Dell PowerScale and Cloudera Data Platform Private Cloud Base QATS Certification Report

Cloudera Quality Assurance Test Suite (QATS) program certifies Dell PowerScale OneFS 8.2.2 using HDFS protocol as Primary Filesystem certified with Cloudera Data Platform (CDP) Private Base 7.1.6.

Abstract

This document provides execution summary report of the Cloudera Quality Assurance Test Suite (QATS) certification of Cloudera Data Platform (CDP) Private Cloud Base 7.1.6 on Dell PowerScale powered by Dell PowerScale OneFS 8.2.2.

April 2021

Revisions

Date	Description
April 2021	Initial release
April 2021	Added support for "Cruise Control, Schema Registry, Stream Replication Manager, Stream Messaging Manager, DAS" for fresh installs

Acknowledgments

Author: Kirankumar Bhusanurmath, Analytics Solutions Architect, Dell Technologies

Table of contents

Executive summary

This document provides details for functional testing of Cloudera Data Platform (CDP) Private Cloud Base with Dell PowerScale™ through the HDFS protocol as a primary file system. It provides observations to the Dell and Cloudera field and sales teams as certified components, caveats, and guardrails when engaging with customers who are interested in using Dell PowerScale with HDFS as a primary file system.

The Quality Assurance Test Suite (QATS) appliance from Cloudera was used for functional testing and certifying the functionality and operations of the latest CDP private base 7.1.6 with Dell PowerScale powered by OneFS 8.2.2. QATS certifies only the fresh install of CDP Private Cloud Base on the PowerScale OneFS and does not cover Cloudera upgrades and replication scenarios.

Note: The Cloudera QATS appliance does only functional validation and certification. Performance testing, Cloudera upgrades and replication scenarios were not conducted in this exercise. The Cloudera QATS is certified only for the new installation of Cloudera CDP Private Base 7.1.6 on PowerScale OneFS 8.2.2.

1 QATS test environment

This section discusses the software and hardware components used to setup Cloudera Data Platform, the QATS test environment, and the PowerScale OneFS cluster. Table 1 details the test environment.

Table 1   QATS test environment

Component	Version and resource
CDP private cloud base Version	Cloudera Manager 7.3.1: See download information here
	Cloudera Runtime 7.1.6: See download information here
PowerScale CSD	PowerScale-1.0.0.jar
Isilon Node	A200-4U-Single-16GB-1x1GE-2X10GE SFP+-30TB-400GB SSD
OneFS operating system	Isilon OneFS v8.2.2.0 B_8_2_2_007(RELEASE):
Compute and Operating System	Virtual Machines
	CentOS Linux release 7.7.1908 (Core)
	Memory: 64G
	CPU: 32

Note: CDP Ranger had an issue described at RANGER-2626. A OneFS custom patch was developed to fix this Ranger issue for the QATS testing, same fix is incorporated in the *"[PSP-1091]"* patch will be officially made available as a Roll Up Patch (RUP) in May 2021.

PowerScale Custom Service Descriptors (CSD) are built for Dell PowerScale, and you can download CSD from product download page and extract the contents on to the Cloudera Manager server.

The Cloudera full release documentation is available here including the Cloudera Manager Release Notes and Cloudera Runtime Release Notes.

2   Cloudera CDP components supported on Dell PowerScale

This section discusses the Cloudera CDP Private cloud base 7.1.6 components that successfully completed all the functional test cases of the Cloudera QATS Appliance. This component-specific test success determines the supportability of the components on Dell PowerScale OneFS 8.2.2.

Table 2   CDP components supported

Components	Version	Supported
HDFS on Isilon		YES
Apache YARN	3.1.1.7.1.6.0-297	YES
Apache MapReduce	3.1.1.7.1.6.0-297	YES
Apache ZooKeeper	3.5.5.7.1.6.0-297	YES
Apache HBase	2.2.3.7.1.6.0-297	YES
Apache Spark	2.4.5.7.1.6.0-297	YES
Apache Ranger	2.1.0.7.1.6.0-297	YES
Apache Atlas	2.1.0.7.1.6.0-297	YES
Apache Hive	3.1.3000.7.1.6.0-297	YES
Apache Tez	0.9.1.7.1.6.0-297	YES
Hive on Tez	1.0.0	YES
Apache Impala	3.4.0.7.1.6.0-297	YES
Apache Sqoop	1.4.7.7.1.6.0-297	YES
Apache Knox	1.3.0.7.1.6.0-297	YES
Apache Oozie	5.1.0.7.1.6.0-297	YES
Hue	4.5.0	YES
Apache Phoenix	5.1.0.7.1.6.0-297	YES
Apache Zeppelin	0.8.2.7.1.6.0-297	YES
Apache Livy	0.6.0.7.1.6.0-297	YES
Apache Kafka	2.5.0.7.1.6.0-297	YES
Apache Solr	8.4.1.7.1.6.0-297	YES
Cruise Control	2.0.100	YES
Schema Registry	0.10.0.7.1.6.0	YES
Streams Replication Manager	1.0.0	YES

