

# Hortonworks Data Platform

## HDP-2.3.6 Release Notes

(June 28, 2016)

## Hortonworks Data Platform: HDP-2.3.6 Release Notes

Copyright © 2012-2016 Hortonworks, Inc. Some rights reserved.

The Hortonworks Data Platform, powered by Apache Hadoop, is a massively scalable and 100% open source platform for storing, processing and analyzing large volumes of data. It is designed to deal with data from many sources and formats in a very quick, easy and cost-effective manner.

The Hortonworks Data Platform consists of the essential set of Apache Software Foundation projects that focus on the storage and processing of Big Data, along with operations, security, and governance for the resulting system. This includes Apache Hadoop – which includes MapReduce, Hadoop Distributed File System (HDFS), and Yet Another Resource Negotiator (YARN) – along with Ambari, Falcon, Flume, HBase, Hive, Kafka, Knox, Oozie, Phoenix, Pig, Ranger, Slider, Spark, Sqoop, Storm, Tez, and ZooKeeper. Hortonworks is the major contributor of code and patches to many of these projects. These projects have been integrated and tested as part of the Hortonworks Data Platform release process and installation and configuration tools have also been included.

Unlike other providers of platforms built using Apache Hadoop, Hortonworks contributes 100% of our code back to the Apache Software Foundation. The Hortonworks Data Platform is Apache-licensed and completely open source. We sell only expert technical support, [training](#) and partner-enablement services. All of our technology is, and will remain, free and open source.

Please visit the [Hortonworks Data Platform](#) page for more information on Hortonworks technology. For more information on Hortonworks services, please visit either the [Support](#) or [Training](#) page. Feel free to [contact us](#) directly to discuss your specific needs.



Except where otherwise noted, this document is licensed under  
**Creative Commons Attribution ShareAlike 4.0 License.**  
<http://creativecommons.org/licenses/by-sa/4.0/legalcode>

# Table of Contents

1. HDP 2.3.6 Release Notes .....	1
1.1. New Features .....	2
1.2. Unsupported Features .....	2
1.2.1. Technical Preview Features .....	2
1.2.2. Community Features .....	3
1.3. Upgrading from HDP 2.3.x to HDP 2.3.6 .....	4
1.3.1. Before you begin .....	5
1.3.2. Upgrade Procedure .....	5
1.4. Behavioral Changes .....	8
1.5. Apache Patch Information .....	11
1.5.1. Hadoop .....	11
1.5.2. Accumulo .....	31
1.5.3. Atlas .....	34
1.5.4. Calcite .....	36
1.5.5. Falcon .....	38
1.5.6. Flume .....	42
1.5.7. HBase .....	44
1.5.8. Hive .....	48
1.5.9. Kafka .....	61
1.5.10. Knox .....	65
1.5.11. Mahout .....	66
1.5.12. Oozie .....	67
1.5.13. Phoenix .....	67
1.5.14. Pig .....	70
1.5.15. Ranger .....	71
1.5.16. Slider .....	78
1.5.17. Spark .....	79
1.5.18. Sqoop .....	83
1.5.19. Storm .....	83
1.5.20. Tez .....	85
1.5.21. ZooKeeper .....	91
1.6. Common Vulnerabilities and Exposures .....	92
1.7. Third-Party Licenses .....	93
1.8. Fixed Issues .....	93
1.9. Known Issues .....	102
1.10. Documentation Errata .....	105
1.10.1. Flume: Kafka Sink .....	106
1.10.2. Hive Sink .....	107
1.10.3. Configuring Pig Scripts to Use HCatalog in Oozie Workflows .....	110
1.10.4. Configuring a Sqoop Action to Use Tez to Load Data into a Hive Table .....	112

# List of Tables

- 1.1. Technical Previews ..... 2
- 1.2. Community Features ..... 3
- 1.3. Behavioral Changes ..... 8

# 1. HDP 2.3.6 Release Notes

This document provides you with the latest information about Hortonworks Data Platform (HDP) 2.3.6 and its product documentation.

## Component Versions

Hortonworks Data Platform (HDP) 2.3.6 is the first maintenance release that aligns the 2.3.x core with the 2.4.x core; see [New Features](#) for details.

The Hortonworks approach is to provide patches only when necessary, to ensure the interoperability of components. Unless you are explicitly directed by Hortonworks Support to take a patch update, each of the HDP components should remain at the following package version levels, to ensure a certified and supported copy of HDP 2.3.6.

Official Apache component versions for HDP 2.3.6 are the following:

- Apache Accumulo 1.7.0
- Apache Atlas 0.5.0
- Apache Calcite 1.2.0
- Apache DataFu 1.2.0
- Apache Falcon 0.6.1
- Apache Flume 1.5.2
- Apache Hadoop 2.7.1
- Apache HBase 1.1.2
- Apache Hive 1.2.1
- Apache Kafka 0.9.0.1
- Apache Knox 0.6.0
- Apache Mahout 0.9.0+
- Apache Oozie 4.2.0
- Apache Phoenix 4.4.0
- Apache Pig 0.15.0
- Apache Ranger 0.5.2
- Apache Slider 0.80.0
- Apache Solr 5.5.0
- Apache Spark 1.5.2
- Apache Sqoop 1.4.6

- Apache Storm 0.10.0
- Apache Tez 0.7.0
- Apache ZooKeeper 3.4.6

The following are additional supported component versions:

- Cascading 3.0.1
- Hue 2.6.1

## 1.1. New Features

Hortonworks Data Platform (HDP) 2.3.6 is the first maintenance release that aligns the 2.3.x core with the 2.4.x core; Apache component versions are the same across HDP 2.3.6 and HDP 2.4.2\*. HDP 2.3.6 includes:

- All 2.4.0 and 2.4.2 bug fixes and Apache patches\*
- Additional Apache patches beyond those ported from the common components across HDP 2.3.x and 2.4.x (see [Apache Patche Information](#))



### Note

\* Except Spark and Hive. For HDP 2.3.x, Spark 1.5.2 and Hive 1.2.1 are provided.

## 1.2. Unsupported Features

Although the following features exist within HDP 2.3.6, Hortonworks does not currently support these specific capabilities:

- [Technical Preview Features \[2\]](#)
- [Community Features \[3\]](#)

### 1.2.1. Technical Preview Features

The following features are available within HDP 2.3.6 but are not ready for production deployment. Hortonworks encourages you to explore these technical preview features in nonproduction environments and provide feedback on your experiences through the [Hortonworks Community Forums](#).

**Table 1.1. Technical Previews**

Apache Component	Feature
HBase and Phoenix	<ul style="list-style-type: none"> <li>• Phoenix Query Server</li> <li>• Phoenix Query Server (<a href="#">PHOENIX-971</a>)</li> <li>• Phoenix-Spark Integration</li> <li>• RPC Throttling</li> <li>• Support for <code>init.d</code> scripts</li> </ul>

Apache Component	Feature
Hive	<ul style="list-style-type: none"> <li>• Hive Streaming</li> <li>• ACID support</li> </ul>
Slider	<ul style="list-style-type: none"> <li>• Support for Docker-based application packaging (<a href="#">SLIDER-780</a>)</li> </ul>
Spark	<ul style="list-style-type: none"> <li>• GraphX</li> <li>• Dynamic Executor Allocation</li> </ul>
YARN	<ul style="list-style-type: none"> <li>• Support for network I/O isolation and scheduling for containers (<a href="#">YARN-2140</a>)</li> <li>• NodeManager cgroup support for disk I/O isolation (<a href="#">YARN-2619</a>)</li> </ul>
Zeppelin	<ul style="list-style-type: none"> <li>• Zeppelin 0.6.0</li> <li>• Installation and management of Zeppelin with Ambari 2.2.2</li> <li>• Improvements to PySpark support in Zeppelin</li> <li>• One-way Zeppelin support for SSL</li> <li>• Zeppelin support for form-based LDAP authentication</li> </ul>

## 1.2.2. Community Features

The following features are developed and tested by the Hortonworks community but are not officially supported by Hortonworks. These features are excluded for a variety of reasons, including insufficient reliability or incomplete test case coverage, declaration of nonproduction readiness by the community at large, and feature deviation from Hortonworks best practices. Do not use these features in your production environments.

**Table 1.2. Community Features**

Apache Component	Feature
Falcon	<ul style="list-style-type: none"> <li>• Prism Server</li> <li>• User Recipes</li> </ul>
HBase	<ul style="list-style-type: none"> <li>• HBase Column Family Encryption</li> <li>• Use HDFS data at rest encryption instead.</li> <li>• Use of Memcached as block cache (<a href="#">HBASE-13170</a>)</li> <li>• Region assignment without using ZooKeeper</li> <li>• Region size balancing (<a href="#">HBASE-13103</a>)</li> </ul>
HDFS	<ul style="list-style-type: none"> <li>• Block-volume device choosing (<a href="#">HDFS-1804</a>)</li> <li>• NameNode Federation (<a href="#">HDFS-1052</a>)</li> <li>• viewFS (<a href="#">HADOOP-7257</a>)</li> <li>• Disk Hot Swap (<a href="#">HDFS-1362</a>)</li> </ul>
Kafka	<ul style="list-style-type: none"> <li>• New Consumer API</li> </ul>
Knox	<ul style="list-style-type: none"> <li>• Storm REST APIs</li> </ul>
Oozie	<ul style="list-style-type: none"> <li>• Spark action (<a href="#">OOZIE-1983</a>)</li> </ul>
Slider	<ul style="list-style-type: none"> <li>• Simplified Application Packaging</li> </ul>
Spark	<ul style="list-style-type: none"> <li>• Spark Standalone</li> </ul>

Apache Component	Feature
YARN	<ul style="list-style-type: none"> <li>• Fair Scheduler</li> <li>• MapReduce Eclipse Plug-in</li> <li>• MapReduce Uber AM</li> </ul>

## 1.3. Upgrading from HDP 2.3.x to HDP 2.3.6

- [Before you begin \[5\]](#)
- [Upgrade Procedure \[5\]](#)



Hortonworks Data Platform (HDP) 2.3.6 is the first maintenance release that aligns the 2.3.x core with the 2.4.x core; see [New Features](#) for details.

These Release Notes include abbreviated upgrade instructions; for full upgrade instructions, use the [Non-Ambari Upgrade Guide](#).


HDP 2.3.6 is a maintenance release of HDP 2.3.x. If you already have HDP 2.3.x installed, upgrading your cluster to HDP 2.3.6 requires the following actions:

- Keeping the same configuration files you used for HDP 2.3.x
- Keeping the same data and metadata in the same location you used for HDP 2.3.x
- Installing any new components (added for the first time in HDP 2.3.6) side-by-side with existing components

The following table summarizes HDP 2.2.x-to-2.3.6 upgrade options:

Cluster Management	Supporting Doc	
Cluster managed manually (HDP 2.1 and earlier)	<a href="#">Non-Ambari Upgrade Guide</a>	
Cluster managed manually (HDP 2.3 and later)	<a href="#">Before you begin [5]</a>	
Cluster managed by Ambari 1.7.0	<a href="#">Before you begin [5]</a>	
Cluster managed by Ambari 2.0	<a href="#">Upgrading Ambari Guide</a>	<p>Ambari 2.0 supports rolling upgrade between HDP 2.2.x and HDP 2.3.6.</p> <p> <b>Note</b></p> <p>Ambari does not support rolling upgrade between HDP 2.1 and 2.3 for Falcon. Use <a href="#">Configure and Validate Falcon</a> to upgrade this component.</p>
Cluster managed by Ambari 2.1	<a href="#">Upgrading Ambari Guide</a>	<p>Ambari 2.1 supports rolling upgrade between HDP 2.3.x and HDP 2.3.6.</p> <p> <b>Note</b></p> <p>Ambari does not support rolling upgrade between HDP 2.1 and 2.3 for Falcon. Use</p>



Cluster Management	Supporting Doc	
		<a href="#">Configure and Validate Falcon</a> to upgrade this component.
Cluster managed by Ambari 2.2	<a href="#">Upgrading Ambari Guide</a>	<p>Ambari 2.2 supports rolling upgrade between HDP 2.3.x and HDP 2.3.6.</p> <p> <b>Note</b></p> <p>Ambari does not support rolling upgrade between HDP 2.1 and 2.3 for Falcon. Use <a href="#">Configure and Validate Falcon</a> to upgrade this component.</p>

### 1.3.1. Before you begin

- Ensure that you know what HDP components need to be upgraded at your installation.
- Decide whether you are going to upgrade using a [local repository](#) or a [remote repository](#).

### 1.3.2. Upgrade Procedure

To upgrade your cluster from HDP 2.3.x to HDP 2.3.6, follow these steps:

1. Download one of the following:

The appropriate HDP 2.3.6 `hdp.repo` file for your OS:

Operating System	Repository Location
Debian 6 (deprecated)	<a href="http://public-repo-1.hortonworks.com/HDP/debian6/2.x/updates/2.3.6.0/hdp.list">http://public-repo-1.hortonworks.com/HDP/debian6/2.x/updates/2.3.6.0/hdp.list</a>
Debian 7	<a href="http://public-repo-1.hortonworks.com/HDP/debian7/2.x/updates/2.3.6.0/hdp.list">http://public-repo-1.hortonworks.com/HDP/debian7/2.x/updates/2.3.6.0/hdp.list</a>
RHEL/CentOS/Oracle LINUX 6	<a href="http://public-repo-1.hortonworks.com/HDP/centos6/2.x/updates/2.3.6.0/hdp.repo">http://public-repo-1.hortonworks.com/HDP/centos6/2.x/updates/2.3.6.0/hdp.repo</a>
RHEL/CentOS/Oracle LINUX 7	<a href="http://public-repo-1.hortonworks.com/HDP/centos7/2.x/updates/2.3.6.0/hdp.repo">http://public-repo-1.hortonworks.com/HDP/centos7/2.x/updates/2.3.6.0/hdp.repo</a>
SLES 11 SP3/SP4	<a href="http://public-repo-1.hortonworks.com/HDP/suse11sp3/2.x/updates/2.3.6.0/hdp.repo">http://public-repo-1.hortonworks.com/HDP/suse11sp3/2.x/updates/2.3.6.0/hdp.repo</a>
Ubuntu 12	<a href="http://public-repo-1.hortonworks.com/HDP/ubuntu12/2.x/updates/2.3.6.0/hdp.list">http://public-repo-1.hortonworks.com/HDP/ubuntu12/2.x/updates/2.3.6.0/hdp.list</a>
Ubuntu 14	<a href="http://public-repo-1.hortonworks.com/HDP/ubuntu14/2.x/updates/2.3.6.0/hdp.list">http://public-repo-1.hortonworks.com/HDP/ubuntu14/2.x/updates/2.3.6.0/hdp.list</a>

The HDP RPM single repository tarball:

Operating System	Tarball Location
Debian 6 (deprecated)	<a href="http://public-repo-1.hortonworks.com/HDP/debian6/2.x/updates/2.3.6.0/HDP-2.3.6.0-debian6-deb.tar.gz">http://public-repo-1.hortonworks.com/HDP/debian6/2.x/updates/2.3.6.0/HDP-2.3.6.0-debian6-deb.tar.gz</a>
Debian 7	<a href="http://public-repo-1.hortonworks.com/HDP/debian7/2.x/updates/2.3.6.0/HDP-2.3.6.0-debian7-deb.tar.gz">http://public-repo-1.hortonworks.com/HDP/debian7/2.x/updates/2.3.6.0/HDP-2.3.6.0-debian7-deb.tar.gz</a>

Operating System	Tarball Location
RHEL/CentOS/Oracle LINUX 6	<a href="http://public-repo-1.hortonworks.com/HDP/centos6/2.x/updates/2.3.6.0/HDP-2.3.6.0-centos6-rpm.tar.gz">http://public-repo-1.hortonworks.com/HDP/centos6/2.x/updates/2.3.6.0/HDP-2.3.6.0-centos6-rpm.tar.gz</a>
RHEL/CentOS/Oracle LINUX 7	<a href="http://public-repo-1.hortonworks.com/HDP/centos7/2.x/updates/2.3.6.0/HDP-2.3.6.0-centos7-rpm.tar.gz">http://public-repo-1.hortonworks.com/HDP/centos7/2.x/updates/2.3.6.0/HDP-2.3.6.0-centos7-rpm.tar.gz</a>
SLES 11 SP3/SP4	<a href="http://public-repo-1.hortonworks.com/HDP/suse11sp3/2.x/updates/2.3.6.0/HDP-2.3.6.0-suse11sp3-rpm.tar.gz">http://public-repo-1.hortonworks.com/HDP/suse11sp3/2.x/updates/2.3.6.0/HDP-2.3.6.0-suse11sp3-rpm.tar.gz</a>
Ubuntu 12	<a href="http://public-repo-1.hortonworks.com/HDP/ubuntu12/2.x/updates/2.3.6.0/HDP-2.3.6.0-ubuntu12-deb.tar.gz">http://public-repo-1.hortonworks.com/HDP/ubuntu12/2.x/updates/2.3.6.0/HDP-2.3.6.0-ubuntu12-deb.tar.gz</a>
Ubuntu 14	<a href="http://public-repo-1.hortonworks.com/HDP/ubuntu14/2.x/updates/2.3.6.0/HDP-2.3.6.0-ubuntu14-deb.tar.gz">http://public-repo-1.hortonworks.com/HDP/ubuntu14/2.x/updates/2.3.6.0/HDP-2.3.6.0-ubuntu14-deb.tar.gz</a>

For information on how to install the repositories, see the [local repository](#) instructions.

## 2. Run an update:

```
apt-get update
```

## 3. Install the HDP 2.3.6 bits:

Operating System	Commands
RHEL/CentOS/Oracle LINUX	<p>Install HDP 2.3.6 components on relevant nodes, according to the services that run on those hosts:</p> <pre>yum install "hadoop_2_3_6_3796*" "oozie_2_3_6_3796*" "pig_2_3_6_3796*" "sqoop_2_3_6_3796*" "zookeeper_2_3_6_3796*" "hbase_2_3_6_3796*" "hive_2_3_6_3796*" "tez_2_3_6_3796*" "storm_2_3_6_3796*" "falcon_2_3_6_3796*" "flume_2_3_6_3796*" "phoenix_2_3_6_3796*" "accumulo_2_3_6_3796*" "mahout_2_3_6_3796*"</pre>
SLES	<p>Install HDP 2.3.6 components on relevant nodes, according to the services that run on those hosts:</p> <pre>zypper install "hadoop_2_3_6_3796*" "oozie_2_3_6_3796*" "pig_2_3_6_3796*" "sqoop_2_3_6_3796*" "zookeeper_2_3_6_3796*" "hbase_2_3_6_3796*" "hive_2_3_6_3796*" "tez_2_3_6_3796*" "storm_2_3_6_3796*" "falcon_2_3_6_3796*" "flume_2_3_6_3796*" "phoenix_2_3_6_3796*" "accumulo_2_3_6_3796*" "mahout_2_3_6_3796*"</pre>
Ubuntu/Debian	<p>Install HDP 2.3.6 components on relevant nodes, according to the services that run on those hosts:</p> <pre>apt-get install "hadoop-2-3-6-3796*" "oozie-2-3-6-3796*" "pig-2-3-6-3796*" "sqoop-2-3-6-3796*" "zookeeper-2-3-6-3796*" "hbase-2-3-6-3796*" "hive-2-3-6-3796*" "tez-2-3-6-3796*" "storm-2-3-6-3796*" "falcon-2-3-6-3796*" "flume-2-3-6-3796*" "phoenix-2-3-6-3796*" "accumulo-2-3-6-3796*" "mahout-2-3-6-3796*"</pre>

## 4. Stop all HDP 2.3.x Services for your scenario:

- For non-Ambari managed clusters:

- a. Stop all HDP 2.3.x services using the [Stopping HDP Services](#) section of the *HDP Reference Guide*.
- For Ambari 1.7.0-managed clusters:
  - a. Open Ambari Web.
  - b. Browse to **Services**.
  - c. Use **Service Actions** to stop each service.
5. **For all services, switch the active version to HDP 2.3.6.**

On each host in the cluster, use `hdps-select` to switch all services to the HDP 2.3.6 version:

```
hdps-select set all 2.3.6.0-3796
```
6. **For Ambari 1.7.0-managed clusters, complete the Stack Upgrade:**

Update the repository base URLs to use the HDP 2.3.6 repositories for HDP and HDP-UTILS:

  - a. Open Ambari Web.
  - b. Browse to **Admin > Repositories**.
  - c. Edit the base URLs.
7. **Start all HDP 2.3.6 services, in the following order:**
  - a. **ZooKeeper**

```
su - zookeeper export ZOOCFGDIR=/usr/hdp/current/zookeeper-server/conf ; export ZOOCFG=zoo.cfg; source /usr/hdp/current/zookeeper-server/conf/zookeeper-env.sh ; /usr/hdp/current/zookeeper-server/bin/zkServer.sh start
```
  - b. **(HA NameNode upgrade only) ZooKeeper Failover Controller Daemons**

```
/usr/hdp/current/hadoop-hdfs-namenode/..../hadoop/sbin/hadoop-daemon.sh start zkfc
```
  - c. **(HA NameNode upgrade only) JournalNodes**

```
su - hdfs /usr/hdp/current/hadoop-hdfs-journalnode/..../hadoop/sbin/hadoop-daemon.sh start journalnode
```
  - d. **HDFS NameNodes**

Start the HDFS NameNodes. Because there is no metadata schema update for this upgrade, start the NameNodes in normal mode:

```
su - hdfs /usr/hdp/current/hadoop-hdfs-namenode/..../hadoop/sbin/hadoop-daemon.sh start namenode
```

### e. Remaining Services

Start the rest of the HDP services. On each host in the cluster, start the services that are relevant to that cluster. To identify the start commands for all services, see "Controlling HDP Services Manually" in the *HDP Reference Guide*.

You now have an upgraded cluster. Ensure that your workloads run correctly on this upgraded cluster.

## 1.4. Behavioral Changes

Behavioral changes denote a marked change in behavior from the previously released version to this version of software. In HDP 2.3.6, behavioral changes affect the following components.

**Table 1.3. Behavioral Changes**

Hortonworks Bug ID	Apache Component	Apache JIRA	Summary	Details	Source
			Debian 6 support deprecated.	HDP support for Debian 6 is deprecated with HDP 2.4.0.  Support for Debian 6 will be removed in a future HDP release.	2.4.0
BUG-43020	Hue		CherryPy is not the default webserver	<p><b>Scenario:</b> In previous releases of Hue, CherryPy was not the default web-server, but spawning webserver, which is generally not advisable as CherryPy is multithreaded and scales better than spawning webserver.</p> <p><b>Previous Behavior:</b> By not commenting out <code>use_cherry_server=false</code> in <code>hue.ini</code>, Hue used spawning webserver instead of Hue.</p> <p><b>New Behavior:</b> CherryPy is the default webserver for Hue.</p> <p><b>Workaround/Expected</b></p>	2.4.0

Hortonworks Bug ID	Apache Component	Apache JIRA	Summary	Details	Source
				<p><b>Customer Action:</b> No action required, unless spawning web server is the desired webserver for Hue. In this case, modify <code>use_cherrypy_server</code> accordingly.</p>	
BUG-50292	Hue		HUE DataTables Warning in partition columns tab.	<p><b>Scenario:</b> In previous releases of Hue, an error message appeared when <b>Partition Columns</b> tab was clicked in the Beeswax describe table view.</p> <p><b>Previous Behavior:</b> When clicking the <b>Partition Columns</b> tab in the Beeswax describe table view, an error message appeared.</p> <p><b>New Behavior:</b> No error message appears.</p> <p><b>Workaround:</b> None.</p>	2.4.0
BUG-50775	YARN		Enable collection of debug information by default	<p><b>Scenario:</b> Container failures during job execution.</p> <p><b>Previous Behavior:</b> New feature - no previous behaviour.</p> <p><b>New Behavior:</b> This feature(enabled by default) enables generation of additional information which is useful for debugging container failures. The extra information generated is:</p> <ul style="list-style-type: none"> <li>• A copy of the <code>launch_container.sh</code> script.</li> </ul>	2.4.0

Hortonworks Bug ID	Apache Component	Apache JIRA	Summary	Details	Source
				<ul style="list-style-type: none"> <li>A listing of the contents of the container work directory.</li> </ul> <p>This information is automatically collected if log-aggregation is enabled, or ignored and discarded if log-aggregation is disabled. The feature can be disabled by setting <code>yarn.nodemanager.log-container-debug-info.enabled</code> to false in <code>yarn-site.xml</code>.</p>	
BUG-50786	Tez		Increase the default value for MAX_COUNTERS	<p>The default value of the max counter limit has been increased from 2000 to:</p> <ul style="list-style-type: none"> <li><code>tez.counters.max: 10000</code></li> <li><code>tez.counters.maxgroups: 3000</code></li> </ul>	2.4.0
BUG-55056	YARN	<a href="#">YARN-4414</a>	Nodemanager connection errors are retried at multiple levels.	<p><b>Scenario:</b> During NodeManager restart, all YARN clients connect to NodeManager need to retry the connection to NodeManager.</p> <p>We formerly had a two layer retry for YARN Client connection with NodeManager: one was at the RPC layer, and the other was at the NodeManager proxy layer. This caused the log of retry activities to expand logarithmically. After this fix, we consolidated the RPC layer retry and NodeManager proxy retry; the retry log no longer has duplicate entries.</p>	2.4.2

## 1.5. Apache Patch Information

The following sections list patches in each HDP 2.3.6 component beyond what was fixed in the base version of the Apache component.

Please see the [Fixed Issues](#) section for customer-reported issues.

