data Platform (CDP) 7.1.6 with PowerFlex and PowerScale solution. It also provides best practices guidance that Dell Technologies followed when deploying and validating the solution.

Hardware and Software components

Table 10 describes the PowerFlex components that are used in this solution.

Table 10. PowerFlex components

| PowerFlex component | Configuration |
|---------------------|--|
| Node | PowerEdge R640 |
| CPU | 2x Intel Xeon Platinum 8153 CPU @2 GHz ,16 cores |
| Memory | 24x16 GB DDR-4 |
| NIC | 2 x Mellanox ConnectX-4 LX 25 GbE SFP Adapter |
| Storage | Dell HBA330 controller 10 x 960 GB SATA SSD |
| PowerFlex | 3.5 |

Table 11 describes the PowerScale components that are used in this solution.

Table 11. PowerScale components

| PowerScale component | Configuration |
|----------------------|----------------------|
| PowerScale model | H600 |
| OneFS version | 9.1.2 |
| PowerScale nodes | 4 |
| Network | 1x1 GbE-2x40 GbE SFP |

Best practices

This section outlines the performance best practices that are followed in this solution.

PowerFlex rack

PowerFlex rack best practices include:

- The minimum number of nodes for production workloads is eight.
- Dell Technologies recommends homogenous node types for predictable performance.
- The maximum number of devices in a storage pool is 300.
- Change the passwords for all default accounts.
- Use secure communication - HTTPS (TCP port 443) - to remotely access PowerFlex nodes.
- Ensure that PowerFlex rack is compliant to an RCM.

PowerFlex

PowerFlex best practices include:

- Configure the high-performance profile for MDM, SDS, and SDC.
- Disable Read Flash cache and Read RAM cache for all flash clusters.
- Sometimes, increasing device queue length value to 256 can improve the I/O concurrency.

See [PowerFlex Tuning Parameters](#) for more details on PowerFlex performance tuning and best practices.

PowerFlex rack network

See [PowerFlex Networking Best Practices](#) for detailed networking best practices.

PowerScale

See the [Dell EMC PowerScale Hadoop Data Storage best practice guide](#) for PowerScale performance tuning and best practices.

Cloudera Data Platform

CDP best practices include:

Minimize anonymous page faults

Minimize anonymous page faults by setting `vm.swappiness=1`, freeing from page cache before “swapping” application pages. This setting reduces the OOM-killer invocation.

1. Edit `/etc/sysctl.conf` in your editor of choice, and add following line:

```
vm.swappiness=1
```

2. Then, run the following commands:

```
# sysctl -p
# sysctl -a|grep "vm.swappiness"
```

Disable transparent huge-page compaction

Add the following command to `/etc/rc.local` to ensure that transparent huge page compaction remains disabled across reboots:

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

Disable transparent huge-page defragmentation

Add the following command to `/etc/rc.local` to ensure that defragmentation remains disabled across reboots:

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

Network

The backend network between PowerFlex nodes and PowerScale should be 40 Gbps, with Jumbo Frames set (MTU=9162) for the NICs and the switch ports.

Appendix B References

Additional information can be obtained at the [Dell Technologies Info Hub for PowerFlex](#). If you need additional services or implementation help, contact your Dell Technologies sales representative.

Grafana dashboards

The Grafana Dashboards can be found by searching for the term, “Isilon” on the [Grafana website](#). For example, <https://grafana.com/grafana/dashboards?search=isilon>.

Dell Technologies documentation

The following Dell Technologies documentation provides additional and relevant information. Access to these documents depends on your login credentials. If you do not have access to a document, contact your Dell Technologies representative.

- [PowerFlex InfoHub](#)
- [PowerFlex Networking Best Practices](#)
- [PowerFlex Tuning Parameters](#)
- [OneFS best practice guide](#)
- [PowerScale CDP private base Install guide](#)
- [Isilon External Network Connectivity Guide](#)
- [Dell EMC PowerScale network design best practice](#)

Cloudera documentation

The following Cloudera 7.1.6 documentation provides additional and relevant information:

- [Cloudera Private Cloud Base 7.1.6](#)
- [Cloudera runtime components](#)