Streams Messaging Manager	2.1.0	YES
Data Analytics Studio	1.4.2	YES

See Cloudera runtime components versions for more details.

2.1   CDP components planned in future release cycle

The following table list the CDP Private Cloud Base 7.1.6 components that are in the QATS certification scope and will be tested in the future release cycle.

Table 3   QATS testing planned for future release

Components	Version	Supported
Apache Kudu	NA	Pending
TDE	NA	Pending
TLS	NA	Pending

3 Specific features within a component not supported on Dell Isilon OneFS 8.2.2

This section discusses and lists some of the features of the specific components of Cloudera Data Platform Private Base 7.1.6 that were excluded from the Cloudera QATS testing. These features are marked as "Not Supported Features of a component on Dell PowerScale OneFS 8.2.2." Detailed explanations and alternative workarounds are provided, which are supported by Dell PowerScale.

3.1 Features not supported in OneFS as HDFS

The following table describes the features not supported in OneFS as HDFS.

Table 4   Features not supported in OneFS as HDFS

Feature	Remark
HDFS Extended ACLs	OneFS ACLs provides rich features, for more information see Dell Isilon: Access Control Lists on HDFS and Isilon OneFS
HDFS Cache	Isilon has internal caching within OneFS, for more information see Dell PowerScale OneFS SmartFlash
Erasure Coding - HDP getErasureCodingPolicy no supported	OneFS has Erasure Coding natively in OneFS, for information see High Availability and Data Protection with Dell PowerScale
fsck NameNode commands	NA. For more information see Dell PowerScale HealthCheck service packs
HDFS NFS Support	Dell PowerScale OneFS supports multi protocols, for more information on NFS support see Dell PowerScale: OneFS NFS Design Considerations and Best Practices.
Trash Behavior	NA
Distcp with preserve checksum (-pc) option.	Skip CRC checks when using PowerScale. PowerScale maintains its own integrity checking internally in OneFS.
HDFS Snapshot	PowerScale SnapshotIQ provides data protection. For more information see Data Protection with Dell PowerScale SnapshotIQ
HDFS ShortcircuitReads	NA
HDFS Quota	Dell EMS PowerScale provides Scale-out storage quota management and thin provisioning, for more information see Storage Quota Management and Provisioning with Dell PowerScale SmartQuotas
NameNode command: dfs -report	NA, Dell PowerScale has inbuilt customizable analytics platform to accelerate workflows and applications on Isilon cluster, for more information see Dell Isilon InsightIQ

3.2 Other features not supported in OneFS 8.2.2

None

4 Setup, workarounds, and issues

4.1 Environment setup

This section discusses the standard process used for setting up the Cloudera Data Platform Private Cloud Base 7.1.6 and any service-specific work-around employed during the Cloudera QATS testing.

1. Set up CDP 7.1.6 cluster using the CDP Private Cloud Base Installation Guide

2. Set up CDP on PowerScale. Dell Technologies used the following installation guide: Download "Dell EMC PowerScale OneFS 8.2.2 and CDP Private Base 7.1.6 install guide" from the product download page.

4.2 Manual workaround

Issue	Workaround
Failed to upload Tez jar file during installation	Manually upload the Tez jar
Zeppelin fails on first run	`chown zeppelin:zeppelin /var/lib/zeppelin`

4.3 Other issues and fixes

Ranger issue:

The Apache Ranger 2.0 had an issue which is fixed in RANGER-2626. The Dell Technologies team provided a OneFS 8.2.2 custom patch for the RANGER-2626 issue for the QATS testing. The same patch will be officially released as a *"[PSP-1091]"* Roll Up Patch in May 2021.

A   Technical support and resources

Dell.com/support is focused on meeting customer needs with proven services and support.

Storage technical documents and videos provide expertise that helps to ensure customer success on Dell storage platforms.

DELLTechnologies