- [Hadoop \[11\]](#)
- [Accumulo \[31\]](#)
- [Atlas \[34\]](#)
- [Calcite \[36\]](#)
- [Falcon \[38\]](#)
- [Flume \[42\]](#)
- [HBase \[44\]](#)
- [Hive \[48\]](#)
- [Kafka \[61\]](#)
- [Knox \[65\]](#)
- [Mahout \[66\]](#)
- [Oozie \[67\]](#)
- [Phoenix \[67\]](#)
- [Pig \[70\]](#)
- [Ranger \[71\]](#)
- [Slider \[78\]](#)
- [Spark \[79\]](#)
- [Sqoop \[83\]](#)
- [Storm \[83\]](#)
- [Tez \[85\]](#)
- [ZooKeeper \[91\]](#)

### 1.5.1. Hadoop

In addition to any Apache patches ported over [from the 2.4.x core](#), HDP 2.3.6 provides the following Apache patches:

- [HADOOP-11252](#): RPC client does not time out by default.

- [HADOOP-11785](#): Reduce the number of listStatus operation in distcp buildListing.
- [HADOOP-11827](#): Speed-up distcp buildListing() using threadpool.
- [HADOOP-11876](#): Refactor code to make it more readable, minor maybePrintStats bug.
- [HADOOP-11901](#): BytesWritable fails to support 2G chunks due to integer overflow.
- [HADOOP-12423](#): Handle failure of registering shutdownhook by ShutdownHookManager in static block.
- [HADOOP-12559](#): KMS connection failures should trigger TGT renewal.
- [HADOOP-12672](#): RPC timeout should not override IPC ping interval.
- [HADOOP-12716](#): KerberosAuthenticator#doSpnegoSequence use incorrect class to determine isKeyTab in JDK8.
- [HADOOP-12847](#): Hadoop daemonlog should support HTTPS and SPNEGO for Kerberized cluster.
- [HADOOP-12950](#): ShutdownHookManager should have a timeout for each of the Registered shutdown hook.
- [HADOOP-12984](#): Add GenericTestUtils.getTestDir method and use it for temporary directory in tests.
- [HADOOP-12993](#): Change ShutdownHookManger complete shutdown log from INFO to DEBUG.
- [HADOOP-13008](#): Add XFS Filter for UIs to Hadoop Common.
- [HADOOP-13039](#): Add documentation for configuration property ipc.maximum.data.length.
- [HADOOP-13098](#): Dynamic LogLevel setting page should accept case-insensitive log level string.
- [HADOOP-13103](#): Group resolution from LDAP may fail on javax.naming.ServiceUnavailableException.
- [HADOOP-1540](#): breaks backward compatibility.
- [HADOOP-1540](#): Support file exclusion list in distcp.
- [HDFS-10178](#): Permanent write failures can happen if pipeline recoveries occur for the first packet.
- [HDFS-10216](#): Distcp -diff throws exception when handling relative path.
- [HDFS-10223](#): peerFromSocketAndKey performs SASL exchange before setting connection timeouts.
- [HDFS-10312](#): Large block reports may fail to decode at NameNode due to 64 MB protobuf maximum length restriction.



- [HDFS-10313](#): Distcp need to enforce the order of snapshot names passed to -diff.
- [HDFS-10335](#): Mover\$Processor#chooseTarget() always chooses the first matching target storage group.
- [HDFS-10341](#): Add a metric to expose the timeout number of pending replication blocks.
- [HDFS-10347](#): NameNode report bad block method doesn't log the bad block or datanode.
- [HDFS-10397](#): Distcp should ignore -delete option if -diff option is provided instead of exiting.
- [HDFS-10424](#): DatanodeLifelineProtocol not able to use under security cluster.
- [HDFS-10438](#): When NameNode HA is configured to use the lifeline RPC server, it should log the address of that server.
- [HDFS-6951](#): Correctly persist raw namespace xattrs to edit log and fsimage.
- [HDFS-7163](#): WebHdfsFileSystem should retry reads according to the configured retry policy.
- [HDFS-7916](#): 'reportBadBlocks' from datanodes to standby Node BPSERVICEACTOR goes for infinite loop.
- [HDFS-8512](#): WebHDFS : GETFILESTATUS should return LocatedBlock with storage type info.
- [HDFS-8828](#): Utilize Snapshot diff report to build diff copy list in distcp.
- [HDFS-8887](#): Expose storage type and storage ID in BlockLocation.
- [HDFS-9198](#): Coalesce IBR processing in the NN.
- [HDFS-9476](#): TestDFSUpgradeFromImage#testUpgradeFromRel1BBWImage occasionally fail.
- [HDFS-9612](#): DistCp worker threads are not terminated after jobs are done.
- [HDFS-9764](#): DistCp doesn't print value for several arguments including -numListstatusThreads.
- [HDFS-9902](#): Support different values of dfs.datanode.du.reserved per storage type.
- [HDFS-9917](#): IBR accumulate more objects when SNN was down for sometime.
- [HDFS-9958](#): BlockManager#createLocatedBlocks can throw NPE for corruptBlocks on failed storages.
- [MAPREDUCE-6514](#): Fixed MapReduce ApplicationMaster to properly updated resources ask after ramping down of all reducers avoiding job hangs.
- [MAPREDUCE-6689](#): MapReduce job can infinitely increase number of reducer resource requests.

- [YARN-3021](#): YARN's delegation-token handling disallows certain trust setups to operate properly over DistCp.
- [YARN-3602](#): TestResourceLocalizationService.testPublicResourceInitializesLocalDir fails intermittently due to IOException from cleanup.
- [YARN-3846](#): RM Web UI queue filter is not working for sub queue.
- [YARN-4596](#): SystemMetricPublisher should not swallow error messages from TimelineClient#putEntities.
- [YARN-4717](#): TestResourceLocalizationService.testPublicResourceInitializesLocalDir fails Intermittently due to IllegalArgumentException from cleanup.
- [YARN-4820](#): ResourceManager web redirects in HA mode drops query parameters.
- [YARN-4844](#): Scheduler keeps skipping scheduling because large amount of resources asked which causes pending resource overflows.
- [YARN-5048](#): DelegationTokenRenewer#skipTokenRenewal may throw NPE.
- [YARN-5076](#): YARN web interfaces lack XFS (Cross-Frame Script) protection.
- [YARN-5112](#): Excessive log warnings for directory permission issue on NM recovery.

HDP 2.3.6 ports the following Apache patches [from the 2.4.x core](#):

- [HADOOP-10365](#): BufferedOutputStream in FileUtil#unpackEntries() should be closed in finally block.
- [HADOOP-10406](#): TestIPC.testIpcWithReaderQueuing may fail.
- [HADOOP-11212](#): NetUtils.wrapException to handle SocketException explicitly.
- [HADOOP-12100](#): ImmutableFsPermission should not override applyUmask since that method doesn't modify the FsPermission.
- [HADOOP-12103](#): Small refactoring of DelegationTokenAuthenticationFilter to allow code sharing.
- [HADOOP-12107](#): Long running apps may have a huge number of StatisticsData instances under FileSystem.
- [HADOOP-12161](#): Add getStoragePolicy API to the FileSystem interface.
- [HADOOP-12191](#): Bzip2Factory is not thread safe.
- [HADOOP-12213](#): Interrupted exception can occur when Client#stop is called.
- [HADOOP-12348](#): MetricsSystemImpl creates MetricsSourceAdapter with wrong time unit parameter.
- [HADOOP-12374](#): Description of HDFS expunge command is confusing.
- [HADOOP-12426](#): Add Entry point for Kerberos health check.

- [HADOOP-12464](#): Interrupted client may try to fail-over and retry Interrupted client may try to fail-over and retry.
- [HADOOP-12482](#): Race condition in JMX cache update.
- [HADOOP-12551](#): Introduce FileNotFoundException for WASB FileSystem API.
- [HADOOP-12589](#): Fix intermittent test failure of TestCopyPreserveFlag.
- [HADOOP-12608](#): Fix exception message in WASB when connecting with anonymous credential.
- [HADOOP-12609](#): Fix intermittent failure of TestDecayRpcScheduler.
- [HADOOP-12678](#): Handle empty rename pending metadata file during atomic rename in redo path.
- [HADOOP-12699](#): TestKMS#testKMSProvider intermittently fails during 'test rollover draining'.
- [HADOOP-12706](#): TestLocalFsFCStatistics#testStatisticsThreadLocalDataCleanUp times out occasionally.
- [HADOOP-12752](#): Improve diagnostics/use of envvar/sysprop credential propagation.
- [HADOOP-12787](#): KMS SPNEGO sequence does not work with WEBHDFS.
- [HADOOP-12795](#): KMS does not log detailed stack trace for unexpected errors.
- [HADOOP-12825](#): Log slow name resolutions.
- [HADOOP-12829](#): StatisticsDataReferenceCleaner swallows interrupt exceptions.
- [HADOOP-12851](#): S3FileSystem Uptake of ProviderUtils.excludeIncompatibleCredentialProviders.
- [HADOOP-12903](#): IPC Server should allow suppressing exception logging by type, not log 'server too busy' messages.
- [HADOOP-12958](#): PhantomReference for filesystem statistics can trigger OOM.
- [HADOOP-13026](#): Should not wrap IOExceptions into a AuthenticationException in KerberosAuthenticator.
- [HDFS-10199](#): Unit tests TestCopyFiles, TestDistCh, TestLogalyzer under org.apache.hadoop.tools are failing.
- [HDFS-10270](#): TestJMXGet:testNameNode() fails.
- [HDFS-10281](#): TestPendingCorruptDnMessages fails intermittently.
- [HDFS-10283](#): o.a.h.hdfs.server.namenode.TestFSImageWithSnapshot#testSaveLoadImageWithAppending fails intermittently.

- [HDFS-6101](#): TestReplaceDatanodeOnFailure fails occasionally.
- [HDFS-8113](#): Add check for null BlockCollection pointers in BlockInfoContiguous structures.
- [HDFS-8337](#): Accessing HttpFS via WebHDFS doesn't work from a jar with Kerberos.
- [HDFS-8647](#): Abstract BlockManager's rack policy into BlockPlacementPolicy.
- [HDFS-8659](#): Block scanner INFO message is spamming logs.
- [HDFS-8676](#): Delayed rolling upgrade finalization can cause heartbeat expiration.
- [HDFS-8729](#): Fix TestFileTruncate#testTruncateWithDataNodesRestartImmediately which occasionally failed.
- [HDFS-8772](#): Fix TestStandbyIsHot#testDatanodeRestarts which occasionally fails.
- [HDFS-8806](#): Inconsistent metrics: number of missing blocks with replication factor 1 not properly cleared.
- [HDFS-8815](#): DFS getStoragePolicy implementation using single RPC call.
- [HDFS-8891](#): HDFS concat should keep srcs order.
- [HDFS-9072](#): Fix random failures in TestJMXGet.
- [HDFS-9130](#): Use GenericTestUtils#setLogLevel to the logging level.
- [HDFS-9221](#): HdfsServerConstants#ReplicaState#getState should avoid calling values() since it creates a temporary array.
- [HDFS-9239](#): DataNode Lifeline Protocol: an alternative protocol for reporting DataNode liveness.
- [HDFS-9289](#): Make DataStreamer#block thread safe and verify genStamp in commitBlock.
- [HDFS-9290](#): DFSClient#callAppend() is not backward compatible for slightly older NameNodes.
- [HDFS-9313](#): Possible NullPointerException in BlockManager if no excess replica can be chosen.
- [HDFS-9314](#): Improve BlockPlacementPolicyDefault's picking of excess replicas.
- [HDFS-9347](#): Invariant assumption in TestQuorumJournalManager.shutdown() is wrong.
- [HDFS-9358](#): TestNodeCount#testNodeCount timed out.
- [HDFS-9383](#): TestByteArrayManager#testByteArrayManager fails.
- [HDFS-9402](#): Switch DataNode.LOG to use slf4j.
- [HDFS-9406](#): FSImage may get corrupted after deleting snapshot.
- [HDFS-9431](#): DistributedFileSystem#concat fails if the target path is relative.

- [HDFS-9434](#): Recommission a datanode with 500k blocks may pause NN for 30 seconds for printing info log messages.
- [HDFS-9445](#): Datanode may deadlock while handling a bad volume.
- [HDFS-9478](#): Reason for failing ipc.FairCallQueue construction should be thrown.
- [HDFS-9534](#): Add CLI command to clear storage policy from a path.
- [HDFS-9557](#): Reduce object allocation in PB conversion.
- [HDFS-9572](#): Prevent DataNode log spam if a client connects on the data transfer port but sends no data.
- [HDFS-9574](#): Reduce client failures during datanode restart.
- [HDFS-9600](#): Do not check replication if the block is under construction.
- [HDFS-9625](#): Set replication for empty file failed when set storage policy.
- [HDFS-9655](#): NN should start JVM pause monitor before loading fsimage.
- [HDFS-9661](#): Deadlock in DN.FsDatasetImpl between moveBlockAcrossStorage and createRbw.
- [HDFS-9672](#): o.a.h.hdfs.TestLeaseRecovery2 fails intermittently.
- [HDFS-9710](#): DN can be configured to send block receipt IBRs in batches.
- [HDFS-9724](#): Degraded performance in WebHDFS listing as it does not reuse ObjectMapper.
- [HDFS-9726](#): Refactor IBR code to a new class.
- [HDFS-9740](#): Use a reasonable limit in DFSTestUtil.waitForMetric().
- [HDFS-9743](#): Fix TestLazyPersistFiles#testFallbackToDiskFull in branch-2.7.
- [HDFS-9752](#): Permanent write failures may happen to slow writers during datanode rolling upgrades.
- [HDFS-9768](#): Reuse ObjectMapper instance in HDFS to improve the performance.
- [HDFS-9790](#): HDFS Balancer should exit with a proper message if upgrade is not finalized.
- [HDFS-9839](#): Reduce verbosity of processReport logging.
- [HDFS-9851](#): NameNode throws NPE when setPermission is called on a path that does not exist.
- [HDFS-9854](#): Log cipher suite negotiation more verbosely.
- [HDFS-9906](#): Remove spammy log spew when a datanode is restarted.
- [HDFS-9941](#): Do not log StandbyException on NN, other minor logging fixes.

- [MAPREDUCE-6436](#): JobHistory cache issue.
- [MAPREDUCE-6460](#): TestRMContainerAllocator.testAttemptNotFoundCausesRMCommunicatorException fails.
- [MAPREDUCE-6492](#): AsyncDispatcher exit with NPE on TaskAttemptImpl#sendJHStartEventForAssignedFailTask.
- [MAPREDUCE-6566](#): Add retry support to MapReduce CLI tool.
- [MAPREDUCE-6577](#): MR AM unable to load native library without MR\_AM\_ADMIN\_USER\_ENV set.
- [MAPREDUCE-6618](#): YarnClientProtocolProvider leaking the YarnClient thread.
- [MAPREDUCE-6621](#): Memory Link in JobClient#submitJobInternal().
- [MAPREDUCE-6635](#): Unsafe long to int conversion in UncompressedSplitLineReader and IndexOutOfBoundsException.
- [MAPREDUCE-6670](#): TestJobListCache#testEviction sometimes fails on Windows with timeout.
- [MAPREDUCE-6680](#): JHS UserLogDir scan algorithm sometime could skip directory with update in CloudFS.
- [YARN-2046](#): Out of band heartbeats are sent only on container kill and possibly too early.
- [YARN-2871](#): TestRMRestart#testRMRestartGetApplicationList sometimes fails in trunk.
- [YARN-3102](#): Decommissioned Nodes not listed in Web UI.
- [YARN-3480](#): Remove attempts that are beyond max-attempt limit from state store.
- [YARN-3695](#): ServerProxy (NMProxy, etc.) shouldn't retry forever for non network exception.
- [YARN-3769](#): Consider user limit when calculating total pending resource for preemption policy in Capacity Scheduler.
- [YARN-4155](#): TestLogAggregationService.testLogAggregationServiceWithInterval failing.
- [YARN-4309](#): Add container launch related debug information to container logs when a container fails.
- [YARN-4365](#): FileSystemNodeLabelStore should check for root dir existence on startup.
- [YARN-4414](#): Nodemanager connection errors are retried at multiple levels.
- [YARN-4422](#): Generic AHS sometimes doesn't show started, node, or logs on App page.
- [YARN-4428](#): Redirect RM page to AHS page when AHS turned on and RM page is not available.
- [YARN-4439](#): Clarify NMContainerStatus#toString method.

- [YARN-4497](#): RM might fail to restart when recovering apps whose attempts are missing.
- [YARN-4546](#): ResourceManager crash due to scheduling opportunity overflow.
- [YARN-4565](#): Sometimes when sizeBasedWeight FairOrderingPolicy is enabled, under stress appears that cluster is virtually in deadlock.
- [YARN-4584](#): RM startup failure when AM attempts greater than max-attempts.
- [YARN-4598](#): Invalid event: RESOURCE\_FAILED at CONTAINER\_CLEANEDUP\_AFTER\_KILL.
- [YARN-4610](#): Reservations continue looking for one app causes other apps to starve.
- [YARN-4623](#): TestSystemMetricsPublisher#testPublishAppAttemptMetricsForUnmanagedAM fails with NPE on branch-2.7.
- [YARN-4625](#): ApplicationSubmissionContext and ApplicationSubmissionContextInfo more consistent.
- [YARN-4633](#): Fix random test failure in TestRMRestart#testRMRestartAfterPreemption.
- [YARN-4680](#): Fix TimerTasks leak in Application Timeline Server (ATS) v1.5 Writer.
- [YARN-4696](#): TimelineClient to add flush operation for deterministic writes (including testing) and Changes to EntityGroupFSTimelineStore for testability.
- [YARN-4709](#): NMWebServices produces incorrect JSON for containers.
- [YARN-4723](#): NodesListManager\$UnknownNodeId ClassCastException.
- [YARN-4737](#): Add CSRF filter support in YARN.
- [YARN-4769](#): Add support for CSRF header in the dump capacity scheduler logs and kill app buttons in RM web UI.
- [YARN-4785](#): Inconsistent value type of the "type" field for LeafQueueInfo in response of RM REST API.
- [YARN-4814](#): ATS 1.5 timelineclient impl call flush after every event write.
- [YARN-4815](#): ATS 1.5 timelineclient impl try to create attempt directory for every event call.
- [YARN-4817](#): TimelineClient ATsv1.5 logging is very noisy.
- [YARN-4916](#): TestNMProxy.testNMProxyRPCRetry fails.
- [YARN-4928](#): Some yarn.server.timeline.\* tests fail on Windows attempting to use a test root path containing a colon.
- [YARN-4954](#): TestYarnClient.testReservationAPIs fails on machines with less than 4 GB available memory.
- [YARN-4955](#): Add retry for SocketTimeoutException in TimelineClient.

- [YARN-4965](#): Distributed shell AM failed due to ClientHandlerException thrown by jersey.
- [YARN-4968](#): Fix two scheduler related UTs in YARN.

HDP 2.3.4.7 provided no additional Apache patches.

HDP 2.3.4 provided the following Apache patches:

- [HADOOP-11098](#): [JDK8] Max Non Heap Memory default changed between JDK7 and 8.
- [HADOOP-11628](#): SPNEGO auth does not work with CNAMEs in JDK8.
- [HADOOP-11685](#): StorageException complaining "no lease ID" during HBase distributed log splitting.
- [HADOOP-11918](#): Listing an empty s3a root directory throws FileNotFound.
- [HADOOP-11932](#): MetricsSinkAdapter may hang when being stopped.
- [HADOOP-12049](#) Control http authentication cookie persistence via configuration.
- [HADOOP-12089](#): StorageException complaining " no lease ID" when updating FolderLastModifiedTime in WASB.
- [HADOOP-12186](#) ActiveStandbyElector shouldn't call monitorLockNodeAsync multiple times.
- [HADOOP-12239](#): StorageException complaining " no lease ID" when updating FolderLastModifiedTime in WASB.
- [HADOOP-12324](#): Better exception reporting in SaslPlainServer.
- [HADOOP-12334](#): Change Mode Of Copy Operation of HBase WAL Archiving to bypass Azure Storage Throttling after retries.
- [HADOOP-12350](#): WASB Logging; Improve WASB Logging around deletes, reads, and writes.
- [HADOOP-12350](#) WASB Logging: Improve WASB Logging around deletes, reads and writes.
- [HADOOP-12407](#): Test failing; hadoop.ipc.TestSaslRPC.
- [HADOOP-12413](#): AccessControlList should avoid calling getGroupNames in isUserInList with empty groups.
- [HADOOP-12437](#) Allow SecurityUtil to lookup alternate hostnames.
- [HADOOP-12438](#): TestLocalFileSystem tests can fail on Windows after HDFS-8767 fix for handling pipe.
- [HADOOP-12440](#): TestRPC#testRPCServerShutdown did not produce the desired thread states before shutting down.
- [HADOOP-12441](#): Fixed kill-command behavior to work correctly across OSes by using bash shell built-in.



- [HADOOP-12463](#) Fix TestShell.testGetSignalKillCommand failure on windows.
- [HADOOP-12484](#): Single File Rename Throws Incorrectly In Potential Race Condition Scenarios.
- [HADOOP-12508](#): delete fails with exception when lease is held on blob.
- [HADOOP-12533](#): Introduce FileNotFoundException in WASB for read and seek API.
- [HADOOP-12540](#): TestAzureFileSystemInstrumentation#testClientErrorMetrics fails intermittently due to assumption that a lease error will be thrown.
- [HADOOP-12542](#): TestDNS fails on Windows after HADOOP-12437.
- [HADOOP-12577](#) Bump up commons-collections version to 3.2.2 to address a security flaw.
- [HADOOP-12617](#) SPNEGO authentication request to non-default realm gets default realm name inserted in target server principal.
- [HBASE-268](#) Rack locality improvement.
- [HDFS-4015](#): Safemode should count and report orphaned blocks.
- [HDFS-4015](#) Safemode should count and report orphaned blocks.
- [HDFS-4366](#) Block Replication Policy Implementation May Skip Higher-Priority Blocks for Lower-Priority Blocks.
- [HDFS-4937](#) ReplicationMonitor can infinite-loop in BlockPlacementPolicyDefault#chooseRandom.
- [HDFS-6481](#) DatanodeManager#getDatanodeStorageInfos() should check the length of storageIDs.
- [HDFS-6581](#) Support for writing to single replica in RAM.
- [HDFS-7390](#) Provide JMX metrics per storage type.
- [HDFS-7483](#) Display information per tier on the NameNode UI.
- [HDFS-7725](#) Incorrect "nodes in service" metrics caused all writes to fail.
- [HDFS-7858](#) Improve HA NameNode Failover detection on the client.
- [HDFS-7928](#): Scanning blocks from disk during rolling upgrade startup takes a lot of time if disks are busy.
- [HDFS-7928](#) Scanning blocks from disk during rolling upgrade startup takes a lot of time if disks are busy.
- [HDFS-8099](#) Change "DFSInputStream has been closed already" message to debug log level.
- [HDFS-8209](#) Support different number of datanode directories in MiniDFSCluster.
- [HDFS-8554](#): TestDatanodeLayoutUpgrade fails on Windows.

- [HDFS-8656](#): Preserve compatibility of ClientProtocol#rollingUpgrade after finalization.
- [HDFS-8696](#): Make the lower and higher watermark in the DN Netty server configurable.
- [HDFS-8778](#): TestBlockReportRateLimiting#testLeaseExpiration can deadlock.
- [HDFS-8785](#) TestDistributedFileSystem is failing in trunk.
- [HDFS-8809](#): HDFS fsck reports under construction blocks as CORRUPT.
- [HDFS-8829](#) Make SO\_RCVBUF and SO\_SNDBUF size configurable for DataTransferProtocol sockets and allow configuring auto-tuning.
- [HDFS-8846](#): Add a unit test for INotify functionality across a layout version upgrade.
- [HDFS-8855](#): Webhdfs client leaks active NameNode connections.
- [HDFS-8930](#): Block report lease may leak if the 2nd full block report comes when NN is still in safemode.
- [HDFS-8950](#) NameNode refresh doesn't remove DataNodes that are no longer in the allowed list.
- [HDFS-8965](#): Harden edit log reading code against out of memory errors.
- [HDFS-8965](#) Harden edit log reading code against out of memory errors.
- [HDFS-8969](#): Clean up findbugs warnings for HDFS-8823 and HDFS-8932.
- [HDFS-9008](#): Balancer#Parameters class could use a builder pattern.
- [HDFS-9019](#): Adding informative message to sticky bit permission denied exception.
- [HDFS-9063](#): Correctly handle snapshot path for getContentSummary.
- [HDFS-9082](#): Change the log level in WebHdfsFileSystem.initialize() from INFO to DEBUG.
- [HDFS-9083](#): Replication violates block placement policy.
- [HDFS-9107](#): Prevent NNs unrecoverable death spiral after full GC.
- [HDFS-9112](#): Improve error message for Haadmin when multiple name service IDs are configured.
- [HDFS-9112](#) Improve error message for Haadmin when multiple name service IDs are configured,
- [HDFS-9128](#): TestWebHdfsFileContextMainOperations and TestSWebHdfsFileContextMainOperations fail due to invalid HDFS path on Windows.
- [HDFS-9142](#): Separating Configuration object for NameNodes in MiniDFSCluster.
- [HDFS-9175](#): Change scope of 'AccessTokenProvider.getAccessToken()' and 'CredentialBasedAccessTokenProvider.getCredential()' abstract methods to public.
- [HDFS-9178](#) Slow datanode I/O can cause a wrong node to be marked bad.

- [HDFS-9184](#) Logging HDFS operation's caller context into audit logs.
- [HDFS-9205](#): Do not schedule corrupt blocks for replication.
- [HDFS-9220](#): Reading small file greater than 512 bytes that is open for append fails due to incorrect checksum.
- [HDFS-9273](#): ACLs on root directory may be lost after NN restart.
- [HDFS-9294](#): DFSClient deadlock when close file and failed to renew lease.
- [HDFS-9305](#) Delayed heartbeat processing causes Storm of subsequent heartbeats.
- [HDFS-9311](#): Support optional offload of NameNode HA service health checks to a separate RPC server.
- [HDFS-9343](#) Empty caller context considered invalid.
- [HDFS-9354](#): Fix TestBalancer#testBalancerWithZeroThreadsForMove on Windows.
- [HDFS-9362](#): TestAuditLogger#testAuditLoggerWithCallContext assumes Unix line endings, fails on Windows.
- [HDFS-9364](#) Unnecessary DNS resolution attempts when creating NameNodeProxies.
- [HDFS-9384](#): TestWebHdfsContentLength intermittently hangs and fails due to TCP conversation mismatch between client and server.
- [HDFS-9397](#) Fix typo for readChecksum() LOG.warn in BlockSender.java.
- [HDFS-9413](#): getContentSummary() on standby should throw StandbyException.
- [HDFS-9426](#): Rollingupgrade finalization is not backward compatible.
- [HDFS-9434](#): Recommission a datanode with 500k blocks may pause NN for 30 seconds for printing info log messages.
- [MAPREDUCE-5485](#) Allow repeating job commit by extending OutputCommitter API.
- [MAPREDUCE-6273](#) HistoryFileManager should check whether summaryFile exists to avoid FileNotFoundException causing HistoryFileInfo into MOVE\_FAILED state.
- [MAPREDUCE-6302](#) Backport preempt reducers after a configurable timeout irrespective of headroom.
- [MAPREDUCE-6549](#) Multibyte delimiters with LineRecordReader cause duplicate records.
- [YARN-2194](#) Fix bug causing CGroups functionality to fail on RHEL7.
- [YARN-2571](#) RM to support YARN registry.
- [YARN-3467](#) Expose allocatedMB, allocatedVCores, and runningContainers metrics on running Applications in RM Web UI.
- [YARN-3600](#) AM container link is broken (on a killed application, at least).

- [YARN-3727](#) For better error recovery, check if the directory exists before using it for localization.
- [YARN-3751](#) Fixed AppInfo to check if used resources are null.
- [YARN-3766](#) Fixed the apps table column error of generic history web UI.
- [YARN-3849](#) Too much of preemption activity causing continuous killing of containers across queues.
- [YARN-4140](#) RM container allocation delayed in case of app submitted to Nodelabel partition.
- [YARN-4233](#) YARN Timeline Service plugin: ATS v1.5.
- [YARN-4285](#) Display resource usage as percentage of queue and cluster in the RM UI.
- [YARN-4287](#) Rack locality improvement.
- [YARN-4288](#) Fixed RMProxy to retry on IOException from local host.
- [YARN-4313](#) Race condition in MiniMRYarnCluster when getting history server address.
- [YARN-4345](#) YARN radmin -updateNodeResource doesn't work.
- [YARN-4347](#) Resource manager fails with Null pointer exception.
- [YARN-4349](#) YARN\_APPLICATION call to ATS does not have YARN\_APPLICATION\_CALLER\_CONTEXT.
- [YARN-4384](#) updateNodeResource CLI should not accept negative values for resource.
- [YARN-4405](#) Support node label store in non-appendable file system.

HDP 2.3.2 provided the following Apache patches:

#### NEW FEATURES

- [HDFS-8155](#) Support OAuth2 in WebHDFS.

#### IMPROVEMENTS

- [HADOOP-10597](#) RPC Server signals backoff to clients when all request queues are full.
- [HADOOP-11960](#) Enable Azure-Storage Client Side logging.
- [HADOOP-12325](#) RPC Metrics: Add the ability track and log slow RPCs.
- [HADOOP-12358](#) Add -safely flag to rm to prompt when deleting many files.
- [HDFS-4185](#) Add a metric for number of active leases.
- [HDFS-4396](#) Add START\_MSG/SHUTDOWN\_MSG for ZKFC.
- [HDFS-6860](#) BlockStateChange logs are too noisy.

- [HDFS-7923](#) The DataNodes should rate-limit their full block reports by asking the NN on heartbeat messages.
- [HDFS-8046](#) Allow better control of getContentSummary.
- [HDFS-8180](#) AbstractFileSystem Implementation for WebHdfs.
- [HDFS-8278](#) When computing max-size-to-move in Balancer, count only the storage with remaining  $\geq$  default block size.
- [HDFS-8432](#) Introduce a minimum compatible layout version to allow downgrade in more rolling upgrade use cases.
- [HDFS-8435](#) Support CreateFlag in WebHDFS.
- [HDFS-8549](#) Abort the balancer if an upgrade is in progress.
- [HDFS-8797](#) WebHdfsFileSystem creates too many connections for pread.
- [HDFS-8818](#) Changes the global moveExecutor to per datanode executors and changes MAX\_SIZE\_TO\_MOVE to be configurable.
- [HDFS-8824](#) Do not use small blocks for balancing the cluster.
- [HDFS-8826](#) In Balancer, add an option to specify the source node list so that balancer only selects blocks to move from those nodes.
- [HDFS-8883](#) NameNode Metrics: Add FSNameSystem lock Queue Length.
- [HDFS-8911](#) NameNode Metric Add Editlog counters as a JMX metric.
- [HDFS-8983](#) NameNode support for protected directories.
- [HDFS-8983](#) NameNode support for protected directories.
- [YARN-2513](#) Host framework UIs in YARN for use with the ATS.
- [YARN-3197](#) Confusing log generated by CapacityScheduler.
- [YARN-3357](#) Move TestFifoScheduler to FIFO package.
- [YARN-3360](#) Add JMX metrics to TimelineDataManager.
- [YARN-3579](#) CommonNodeLabelsManager should support NodeLabel instead of string label name when getting node-to-label/label-to-label mappings.
- [YARN-3978](#) Configurably turn off the saving of container info in Generic AHS.
- [YARN-4082](#) Container shouldn't be killed when node's label updated.
- [YARN-4101](#) RM should print alert messages if ZooKeeper and ResourceManager gets connection issue.
- [YARN-4149](#) YARN logs -am should provide an option to fetch all the log files.

#### BUG FIXES

- [HADOOP-11802](#) DomainSocketWatcher thread terminates sometimes after there is an I/O error during requestShortCircuitShm.
- [HADOOP-12052](#) IPC client downgrades all exception types to IOE, breaks callers trying to use them.
- [HADOOP-12073](#) Azure FileSystem PageBlobInputStream does not return -1 onEOF.
- [HADOOP-12095](#) org.apache.hadoop.fs.shell.TestCount fails.
- [HADOOP-12304](#) Applications using FileContext fail with the default filesystem configured to be wasb/s3/etc.
- [HADOOP-8151](#) Error handling in Snappy decompressor throws invalidexceptions.
- [HDFS-6945](#) BlockManager should remove a block from excessReplicateMap anddecrement ExcessBlocks metric when the block is removed.
- [HDFS-7608](#) HDFS dfsclient newConnectedPeer has nowrite timeout.
- [HDFS-7609](#) Avoid retry cache collision when Standby NameNode loading edits.
- [HDFS-8309](#) Skip unit test using DataNodeTestUtils#injectDataDirFailure() on Windows.
- [HDFS-8310](#) Fix TestCLI.testAll "help for find" on Windows.
- [HDFS-8311](#) DataStreamer.transfer() should timeout the socket InputStream.
- [HDFS-8384](#) Allow NN to startup if there are files having a lease but are notunder construction.
- [HDFS-8431](#) HDFS crypto class not found in Windows.
- [HDFS-8539](#) HDFS doesn't have class 'debug' in windows.
- [HDFS-8542](#) WebHDFS getHomeDirectory behavior does not match specification.
- [HDFS-8593](#) Calculation of effective layout version mishandles comparison to current layout version in storage.
- [HDFS-8767](#) RawLocalFileSystem.listStatus() returns null for UNIX pipefile.
- [HDFS-8850](#) VolumeScanner thread exits with exception if there is no blockpool to be scanned but there are suspicious blocks.
- [HDFS-8863](#) The remaining space check in BlockPlacementPolicyDefault is flawed.
- [HDFS-8879](#) Quota by storage type usage incorrectly initialized upon namenoderestart.
- [HDFS-8885](#) ByteRangeInputStream used in webhdfs does not overrideavailable().
- [HDFS-8932](#) NPE thrown in NameNode when try to get TotalSyncCount metricbefore editLogStream initialization.
- [HDFS-8939](#) Test(S)WebHdfsFileContextMainOperations failing on branch-2.

- [HDFS-8969](#) Clean up findbugs warnings for HDFS-8823 and HDFS-8932.
- [HDFS-8995](#) Flaw in registration bookkeeping can make DN die on reconnect.
- [HDFS-9009](#) Send metrics logs to NullAppender by default.
- [YARN-3413](#) Changed Nodelabel attributes (like exclusivity) to be settable only via addToClusterNodeLabelsbut not changeable at runtime.
- [YARN-3885](#) ProportionalCapacityPreemptionPolicy doesn't preempt if queue is more than 2 level.
- [YARN-3894](#) RM startup should fail for wrong CS xml NodeLabel capacity configuration.
- [YARN-3896](#) RMNode transitioned from RUNNING to REBOOTED because its response idhas not been reset synchronously.
- [YARN-3932](#) SchedulerApplicationAttempt#getResourceUsageReport and UserInfo should based on total-used-resources.
- [YARN-3971](#) Skip RMNodeLabelsManager#checkRemoveFromClusterNodeLabelsOfQueue on nodelabel recovery.
- [YARN-4087](#) Followup fixes after YARN-2019 regarding RM behavior when state-store error occurs.
- [YARN-4092](#) Fixed UI redirection to print useful messages when both RMs are in standby mode.

#### OPTIMIZATION

- [HADOOP-11772](#) RPC Invoker relies on static ClientCache which has synchronized(this) blocks.
- [HADOOP-12317](#) Applications fail on NM restart on some Linux distro because NM container recovery declares AM container as LOST.
- [HADOOP-7713](#) dfs -count -q should label output column.
- [HDFS-8856](#) Make LeaseManager#countPath O(1).
- [HDFS-8867](#) Enable optimized block reports.

HDP 2.3.0 provided the following Apache patches:

#### NEW FEATURES

- [HDFS-8008](#) Support client-side back off when the datanodes are congested.
- [HDFS-8009](#) Signal congestion on the DataNode.
- [YARN-1376](#) NM need to notify the log aggregation status to RM through heartbeat.
- [YARN-1402](#) Update related Web UI and CLI with exposing client API to check log aggregation status.

- [YARN-2498](#) Respect labels in preemption policy of capacity scheduler for inter-queue preemption.
- [YARN-2571](#) RM to support YARN registry
- [YARN-2619](#) Added NodeManager support for disk IO isolation through cgroups.
- [YARN-3225](#) New parameter of CLI for decommissioning node gracefully in RMAdmin CLI.
- [YARN-3318](#) Create Initial OrderingPolicy Framework and FifoOrderingPolicy.
- [YARN-3319](#) Implement a FairOrderingPolicy.
- [YARN-3326](#) Support RESTful API for getLabelsToNodes.
- [YARN-3345](#) Add non-exclusive node label API.
- [YARN-3347](#) Improve YARN log command to get AMContainer logs as well as running containers logs.
- [YARN-3348](#) Add a 'YARN top' tool to help understand cluster usage.
- [YARN-3354](#) Add node label expression in ContainerTokenIdentifier to support RM recovery.
- [YARN-3361](#) CapacityScheduler side changes to support non-exclusive node labels.
- [YARN-3365](#) Enhanced NodeManager to support using the 'tc' tool via container-executor for outbound network traffic control.
- [YARN-3366](#) Enhanced NodeManager to support classifying/shaping outgoing network bandwidth traffic originating from YARN containers
- [YARN-3410](#) YARN admin should be able to remove individual application records from RMStateStore.
- [YARN-3443](#) Create a 'ResourceHandler' subsystem to ease addition of support for new resource types on the NM.
- [YARN-3448](#) Added a rolling time-to-live LevelDB timeline store implementation.
- [YARN-3463](#) Integrate OrderingPolicy Framework with CapacityScheduler.
- [YARN-3505](#) Node's Log Aggregation Report with SUCCEED should not be cached in RMApps.
- [YARN-3541](#) Add version info on timeline service / generic history web UI and REST API.

#### IMPROVEMENTS

- [HADOOP-10597](#) RPC Server signals backoff to clients when all request queues are full.
- [YARN-1880](#) Cleanup TestApplicationClientProtocolOnHA
- [YARN-2495](#) Allow admin specify labels from each NM (Distributed configuration for node label).



- [YARN-2696](#) Queue sorting in CapacityScheduler should consider node label.
- [YARN-2868](#) FairScheduler: Metric for latency to allocate first container for an application.
- [YARN-2901](#) Add errors and warning metrics page to RM, NM web UI.
- [YARN-3243](#) CapacityScheduler should pass headroom from parent to children to make sure ParentQueue obey its capacity limits.
- [YARN-3248](#) Display count of nodes blacklisted by apps in the web UI.
- [YARN-3293](#) Track and display capacity scheduler health metrics in web UI.
- [YARN-3294](#) Allow dumping of Capacity Scheduler debug logs via web UI for a fixed time period.
- [YARN-3356](#) Capacity Scheduler FiCaSchedulerApp should use ResourceUsage to track used-resources-by-label.
- [YARN-3362](#) Add node label usage in RM CapacityScheduler web UI.
- [YARN-3394](#) Enrich WebApplication proxy documentation.
- [YARN-3397](#) YARN radmin should skip -failover.
- [YARN-3404](#) Display queue name on application page.
- [YARN-3406](#) Display count of running containers in the RM's Web UI.
- [YARN-3451](#) Display attempt start time and elapsed time on the web UI.
- [YARN-3494](#) Expose AM resource limit and usage in CS QueueMetrics.
- [YARN-3503](#) Expose disk utilization percentage and bad local and log dir counts in NM metrics.
- [YARN-3511](#) Add errors and warnings page to ATS.
- [YARN-3565](#) NodeHeartbeatRequest/RegisterNodeManagerRequest should use NodeLabel object instead of String.
- [YARN-3581](#) Deprecate -directlyAccessNodeLabelStore in RMAAdminCLI.
- [YARN-3583](#) Support of NodeLabel object instead of plain String in YarnClient side.
- [YARN-3593](#) Add label-type and Improve "DEFAULT\_PARTITION" in Node Labels Page.
- [YARN-3700](#) Made generic history service load a number of latest applications according to the parameter or the configuration.

#### BUG FIXES

- [HADOOP-11859](#) PseudoAuthenticationHandler fails with httpcomponents v4.4.
- [HADOOP-7713](#) dfs -count -q should label output column

- [HDFS-27](#) HDFS CLI with `--config` set to default config complains log file not found error.
- [HDFS-6666](#) Abort NameNode and DataNode startup if security is enabled but block access token is not enabled.
- [HDFS-7645](#) Fix CHANGES.txt
- [HDFS-7645](#) Rolling upgrade is restoring blocks from trash multiple times
- [HDFS-7701](#) Support reporting per storage type quota and usage with Hadoop/HDFS shell.
- [HDFS-7890](#) Improve information on Top users for metrics in RollingWindowsManager and lower log level.
- [HDFS-7933](#) fsck should also report decommissioning replicas.
- [HDFS-7990](#) IBR delete ack should not be delayed.
- [HDFS-8008](#) Support client-side back off when the datanodes are congested.
- [HDFS-8009](#) Signal congestion on the DataNode.
- [HDFS-8055](#) NullPointerException when topology script is missing.
- [HDFS-8144](#) Split TestLazyPersistFiles into multiple tests.
- [HDFS-8152](#) Refactoring of lazy persist storage cases.
- [HDFS-8205](#) CommandFormat#parse() should not parse option as value of option.
- [HDFS-8211](#) DataNode UUID is always null in the JMX counter.
- [HDFS-8219](#) setStoragePolicy with folder behavior is different after cluster restart.
- [HDFS-8229](#) LAZY\_PERSIST file gets deleted after NameNode restart.
- [HDFS-8232](#) Missing datanode counters when using Metrics2 sink interface.
- [HDFS-8276](#) LazyPersistFileScrubber should be disabled if scrubber interval configured zero.
- [YARN-2666](#) TestFairScheduler.testContinuousScheduling fails Intermittently.
- [YARN-2740](#) Fix NodeLabelsManager to properly handle node label modifications when distributed node label configuration enabled.
- [YARN-2821](#) Fixed a problem that DistributedShell AM may hang if restarted.
- [YARN-3110](#) Few issues in ApplicationHistory web UI.
- [YARN-3136](#) Fixed a synchronization problem of AbstractYarnScheduler#getTransferredContainers.
- [YARN-3266](#) RMContext#inactiveNodes should have NodeId as map key.

- [YARN-3269](#) Yarn.nodemanager.remote-app-log-dir could not be configured to fully qualified path.
- [YARN-3305](#) Normalize AM resource request on app submission.
- [YARN-3343](#) Increased TestCapacitySchedulerNodeLabelUpdate#testNodeUpdate timeout.
- [YARN-3383](#) AdminService should use "warn" instead of "info" to log exception when operation fails.
- [YARN-3387](#) Previous AM's container completed status couldn't pass to current AM if AM and RM restarted during the same time.
- [YARN-3425](#) NPE from RMNodeLabelsManager.serviceStop when NodeLabelsManager.servicelnit failed.
- [YARN-3435](#) AM container to be allocated Appattempt AM container shown as null.
- [YARN-3459](#) Fix failure of TestLog4jWarningErrorMetricsAppender.
- [YARN-3517](#) RM web UI for dumping scheduler logs should be for admins only
- [YARN-3530](#) ATS throws exception on trying to filter results without otherinfo.
- [YARN-3552](#) RM Web UI shows -1 running containers for completed apps
- [YARN-3580](#) [JDK8] TestClientRMService.testGetLabelsToNodes fails.
- [YARN-3632](#) Ordering policy should be allowed to reorder an application when demand changes.
- [YARN-3654](#) ContainerLogsPage web UI should not have meta-refresh.
- [YARN-3707](#) RM Web UI queue filter doesn't work.
- [YARN-3740](#) Fixed the typo in the configuration name: APPLICATION\_HISTORY\_PREFIX\_MAX\_APPS.

## 1.5.2. Accumulo

HDP 2.3.6 provides Accumulo 1.7.0 with no additional Apache patches beyond any ported over [from the 2.4.x core](#).

HDP 2.3.6 provides Accumulo 1.7.0 and ports the following Apache patches [from the 2.4.x core](#):

- [ACCUMULO-4080](#) TabletServers should be less aggressively "monitoring RO filesystems".
- [ACCUMULO-4185](#) Use SYNC durability and expect the update to be durable in ProxyDurabilityIT.

HDP 2.3.4.7 provided Accumulo 1.7.0 with no additional Apache patches.

HDP 2.3.4 provided Accumulo 1.7.0 and the following Apache patches:

- [ACCUMULO-3963](#) Incremental backoff on inability to write to HDFS
- [ACCUMULO-4053](#) ReplicationOperations.drain() is returning too quickly
- [ACCUMULO-4060](#) Transient ZooKeeper connection issues kills FATE Runner threads
- [ACCUMULO-4065](#) Oneway Thrift calls leave message on client's InputStream
- [ACCUMULO-4069](#) Services failing to renew Kerberos ticket
- [ACCUMULO-4071](#) BulkImportSequentialRowsIT fails on standalone cluster with HDFS permission errors

HDP 2.3.2 provided Accumulo 1.7.0 and the following Apache patches:

- [ACCUMULO-3890](#) Use of CredentialProvider results in a lot of NN ops
- [ACCUMULO-3957](#) Consider moving off getContentSummary in the monitor
- [ACCUMULO-3967](#) bulk import loses records when loading pre-split table
- [ACCUMULO-3973](#) ShellServerIT.addauths fails to correctly deal with cached authorizations
- [ACCUMULO-4001](#) BulkImportSequentialRowsIT fails when using HDFS

HDP 2.3.0 provided Accumulo 1.7.0 and the following Apache patches:

- [ACCUMULO-3809](#) Table problem report has bogus table name for user table
- [ACCUMULO-3810](#) RandomWalk test, MultiTable fails throwing java.lang.NullPointerException w/ Kerberos on
- [ACCUMULO-3812](#) T\*ProxyIT classes need cleanup
- [ACCUMULO-3814](#) StandaloneAccumuloClusterControl doesn't set provided ACCUMULO\_CONF\_DIR on SetGoalState
- [ACCUMULO-3815](#) StandaloneClusterControl shouldn't use canonical paths
- [ACCUMULO-3816](#) rpc.sasl.qop not mentioned in Kerberos server-configuration user manual section
- [ACCUMULO-3821](#) CleanTmpIT fails on dfs.permission enabled HDFS instance
- [ACCUMULO-3822](#) ImportExportIT fails to write to export directory in HDFS due to permissions
- [ACCUMULO-3823](#) Support separate client and server ACCUMULO\_CONF\_DIRS for StandaloneCluster ITs
- [ACCUMULO-3826](#) User manual accidentally references commerical product
- [ACCUMULO-3827](#) Default store types for monitor SSL are broken
- [ACCUMULO-3828](#) SimpleProxyBase ITs failing due to constraint propagation

- [ACCUMULO-3834](#) ConstraintIT occasionally failing
- [ACCUMULO-3838](#) ReplicationIT.replicationEntriesPrecludeWalDeletion failed because it missed an expected WAL
- [ACCUMULO-3839](#) Nonsense error when configuring instance.volumes.replacements
- [ACCUMULO-3845](#) DurabilityIT failed
- [ACCUMULO-3846](#) Allow override of C++ compiler through Maven build
- [ACCUMULO-3847](#) StandaloneClusterControl needs to launch MR jobs locally
- [ACCUMULO-3849](#) Proxy sets incorrect primary for SASL server transport
- [ACCUMULO-3850](#) Improve logging in replication code path
- [ACCUMULO-3852](#) NPE in WorkMaker for non-existent table
- [ACCUMULO-3853](#) Contention around ConcurrentLinkedQueue.size() in AsyncSpanReceiver
- [ACCUMULO-3856](#) ProxyServer.updateAndFlush leaks BatchWriter
- [ACCUMULO-3858](#) WatchTheWatchCountIT failed with too few watchers
- [ACCUMULO-3859](#) TabletServer never acknowledged constraint
- [ACCUMULO-3861](#) DurabilityIT might actually see all results with durability=none
- [ACCUMULO-3862](#) Improve how AsyncSpanReceiver drops short spans
- [ACCUMULO-3870](#) Loads of warnings from ClientConfiguration delimiter parsing w/ Kerberos
- [ACCUMULO-3874](#) Wrong username in exception when user doesn't exist
- [ACCUMULO-3877](#) TableOperationsIT failed in testCompactEmptyTableWithGeneratorIterator\_Splits\_Cancel
- [ACCUMULO-3878](#) Hunt down ClientConfiguration warnings
- [ACCUMULO-3879](#) MultiInstanceReplicationIT.dataWasReplicatedToThePeer failed
- [ACCUMULO-3880](#) Malformed Configuration Causes tservers To Shutdown
- [ACCUMULO-3881](#) T\*ProxyITs fail with useKrbForIT=true
- [ACCUMULO-3882](#) AccumuloOutputFormatIT loads installed client.conf instead of minicluster's
- [ACCUMULO-3883](#) ITs should not load default ClientConfiguration
- [ACCUMULO-3886](#) Boolean values in SiteConfiguration must use lower-case starting characters
- [ACCUMULO-3887](#) Lack of insight into `accumulo admin stop \$tserver`

- [ACCUMULO-3893](#) ReadWriteIT#sunnyDay fails against Monitor w/ SSL enabled
- [ACCUMULO-3894](#) KerberosProxyIT too aggressive in waiting for proxy to start

### 1.5.3. Atlas

HDP 2.3.6 provides Atlas 0.5.0 with no additional Apache patches beyond any ported over from the [2.4.x core](#).

HDP 2.3.6 provided Atlas 0.5.0 and ports the following Apache patches [from the 2.4.x core](#):

- [ATLAS-448](#): Hive IllegalArgumentException with Atlas hook enabled on SHOW TRANSACTIONS AND SHOW COMPACTIONS.

HDP 2.3.4.7 provided Atlas 0.5.0 with no additional Apache patches.

HDP 2.3.4 provided Atlas 0.5.0 and the following Apache patches:

- [ATLAS-102](#): Issue with SolrIndex.
- [ATLAS-114](#): Upgrade HBase client to 1.1.2.
- [ATLAS-117](#): Build fails on the latest commit.
- [ATLAS-118](#): rename Atlas log4j configuration files to avoid config collisions with other projects.
- [ATLAS-128](#): DSL - Add support for comparisons on list type.
- [ATLAS-16](#): jersey jaxb exception.
- [ATLAS-168](#): Atlas UI - Max column in Hive 4.
- [ATLAS-17](#): Parameterize schema API query per typeName.
- [ATLAS-179](#): Atlas hook causes mem leak and Hive server 2 crashes.
- [ATLAS-180](#): Cleanup Atlas doc packaging.
- [ATLAS-195](#): Document HBase configs.
- [ATLAS-196](#): Fix Solr documentation.
- [ATLAS-198](#): Atlas UI Requires Internet Access.
- [ATLAS-238](#): atlas\_start.py- the Atlas server won't restart after improper shutdown.
- [ATLAS-279](#): UI not displaying results for certain successful "select" search queries.
- [ATLAS-31](#): Fixed ATLAS build fails with clean repo.
- [ATLAS-33](#): Atlas restart fails.
- [ATLAS-334](#): Update documentation to reflect copying required Atlas file on Solr installation.

- [ATLAS-335](#): Kerberized cluster: Atlas fails to come up with HBase as backend.
- [ATLAS-344](#): Document HBase permissions for secure cluster.
- [ATLAS-344](#): Document HBase permissions for secure cluster.
- [ATLAS-350](#): Document jaas config details for Atlas.
- [ATLAS-350](#): Document jaas config details for Atlas.
- [ATLAS-351](#): Solr as indexing backend is not picked up by titan when storage backend is set to HBase.
- [ATLAS-352](#): Fix performance issues with type and entity creation with HBase as storage backend.
- [ATLAS-355](#): Kerberized cluster: client.properties does not have correct values for the properties.
- [ATLAS-36](#): Need separate persisted properties for HTTP and HTTPS ports.
- [ATLAS-361](#): Add validation when index backends are switched in ATLAS configuration.
- [ATLAS-37](#): Atlas repository, webapp, hive-bridge tests fails with HBase and Solr as Titan storage backend.
- [ATLAS-373](#): Renew TGT from KeyTab for ATLAS service principal.
- [ATLAS-45](#): Entity submit fails.
- [ATLAS-54](#): Rename configs in Hive hook.
- [ATLAS-67](#): Validate with secure ZooKeeper server for titan interaction with HBase and solr.
- [ATLAS-81](#): Atlas debian packaging fails in maven build.
- [ATLAS-86](#): Jenkins build failing as of build #41.
- [ATLAS-91](#): Add Solr configuration and documentation.
- [ATLAS-92](#): import-hive.sh failed to find HiveMetaStoreBridge.

HDP 2.3.2 provided Atlas 0.5.0 and the following Apache patches:

- [ATLAS-15](#) remove specific version string as default property value
- [ATLAS-19](#) remove unnecessary docs dir
- [ATLAS-29](#) create configuration that inherits existing Hadoop config
- [ATLAS-31](#) Fixed ATLAS build fails with clean repo
- [ATLAS-31](#) Fixed Mixed Index creation fails with Date types
- [ATLAS-32](#) create HTTP connection in context of invoking user in secure cluster

- [ATLAS-54](#) Rename configs in Hive hook

HDP 2.3.0 provided Atlas 0.5.0, with no additional Apache patches.

## 1.5.4. Calcite

HDP 2.3.6 provides Calcite 1.2.0 with no additional Apache patches beyond any ported over [from the 2.4.x core](#).

HDP 2.3.6 provides Calcite 1.2.0 and ports the following Apache patches [from the 2.4.x core](#):

- [CALCITE-1128](#): Support addBatch()/executeBatch() in remote driver.
- [CALCITE-1180](#): Support clearBatch() in remote driver.
- [CALCITE-984](#): NPE in ErrorResponse construction.

HDP 2.3.4.7 provided Calcite 1.2.0 with no additional Apache patches.

HDP 2.3.4 provided Calcite 1.2.0 and the following Apache patches:

- [CALCITE-259](#): Using sub-queries in CASE statement against JDBC tables generates invalid Oracle SQL.
- [CALCITE-429](#): Add statistics SPI for lattic optimization algorithm.
- [CALCITE-522](#): In remote JDBC driver, transmit static database properties as a map.
- [CALCITE-585](#): Avatica JDBC methods should throw SQLFeatureNotSupportedException.
- [CALCITE-645](#): Pass server-side exceptions back to the client.
- [CALCITE-661](#): Remote fetch in Calcite JDBC driver.
- [CALCITE-671](#): ByteString does not deserialize properly as a FetchRequest parameterValue.
- [CALCITE-677](#): RemoteDriverTest.testTypeHandling fails east of Greenwich.
- [CALCITE-687](#): Make RemoteDriverTest thread-safe.
- [CALCITE-699](#): In Avatica, synchronize access to Calendar.
- [CALCITE-705](#): DML in Avatica, and split Execute out from Fetch request.
- [CALCITE-708](#): Avatica and Calcite to support DatabaseMetaData getTypeInfo.
- [CALCITE-712](#): Avatica statement execute return all resultset instead of MaxRows from setMaxRows.
- [CALCITE-717](#): Compare BINARY and VARBINARY on unsigned byte values.
- [CALCITE-718](#): Enable fetch to work for Statement.execute\*()&#41;.
- [CALCITE-728](#): Test suite hangs on Windows.



- [CALCITE-730](#): ClassCastException in table from CloneSchema.
- [CALCITE-741](#): Dependencies should not be empty.
- [CALCITE-765](#): RPC server returns JSON data with Content-Type set to text/html.
- [CALCITE-780](#): HTTP error 413 when sending a long string to the Avatica server.
- [CALCITE-789](#): MetaImpl.MetaCatalog should expose TABLE\_CAT instead of TABLE\_CATALOG.
- [CALCITE-795](#): Loss of precision when sending a decimal number via the remote JSON service.
- [CALCITE-813](#): Upgrade updateCount, maxRows from int to long.
- [CALCITE-825](#): Allow user to specify sort order of an ArrayTable.
- [CALCITE-840](#): Protobuf transport for Avatica.
- [CALCITE-843](#): AvaticaConnection.getAutoCommit throws NullPointerException.
- [CALCITE-865](#): NullPointerException in getTables with PostgreSQL.
- [CALCITE-866](#): Create documentation for RPC message format.
- [CALCITE-871](#): JdbcResultSet returns incomplete Frame with "default" statement ID.
- [CALCITE-903](#): Enable client to recover from missing server-side state.
- [CALCITE-905](#): getTables returns empty result in JdbcMeta.
- [CALCITE-906](#): Avatica JdbcMeta statement IDs are not unique.
- [CALCITE-908](#): Bump protobuf dependency to protobuf-3.0.0-beta-1.
- [CALCITE-910](#): Improve handling of ARRAY, MULTISSET, STRUCT types.
- [CALCITE-912](#): Pass URL connection properties to avatica server.
- [CALCITE-913](#): Avatica; transport of array fields fails.
- [CALCITE-914](#): Missing JsonSubType for ExecuteResponse.
- [CALCITE-919](#): ArithmeticException when querying against decimal field.
- [CALCITE-921](#): ClassCastException byte[] to String with protobuf.
- [CALCITE-927](#): ColumnsRequest Service call doesn't "fix" ResultSetResponse.
- [CALCITE-951](#): Print the server-side stack in the local exception.
- [CALCITE-962](#): Server-side exception & stack trace; not propagated in JdbcMeta.propagate.
- [CALCITE-983](#): NPE in ErrorResponse construction.

- [CALCITE-989](#): Provide generic server metadata in responses.

HDP 2.3.2 provided Calcite 1.2.0, with no additional Apache patches.

## 1.5.5. Falcon

In addition to any Apache patches ported over [from the 2.4.x core](#), HDP 2.3.6 provides Falcon 0.6.1 and the following Apache patches:

- [FALCON-1372](#) Retention does not work in corner cases.
- [FALCON-1595](#) In secure cluster, Falcon server loses ability to communicate with HDFS over time.
- [FALCON-1647](#) Unable to create feed: FilePermission error under cluster staging directory.
- [FALCON-1877](#) Falcon webUI returns 413 (Full head - Request entity too large) error when TLS is enabled in a secure cluster with AD integration.

HDP 2.3.6 provides Falcon 0.6.1 and ports no additional Apache patches [from the 2.4.x core](#).

HDP 2.3.4.7 provided Falcon 0.6.1 with no additional Apache patches.

HDP 2.3.2 provided Falcon 0.6.1 and the following Apache patches:

### NEW FEATURES

- [FALCON-1039](#) Add instance dependency API in Falcon.
- [FALCON-1188](#) Falcon support for Hive Replication.
- [FALCON-1325](#) Falcon UI.
- [FALCON-796](#) Enable users to triage data processing issues through Falcon.

### IMPROVEMENTS

- [FALCON-1060](#) Handle transaction failures in Lineage.
- [FALCON-1147](#) Allow \_ in the names for name value pair.
- [FALCON-1174](#) Ability to disable Oozie dryrun while scheduling or updating the Falcon entity.
- [FALCON-1186](#) Add filtering capability to result of instance summary.
- [FALCON-1204](#) Expose default configs for feed late data handling in runtime.properties.
- [FALCON-1317](#) Inconsistent JSON serialization.
- [FALCON-1322](#) Add prefix in runtime.properties.
- [FALCON-1324](#) Pagination API breaks backward compatibility.
- [FALCON-1359](#) Improve output format for Feed Instance Listing.

- [FALCON-1361](#) Default end date should be now.
- [FALCON-1368](#) Improve Falcon server restart time.
- [FALCON-1374](#) Remove the cap on numResults.
- [FALCON-1378](#) Falcon documentation lacks information on how to run Falcon on standalone Oozie/Hadoop setup.
- [FALCON-668](#) FeedReplicator improvement to include more DistCP options.
- [FALCON-676](#) Enable metrics for Titan.
- [FALCON-75](#) Falcon CLI for deleting entities should inform user if entity does not exist.

#### BUG FIXES

- [FALCON-1038](#) Log mover fails for map-reduce action.
- [FALCON-1101](#) Cluster submission in Falcon does not create an owned-by edge.
- [FALCON-1104](#) Exception while adding process instance to graphdb when feed has partition expression.
- [FALCON-1121](#) Backend support for free-text entity search.
- [FALCON-1129](#) In a secure cluster, feed replication fails because of Authentication issues.
- [FALCON-1141](#) Reverse Lookup for feed in prism fails with BadRequest.
- [FALCON-1143](#) Correcting order of entities on reload.
- [FALCON-1144](#) Dynamic partitions not getting registered in Hcat.
- [FALCON-1146](#) feed retention policy deleted everything all the way up to the root.
- [FALCON-1153](#) Instance kill fails intermittently.
- [FALCON-1162](#) Cluster submit succeeds when staging HDFS dir does not have 777.
- [FALCON-1165](#) Falcon restart failed, if defined service in cluster entity is unreachable.
- [FALCON-1244](#) numResults query param in listInstances is ignored when start and end params are not specified.
- [FALCON-1252](#) The parameter "tagkey" should be "tagkeys" in EntityList and FalconCLI twiki.
- [FALCON-1260](#) Instance dependency API produces incorrect results.
- [FALCON-1268](#) Instance Dependency API failure message is not intuitive in distributed mode.
- [FALCON-1282](#) Incorrect HDFS servers property for feed replication in secured environment.

- [FALCON-1310](#) Falcon build fails with Oozie-4.2.0.
- [FALCON-1311](#) Instance dependency API produces inconsistent results in some scenarios.
- [FALCON-1312](#) Falcon post processing action should use Oozie prepared configuration.
- [FALCON-1323](#) Reverse lookup of feeds causes NPE.
- [FALCON-1325](#) Triage API on prism, for an instance at which a process does not exist sends incorrect message.
- [FALCON-1327](#) When using triage on a server for a process which does not exist on that server, a NullPointerException is encountered.
- [FALCON-1328](#) Error in Triage documentation.
- [FALCON-1329](#) Falcon's idempotent behaviour breaks in some cases.
- [FALCON-1344](#) EntityGraph returns null in list of dependent entities.
- [FALCON-1363](#) Fix retry policy example in documentation.
- [FALCON-1398](#) CrossEntityValidations contains incorrect validations.
- [FALCON-1399](#) Property for default number of results is not loaded dynamically.
- [FALCON-1409](#) Update API throws NullPointerException.
- [FALCON-1412](#) Process waits indefinitely and finally timed out even though missing dependencies are met.
- [FALCON-1487](#) In secure cluster setup Hcat process/feed scheduling or replication fails.
- [FALCON-954](#) Secure Kerberos setup: Falcon should periodically revalidate auth token.
- [FALCON-99](#) Adding late data to process doesn't create new coord.

HDP 2.3.0 provided Falcon 0.6.1 and the following Apache patches:

#### NEW FEATURES

- [FALCON-1039](#) Add instance dependency API in Falcon
- [FALCON-1188](#) Falcon support for Hive Replication
- [FALCON-790](#) Falcon UI to enable entity/process/feed edits and management
- [FALCON-796](#) Enable users to triage data processing issues through Falcon

#### IMPROVEMENTS

- [FALCON-1060](#) Handle transaction failures in Lineage
- [FALCON-1147](#) Allow \_ in the names for name value pair
- [FALCON-1174](#) Ability to disable Oozie dryrun while scheduling or updating the Falcon entity

- [FALCON-1186](#) Add filtering capability to result of instance summary
- [FALCON-1204](#) Expose default configs for feed late data handling in runtime.properties
- [FALCON-1317](#) Inconsistent JSON serialization
- [FALCON-1322](#) Add prefix in runtime.properties
- [FALCON-1324](#) Pagination API breaks backward compatibility.
- [FALCON-1359](#) Improve output format for Feed Instance Listing
- [FALCON-1361](#) Default end date should be now
- [FALCON-1368](#) Improve Falcon server restart time
- [FALCON-1374](#) Remove the cap on numResults
- [FALCON-1378](#) Falcon documentation lacks information on how to run Falcon on standalone Oozie/Hadoop setup
- [FALCON-668](#) FeedReplicator improvement to include more DistCP options
- [FALCON-676](#) Enable metrics for Titan
- [FALCON-75](#) Falcon CLI for deleting entities should inform user if entity does not exist

#### BUG FIXES

- [FALCON-1101](#) Cluster submission in Falcon does not create an owned-by edge
- [FALCON-1104](#) Exception while adding process instance to graphdb when feed has partition expression
- [FALCON-1121](#) Backend support for free-text entity search
- [FALCON-1129](#) In a secure cluster, feed replication fails because of Authentication issues
- [FALCON-1141](#) Reverse Lookup for feed in prism fails with BadRequest
- [FALCON-1143](#) Correcting order of entities on reload
- [FALCON-1144](#) Dynamic partitions not getting registered in Hcat
- [FALCON-1146](#) feed retention policy deleted everything all the way up to the root
- [FALCON-1153](#) Instance kill fails intermittently
- [FALCON-1162](#) Cluster submit succeeds when staging HDFS dir does not have 777
- [FALCON-1165](#) Falcon restart failed, if defined service in cluster entity is unreachable
- [FALCON-1244](#) numResults query param in listInstances is ignored when start and end params are not specified
- [FALCON-1252](#) The parameter "tagkey" should be "tagkeys" in EntityList and FalconCLI twiki

- [FALCON-1260](#) Instance dependency API produces incorrect results
- [FALCON-1268](#) Instance Dependency API failure message is not intuitive in distributed mode
- [FALCON-1282](#) Incorrect HDFS servers property for feed replication in secured environment
- [FALCON-1310](#) Falcon build fails with Oozie-4.2.0
- [FALCON-1311](#) Instance dependency API produces inconsistent results in some scenarios
- [FALCON-1312](#) Falcon post processing action should use Oozie prepared configuration
- [FALCON-1323](#) Reverse lookup of feeds causes NPE
- [FALCON-1325](#) Triage API on prism, for an instance at which a process does not exist sends incorrect message
- [FALCON-1327](#) When using triage on a server for a process which does not exist on that server, a NullPointerException is encountered
- [FALCON-1328](#) Error in Triage documentation
- [FALCON-1329](#) Falcon's idempotent behaviour breaks in some cases
- [FALCON-1344](#) EntityGraph returns null in list of dependent entities
- [FALCON-1363](#) Fix retry policy example in documentation
- [FALCON-1398](#) CrossEntityValidations contains incorrect validations
- [FALCON-1399](#) Property for default number of results is not loaded dynamically
- [FALCON-1409](#) Update API throws NullPointerException
- [FALCON-1412](#) Process waits indefinitely and finally timed out even though missing dependencies are met
- [FALCON-1487](#) In secure cluster setup Hcat process/feed scheduling or replication fails
- [FALCON-954](#) Secure Kerberos setup : Falcon should periodically revalidate auth token
- [FALCON-99](#) Adding late data to process doesn't create new coord  
[FALCON-1038](#) Log mover fails for map-reduce action

## 1.5.6. Flume

HDP 2.3.6 provides Flume 1.5.2 with no additional Apache patches beyond any ported over [from the 2.4.x core](#).

HDP 2.3.6 provides Flume 1.5.2 and ports the following Apache patches [from the 2.4.x core](#):

- [FLUME-2865](#) Upgrade thrift version (0.9.2).

- [FLUME-2889](#) Fixes to DateTime computations.

HDP 2.3.4.7 provided Flume 1.5.2 with no additional Apache patches.

HDP 2.3.4 provided Flume 1.5.2 and the following Apache patches:

- [FLUME-2841](#) Upgrade commons-collections to 3.2.2.
- [FLUME-2854](#) Parameterizing jetty version.

HDP 2.3.2 provided Flume 1.5.2, with no additional Apache patches.

HDP 2.3.0 provided Flume 1.5.2 and the following Apache patches:

#### NEW FEATURES

- [FLUME-1734](#) Hive Sink based on the new Hive Streaming support
- [FLUME-2442](#) Need an alternative to providing clear text passwords in Flume config

#### Kafka Sink (preview)

- [FLUME-2251](#) Add support for Kafka Sink
- [FLUME-2454](#) Support batchSize to allow multiple events per transaction to the Kafka Sink
- [FLUME-2455](#) Documentation update for Kafka Sink
- [FLUME-2470](#) Kafka Sink and Source must use camel case for all configs.
- [FLUME-2499](#) Include Kafka Message Key in Event Header, Updated Comments

#### Kafka Source

- [FLUME-2250](#) Add support for Kafka Source

#### IMPROVEMENTS

- [FLUME-2095](#) JMS source with TIBCO (patch-1)
- [FLUME-2226](#) Refactor BlobHandler out of morphline sink and into HTTP source
- [FLUME-2227](#) Move BlobDeserializer from Morphline Sink to flume-ng-core
- [FLUME-2337](#) export JAVA\_HOME in flume-env.sh.template and increase heap size
- [FLUME-2450](#) Improve replay index insertion speed
- [FLUME-2511](#) Allow configuration of enabled protocols in Avro source and Rpc client
- [FLUME-2586](#) HDFS Sink should have an option to try rename even if close fails
- [FLUME-2595](#) Add option to checkpoint on file channel shutdown
- [FLUME-2624](#) Streaming ingest performance improvement

- [FLUME-2662](#) Upgrade to Commons-IO 2.4
- [FLUME-2663](#) Address Build warnings of duplicate dependencies listed
- [FLUME-2665](#) Update documentation for `hdfs.closeTries` based on [FLUME-2586](#)

#### BUG FIXES

- [FLUME-2122](#) Minor cleanups of User guide
- [FLUME-2123](#) Morphline Solr sink missing short type name
- [FLUME-2162](#) `TestHDFSEventSinkOnMiniCluster.maxUnderReplicationTest` fails on `hadoop2`
- [FLUME-2175](#) Update Developer Guide with notes on how to upgrade Protocol Buffer version
- [FLUME-2358](#) File Channel needs to close `BackingStore` and `EventQueue` before deleting files in checkpoint directory
- [FLUME-2402](#) Warning seen when overflow is disabled for Spillable Channel
- [FLUME-2407](#) Spillable Channel sometimes fails on reconfigure
- [FLUME-2412](#) Improve Logging in Spillable Channel
- [FLUME-2441](#) Unit test `TestHTTPSource.java` failed with IBM JDK 1.7
- [FLUME-2451](#) HDFS Sink Cannot Reconnect After NameNode Restart
- [FLUME-2501](#) Updating `HttpClient` lib version to ensure compatibility with Solr
- [FLUME-2520](#) HTTP Source should be able to block a prefixed set of protocols.
- [FLUME-2530](#) Resource leaks found by Coverity tool
- [FLUME-2533](#) HTTPS tests fail on Java 6
- [FLUME-2541](#) Bug in `TestBucketWriter.testSequenceFileCloseRetries`

## 1.5.7. HBase

In addition to any Apache patches ported over [from the 2.4.x core](#), HDP 2.3.6 provides HBase 1.1.2 and the following Apache patches:

- [HBASE-13318](#): `RpcServer.getListenerAddress` should handle when the accept channel is closed.
- [HBASE-13639](#): `SyncTable` - `rsync` for HBase tables.
- [HBASE-14347](#): Add a switch to `DynamicClassLoader` to disable it.
- [HBASE-14604](#): Improve `MoveCostFunction` in `StochasticLoadBalancer`.



- [HBASE-14606](#): TestSecureLoadIncrementalHFiles tests timed out in trunk build on Apache.
- [HBASE-14655](#): Narrow the scope of doAs() calls to region observer notifications for compaction.
- [HBASE-14674](#): Rpc handler / task monitoring seems to be broken after 0.98.
- [HBASE-14840](#): Sink cluster reports data replication request as success though the data is not replicated.
- [HBASE-14941](#): locate\_region shell command.
- [HBASE-14976](#): Add RPC call queues to the web UI.
- [HBASE-15039](#): HMaster and RegionServers should try to refresh token keys from zk when facing InvalidToken.
- [HBASE-15066](#): Small improvements to Canary tool.
- [HBASE-15125](#): HBaseFsck's adoptHdfsOrphan function creates region with wrong end key boundary.
- [HBASE-15128](#): Disable region splits and merges switch in master.
- [HBASE-15145](#): Default hbase-env.sh breaks add peer command that writes to /hbase-secure/replication znode for kerberized clusters.
- [HBASE-15147](#): Shell should use Admin.listTableNames() instead of Admin.listTables().
- [HBASE-15252](#): Data loss when replaying wal if HDFS timeout.
- [HBASE-15295](#): MutateTableAccess.multiMutate() does not get high priority causing a deadlock.
- [HBASE-15592](#): Print Procedure WAL content.
- [HBASE-15613](#): TestNamespaceCommand times out.
- [HBASE-15622](#): Superusers does not consider the keytab credentials.
- [HBASE-15720](#): Print row locks at the debug dump page.
- [HBASE-15811](#): Batch Get after batch Put does not fetch all Cells We were not waiting on all executors in a batch to complete. The test for no-more-executors was damaged by the 0.99/0.98.4 fix "HBASE-11403 Fix race conditions around Object#notify".
- [HBASE-15837](#): MemStore size accounting is wrong if postBatchMutate() throws exception.

HDP 2.3.6 provides HBase 1.1.2 and ports the following Apache patches [from the 2.4.x core](#):

- [HBASE-14107](#): Administrative Task: Provide an API to List all procedures.

- [HBASE-14108](#): Administrative Task: provide an API to abort a procedure.
- [HBASE-14283](#): Reverse scan doesn't work with HFile inline index/bloom blocks.
- [HBASE-14307](#): Incorrect use of positional read API in HFileBlock.
- [HBASE-14432](#): Enforce ACL on procedure admin tasks.
- [HBASE-14468](#): Compaction improvements: FIFO compaction policy.
- [HBASE-14471](#): Thrift - HTTP Error 413 full HEAD if using Kerberos authentication.
- [HBASE-14487](#): Shell command to list all procedures.
- [HBASE-14488](#): Shell command to abort a procedure.
- [HBASE-14575](#): Relax region read lock for compactions.
- [HBASE-14761](#): Deletes with and without visibility expression do not delete the matching mutation.
- [HBASE-14812](#): Fix ResultBoundedCompletionService deadlock.
- [HBASE-14866](#): VerifyReplication should use peer configuration in peer connection.
- [HBASE-14883](#): TestSplitTransactionOnCluster#testFailedSplit flakey.
- [HBASE-14928](#): Start row should be set for query through HBase REST gateway involving globbing option.
- [HBASE-14963](#): Remove Guava dependency from HBase client code.
- [HBASE-14987](#): Compaction marker whose region name doesn't match current region's needs to be handled.
- [HBASE-15014](#): Fix filterCellByStore in WALsplitter is awful for performance.
- [HBASE-15022](#): replication\_admin.rb throws undefined method 'getZooKeeperClusterKey' for ZKUtil.
- [HBASE-15035](#): bulkloading HFiles with tags that require splits do not preserve tags.
- [HBASE-15220](#): Change two logs in SimpleRegionNormalizer to INFO level.
- [HBASE-15221](#): Reload the cache on re-tried puts in HTableMultiplexer and adds a close() method to HTableMultiplexer.
- [HBASE-15232](#): Handle region location cache management in AsyncProcess for multi()'s.
- [HBASE-15621](#): Suppress HBase SnapshotHFile cleaner error messages when a snapshot is going on.

HDP 2.3.4.7 provided HBase 1.1.2 with no additional Apache patches.

HDP 2.3.4 provided HBase 1.1.2 and the following Apache patches:

- [HBASE-13103](#): [ergonomics] add region size balancing as a feature of master.
- [HBASE-13250](#): chown of ExportSnapshot does not cover all path and files.
- [HBASE-14207](#): Region was hijacked and remained in transition when RS failed to open a region and later regionplan changed to new RS on retry.
- [HBASE-14268](#): Improve KeyLocker.
- [HBASE-14280](#): Bulk Upload from HA cluster to remote HA HBase cluster fails.
- [HBASE-14309](#): Allow load balancer to operate when there is region in transition by adding force flag.
- [HBASE-14314](#): Metrics for block cache should take region replicas into account.
- [HBASE-14342](#): Recursive call in RegionMergeTransactionImpl.getJournal().
- [HBASE-14359](#): HTable#close will hang forever if unchecked error/exception thrown in AsyncProcess#sendMultiAction.
- [HBASE-14361](#): ReplicationSink should create Connection instances lazily.
- [HBASE-14367](#): Add normalization support to shell.
- [HBASE-14445](#): ExportSnapshot does not honor -chmod option.
- [HBASE-14463](#): Severe performance downgrade when parallel reading a single key from BucketCache.
- [HBASE-14475](#): Region split requests are always audited with HBase user rather than request user.
- [HBASE-14497](#): Reverse Scan threw StackOverflow caused by readPt checking.
- [HBASE-14536](#): Balancer and SSH interfering with each other leading to unavailability.
- [HBASE-14565](#): Make ZK connection timeout configurable in MiniZooKeeperCluster.
- [HBASE-14578](#): URISyntaxException during snapshot restore for table with user defined namespace.
- [HBASE-14581](#): Znode cleanup throws auth exception in secure mode.
- [HBASE-14591](#): Region with reference HFile may split after a forced split in IncreasingToUpperBoundRegionSplitPolicy.
- [HBASE-14605](#): Split fails due to 'No valid credentials' error when SecureBulkLoadEndpoint#start tries to access HDFS.
- [HBASE-14621](#): ReplicationLogCleaner stuck on RS crash.
- [HBASE-14624](#): BucketCache.freeBlock is too expensive.
- [HBASE-14631](#): Region merge request should be audited with request user through proper scope of doAs() calls to region observer notifications.

- [HBASE-14632](#): Region server aborts due to unguarded dereference of Reader.
- [HBASE-14680](#): Two configs for snapshot timeout and better defaults.
- [HBASE-14706](#): RegionLocationFinder should return multiple servernames by top host.
- [HBASE-14759](#): Avoid using Math.abs when selecting SyncRunner in FSHLog.
- [HBASE-14788](#): Splitting a region does not support the hbase.rs.evictblocksonclose config when closing source region.
- [HBASE-14799](#): Commons-collections object deserialization remote command execution vulnerability.
- [HBASE-14809](#): Grant / revoke Namespace admin permission to group.
- [HBASE-14885](#): NullPointerException in HMaster#normalizeRegions() due to missing TableDescriptor.
- [HBASE-14893](#): Race between mutation on region and region closing operation leads to NotServingRegionException.

HDP 2.3.2 provided HBase 1.1.2 and the following Apache patches:

- [HBASE-14258](#) Make region\_mover.rb script case insensitive with regard to hostname
- [HBASE-14258](#) Make region\_mover.rb script case insensitive with regard to hostname
- [HBASE-14269](#) FuzzyRowFilter omits certain rows when multiple fuzzy key exist
- [HBASE-14302](#) TableSnapshotInputFormat should not create back references when restoring snapshot
- [HBASE-14313](#) After a Connection sees ConnectionClosingException it never recovers
- [HBASE-14449](#) Rewrite deadlock prevention for concurrent connection close
- [HBASE-14474](#) DeadLock in RpcClientImpl.Connection.close()

HDP 2.3.0 provided HBase 1.1.1 and the following Apache patches:

- [HBASE-11658](#) Piped commands to HBase shell should return non-zero if shell command failed
- [HBASE-11940](#) Add utility scripts for snapshotting / restoring all tables in cluster

## 1.5.8. Hive

In addition to any Apache patches ported over [from the 2.4.x core](#), HDP 2.3.6 provides Hive 1.2.1 and the following Apache patches:

- [HIVE-10061](#): HiveConf Should not be used as part of the HS2 client side code.
- [HIVE-10729](#): Query failed when select complex columns from joined table (Tez map join only).

- [HIVE-10729](#): Query failed when select complex columns from joined table (Tez map join only).
- [HIVE-11097](#): HiveInputFormat uses String.startsWith to compare splitPath and PathToAliases.
- [HIVE-11470](#): LLAP: hadoop-1 build is broken as MiniLlapCluster cannot compile on hadoop-1.
- [HIVE-11583](#): When PTF is used over a large partitions result could be corrupted.
- [HIVE-11768](#): java.io.DeleteOnExitHook leaks memory on long running Hive Server2 Instances.
- [HIVE-11816](#): Upgrade groovy to 2.4.4.
- [HIVE-11841](#): KeyValuesInputMerger creates huge logs.
- [HIVE-11891](#): Add basic performance logging to Metastore calls.
- [HIVE-12165](#): wrong result when hive.optimize.sampling.orderby=true with some aggregate functions.
- [HIVE-12189](#): The list in pushdownPreds of ppd.ExprWalkerInfo should not be allowed to grow very large.
- [HIVE-12279](#): Test case to verify session temporary files are removed after HIVE-11768.
- [HIVE-12435](#): SELECT COUNT(CASE WHEN...) GROUPBY returns 1 for 'NULL' in a case of ORC and vectorization is enabled..
- [HIVE-12563](#): NullPointerException with 3-way Tez merge join.
- [HIVE-12660](#): HS2 memory leak with .hiverc file use.
- [HIVE-12749](#): Constant propagate returns string values in incorrect format.
- [HIVE-12766](#): TezTask does not close DagClient after execution.
- [HIVE-12768](#): Thread safety: binary sortable serde decimal deserialization.
- [HIVE-12784](#): Group by SemanticException: Invalid column reference.
- [HIVE-12875](#): Verify sem.getInputs() and sem.getOutputs().
- [HIVE-12879](#): RowResolver of Semijoin not updated in CalcitePlanner.
- [HIVE-12947](#): SMB join in Tez has ClassCastException when container reuse is on.
- [HIVE-12992](#): Hive on Tez: Bucket map join plan is incorrect.
- [HIVE-13043](#): Reload function has no impact to function registry.
- [HIVE-13051](#): Deadline class has numerous issues.
- [HIVE-13090](#): Hive Metastore crashes on NPE with ZooKeeperTokenStore.

- [HIVE-13092](#): Vectorized java.lang.ClassCastException: org.apache.hadoop.hive.serde2.typeinfo.ListTypeInfo cannot be cast to org.apache.hadoop.hive.serde2.typeinfo.PrimitiveTypeInfo.
- [HIVE-13120](#): propagate doAs when generating ORC splits.
- [HIVE-13126](#): Clean up MapJoinOperator properly to avoid object cache reuse with unintentional states.
- [HIVE-13144](#): HS2 can leak ZK ACL objects when curator retries to create the persistent ephemeral node.
- [HIVE-13174](#): Remove Vectorizer noise in logs.
- [HIVE-13200](#): Aggregation functions returning empty rows on partitioned columns.
- [HIVE-13242](#): DISTINCT keyword is dropped by the parser for windowing.
- [HIVE-13263](#): Vectorization: Unable to vectorize regexp\_extract/regexp\_replace " Udf": GenericUDFBridge, is not supported".
- [HIVE-13285](#): Orc concatenation may drop old files from moving to final path.
- [HIVE-13296](#): Add vectorized Q test with complex types showing count(\*) etc work correctly.
- [HIVE-13298](#): nested join support causes undecipherable errors in SemanticAnalyzer.
- [HIVE-13325](#): Excessive logging when ORC PPD fails type conversions.
- [HIVE-13330](#): ORC vectorized string dictionary reader does not differentiate null vs empty string dictionary.
- [HIVE-13372](#): Hive Macro overwritten when multiple macros are used in one column.
- [HIVE-13394](#): Analyze table fails in Tez on empty partitions/files/tables.
- [HIVE-13561](#): HiveServer2 is leaking ClassLoaders when add jar / temporary functions are used.
- [HIVE-13602](#): TPCQ q16 return wrong result when CBO is on.
- [HIVE-13609](#): Fix UDTFs to allow local fetch task to fetch rows forwarded by GenericUDTF.close().
- [HIVE-13619](#): Bucket map join plan is incorrect.
- [HIVE-13693](#): Multi-insert query drops Filter before file output when there is a.val <> b.val.
- [HIVE-13767](#): Wrong type inferred in Semijoin condition leads to AssertionError.
- [HIVE-13831](#): Error pushing predicates to HBase storage handler.

HDP 2.3.6 provides Hive 1.2.1 and ports no additional Apache patches [from the 2.4.x core](#).

HDP 2.3.4.7 provided Hive 1.2.1 and the following Apache patch:

- [HIVE-12766](#): Thread leaks from TimelineClient.

HDP 2.3.4 provided Hive 1.2.1 and the following Apache patches:

- [HIVE-10397](#): Fix Split Computation for ACID tables with Delta files.
- [HIVE-10592](#): query fails on ORC ppd on timestamp datatype on stripes with all null on the column.
- [HIVE-10752](#): Revert HIVE-5193.
- [HIVE-10755](#): Rework on HIVE-5193 to enhance the column oriented table access.
- [HIVE-10778](#): LLAP: Utilities::gWorkMap needs to be cleaned in HiveServer2.
- [HIVE-10807](#): Invalidate basic stats for insert queries if autogather=false.
- [HIVE-10980](#): Merge of dynamic partitions loads all data to default partition.
- [HIVE-11008](#): WebHCat GET /jobs retries on getting job details from history server is too aggressive.
- [HIVE-11016](#): MiniTez mergejoin test fails with Tez input error (issue in merge join under certain conditions).
- [HIVE-11123](#): Fix how to confirm the RDBMS product name at Metastore.
- [HIVE-11149](#): sometimes HashMap in PerfLogger.java hangs.
- [HIVE-11312](#): ORC format: where clause with CHAR data type not returning any rows.
- [HIVE-11371](#): Null pointer exception for nested table query when using ORC versus text.
- [HIVE-11372](#): join with between predicate comparing integer types returns no rows when ORC format used.
- [HIVE-11398](#): Parse wide OR and wide AND trees to flat OR/AND trees.
- [HIVE-11422](#): Join a ACID table with non-ACID table fail with MR.
- [HIVE-11428](#): Performance: Struct IN() clauses are extremely slow (~ 10x slower).
- [HIVE-11432](#): Hive macro give same result for different arguments.
- [HIVE-11448](#): Support vectorization of Multi-OR and Multi-AND.
- [HIVE-11461](#): Transform flat AND/OR into IN struct clause.
- [HIVE-11462](#): GenericUDFStruct should constant fold at compile time.
- [HIVE-11468](#): Vectorize: Struct IN() clauses.
- [HIVE-11497](#): Make sure -orcfiledump utility includes OrcRecordUpdate.AcidStats.
- [HIVE-11499](#): Hiveserver2 failing with OOM PermGen when using temporary functions due to DataNucleus caching classloaders.

- [HIVE-11510](#): Metatool updateLocation warning on views.
- [HIVE-11517](#): Vectorized auto\_smb\_mapjoin\_14.q produces different results.
- [HIVE-11523](#): org.apache.hadoop.hive ql.io.orc.FileDump should handle errors.
- [HIVE-11540](#): Too many delta files during Compaction OOM.
- [HIVE-11546](#): Projected columns read size should be scaled to split size for ORC Splits.
- [HIVE-11573](#): PointLookupOptimizer can be pessimistic at a low nDV.
- [HIVE-11583](#): When PTF is used over a large partitions result could be corrupted.
- [HIVE-11613](#): schematool should return non zero exit status for info command, if state is inconsistent.
- [HIVE-11634](#): Gen Plan Changes to support multi col in clause.
- [HIVE-11668](#): make sure directsql calls pre-query init when needed.
- [HIVE-11669](#): OrcFileDump service should support directories.
- [HIVE-11698](#): Add additional test for PointLookupOptimizer.
- [HIVE-11712](#): Duplicate groupby keys cause ClassCastException.
- [HIVE-11720](#): Allow customers set a custom request/response header size for hiveserver2 using http.
- [HIVE-11723](#): Incorrect string literal escaping.
- [HIVE-11724](#): WebHcat get jobs to order jobs on time order with latest at top.
- [HIVE-11737](#): IndexOutOfBounds compiling query with duplicated groupby keys.
- [HIVE-11745](#): Alter table Exchange partition with multiple partition\_spec is not working.
- [HIVE-11748](#): HivePreparedStatement's setTimestamp() does not quote value as required.
- [HIVE-11807](#): Adjust compression buffer size to avoid creation of too many small stripes.
- [HIVE-11825](#): get\_json\_object(col,'\$.a') is null in where clause didn't work.
- [HIVE-11831](#): TXN tables in Oracle should be created with ROWDEPENDENCIES.
- [HIVE-11835](#): Type decimal(1,1) reads 0.0, 0.00, etc from text file as NULL.
- [HIVE-11892](#): UDTF run in local fetch task does not return rows forwarded during GenericUDTF.close().
- [HIVE-11901](#): StorageBasedAuthorizationProvider requires write permission on table for SELECT statements.
- [HIVE-11902](#): Abort txn cleanup thread throws SyntaxErrorException.



- [HIVE-11914](#): When transactions gets a heartbeat, it doesn't update the lock heartbeat.
- [HIVE-11915](#): BoneCP returns closed connections from the pool.
- [HIVE-11916](#): TxnHandler.getOpenTxnsInfo() and getOpenTxns() may produce inconsistent result.
- [HIVE-11919](#): Hive Union Type Mismatch.
- [HIVE-11920](#): ADD JAR failing with URL schemes other than file/ivy/HDFS.
- [HIVE-11934](#): Transaction lock retry logic results in infinite loop.
- [HIVE-11939](#): TxnDbUtil should turn off jdbc auto commit.
- [HIVE-11940](#): 'INSERT OVERWRITE' query is very slow because it creates one distcp per file to copy data from staging directory to target directory.
- [HIVE-11948](#): Investigate TxnHandler and CompactionTxnHandler to see where we improve concurrency.
- [HIVE-11950](#): WebHCat status file doesn't show UTF8 character.
- [HIVE-11960](#): braces in join conditions are not supported.
- [HIVE-11964](#): RelOptHiveTable.hiveColStatsMap might contain mismatched column stats.
- [HIVE-11975](#): mssql scripts contains invalid 'GO' statement.
- [HIVE-11977](#): Hive should handle an external Avro table with zero length files present.
- [HIVE-11983](#): Hive streaming API uses incorrect logic to assign buckets to incoming records.
- [HIVE-11983](#): Hive streaming API : uses incorrect logic to assign buckets to incoming records
- [HIVE-11988](#): SUMMARY-[security issue with Hive & Ranger for import table command].
- [HIVE-11990](#): Humboldt: loading data inpath from a temporary table dir fails.
- [HIVE-11990](#): loading data inpath from a temporary table dir fails.
- [HIVE-11995](#): Remove repetitively setting permissions in insert/load overwrite partition.
- [HIVE-11997](#): Add ability to send Compaction Jobs to specific queue.
- [HIVE-11998](#): Improve Compaction process logging.
- [HIVE-12003](#): HIVE-12276: Fix messages in InvalidTable.
- [HIVE-12003](#): Hive Streaming API Add check to ensure table is transactional.
- [HIVE-12011](#): unable to create temporary table using CTAS if regular table with that name already exists.

- [HIVE-12012](#): select query on json table with map containing numeric values fails.
- [HIVE-12021](#): wrong results: HivePreFilteringRule may introduce wrong common operands.
- [HIVE-12057](#): ORC sarg is logged too much.
- [HIVE-12076](#): WebHCat listing jobs after the given JobId even when templeton.jobs.listorder is set to lexicographicaldesc.
- [HIVE-12083](#): HIVE-10965 introduces thrift error if partNames or colNames are empty.
- [HIVE-12084](#): Hive queries with ORDER BY and large LIMIT fails with OutOfMemoryError Java heap space.
- [HIVE-12156](#): expanding view doesn't quote reserved keyword.
- [HIVE-12179](#): Add option to not add spark-assembly.jar to Hive classpath.
- [HIVE-12196](#): NPE when converting bad timestamp value.
- [HIVE-12201](#): Tez settings need to be shown in set -v output when execution engine is tez.
- [HIVE-12204](#): Tez queries stopped running with ApplicationNotRunningException.
- [HIVE-12206](#): using UDF in same session after other queries fails with ClassNotFound error.
- [HIVE-12223](#): Filter on Grouping\_\_ID does not work properly.
- [HIVE-12230](#): custom UDF configure() not called in Vectorization mode.
- [HIVE-12232](#): BucketingSortingReduceSinkOptimizer throws IOB exception for duplicate columns.
- [HIVE-12235](#): a way for admin and jdbc/odbc client to see which hiveserver2 connected in ZooKeeper discovery service.
- [HIVE-12236](#): Enable SimpleFetchOptimizer for more query types.
- [HIVE-12249](#): Improve logging with tez.
- [HIVE-12250](#): query on Hive table on hbasestoragehandler leak ZooKeeper connection.
- [HIVE-12252](#): Streaming API HiveEndPoint can be created w/o partitionVals for partitioned table.
- [HIVE-12254](#): Improve logging with YARN/HDFS.
- [HIVE-12257](#): Enhance ORC FileDump utility to handle flush\_length files.
- [HIVE-12261](#): schematool version info exit status should depend on compatibility, not equality.
- [HIVE-12262](#): Session log dir cannot be created in some cases.
- [HIVE-12266](#): When client exists abnormally, it doesn't release ACID locks.

- [HIVE-12273](#): Improve user level explain.
- [HIVE-12276](#): Fix messages in InvalidTable.
- [HIVE-12277](#): Hive macro results on macro\_duplicate.q different after adding ORDER BY.
- [HIVE-12280](#): HiveConnection does not try other HS2 after failure for service discovery.
- [HIVE-12282](#): Beeline update command printing in verbose mode.
- [HIVE-12295](#): change some logs from info to debug.
- [HIVE-12307](#): Streaming API TransactionBatch.close() must abort any remaining transactions in the batch.
- [HIVE-12312](#): Excessive logging in PPD code.
- [HIVE-12318](#): QTest failing due to NPE in logStats.
- [HIVE-12327](#): WebHCat e2e tests TestJob\_1 and TestJob\_2 fail.
- [HIVE-12344](#): Wrong types inferred for SemiJoin generation in CBO.
- [HIVE-12345](#): Follow up for HIVE-9013 Hidden conf vars still visible through Beeline.
- [HIVE-12357](#): Allow user to set Tez job name.
- [HIVE-12364](#): insert into directory query fail, using distcp when data size > hive.exec.copyfile.maxsize (default 32MB).
- [HIVE-12384](#): Union Operator may produce incorrect result on TEZ.
- [HIVE-12387](#): Bug with logging improvements in ATS.
- [HIVE-12387](#): Issues in Hive's use of CallerContext.
- [HIVE-12387](#): When using Hive on Tez or MR, no caller context shows up in YARN audit logs.
- [HIVE-12389](#): CompactionTxnHandler.cleanEmptyAbortedTxns() should safeguard against huge IN clauses.
- [HIVE-12396](#): BucketingSortingReduceSinkOptimizer may still throw IOB exception for duplicate columns.
- [HIVE-12399](#): Filter out NULLs in the Native Vector MapJoin operators.
- [HIVE-12418](#): HiveHBaseTableInputFormat.getRecordReader() causes ZooKeeper connection leak.
- [HIVE-12437](#): SMB join in Tez fails when one of the tables is empty.
- [HIVE-12444](#): Global Limit optimization on ACID table without base directory may throw exception.
- [HIVE-12450](#): OrcFileMergeOperator does not use correct compression buffer size.

- [HIVE-12465](#): Hive might produce wrong results when (outer) joins are merged.
- [HIVE-12469](#): Bump Commons-Collections dependency from 3.2.1 to 3.2.2. to address vulnerability.
- [HIVE-12476](#): Oracle directSQL NPE error when fetching empty to null for serdeproperties.
- [HIVE-12498](#): ACID: Setting OrcRecordUpdater.OrcOptions.tableProperties() has no effect.
- [HIVE-12500](#): JDBC driver not overlaying params supplied via properties object when reading params from ZK.
- [HIVE-12522](#): Wrong FS error during Tez merge files when warehouse and scratchdir are on different FS.
- [HIVE-12523](#): ATS HIVE\_QUERY\_ID api returns different dagName than TEZ\_DAG\_ID api.
- [HIVE-12529](#): HiveTxnManager.acquireLocks() should not block forever.
- [HIVE-12529](#): HiveTxnManager.acquireLocks() should not block forever.
- [HIVE-12556](#): Have an option to kill DAG when user cancels query in HiveServer2.
- [HIVE-12563](#): NullPointerException with 3-way Tez merge join.
- [HIVE-12565](#): VectorUDAFCount.aggregateInputSelection does not handle isRepeated case.
- [HIVE-12567](#): Enhance TxnHandler retry logic to handle ORA-08176.
- [HIVE-12578](#): Hive query failing with error ClassCastException org.apache.hadoop.hive.ql.plan.ExprNodeConstantDesc cannot be cast to org.apache.hadoop.hive.ql.plan.ExprNodeColumnDesc.
- [HIVE-12583](#): HS2 ShutdownHookManager holds extra of Driver instance.
- [HIVE-12584](#): join on char cols with different length returns empty result with vectorization and Tez on.
- [HIVE-12585](#): fix TxnHandler connection leak.
- [HIVE-5623](#): ORC accessing array column that's empty will fail with java out of bound exception.
- [HIVE-7723](#): Explain plan for complex query with lots of partitions is slow due to inefficient collection used to find a matching ReadEntity.
- [HIVE-9013](#): Beeline (hiveserver2 client) exposes sensitive Metastore DB connection info (connection, password).
- [HIVE-9695](#): Redundant filter operator in reducer Vertex when CBO is disabled.

HDP 2.3.2 provided Hive 1.2.1 and the following Apache patches:

#### IMPROVEMENTS

- [HIVE-11037](#) HiveOnTez: make explain user level = true as default

#### BUG FIXES

- [HIVE-10140](#): Window boundary is not compared correctly
- [HIVE-10453](#): Reverted
- [HIVE-10569](#): Hive CLI gets stuck when hive.exec.parallel=true; and some exception happens during SessionState.start
- [HIVE-10571](#): HiveMetaStoreClient should close existing thrift connection before its reconnect
- [HIVE-10620](#): ZooKeeperHiveLock overrides equal() method but not hashCode()
- [HIVE-10646](#): ColumnValue does not handle NULL\_TYPE
- [HIVE-10651](#): ORC file footer cache should be bounded
- [HIVE-10698](#): query on view results fails with table not found error if view is created with subquery alias (CTE).
- [HIVE-10714](#): Bloom filter column names specification should be case insensitive
- [HIVE-10722](#): external table creation with msck in Hive can create unusable partition
- [HIVE-10726](#): Hive JDBC setQueryTimeout should not throw exception to make it work with JMeter
- [HIVE-10731](#): NullPointerException in HiveParser.g
- [HIVE-10732](#): Hive JDBC driver does not close operation for metadata queries
- [HIVE-10771](#): "separatorChar" has no effect in "CREATE TABLE AS SELECT" statement
- [HIVE-10781](#): HadoopJobExecHelper Leaks RunningJobs
- [HIVE-10790](#): orc write on viewFS throws exception
- [HIVE-10793](#): HIVE-11587 Hybrid Grace Hash Join: Don't allocate all hash table memory upfront
- [HIVE-10802](#): Table join query with some constant field in select fails
- [HIVE-10808](#): Inner join on Null throwing Cast Exception
- [HIVE-10835](#): Concurrency issues in JDBC driver
- [HIVE-10880](#): The bucket number is not respected in insert overwrite.
- [HIVE-10925](#): Non-static threadlocals in Metastore code can potentially cause memory leak
- [HIVE-10963](#): Hive throws NPE rather than meaningful error message when window is missing

- [HIVE-10972](#): DummyTxnManager always locks the current database in shared mode, which is incorrect.
- [HIVE-11013](#): MiniTez tez\_join\_hash test on the branch fails with NPE (initializeOp not called?)
- [HIVE-11024](#): Error inserting a date value via parameter marker (PreparedStatement.setDate)
- [HIVE-11029](#): hadoop.proxyuser.mapr.groups does not work to restrict the groups that can be impersonated
- [HIVE-11054](#): Read error: Partition Varchar column cannot be cast to string
- [HIVE-11079](#): Fix qfile tests that fail on Windows due to CR/character escape differences
- [HIVE-11087](#): DbTxnManager exceptions should include txnid
- [HIVE-11090](#): ordering issues with windows unit test runs
- [HIVE-11095](#): SerDeUtils another bug ,when Text is reused
- [HIVE-11102](#): ReaderImpl: getColumnIndicesFromNames does not work for some cases
- [HIVE-11112](#): ISO-8859-1 text output has fragments of previous longer rows appended
- [HIVE-11135](#): Fix the Beeline set and save command in order to avoid the NullPointerException
- [HIVE-11151](#): Calcite transitive predicate inference rule should not transitively add not null filter on non-nullable input
- [HIVE-11152](#): Swapping join inputs in ASTConverter
- [HIVE-11157](#): Hive.get(HiveConf) returns same Hive object to different user sessions
- [HIVE-11171](#): Join reordering algorithm might introduce projects between joins
- [HIVE-11172](#): Vectorization wrong results for aggregate query with where clause without group by
- [HIVE-11174](#): Hive does not treat floating point signed zeros as equal (-0.0 should equal 0.0 according to IEEE floating point spec)
- [HIVE-11176](#): Caused by: java.lang.ClassCastException: org.apache.hadoop.hive.serde2.lazybinary.LazyBinaryStruct cannot be cast to [Ljava.lang.Object;
- [HIVE-11193](#): ConstantPropagateProcCtx should use a Set instead of a List to hold operators to be deleted
- [HIVE-11198](#): Fix load data query file format check for partitioned tables
- [HIVE-11203](#): Beeline force option doesn't force execution when errors occurred in a script.

- [HIVE-11211](#): Reset the fields in JoinStatsRule in StatsRulesProcFactory
- [HIVE-11216](#): UDF GenericUDFMapKeys throws NPE when a null map value is passed in
- [HIVE-11221](#): In Tez mode, alter table concatenate orc files can intermittently fail with NPE
- [HIVE-11255](#): get\_table\_objects\_by\_name() in HiveMetaStore.java needs to retrieve table objects in multiple batches
- [HIVE-11258](#): The function drop\_database\_core() of HiveMetaStore.java may not drop all the tables
- [HIVE-11271](#): java.lang.IndexOutOfBoundsException when union all with if function
- [HIVE-11301](#): thrift Metastore issue when getting stats results in disconnect
- [HIVE-11303](#): Getting Tez LimitExceededException after dag execution on large query
- [HIVE-11317](#): ACID: Improve transaction Abort logic due to timeout
- [HIVE-11320](#): ACID enable predicate pushdown for insert-only delta file
- [HIVE-11344](#): HIVE-9845 makes HCatSplit.write modify the split so that PartInfo objects are unusable after it
- [HIVE-11356](#): SMB join on Tez fails when one of the tables is empty
- [HIVE-11357](#): ACID enable predicate pushdown for insert-only delta file 2
- [HIVE-11375](#): Broken processing of queries containing NOT (x IS NOT NULL and x 0)
- [HIVE-11407](#): JDBC DatabaseMetaData.getTables with large no of tables call leads to HS2 OOM
- [HIVE-11429](#): Increase default JDBC result set fetch size (# rows it fetches in one RPC call) to 1000 from 50
- [HIVE-11433](#): NPE for a multiple inner join query
- [HIVE-11442](#): Remove commons-configuration.jar from Hive distribution
- [HIVE-11449](#): HIVE-11587 "Capacity must be a power of two" error when HybridHashTableContainer memory threshold is too low
- [HIVE-11456](#): HCatStorer should honor mapreduce.output.basename
- [HIVE-11467](#): HIVE-11587 WriteBuffers rounding wbSize to next power of 2 may cause OOM
- [HIVE-11493](#): Predicate with integer column equals double evaluates to false
- [HIVE-11502](#): Map side aggregation is extremely slow
- [HIVE-11581](#): HiveServer2 should store connection params in ZK when using dynamic service discovery for simpler client connection string.

- [HIVE-11587](#): Fix memory estimates for mapjoin hashtable
- [HIVE-11592](#): ORC metadata section can sometimes exceed protobuf message size limit
- [HIVE-11600](#): Hive Parser to Support multi col in clause (x,y..) in ((..),..., ())
- [HIVE-11605](#): Incorrect results with bucket map join in tez.
- [HIVE-11606](#): Bucket map joins fail at hash table construction time
- [HIVE-11607](#): Export tables broken for data > 32 MB
- [HIVE-11658](#): Load data file format validation does not work with directories
- [HIVE-11727](#): (BUG-44285). Hive on Tez through Oozie: Some queries fail with fnf exception
- [HIVE-11755](#): Incorrect method called with Kerberos enabled in AccumuloStorageHandler
- [HIVE-11820](#): export tables with size of >32MB throws "java.lang.IllegalArgumentException: Skip CRC is valid only with update options"
- [HIVE-11836](#): ORC SARG creation throws NPE for null constants with void type
- [HIVE-11839](#): Vectorization wrong results with filter of (CAST AS CHAR)
- [HIVE-11849](#): NPE in HiveHBaseTableShapshotInputFormat in query with just count(\*)
- [HIVE-11852](#): numRows and rawDataSize table properties are not replicated
- [HIVE-11875](#): JDBC Driver does not honor delegation token mechanism when readings params from ZooKeeper
- [HIVE-11897](#): JDO rollback can throw pointless exceptions
- [HIVE-11928](#): ORC footer section can also exceed protobuf message limit
- [HIVE-11936](#): Support SQLAnywhere as a backing DB for the Hive Metastore
- [HIVE-5277](#): HBase handler skips rows with null valued first cells when only row key is selected
- [HIVE-6727](#): Table level stats for external tables are set incorrectly
- [HIVE-7476](#): CTAS does not work properly for s3
- [HIVE-8529](#): HiveSessionImpl#fetchResults should not try to fetch operation log when hive.server2.logging.operation.enabled is false.
- [HIVE-9566](#): HiveServer2 fails to start with NullPointerException
- [HIVE-9625](#): Delegation tokens for HMS are not renewed
- [HIVE-9811](#): Hive on Tez leaks WorkMap objects
- [HIVE-9974](#): Sensitive data redaction: data appears in name of MapReduce job



HDP 2.3.0 provided Hive 1.2.1 and the following Apache patches:

#### INCOMPATIBLE CHANGES

- [HIVE-11118](#) Load data query should validate file formats with destination tables

#### NEW FEATURES

- [HIVE-10233](#) Hive on Tez: memory manager for grace hash join

#### IMPROVEMENTS

- [HIVE-11164](#) WebHCat should log contents of HiveConf on startup [HIVE-11037](#)  
HiveOnTez: make explain user level = true as default

#### BUG FIXES

- [HIVE-10251](#) [HIVE-9664](#) makes Hive depend on ivysettings.xml (using [HIVE-10251](#).simple.patch)
- [HIVE-10996](#) Aggregation / Projection over Multi-Join Inner Query producing incorrect results
- [HIVE-11028](#) Tez: table self join and join with another table fails with IndexOutOfBoundsException
- [HIVE-11048](#) Make test cbo\_windowing robust
- [HIVE-11050](#) testCliDriver\_vector\_outer\_join.\* failures in Unit tests due to unstable data creation queries
- [HIVE-11051](#) Hive 1.2.0 MapJoin w/Tez - LazyBinaryArray cannot be cast to [Ljava.lang.Object;
- [HIVE-11059](#) hcatalog-server-extensions tests scope should depend on hive-exec
- [HIVE-11060](#) Make test windowing.q robust
- [HIVE-11066](#) Ensure tests don't share directories on FS
- [HIVE-11074](#) Update tests for [HIVE-9302](#) after removing binaries
- [HIVE-11076](#) Explicitly set hive.cbo.enable=true for some tests
- [HIVE-11083](#) Make test cbo\_windowing robust
- [HIVE-11104](#) Select operator doesn't propagate constants appearing in expressions
- [HIVE-11147](#) MetaTool doesn't update FS root location for partitions with space in name

## 1.5.9. Kafka

In addition to any Apache patches ported over [from the 2.4.x core](#), HDP 2.3.6 provides Kafka 0.9.0.1 and the following Apache patches:

- [KAFKA-3448](#): Support zone index in IPv6 regex.

HDP 2.3.6 provides Kafka 0.9.0.1 and ports the following Apache patches [from the 2.4.x core](#):

- [KAFKA-2803](#): Add hard bounce system test for Kafka Connect.
- [KAFKA-2812](#): Improve consumer integration tests.
- [KAFKA-2862](#): Fix MirrorMaker's message.handler.args description.
- [KAFKA-2872](#): Unite sink nodes with parent nodes in addSink.
- [KAFKA-2877](#): Handle request timeout in sync group.
- [KAFKA-2878](#): Guard against OutOfMemory in Kafka broker.
- [KAFKA-2879](#): Make MiniKDC test service slightly more generic.
- [KAFKA-2880](#): Consumer should handle disconnect/timeout for metadata requests.
- [KAFKA-2881](#): Improve Consumer Configs and API Documentation.
- [KAFKA-2882](#): Add constructor cache for Snappy and LZ4 Output/Input streams in Compressor.java.
- [KAFKA-2892](#): Consumer Docs Use Wrong Method.
- [KAFKA-2899](#): Improve logging when unexpected exceptions thrown in reading local log.
- [KAFKA-2906](#): Fix Connect javadocs, restrict only to api subproject, and clean up javadoc warnings.
- [KAFKA-2913](#): Missing partition check when removing groups from cache.
- [KAFKA-2942](#): Inadvertent auto-commit when pre-fetching can cause message loss.
- [KAFKA-2950](#): Fix performance regression in the producer.

HDP 2.3.4.7 provided Kafka 0.9.0 with no additional Apache patches.

HDP 2.3.4 provided Kafka 0.9.0 and the following Apache patches:

- [KAFKA-2803](#): Add hard bounce system test for Kafka Connect.
- [KAFKA-2812](#): improve consumer integration tests.
- [KAFKA-2862](#): Fix MirrorMaker's message.handler.args description.
- [KAFKA-2872](#): unite sink nodes with parent nodes in addSink.
- [KAFKA-2877](#): handle request timeout in sync group.
- [KAFKA-2878](#): Guard against OutOfMemory in Kafka broker.
- [KAFKA-2879](#): Make MiniKDC test service slightly more generic.

- [KAFKA-2880](#): consumer should handle disconnect/timeout for metadata requests.
- [KAFKA-2881](#): Improve Consumer Configs and API Documentation.
- [KAFKA-2882](#): Add constructor cache for Snappy and LZ4 Output/Input streams in Compressor.java.
- [KAFKA-2892](#): Consumer Docs Use Wrong Method.
- [KAFKA-2899](#): improve logging when unexpected exceptions thrown in reading local log.
- [KAFKA-2906](#): Fix Connect javadocs, restrict only to api subproject, and clean up javadoc warnings.
- [KAFKA-2913](#): missing partition check when removing groups from cache.
- [KAFKA-2942](#): inadvertent auto-commit when pre-fetching can cause message loss.
- [KAFKA-2950](#): Fix performance regression in the producer.

HDP 2.3.2 provided Kafka 0.8.2, with no additional Apache patches.

HDP 2.3.0 provided Kafka 0.8.2 and the following Apache patches:

- [KAFKA-1005](#) Shutdown consumer at the end of consumer performance test.
- [KAFKA-1416](#) Unify sendMessages in TestUtils
- [KAFKA-1461](#) Replica fetcher thread does not implement any back-off behavior
- [KAFKA-1499](#) trivial follow-up (remove unnecessary parentheses)
- [KAFKA-1501](#) Let the OS choose the port in unit tests to avoid collisions
- [KAFKA-1517](#) Messages is a required argument to Producer Performance Test
- [KAFKA-1546](#) Automate replica lag tuning;
- [KAFKA-1634](#) Bump up Offset Commit Request to v2 to add global retention and remove per-partition commit timestamp
- [KAFKA-1664](#) Kafka does not properly parse multiple ZK nodes with non-root chroot
- [KAFKA-1683](#) add Session concept in SocketServer.PlainTextTransportLayer fixes.
- [KAFKA-1684](#) Kerberos/SASL implementation.
- [KAFKA-1688](#) Adding all public entities for adding a pluggable authorizer to Kafka.
- [KAFKA-1755](#) Reject compressed and unkeyed messages sent to compacted topics
- [KAFKA-1809](#) Refactor brokers to allow listening on multiple ports and IPs
- [KAFKA-1824](#) ConsoleProducer - properties key.separator and parse.key no longer work
- [KAFKA-1845](#) KafkaConfig should use ConfigDef

- [KAFKA-1852](#) Reject offset commits to unknown topics
- [KAFKA-1863](#) Add docs for possible thrown exception in Callback;
- [KAFKA-1865](#) Add a flush() method to the producer.
- [KAFKA-1866](#) LogStartOffset gauge throws exceptions after log.delete()
- [KAFKA-1910](#) Refactor new consumer and fixed a bunch of corner cases / unit tests
- [KAFKA-1926](#) Replace kafka.utils.Utils with o.a.k.common.utils.Utils
- [KAFKA-1961](#) Prevent deletion of \_consumer\_offsets topic
- [KAFKA-1973](#) Remove the accidentally created LogCleanerManager.scala.orig
- [KAFKA-1982](#) (add missing files) change kafka.examples.Producer to use the new java producer
- [KAFKA-1982](#) change kafka.examples.Producer to use the new java producer
- [KAFKA-1986](#) Request failure rate should not include invalid message size and offset out of range
- [KAFKA-1988](#) Fix org.apache.kafka.common.utils.Utils.abs and add Partitioner.toPositive
- [KAFKA-1989](#) New purgatory design; patched by Yasuhiro Matsuda
- [KAFKA-1990](#) Add unlimited time-based log retention
- [KAFKA-1992](#) checkEnoughReplicasReachOffset doesn't need to get requiredAcks
- [KAFKA-1994](#) Evaluate performance effect of chroot check on Topic creation
- [KAFKA-1996](#) Fix scaladoc error.
- [KAFKA-1997](#) Refactor MirrorMaker based on [KIP-3](#);
- [KAFKA-2002](#) Mx4JLoader doesn't disable when kafka\_mx4jenable=false.
- [KAFKA-2009](#) Fix two minor bugs in mirror maker.
- [KAFKA-2013](#) benchmark test for the purgatory
- [KAFKA-2016](#) RollingBounceTest takes long
- [KAFKA-2024](#) Log compaction can generate unindexable segments.
- [KAFKA-2033](#) Small typo in documentation
- [KAFKA-2034](#) sourceCompatibility not set in Kafka build.gradle
- [KAFKA-2039](#) Update Scala to 2.10.5 and 2.11.6
- [KAFKA-2042](#) Update topic list of the metadata regardless of cluster information;
- [KAFKA-2043](#) CompressionType is passed in each RecordAccumulator append

- [KAFKA-2044](#) Support requests and responses from o.a.k.common in KafkaApis
- [KAFKA-2047](#) Move the stream creation into concurrent mirror maker threads
- [KAFKA-2048](#) Change lock synchronized to inLock() for partitionMapCond
- [KAFKA-2050](#) Avoid calling .size() on linked list.
- [KAFKA-2056](#) Fix transient testRangePartitionAssignor failure
- [KAFKA-2088](#) kafka-console-consumer.sh should not create ZooKeeper path when no brokers found and chroot was set in zookeeper.connect.
- [KAFKA-2090](#) Remove duplicate check to metadataFetchInProgress
- [KAFKA-2096](#) Enable keepalive socket option for broker to prevent socket leak
- [KAFKA-2099](#) BrokerEndPoint file, methods and object names should match
- [KAFKA-2104](#) testDuplicateListeners() has a typo
- [KAFKA-2109](#) Support retries in KafkaLog4jAppender
- [KAFKA-2112](#) make overflowWheel volatile
- [KAFKA-2113](#) TestPurgatoryPerformance does not compile using IBM JDK
- [KAFKA-2114](#) Unable to change min.insync.replicas default.
- [KAFKA-2115](#) Error updating metrics in RequestChannel
- [KAFKA-2117](#) Use the correct metadata field for reading offset struct
- [KAFKA-2118](#) Cleaner cannot clean after shutdown during replaceSegments.
- [KAFKA-2119](#) ConsumerRecord key() and value() methods should not have throws Exception
- [KAFKA-2121](#) Close internal modules upon client shutdown
- [KAFKA-2122](#) Remove controller.message.queue.size Config
- [KAFKA-2128](#) kafka.Kafka should return non-zero exit code when caught exception.
- [KAFKA-2131](#) Update new producer javadocs with correct documentation links
- [KAFKA-2138](#) Fix producer to honor retry backoff
- [KAFKA-2140](#) Improve code readability
- [KAFKA-527](#) Compression support does numerous byte copies;

## 1.5.10. Knox

HDP 2.3.6 provides Knox 0.6.0 with no additional Apache patches beyond any ported over from the [2.4.x core](#).

HDP 2.3.6 provides Knox 0.6.0 and ports the following Apache patches [from the 2.4.x core](#):

- [KNOX-647](#): Rename LDAP artifacts from test to demo.
- [KNOX-677](#): Upgrade to latest Groovy.
- [KNOX-695](#): Expose configuration of HttpClient's connection and socket timeout settings.

HDP 2.3.4.7 provided Knox 0.6.0 with no additional Apache patches.

HDP 2.3.4 provided Knox 0.6.0 and the following Apache patches:

- [KNOX-566](#): Make the Default Ephemeral DH Key Size 2048 for TLS.
- [KNOX-579](#): Regex based identity assertion provider with static dictionary lookup.
- [KNOX-581](#): Hive dispatch not propagating effective principal name.
- [KNOX-633](#): Upgrade Apache commons-collections.

HDP 2.3.2 provided Knox 0.6.0 and the following Apache patches:

- [KNOX-598](#): Concurrent JDBC clients via KNOX to Kerberized HiveServer2 causes HTTP 401 error (due to Kerberos Replay attack error)
- [KNOX-599](#): Template with {\*\*} in queries are expanded with =null for query params without a value.

HDP 2.3.0 provided Knox 0.6.0 and the following Apache patches:

#### BUG FIXES

- [KNOX-476](#) implementation for X-Forwarded-\* headers support and population
- [KNOX-546](#) Consuming intermediate response during Kerberos request dispatching
- [KNOX-550](#) reverting back to original Hive Kerberos dispatch behavior
- [KNOX-559](#) renaming service definition files

#### IMPROVEMENTS

- [KNOX-545](#) Simplify Keystore Management for Cluster Scaleout
- [KNOX-561](#) Allow Knox pid directory to be configured via the Knox-env.sh file

## 1.5.11. Mahout

In HDP-2.3.x and 2.3.x, instead of shipping a specific Apache release of Mahout, we synchronized to a particular revision point on Apache Mahout trunk. This revision point is after the 0.9.0 release, but before the 0.10.0 release. This provides a large number of bug fixes and functional enhancements over the 0.9.0 release, but provides a stable release of the Mahout functionality before the complete conversion to new Spark-based Mahout in 0.10.0.

The revision point chosen for Mahout in HDP 2.3.x and 2.3.x is from the "mahout-0.10.x" branch of Apache Mahout, as of 19 December 2014, revision 0f037cb03e77c096 in GitHub.

In addition, we have provided the following patches:

- [MAHOUT-1493](#) Port Naive Bayes to Scala DSL.
- [MAHOUT-1589](#) mahout.cmd has duplicated content.

## 1.5.12. Oozie

In addition to any Apache patches ported over [from the 2.4.x core](#), HDP 2.3.6 provides Oozie 4.2.0 and the following Apache patches:

- [OOZIE-2492](#) JSON security issue in js code.

HDP 2.3.6 provides Oozie 4.2.0 and ports no additional Apache patches [from the 2.4.x core](#).

HDP 2.3.4.7 provided Oozie 4.2.0 with no additional Apache patches.

HDP 2.3.4 provided Oozie 4.2.0 with no additional Apache patches.

HDP 2.3.2 provided Oozie 4.2.0 with no additional Apache patches.

HDP 2.3.0 provided Oozie 4.2.0 and the following Apache patches:

- [OOZIE-2289](#) hive-jdbc dependency in core/pom.xml should be compile
- [OOZIE-2290](#) Oozie db version update should happen after all DDL tweak
- [OOZIE-2291](#) Hive2 workflow.xml.security should have "cred" in action tag instead of "hive2" tag

## 1.5.13. Phoenix

In addition to any Apache patches ported over [from the 2.4.x core](#), HDP 2.3.6 provides Phoenix 4.4.0-HBase-1.1 and the following Apache patches:

- [PHOENIX-2295](#): NullPointerException when parsing a query with hints in UNION ALL.
- [PHOENIX-2667](#): Race condition between IndexBuilder and Split for region lock.

HDP 2.3.6 provides Phoenix 4.4.0-HBase-1.1 and ports the following Apache patches [from the 2.4.x core](#):

- [PHOENIX-2601](#): Query result is incorrect when both index hint and limit are used.
- [PHOENIX-2608](#): Incompatibility between Jackson1 version shipped with Phoenix, YARN.
- [PHOENIX-2817](#): Phoenix-Spark plugin doesn't work in secured environment.

HDP 2.3.4.7 provided Phoenix 4.4.0-HBase-1.1 with no additional Apache patches.

HDP 2.3.4 provided Phoenix 4.4.0-HBase-1.1 and the following Apache patches:

- [PHOENIX-1277](#): CSVCommonsLoader not allowing null CHAR values (non PK).
- [PHOENIX-1734](#): Local index improvements.
- [PHOENIX-1975](#): Detect and use HBASE\_HOME when set.
- [PHOENIX-2014](#): WHERE search condition ignored when also using row value constructor in view.
- [PHOENIX-2027](#): Subqueries with no data are raising IllegalStateException.
- [PHOENIX-2029](#): Queries are making two rpc calls for getTable.
- [PHOENIX-2030](#): CsvBulkLoadTool should use logical name of the table for output directory suffix.
- [PHOENIX-2040](#): Mark spark/scala dependencies as 'provided'.
- [PHOENIX-2059](#): MR index build does not handle table with a schema name correctly.
- [PHOENIX-2087](#): Ensure predictable column position during alter table.
- [PHOENIX-2118](#): Remove/modify usages of Guava StopWatch and deprecated ComparisonChain methods.
- [PHOENIX-2125](#): ORDER BY on full PK on salted table does not work.
- [PHOENIX-2139](#): LIKE '%' is not filtering out null columns.
- [PHOENIX-2141](#): ComparisonExpression should return Boolean null if either operand is null.
- [PHOENIX-2149](#): MAX Value of Sequences not honored when closing Connection between calls to NEXT VALUE FOR.
- [PHOENIX-2205](#): Group by a divided value (e.g., time/10) returns NULL.
- [PHOENIX-2207](#): Load scanner caches in parallel when using stats and round robin iterator.
- [PHOENIX-2238](#): Bulk dataload Hive data with \001 delimiter.
- [PHOENIX-2274](#): Sort-merge join could not optimize out the sort on the right table.
- [PHOENIX-2313](#): TypeMismatchException thrown while querying a table that has an index with a Boolean (BUG-46167).
- [PHOENIX-2359](#): Configuration for PQS to use Protobuf serialization instead of JSON.
- [PHOENIX-2369](#): Some UDF tests failed due to not finding Hadoop classpath.
- [PHOENIX-2372](#): Check for null before passing value to Calendar.
- [PHOENIX-2448](#): Fix quoting for sqlline-thin.py on Windows.
- [PHOENIX-2449](#): QueryServer needs Hadoop configuration on classpath with Kerberos.



- [PHOENIX-2477](#): ClassCastException in IndexedWALEditCodec after HBASE-14501.

HDP 2.3.2 provided Phoenix 4.4.0-HBase-1.1 and the following Apache patches:

- [PHOENIX-1659](#) PHOENIXDatabaseMetaData.getColumns does not return REMARKS column.
- [PHOENIX-1978](#) UDF ArgumentTypeMismatchException.
- [PHOENIX-2011](#) Default, min, and max values should not require quotes around it in create function.
- [PHOENIX-2022](#) BaseRegionScanner.next should be abstract.
- [PHOENIX-2066](#) Existing client fails initialization due to upgrade attempting to create column with no name.
- [PHOENIX-2073](#) Two bytes character in LIKE expression is not allowed.
- [PHOENIX-2074](#) StackOverflowError with RoundRobinResultIterator.
- [PHOENIX-2075](#) MR integration uses single mapper unless table is salted.
- [PHOENIX-2096](#) Tweak criteria for when round robin iterator is used.
- [PHOENIX-2131](#) CastParseNode.toSQL omits closing parenthesis.
- [PHOENIX-2151](#) Two different UDFs called on same column return values from first UDF only.
- [PHOENIX-2254](#) zookeeper.znode.parent value is not taking affect in MR integration job.

HDP 2.3.0 provided Phoenix 4.4.0-HBase-1.1 and the following Apache patches:

- [PHOENIX-1395](#) ResultSpooler spill files are left behind in /tmp folder.
- [PHOENIX-1976](#) Exit gracefully if addShutdownHook fails.
- [PHOENIX-1980](#) CsvBulkLoad cannot load hbase-site.xml from classpath
- [PHOENIX-1995](#) client uberjar doesn't support dfs
- [PHOENIX-1996](#) Use ByteString instead of ZeroCopyByteString
- [PHOENIX-2005](#) Connection utilities omit zk client port, parent znode
- [PHOENIX-2005](#) Connection utilities omit zk client port, parent znode (addendum)
- [PHOENIX-2007](#) java.sql.SQLException: Encountered exception in sub plan [0] execution'
- [PHOENIX-2010](#) Properly validate number of arguments passed to the functions in FunctionParseNode#validate
- [PHOENIX-2012](#) RowKeyComparisonFilter logs unencoded data at DEBUG level
- [PHOENIX-2013](#) Apply [PHOENIX-1995](#) to runnable uberjar as well

- [PHOENIX-2027](#) Queries with Hints are raising `IllegalStateException`
- [PHOENIX-2031](#) Unable to process timestamp/Date data loaded via PHOENIX `org.apache.PHOENIX.pig.PHOENIXHBaseLoader`.
- [PHOENIX-2032](#) `psql.py` is broken after [PHOENIX-2013](#)
- [PHOENIX-2033](#) PQS log environment details on launch
- [PHOENIX-2063](#) Row value constructor doesn't work when used in COUNT.
- [PHOENIX-2181](#) HPOENIXHBaseLoader doesn't work with salted tables.
- [PHOENIX-914](#) Native HBase timestamp support to optimize date range queries in Phoenix

## 1.5.14. Pig

In addition to any Apache patches ported over [from the 2.4.x core](#), HDP 2.3.6 provides Pig 0.15.0 and the following Apache patches:

- [PIG-4873](#): `InputSplit.getLocations` return null and result a NPE in Pig.

HDP 2.3.6 provides Pig 0.15.0 and ports the following Apache patches [from the 2.4.x core](#):

- [PIG-4690](#): Union with self replicate join will fail in Tez.
- [PIG-4760](#): `TezDAGStats.convertToHadoopCounters` is not used, but impose MR counter limit.
- [PIG-4790](#): Join after union fail due to `UnionOptimizer`.
- [PIG-4814](#): `AvroStorage` does not take NameNode HA as part of schema file URL.
- [PIG-4816](#): Read a null scalar causing a Tez failure.

HDP 2.3.4.7 provided Pig 0.15.0 with no additional Apache patches.

HDP 2.3.4 provided Pig 0.15.0 and the following Apache patches:

- [PIG-4635](#): NPE while running Pig script in Tez mode( Pig 0.15 with Tez 0.7).
- [PIG-4683](#): Nested order is broken after [PIG-3591](#) in some cases.
- [PIG-4688](#): Limit followed by `POPartialAgg` can give empty or partial results in Tez.
- [PIG-4696](#): Empty map returned by a `streaming_python` udf wrongly contains a null key.
- [PIG-4703](#): `TezOperator.stores` shall not ship to backend.
- [PIG-4707](#): [Pig on Tez] Streaming job hangs with `pig.exec.mapPartAgg=true`.
- [PIG-4712](#): [Pig on Tez] NPE in Bloom UDF after Union.
- [PIG-4714](#): Improve logging across multiple components with `callerId`.

- [PIG-4744](#): Honor tez.staging-dir setting in tez-site.xml.

HDP 2.3.2 provided Pig 0.15.0 and the following Apache patches:

- [PIG-4627](#) [Pig on Tez] Self join does not handle null values correctly
- [PIG-4628](#) Pig 0.14 job with order by fails in MapReduce mode with Oozie
- [PIG-4649](#) [Pig on Tez] Union followed by HCatStorer misses some data
- [PIG-4679](#) Performance degradation due to InputSizeReducerEstimator since [PIG-3754](#)

HDP 2.3.0 provided Pig 0.15.0 and the following Apache patch:

- [PIG-4624](#) Error on ORC empty file without schema

## 1.5.15. Ranger

In addition to any Apache patches ported over [from the 2.4.x core](#), HDP 2.3.6 provides Ranger 0.5.2 and the following Apache patches:

- [RANGER-204](#): Not able to delete user or group if user/group has any policy defined..
- [RANGER-205](#): Delete rest api of User not deleting user completely from system.
- [RANGER-617](#): Handle Search by status on policy search filter.
- [RANGER-843](#): Add indexes to improve Postgres query performance.
- [RANGER-875](#): Restrict Grantor privileges of Ranger db user for Oracle DB Flavour.
- [RANGER-888](#): Provide support to delete Users and Groups from Ranger Admin UI.
- [RANGER-891](#): Audit shutdown hook to be registered with Hadoop ShutdownHookManager, instead of directly with Java Runtime.
- [RANGER-893](#): Ranger ugsync with LDAP is not able to fetch group information when member attribute contains user's short name instead of full DN.
- [RANGER-921](#): Improve implementation of internal SQL calls and make it more generic.
- [RANGER-926](#): Ranger UI validation changes to allow . [dot] in firstname and lastname fields.
- [RANGER-934](#): Ranger should use released version of EclipseLink (EclipseLink 2.6.2).

HDP 2.3.6 provides Ranger 0.5.2 and ports the following Apache patches [from the 2.4.x core](#):

- [RANGER-173](#): Utility scripts to create HDFS audit folders and policies.
- [RANGER-218](#): LDAP Groups incorrectly labelled internal.
- [RANGER-725](#): Add the right .gitignore file to the newly added projects so that directory listing is clean after a build.

- [RANGER-746](#): Addressing suggestions from Review - Add wildcard, multiple CN & SAN support when validating plugins' SSL certs.
- [RANGER-767](#): Refactor UserGroupSink implementation and consolidate performance improvements.
- [RANGER-771](#): Fix 4+ Log entries upon login in X\_AUTH\_SESS.
- [RANGER-772](#): Hive plugin Update Ranger authorizer to mimic changes made by Hive standard authorizer for the case when IMPORT can end up creating a table.
- [RANGER-773](#): Fix newly found Coverity scan issues for Ranger KMS.
- [RANGER-777](#): Remove the maven profile that was added for building Kafka plugin.
- [RANGER-778](#): Fix user update issue.
- [RANGER-789](#): Fix incorrect policy list paging for non-admin users.
- [RANGER-794](#): Ranger policy engine performance measurement.
- [RANGER-798](#): Handle different timezone issue while saving audit logs to Solr.
- [RANGER-799](#): Ranger UI fixes - partial search not working on Policy listing page.
- [RANGER-804](#): Delete groups associated with User causes Exception in UserSync.
- [RANGER-809](#): Audit framework need to cache the getHostName() values to reuse for successive calls.
- [RANGER-809](#): Audit framework need to cache the getHostName() values to reuse for successive calls.
- [RANGER-821](#): Ranger shutdown hook should not only do its processing asynchronously but also terminate itself if it runs more than a configurable amount of time.
- [RANGER-825](#): groupId in Ranger s child poms should be consistent and follow convention.
- [RANGER-831](#): policy version incremented twice when resources are updated.
- [RANGER-833](#): In Ranger UI add support for usernames containing a plus + symbol.
- [RANGER-834](#): Correct the excludes flag's treatment when resource value denotes everything.
- [RANGER-836](#): Optimize policy retrieval.
- [RANGER-844](#): Optimize policy retrieval for non-admin users.
- [RANGER-848](#): Policy Listing page fix: users column is empty for non-admin users.
- [RANGER-857](#): Unify (and update) Tomcat versions.
- [RANGER-863](#): Make parameters like maxHttpHeaderSize configurable for EmbeddedServer.

- [RANGER-882](#): Scrub received policies before policy engine uses it to guard against inadvertent data corruption: remove null policy resource values.
- [RANGER-889](#): Policy engine API to find list of users/groups having access to a resource.
- [RANGER-894](#): Fixing few issues in the ldap tool when user search base or group search base is configured as part of input.

HDP 2.3.4.7 provided Ranger 0.5.0 and the following Apache patch:

- [RANGER-834](#): Modify Exclude logic evaluation to handle wildcard characters.

HDP 2.3.4 provided Ranger 0.5.0 and the following Apache patches:

- [RANGER-246](#): Need to update the current implementation for recent changes in Kafka.
- [RANGER-526](#): Provide REST API to change user role.
- [RANGER-586](#): Ranger plugins should not add dependent libraries to component's CLASSPATH.
- [RANGER-590](#): Escape spaces in the user and group names which are part of rest call URI in UserSync process.
- [RANGER-602](#): Solr client in SolrCloud mode should work with ZooKeeper settings also.
- [RANGER-607](#): Unable to create multiple policyItems for same user or group.
- [RANGER-608](#): Denied access to list a directory does not generate audit.
- [RANGER-652](#): LDAP configuration tool.
- [RANGER-656](#): Ranger UI - KMS Need to handle 404 error when clicked on breadcrumb.
- [RANGER-658](#): Package ranger\_credential\_helper.py with Ranger Usersync assembly.
- [RANGER-661](#): Plugin receives empty policy list though the service has policies.
- [RANGER-663](#): Race condition during policy update causes policy to get in an bad state.
- [RANGER-664](#): Ranger PolicyRefresh REST Client timeout parameter should be configurable.
- [RANGER-665](#): ranger.ldap.ad.referral property is not getting updated in RANGER-admin-site.xml.
- [RANGER-666](#): Ranger to support Azure SQL Database.
- [RANGER-671](#): Add support to retrieve permissions for the logged in user from UserSession rather going to database every time.
- [RANGER-673](#): Setup changes to allow Ranger service to installed using custom service user.
- [RANGER-674](#): Ranger public rest api gives 200 response for wrong credential instead of 401.

- [RANGER-677](#): Ranger Admin fails to render policies referring to groups that contain "." in name.
- [RANGER-680](#): Remove public group by default in default policy for KMS repo.
- [RANGER-681](#): Update default sync intervals for LDAP and UNIX.
- [RANGER-682](#): Ranger to support Azure Blob Datastore as an audit destination via HDFS audit handler.
- [RANGER-684](#): Ranger Usersync - Add Ability to transform user/group names.
- [RANGER-687](#): after each 30 seconds audit is getting updated in plugin tab.
- [RANGER-688](#): Handle scenario where ids of XUser and XPortalUser are not in sync.
- [RANGER-697](#): KeyAdmin role user should see only KMS related audit access logs in Audit tab.
- [RANGER-700](#): Provide a wrapper shell script to run the FileSourceUserGroupBuilder process.
- [RANGER-701](#): Update setup scripts to allow special characters in passwords.
- [RANGER-702](#): Optimize policy download performance.
- [RANGER-705](#): Ranger Usersync should provide summary logs on the sync progress instead of not logging any details after 2000 users.
- [RANGER-706](#): Optimize audit db upgrade patches to minimize timeout issues.
- [RANGER-712](#): Create a new project which can serve as a template to write Ranger extensions.
- [RANGER-713](#): Knox-plugin failed to enable after plugin modification for not to add dependent libraries to component's CLASSPATH.
- [RANGER-714](#): Enhancements to the db admin setup scripts.
- [RANGER-715](#): Fix issues reported by coverity test in Ranger Plugin ClassLoader.
- [RANGER-717](#): Hive and HBase Ranger plugin Audit to DB failed to log after plugin modification for not to add dependent libraries to component's CLASSPATH.
- [RANGER-720](#): Ldap discovery tool doesn't seem to be working as expected.
- [RANGER-724](#): AuditBatchQueue: prevQueueSize not recomputed after initial assignment - static code analyzer flagged issue.
- [RANGER-725](#): Add the right .gitignore file to the newly projects so that directory listing is clean after a build.
- [RANGER-727](#): Knox Plugin failed to AuditToSpool file when Audit Destination is down.
- [RANGER-731](#): Ranger plugin for YARN doesn't seem to be able to write audit to Kerberized HDFS.

- [RANGER-733](#): Implement best coding practices to resolve issues found during code scan.
- [RANGER-739](#): Ranger HBase Plugin returning null for RegionObserver.preCompact calls causing HBase:ACL issue.
- [RANGER-740](#): Kafka Authorizer interface has added close() method. Ranger should also implement it.
- [RANGER-741](#): Fix installation script to skip Audit DB password check if audit source is SOLR.
- [RANGER-742](#): Ranger usersync fails after syncing 500 users from AD or ldap server when paged results is enabled.
- [RANGER-743](#): External users with Admin Role should be allowed to create/update users.
- [RANGER-744](#): Kafka Authorizer has updated how IP/Host is passed.
- [RANGER-745](#): Upgrade Apache commons-collections.
- [RANGER-747](#): RangerAdmin is considering "none" as valid ZK Host Name for Solr.
- [RANGER-748](#): Users in policy got changed after upgrade.
- [RANGER-749](#): Ranger KMS to support multiple KMS instances with keys across multiple clusters.
- [RANGER-754](#): Ranger YARN Plugin lookup and test connection should support SPENGO enabled HTTP Authentication.
- [RANGER-755](#): ldap run.sh script fails since auth directory does not exist.
- [RANGER-756](#): LdapTool fails with -r option to retrieve only users/group/all.
- [RANGER-757](#): [LDAP tool] authentication fails if use -d option to search only users.
- [RANGER-758](#): Handle special characters in passwords starting from -r.
- [RANGER-761](#): Transaction logs not getting generated under audit menu admin tab if policy name is changed.
- [RANGER-766](#): Yarn Plugin Config hadoop.security.authentication should be non-mandatory with default value.
- [RANGER-767](#): Refactor UserGroupSink implementation and consolidate performance improvements.

HDP 2.3.2 provided Ranger 0.5.0 and the following Apache patches:

- [RANGER-551](#) Policy Validation: If resource levels are not valid for any hierarchy then checks about missing mandatory levels should be skipped

#### BUG FIXES

- [RANGER-560](#) Policy validation: Provide user friendly error messages about validation failures

- [RANGER-580](#) HBase plugin: Plugin may not work after upgrade
- [RANGER-584](#) Service validation: Provide user friendly error messages about validation failures
- [RANGER-587](#) ranger-admin-site.xml not getting updated when ranger.authentication.method is changed
- [RANGER-588](#) Take care of Ranger KMS installation even if 'java' is not in PATH
- [RANGER-593](#) Service def validation: Provide user friendly error messages about validation failures
- [RANGER-594](#) Policy Validation: Change the logic to generate friendly error messages to be like used for Service and Service def
- [RANGER-598](#) Update Ranger config migration script to work with Ranger 0.5
- [RANGER-615](#) Audit to db: Truncate all string values of audit record so that writing of audit does not fail
- [RANGER-618](#) KMS gets slower in key creation once Database grows
- [RANGER-621](#) Solr service-def JSON has incorrect impliedGrants for solr\_admin permission
- [RANGER-622](#) Hive plugin: Add jar via Beeline throws NPE
- [RANGER-623](#) Enable plugin scripts should handle file permissions for certain umask value
- [RANGER-624](#) Windows installation broken after SQLAnywhere support
- [RANGER-625](#) Change db flavor input parameter value from SQLAnywhere to SQLA
- [RANGER-627](#) Processing done by Audit Shutdown hooks can confuse someone looking at logs to think that shutdown of a service is held up due to Ranger plugin
- [RANGER-628](#) Make filters for ranger-admin search binds configurable
- [RANGER-630](#) Data consistency across API and UI
- [RANGER-632](#) Policy validation error messages produced by the server are not seen by the user
- [RANGER-637](#) Make REFERRAL property in Ranger User sync configurable
- [RANGER-638](#) Ranger admin should redirect back to login page when session cookies expires
- [RANGER-639](#) Storm plugin - commons-lang is a required dependency and hence should be packaged as part of Storm plugin
- [RANGER-641](#) Ranger kms start fails if java is not set and started using service keyword
- [RANGER-642](#) Update USERSEARCHFILTER for Ranger Authentication on Windows



- [RANGER-653](#) Move delegated admin check to mgr layer from service layer for XPermMap and XAuditMap

HDP 2.3.0 provided Ranger 0.5.0 and the following Apache patches:

- [RANGER-422](#) Add additional database columns to support aggregation
- [RANGER-423](#) Support audit log aggregation in Ranger Admin UI
- [RANGER-513](#) Policy validation: resource hierarchies check does not work with single-node hierarchies as in HDFS
- [RANGER-551](#) Policy Validation: If resource levels are not valid for any hierarchy then checks about missing mandatory levels should be skipped.
- [RANGER-564](#) Add incubating to the release name

#### BUG FIXES

- [RANGER-219](#) Autocomplete behavior of Hive tables/columns
- [RANGER-524](#) HBase plugin: list command should prune the tables returned on user permissions
- [RANGER-529](#) Policy Validation: resources of a policy must match one of the resource hierarchies of the service def.
- [RANGER-533](#) HBase plugin: if user does not have family-level access to any family in a table then user may be incorrectly denied access done at table/family level during get or scan
- [RANGER-539](#) Rolling downgrade changes
- [RANGER-545](#) Fix js error for lower versions of FF (less than 30)
- [RANGER-548](#) Key rollover command fails
- [RANGER-550](#) Hive plugin: Add audit logging support for metadata queries that have filtering support from Hive.
- [RANGER-553](#) Default policy creation during service creation should handle service defs with multiple hierarchies, e.g. Hive properly
- [RANGER-554](#) Ranger KMS keys listing page does not support pagination
- [RANGER-555](#) Policy view page (from access audit page) gives 404 with Oracle DB
- [RANGER-558](#) HBase plugin: unless user has READ access at some level under the table/family being accessed (via scan/get) authorizer should throw an exception and audit
- [RANGER-565](#) Ranger Admin install fails (sometimes) with IO Error when DB used in Oracle
- [RANGER-566](#) Installation of Ranger on Oracle 12c with shared database needs to use private synonym instead of public synonym

- [RANGER-569](#) Enabling Ranger plugin for HBase should not modify hbase.rpc.protection value
- [RANGER-570](#) Knox plugin: after upgrading Ranger from 0.4 to 0.5 the Knox plugin won't work because classes with old names are missing
- [RANGER-571](#) Storm plugin: after upgrading Ranger from 0.4 to 0.5 the plugin won't work because classes with old names are missing
- [RANGER-575](#) Allow KMS policies to be assigned to all users
- [RANGER-576](#) Storm audit not showing access type in the Ranger Admin Audit UI

#### HDP CHANGES

- [RANGER-450](#) Failed to install Ranger component due to Ranger policyManager script failures

## 1.5.16. Slider

HDP 2.3.6 provides Slider 0.80.0 with no additional Apache patches beyond any ported over [from the 2.4.x core](#).

HDP 2.3.6 provides Slider 0.80.0 and ports the following Apache patches [from the 2.4.x core](#):

- [SLIDER-1079](#) Cache metaInfo object in AgentClientProvider.
- [SLIDER-1088](#) Slider client getting UnknownHost exception trying to RPC to AM in different subdomain.
- [SLIDER-657](#) Introduce `-force` switch for Slider destroy command.

HDP 2.3.4.7 provided Slider 0.80.0 with no additional Apache patches.

HDP 2.3.4 provided Slider 0.80.0 and the following Apache patches:

- [SLIDER-128](#) Support graceful stop of component instances.
- [SLIDER-777](#) Provide Slider dependencies as a self contained versioned tarball.
- [SLIDER-916](#) Fix Docker related bugs.

HDP 2.3.2 provided Slider 0.80.0 and the following Apache patches:

#### IMPROVEMENTS

- [SLIDER-812](#) Make component configurations in appConfig available on the SliderAgent side.

#### BUG FIXES

- [SLIDER-481](#) giving registry log messages meaningful text.
- [SLIDER-911](#) remove surplus jax rs jsr311-api JAR.

- [SLIDER-912](#) x-insecure rest API should be off by default.
- [SLIDER-923](#) switch to TryOnceThenFail retry policy on IPC (needed for Hadoop 2.8+).
- [SLIDER-931](#) Security permissions on set up ZK path are too lax.
- [SLIDER-941](#) Add JAAS config templates for HBase.

HDP 2.3.0 provided Slider 0.80.0 and the following Apache patches:

#### IMPROVEMENTS

- [SLIDER-812](#) Making component configurations in appConfig available on the SliderAgent side
- [SLIDER-891](#) Add ability to set Slider AM launch environment during cluster create/start

#### BUG FIXES

- [SLIDER-810](#) YARN config changes to enable partial logs upload for long running services (default include/exclude patterns does not upload any files)
- [SLIDER-877](#) move SLIDER\_HOME assignment to slider.py
- [SLIDER-878](#) Slider cannot support JDK 1.8 for command Slider registry --getconf hbase-site --name hb1
- [SLIDER-888](#) intermittent errors when accessing key store password during localization of cert stores
- [SLIDER-901](#) AgentClientProvider should use File.separator in paths for platform independence
- [SLIDER-902](#) add config to client cert gen command
- [SLIDER-904](#) Resource leak reported by coverity scan results
- [SLIDER-905](#) Container request fails when Slider requests container with node label and host constraints

## 1.5.17. Spark

In addition to any Apache patches ported over [from the 2.4.x core](#), HDP 2.3.6 provides Spark 1.5.2 and the following Apache patches:

- [SPARK-11182](#): HDFS Delegation Token will be expired when calling `UserGroupInformation.getCurrentUser.addCredentials` in HA mode.
- [SPARK-13599](#): Groovy-all ends up in spark-assembly if Hive profile set.

HDP 2.3.6 provides Spark 1.5.2 and ports no Apache patches [from the 2.4.x core](#).

HDP 2.3.4.7 provided Spark 1.5.2 with no additional Apache patches included.

HDP 2.3.4 provided Spark 1.5.2 and the following Apache patches:

- [SPARK-10058](#): CORE, TESTS, Fix the flaky tests in HeartbeatReceiverSuite.
- [SPARK-10389](#): SQL, [1.5 support order by non-attribute grouping expression on Aggregate.
- [SPARK-10515](#): When killing executor, the pending replacement executors should not be lost.
- [SPARK-10534](#): SQL, ORDER BY clause allows only columns that are present in the select projection list.
- [SPARK-10577](#): PYSARK, DataFrame hint for broadcast join.
- [SPARK-10581](#): DOCS, Groups are not resolved in scaladoc in SQL classes.
- [SPARK-10619](#): Can't sort columns on Executor Page.
- [SPARK-10741](#): SQL, Hive Query Having/OrderBy against Parquet table is not working
- [SPARK-10790](#): YARN, Fix initial executor number not set issue and consolidate the codes.
- [SPARK-10812](#): YARN, Fix shutdown of token renewer..
- [SPARK-10812](#): YARN, Spark Hadoop util support switching to YARN.
- [SPARK-10825](#): CORE, TESTS, Fix race conditions in StandaloneDynamicAllocationSuite.
- [SPARK-10829](#): SQL, Fix 2 bugs for filter on partitioned columns.
- [SPARK-10833](#): BUILD, Inline, organize BSD/MIT licenses in LICENSE.
- [SPARK-10845](#): SQL, Makes spark.sql.hive.version a SQLConfEntry.
- [SPARK-10858](#): YARN: archives/jar/files rename with # doesn't work unl.
- [SPARK-10859](#): SQL, fix stats of StringType in columnar cache.
- [SPARK-10871](#): include number of executor failures in error msg.
- [SPARK-10885](#): STREAMING, Display the failed output op in Streaming UI.
- [SPARK-10889](#): STREAMING, Bump KCL to add MillisBehindLatest metric.
- [SPARK-10901](#): YARN, spark.yarn.user.classpath.first doesn't work.
- [SPARK-10904](#): SPARKR, Fix to support `select(df, c("col1", "col2"))`.
- [SPARK-10914](#): UnsafeRow serialization breaks when two machines have different Oops size..
- [SPARK-10932](#): PROJECT INFRA, Port two minor changes to release-build.sh from scripts' old repo.
- [SPARK-10934](#): SQL, handle hashCode of unsafe array correctly.
- [SPARK-10952](#): Only add Hive to classpath if HIVE\_HOME is set..

- [SPARK-10955](#): STREAMING, Add a warning if dynamic allocation for Streaming applications.
- [SPARK-10959](#): PYSPARK, StreamingLogisticRegressionWithSGD does not t...
- [SPARK-10959](#): PYSPARK, StreamingLogisticRegressionWithSGD does not train with given regParam and convergenceTol parameters.
- [SPARK-10960](#): SQL, SQL with windowing function should be able to refer column in inner select.
- [SPARK-10971](#): SPARKR, RRunner should allow setting path to Rscript..
- [SPARK-10973](#): ML, PYTHON, Fix IndexError exception on SparseVector when asked for index after the last non-zero entry.
- [SPARK-10980](#): SQL, fix bug in create Decimal.
- [SPARK-10981](#): SPARKR, SparkR Join improvements.
- [SPARK-11009](#): SQL, fix wrong result of Window function in cluster mode.
- [SPARK-11023](#): YARN, Avoid creating URIs from local paths directly..
- [SPARK-11026](#): YARN, spark.yarn.user.classpath.first does work for 'SPARK-submit -jars hdfs://user/foo.jar'.
- [SPARK-11032](#): SQL, correctly handle having.
- [SPARK-11039](#): DOCS, WEBUI, Document additional UI configurations.
- [SPARK-11047](#): Internal accumulators miss the internal flag when replaying events in the history server.
- [SPARK-11051](#): CORE, Do not allow local checkpointing after the RDD is materialized and checkpointed.
- [SPARK-11056](#): Improve documentation of SBT build..
- [SPARK-11063](#): STREAMING, Change preferredLocations of Receiver's RDD to hosts rather than hostports.
- [SPARK-11066](#): Update DAGScheduler's "misbehaved ResultHandler".
- [SPARK-11094](#): Strip extra strings from Java version in test runner.
- [SPARK-11103](#): SQL, Filter applied on Merged Parquet schema with new column fail.
- [SPARK-11104](#): STREAMING, Fix a deadlock in StreamingContext.stop.
- [SPARK-11126](#): SQL, Fix a memory leak in SQLListener.\_stageIdToStageMetrics.
- [SPARK-11126](#): SQL, Fix the potential flaky test.
- [SPARK-11135](#): SQL, Exchange incorrectly skips sorts when existing ordering is non-empty subset of required ordering.

- [SPARK-11153](#): SQL, Disables Parquet filter push-down for string and binary columns.
- [SPARK-11188](#): SQL, Elide stacktraces in bin/SPARK-sql for AnalysisExceptions.
- [SPARK-11233](#): SQL, register cosh in function registry.
- [SPARK-11244](#): SPARKR, sparkR.stop() should remove SQLContext.
- [SPARK-11246](#): SQL, Table cache for Parquet broken in 1.5.
- [SPARK-11251](#): Fix page size calculation in local mode.
- [SPARK-11264](#): bin/SPARK-class can't find assembly jars with certain GREP\_OPTIONS set.
- [SPARK-11270](#): STREAMING, Add improved equality testing for TopicAndPartition from the Kafka Streaming API.
- [SPARK-11287](#): Fixed class name to properly start TestExecutor from deploy.client.TestClient.
- [SPARK-11294](#): SPARKR, Improve R doc for read.df, write.df, saveAsTable.
- [SPARK-11299](#): DOC, Fix link to Scala DataFrame Functions reference.
- [SPARK-11302](#): MLLIB, 2) Multivariate Gaussian Model with Covariance matrix returns incorrect answer in some cases.
- [SPARK-11303](#): SQL, filter should not be pushed down into sample.
- [SPARK-11417](#): SQL, no @Override in codegen.
- [SPARK-11424](#): Guard against double-close() of RecordReaders.
- [SPARK-11434](#): SQL, Fix test "Filter applied on merged Parquet schema with new column fails".
- [SPARK-5966](#): WIP, SPARK-submit deploy-mode cluster is not compatible with master local>.
- [SPARK-8386](#): SQL, add write.mode for insertIntoJDBC when the parameter overwrite is false.

HDP 2.3.2 provided Spark 1.4.1 and the following Apache patches:

#### NEW FEATURES

- [SPARK-1537](#) Add integration with Yarn's Application Timeline Server.
- [SPARK-6112](#) Provide external block store support through HDFS RAM\_DISK.

#### BUG FIXES

- [SPARK-10623](#) NoSuchElementException thrown when ORC predicate push-down is turned on.

HDP 2.3.0 provided Spark 1.3.1 and the following Apache patches:

## IMPROVEMENTS

- [SPARK-7326](#) (Backport) Performing window() on a WindowedDStream doesn't work all the time JDK 1.7 repackaging

## 1.5.18. Sqoop

In addition to any Apache patches ported over [from the 2.4.x core](#), HDP 2.3.6 provides Spark 1.5.2 and the following Apache patches:

- [SQOOP-2737](#) Cannot import table from Oracle with column with spaces in name.
- [SQOOP-2779](#) Sqoop Metastore doesn't seem to recognize --schema option.

HDP 2.3.6 provides Sqoop 1.4.6 and ports no additional Apache patches [from the 2.4.x core](#).

HDP 2.3.4.7 provided Oozie 4.2.0 with no additional Apache patches.

HDP 2.3.4 provided Sqoop 1.4.6 and the following Apache patches:

- [SQOOP-2399](#) Sqoop import into HBase - BigDecimalSplitter java.lang.ArrayIndexOutOfBoundsException.

HDP 2.3.2 provided Sqoop 1.4.6 and the following Apache patches:

## IMPROVEMENTS

- [SQOOP-2387](#) Sqoop should support importing from table with column names containing some special character.
- [SQOOP-2457](#) Add option to automatically compute statistics after loading data into a Hive table.

HDP 2.3.0 provided Sqoop 1.4.6 and the following Apache patches:

## IMPROVEMENTS

- [SQOOP-2370](#) Netezza - need to support additional options for full control character handling.

## BUG FIXES

- [SQOOP-2326](#) Fix Netezza trunc-string option handling and unnecessary log directory during imports.

## 1.5.19. Storm

In addition to any Apache patches ported over [from the 2.4.x core](#), HDP 2.3.6 provides Storm 0.10.0-beta and the following Apache patches:

- [STORM-1689](#): Set request headersize for logviewer.
- [STORM-969](#): HDFS Bolt can end up in an unrecoverable state.

HDP 2.3.6 provides Storm 0.10.0-beta and ports the following Apache patches [from the 2.4.x core](#):

- [STORM-1001](#): Undefined STORM\_EXT\_CLASSPATH adds ':' to classpath of workers.
- [STORM-1005](#): Supervisor do not get running workers after restart.
- [STORM-1024](#): log4j changes leaving `${sys:storm.log.dir}` under STORM\_HOME dir.
- [STORM-1027](#): Use overflow buffer for emitting metrics.
- [STORM-1037](#): Do not remove STORM-code in supervisor until kill job.
- [STORM-1044](#): Setting dop to zero does not raise an error.
- [STORM-1096](#): Fix some issues with impersonation on the UI.
- [STORM-1108](#): Fix NPE in simulated time.
- [STORM-1121](#): Deprecate test only configuration `nimbus.reassign`.
- [STORM-139](#): `hashCode` does not work for `byte[]`.
- [STORM-1476](#): Filter `-c` options from args and add them as part of `storm.options`.
- [STORM-1481](#): Avoid `Math.abs(Integer)` get a negative value.
- [STORM-1482](#): Add missing 'break' for `RedisStoreBolt`.
- [STORM-1521](#): When using Kerberos login from keytab with multiple bolts/executors ticket is not renewed.
- [STORM-422](#): Allow more arguments to be passed to Storm jar.
- [STORM-584](#): Fix logging for `LoggingMetricsConsumer` `metrics.log` file.
- [STORM-745](#): Fix `storm.cmd` to evaluate 'shift' correctly with 'Storm jar'.
- [STORM-793](#): Made change to `logviewer.clj` in order to remove the invalid http 500 response.
- [STORM-810](#): `PartitionManager` in STORM-kafka should commit latest offset before close.
- [STORM-837](#): `HdfsState` ignores commits.
- [STORM-866](#): Use `storm.log.dir` instead of `storm.home` in `log4j2` config.
- [STORM-966](#): `ConfigValidation.DoubleValidator` doesn't really validate whether the type of the object is a double.
- [STORM-977](#): Incorrect signal (-9) when `as-user` is true.
- [STORM-992](#): A bug in the `timer.clj` might cause unexpected delay to schedule new event.

HDP 2.3.4.7 provided Storm 0.10.0-beta with no additional Apache patches.



HDP 2.3.4 provided Storm 0.10.0-beta and the following Apache patches:

- [STORM-1017](#): If ignoreZkOffsets set true, KafkaSpout will reset zk offset when recover from failure.
- [STORM-1030](#): Hive Connector Fixes.
- [STORM-1102](#): Add a default flush interval for HiveBolt.
- [STORM-1147](#): STORM JDBC Bolt should add validation to ensure either insertQuery or table name is specified and not both.
- [STORM-1203](#): Worker metadata file creation doesn't use STORM.log.dir config.
- [STORM-139](#): hashCode does not work for byte[[]].
- [STORM-412](#): Allow users to modify logging levels of running topologies.
- [STORM-742](#): Let ShellBolt treat all messages to update heartbeat.
- [STORM-992](#): A bug in the timer.clj might cause unexpected delay to schedule new event.
- [STORM-996](#): netty-unit-tests/test-batch demonstrates out-of-order delivery.

HDP 2.3.2 provided Storm 0.10.0-beta, with no additional Apache patches.

HDP 2.3.0 provided Storm 0.10.0-beta and the following Apache patches:

- [STORM-166](#) Highly Available Nimbus
- [STORM-583](#) Add Microsoft Azure Event Hub spout implementations
- [STORM-691](#) Add basic lookup / persist bolts
- [STORM-703](#) With hash key option for RedisMapState, only get values for keys in batch
- [STORM-708](#) CORS support for STORM UI.
- [STORM-711](#) All connectors should use collector.reportError and tuple anchoring.
- [STORM-713](#) Include topic information with Kafka metrics.
- [STORM-714](#) Make CSS more consistent with self, prev release
- [STORM-724](#) Document RedisStoreBolt and RedisLookupBolt which is missed.
- [STORM-727](#) Storm tests should succeed even if a Storm process is running locally.
- [STORM-741](#) Allow users to pass a config value to perform impersonation.

## 1.5.20. Tez

In addition to any Apache patches ported over [from the 2.4.x core](#), HDP 2.3.6 provides Tez 0.7.0 and the following Apache patches:

- [TEZ-2097](#): TEZ-UI Add dag logs backend support.
- [TEZ-3114](#): Shuffle OOM due to EventMetaData flood.
- [TEZ-3276](#): Tez Example MRRSleep job fails when tez.staging-dir fs is not same as default FS.
- [TEZ-3289](#): Tez Example MRRSleep job does not set Staging dir correctly on secure cluster.
- [TEZ-3295](#): TestOrderedWordCount should handle relative input/output paths.

HDP 2.3.6 provides Tez 0.7.0 and ports the following Apache patches [from the 2.4.x core](#):

- [TEZ-1961](#): Remove misleading exception "No running dag" from AM logs.
- [TEZ-2307](#): Possible wrong error message when submitting new DAG.
- [TEZ-2863](#): Container, node, and logs not available in UI for tasks that fail to launch.
- [TEZ-2886](#): Ability to merge AM credentials with DAG credentials.
- [TEZ-2898](#): Tez tools: swimlanes.py is broken.
- [TEZ-2900](#): Ignore V\_INPUT\_DATA\_INFORMATION when vertex is in Failed/Killed/Error.
- [TEZ-2963](#): RecoveryService#handleSummaryEvent exception with HDFS transparent encryption & Kerberos authentication.
- [TEZ-3017](#): HistoryACLManager does not have a close method for cleanup.
- [TEZ-3025](#): InputInitializer creation should use the DAG GUI.
- [TEZ-3032](#): Incorrect start time in different events for DAG history events.
- [TEZ-3037](#): History URL should be set regardless of which history logging service is enabled.
- [TEZ-3066](#): TaskAttemptFinishedEvent ConcurrentModificationException in recovery or history logging services.
- [TEZ-3101](#): Tez UI: Task attempt log link doesn't have the correct protocol.
- [TEZ-3117](#): Deadlock in Edge and Vertex code.
- [TEZ-3123](#): Containers can get re-used even with conflicting local resources.
- [TEZ-3126](#): Log reason for not reducing parallelism.
- [TEZ-3128](#): Avoid stopping containers on the AM shutdown thread.
- [TEZ-3137](#): Tez task failed with illegal state exception.
- [TEZ-3156](#): Tez client keeps trying to talk to RM even if RM does not know about the application.

- [TEZ-3177](#): Non-DAG events should use the session domain or no domain if the data does not need protection.
- [TEZ-3189](#): Pre-warm dags should not be counted in submitted dags count by DAGAppMaster.

HDP 2.3.4.7 provided Tez 0.7.0 and the following Apache patches:

- [TEZ-3017](#): Thread leaks from TimelineClient.
- [TEZ-3025](#): InputInitializer creation should use the dag UGI.

HDP 2.3.4 provided Tez 0.7.0 and the following Apache patches:

- [TEZ-1314](#): Port MAPREDUCE-5821 to Tez.
- [TEZ-2398](#): Flaky test: TestFaultTolerance.
- [TEZ-2436](#): Tez UI: Add cancel button in column selector.
- [TEZ-2538](#): ADDITIONAL\_SPILL\_COUNT wrongly populated for DefaultSorter with multiple partitions.
- [TEZ-2553](#): Tez UI: Tez UI Nits.
- [TEZ-2561](#): Port for TaskAttemptListenerImpTezDag should be configurable.
- [TEZ-2758](#): Remove append API in RecoveryService after TEZ-1909..
- [TEZ-2781](#): Fallback to send only TaskAttemptFailedEvent if taskFailed heartbeat fails.
- [TEZ-2787](#): Tez AM should have java.io.tmpdir=./tmp to be consistent with tasks.
- [TEZ-2808](#): Race condition between preemption and container assignment.
- [TEZ-2829](#): Tez UI: minor fixes to in-progress update of UI from AM.
- [TEZ-2850](#): Tez MergeManager OOM for small Map Outputs.
- [TEZ-2851](#): Support a way for upstream applications to pass in a caller context to Tez.
- [TEZ-2866](#): Tez UI: Newly added columns wont be displayed by default in tables.
- [TEZ-2874](#): Improved logging for caller context.
- [TEZ-2878](#): Tez UI: AM error handling - Make the UI handle cases in which AM returns unexpected/no data.
- [TEZ-2882](#): Consider improving fetch failure handling.
- [TEZ-2885](#): Remove counter logs from AMWebController.
- [TEZ-2887](#): Tez build failure due to missing dependency in pom files.
- [TEZ-2893](#): Tez UI: Retain vertex info displayed in DAG details page even after completion.

- [TEZ-2894](#): Tez UI: Disable sorting for few columns while in progress. Display an alert on trying to sort them.
- [TEZ-2895](#): Tez UI: Add option to enable and disable in-progress.
- [TEZ-2896](#): Fix thread names used during Input/Output initialization.
- [TEZ-2899](#): Backport graphical view fix from TEZ-2899.
- [TEZ-2907](#): NPE in IFile.Reader.getLength during final merge operation.
- [TEZ-2908](#): Tez UI: Errors are logged, but not displayed in the UI when AM fetch fails.
- [TEZ-2909](#): Tez UI: Application link in All DAGs table is disable when applicationhistory is unavailable.
- [TEZ-2910](#): Tez should invoke HDFS Client API to set up caller context.
- [TEZ-2915](#): Tez UI: Getting back to the DAG details page is difficult.
- [TEZ-2923](#): Tez Live UI counters view empty for vertices, tasks, attempts.
- [TEZ-2927](#): Tez UI: Graciously fail when system-metrics-publisher is disabled.
- [TEZ-2929](#): Tez UI: Dag details page displays vertices to be running even when dag have completed.
- [TEZ-2930](#): Tez UI: Parent controller is not polling at times.
- [TEZ-2933](#): Tez UI: Load application details from RM when available.
- [TEZ-2936](#): Support HDFS-based Timeline writer.
- [TEZ-2946](#): Tez UI: At times RM return a huge error message making the yellow error bar to fill the whole screen.
- [TEZ-2947](#): Tez UI: Timeline, RM : AM requests gets into a consecutive loop in counters page without any delay.
- [TEZ-2949](#): Allow duplicate dag names within session for Tez.
- [TEZ-2960](#): Tez UI: Move hardcoded URL namespace to the configuration file.
- [TEZ-2968](#): Counter limits exception causes AM to crash.
- [TEZ-2970](#): Re-localization in TezChild does not use correct UGI.
- [TEZ-2975](#): Bump up Apache commons dependency.
- [TEZ-2995](#): Timeline primary filter should only be on callerId and not type.

HDP 2.3.2 provided Tez 0.7.0 and the following Apache patches:

#### IMPROVEMENTS

- [TEZ-2719](#) Consider reducing logs in unordered fetcher with shared-fetch option.

- [TEZ-2767](#) Make TezMxBeanResourceCalculator the default resource calculator.
- [TEZ-2789](#) Backport events added in TEZ-2612 to branch-0.7.
- [TEZ-2812](#) Tez UI, Update task and attempt tables while in progress.
- [TEZ-2813](#) Tez UI, add counter data for rest api calls to AM Web Services v2.
- [TEZ-2817](#) Tez UI, update in progress counter data for the dag vertices and tasks table.
- [TEZ-2817](#) Tez UI, update in progress counter data for the dag vertices and tasks table.
- [TEZ-2830](#) Backport TEZ-2774 to branch-0.7. Improvements to logging in the AM and part of the runtime.
- [TEZ-2844](#) Backport TEZ-2775 to branch-0.7. Improve and consolidate logging in Runtime components.
- [TEZ-2876](#) Tez UI, Update vertex, task and attempt details page while in progress.

#### BUG FIXES

- [TEZ-2211](#) Tez UI. Allow users to configure timezone.
- [TEZ-2291](#) TEZ UI. Improper vertex name in tables.
- [TEZ-2483](#) Tez UI, Show error if in progress fails due to AM not reachable.
- [TEZ-2549](#) Reduce Counter Load on the Timeline Server.
- [TEZ-2602](#) Throwing EOFException when launching MR job.
- [TEZ-2660](#) Tez UI, need to show application page even if system metrics publish is disabled.
- [TEZ-2663](#) SessionNotRunning exceptions are wrapped in a ServiceException from a dying AM.
- [TEZ-2745](#) ClassNotFoundException in InputInitializer causes AM to crash.
- [TEZ-2752](#) logUnsuccessful completion in Attempt should write original finish time to ATS.
- [TEZ-2754](#) Tez UI. StartTime and EndTime is not displayed with right format in Graphical View.
- [TEZ-2761](#) Addendum fix build failure for java 6.
- [TEZ-2761](#) Tez UI. update the progress on the dag and vertices pages with info from AM.
- [TEZ-2766](#) Tez UI. Add vertex in-progress info in DAG details.
- [TEZ-2768](#) Log a useful error message when the summary stream cannot be closed when shutting down an AM.
- [TEZ-2780](#) Tez UI, Update All Tasks page while in progress.

- [TEZ-2792](#) Add AM web service API for tasks.
- [TEZ-2792](#) Addendum fix build failure for java 6.
- [TEZ-2807](#) Log data in the finish event instead of the start event.
- [TEZ-2812](#) Preemption sometimes does not respect heartbeats between pre-emptions.
- [TEZ-2816](#) Preemption sometimes does not respect heartbeats between pre-emptions.
- [TEZ-2825](#) Report progress in terms of completed tasks to reduce load on AM for Tez UI.
- [TEZ-2834](#) Make Tez preemption resilient to incorrect free resource reported by YARN.
- [TEZ-2842](#) Tez UI, Update Tez App details page while in-progress.
- [TEZ-2847](#) Tez UI, Task details doesn't gets updated on manual refresh after job complete.
- [TEZ-2853](#) Tez UI, task attempt page is coming empty.
- [TEZ-814](#) Improve heuristic for determining a task has failed outputs.
- [TEZ-814](#) Improve heuristic for determining a task has failed outputs.

#### INCOMPATIBLE CHANGES

- [TEZ-2768](#) Log a useful error message when the summary stream cannot be closed when shutting down an AM.

HDP 2.3.0 provided Tez 0.7.0 and the following Apache patches:

#### IMPROVEMENTS

- [TEZ-2076](#) Tez framework to extract/analyze data stored in ATS for specific dag.
- [TEZ-2461](#) tez-history-parser compile fails with hadoop-2.4.

#### BUG FIXES

- [TEZ-1529](#) ATS and TezClient integration in secure Kerberos enabled cluster.
- [TEZ-2391](#) TestVertexImpl timing out at times on Jenkins builds.
- [TEZ-2409](#) Allow different edges to have different routing plugins
- [TEZ-2447](#) Tez UI: Generic changes based on feedbacks.
- [TEZ-2453](#) Tez UI: show the dagInfo is the application has set the same.
- [TEZ-2455](#) Tez UI: Dag view caching, error handling and minor layout changes
- [TEZ-2460](#) Temporary solution for issue due to [YARN-2560](#)
- [TEZ-2474](#) The old taskNum is logged incorrectly when parallelism is changed
- [TEZ-2475](#) Fix a potential hang in Tez local mode caused by incorrectly handled interrupts.

- [TEZ-2478](#) Move OneToOne routing to store events in Tasks.
- [TEZ-2481](#) Tez UI: graphical view does not render properly on IE11
- [TEZ-2482](#) Tez UI: Mouse events not working on IE11
- [TEZ-2489](#) Disable warn log for Timeline ACL error when tez.allow.disabled.timeline-domains set to true.
- [TEZ-2504](#) Tez UI: tables - show status column without scrolling, numeric 0 shown as Not available
- [TEZ-2505](#) PipelinedSorter uses Comparator objects concurrently from multiple threads.
- [TEZ-2509](#) YarnTaskSchedulerService should not try to allocate containers if AM is shutting down.
- [TEZ-2513](#) Tez UI: Allow filtering by DAG ID on All dags table.
- [TEZ-2523](#) Tez UI: derive applicationId from dag/vertex id instead of relying on json data
- [TEZ-2527](#) Tez UI: Application hangs on entering erroneous RegEx in counter table search box
- [TEZ-2528](#) Tez UI: Column selector buttons gets clipped, and table scroll bar not visible in mac.
- [TEZ-2535](#) Tez UI: Failed task attempts link in vertex details page is broken.
- [TEZ-2539](#) Tez UI: Pages are not updating in IE.
- [TEZ-2541](#) DAGClientImpl enable TimelineClient check is wrong.
- [TEZ-2546](#) Tez UI: Fetch Hive query text from timeline if dagInfo is not set.
- [TEZ-2547](#) Tez UI: Download Data fails on secure, cross-origin clusters
- [TEZ-2548](#) TezClient submitDAG can hang if the AM is in the process of shutting down.
- [TEZ-2554](#) Tez UI: View log link does not correctly propagate login credential to read log from YARN web.
- [TEZ-2568](#) auto\_sortmerge\_join\_5 fails in Tez mode

## 1.5.21. ZooKeeper

In addition to any Apache patches ported over [from the 2.4.x core](#), HDP 2.3.6 provides ZooKeeper 3.4.6 and the following Apache patches:

- [ZOOKEEPER-2146](#) BinaryInputArchive readString should check length before allocating memory.

HDP 2.3.6 provides ZooKeeper 3.4.6 and ports no additional Apache patches [from the 2.4.x core](#).

HDP 2.3.4.7 provided ZooKeeper 3.4.6 with no additional Apache patches.

HDP 2.3.4 provided ZooKeeper 3.4.6 and the following Apache patches:

- [ZOOKEEPER-706](#) Large numbers of watches can cause session re-establishment to fail.

HDP 2.3.2 provided ZooKeeper 3.4.6 and the following Apache patches:

- [ZOOKEEPER-1506](#): Re-try DNS hostname -> IP resolution if node connection fails.
- [ZOOKEEPER-1574](#): Set svn:eol-style property to native.
- [ZOOKEEPER-1575](#): adding .gitattributes to prevent CRLF and LF mismatches for source and text files.
- [ZOOKEEPER-1791](#): ZooKeeper package includes unnecessary jars that are part of the package.
- [ZOOKEEPER-1848](#): [WINDOWS] Java NIO socket channels does not work with Windows ipv6 on JDK6.
- [ZOOKEEPER-1876](#): Add support for installing windows services in .cmd scripts.
- [ZOOKEEPER-1888](#): ZkCli.cmd commands fail with "'java' is not recognized as an internal or external command".
- [ZOOKEEPER-1897](#): ZK Shell/CLI not processing commands.
- [ZOOKEEPER-1926](#): Unit tests should only use build/test/data for data.
- [ZOOKEEPER-1952](#): Default log directory and file name can be changed.
- [ZOOKEEPER-2053](#): Make scripts use ZOOKEEPER\_HOME if set.
- [ZOOKEEPER-2064](#): Prevent resource leak in various classes.

HDP 2.3.0 provided ZooKeeper 3.4.6, with no additional Apache patches.

## 1.6. Common Vulnerabilities and Exposures

- [CVE-2015-3253](#): The MethodClosure class in runtime/MethodClosure.java in Apache Groovy 1.7.0 through 2.4.3 allows remote attackers to execute arbitrary code or cause a denial of service via a crafted serialized object.

**Severity:** High

**Vendor:** Hortonworks

**Versions Affected:** All HDP versions with Spark 1.5 and 1.6. (HDP 2.3.4.7+).

**Users Affected:** Users whose Apache Spark applications load data files using SparkContext.objectFile().

**Impact:** See [SPARK-13599](#) If the data file loaded contained a serialized Groovy closure, it is possible to execute code as that user. Therefore, if a Spark job read in a file from an



external source/user with lower rights, using the `objectFile()` API, it would be possible execute code as the user running the spark application.

**Recommended Action:**

1. Upgrade to HDP 2.3.6+.
2. Do not have any versions of Apache Groovy on the classpath of any Spark application —or upgrade to Groovy version 2.4.4+.
3. Never use java serialization as a storage/exchange mechanism for data. There have been other Java serialization-based attacks in the past; avoiding the use of Java serialization and loading files via the `objectFile()` method guarantees that this specific vulnerability or any similar one cannot be exploited by maliciously crafted files.

## 1.7. Third-Party Licenses

The following license globally applies to Apache Hadoop: [Apache 2.0](#)

Apache Component	Subcomponents	License
Accumulo	JCommander	<a href="#">JCommander</a>
Falcon	cern.colt*, cern.jet*, cern.clhep	<a href="#">CERN</a>
Knox	ApacheDS, Groovy	<a href="#">ANTLR</a>
Knox	SL4J	<a href="#">MIT</a>
Knox	Jetty and Jerico	<a href="#">EPL</a>
Knox	ApacheDS	<a href="#">Bouncy Castle</a>
Oozie	JDOM Oro	
Phoenix		<a href="#">EPL</a>
Storm	Logback	<a href="#">EPL</a>

## 1.8. Fixed Issues

Fixed issues represents selected issues previously logged by Hortonworks Support and addressed in the current release. These issues might have been reported in previous versions within the [Known Issues](#) section, meaning that they were reported by customers or identified by Hortonworks Quality Engineering team.

Hortonworks Data Platform (HDP) 2.3.6 is the first maintenance release that aligns the 2.3.x core with the 2.4.x core; see [New Features](#) for details.

**Potential Data Loss**

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary	Source
BUG-52962	<a href="#">HDFS-9289</a>	HDFS	Make DataStreamer#block thread safe and verify genStamp in commitBlock.	2.3.6
BUG-52963	<a href="#">HDFS-9289</a>	HDFS	Make DataStreamer#block	2.4.2

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary	Source
			thread safe and verify genStamp in commitBlock.	
BUG-55474	<a href="#">HDFS-10178</a>	HDFS	Permanent write failures can happen if pipeline recoveries occur for the first packet.	2.3.6

## Security

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary	Source
BUG-46092	<a href="#">HIVE-7193</a>	Hive	We need to filter users who can connect on our HiveServer2 to those in a particular LDAP group (with nested groups).	2.3.6
BUG-46910		Hue	Request to add HUE-1663 to fix LDAP referral handling for authentication.	2.4.0
BUG-48092	<a href="#">RANGER-746</a>	Ranger	Ranger Admin: For validating plugins SSL certs - Add wildcard, multiple CN & SAN support.	2.4.2
BUG-50350	<a href="#">HBASE-14471</a>	HBase	Thrift - HTTP Error 413 full HEAD if using Kerberos authentication.	2.4.0
BUG-51013	<a href="#">TEZ-3025</a>	Tez	InputInitializer creation should use the dag UGI.	2.4.0
BUG-51298	<a href="#">HBASE-15145</a>	HBase	Default hbase-env.sh breaks add peer command that writes to /hbase-secure/ replication znode for kerberized clusters.	2.4.2
BUG-52193	<a href="#">RANGER-798</a>	Ranger	Ranger "Access > Audit" not showing anything with Source "Solr" due to Time filter Issue (GMT).	2.4.2
BUG-52636	<a href="#">STORM-1521</a>	Storm	Kerberos login from keytab with multiple bolts/executors ticket is not renewed for HBase Bolt.	2.4.2
BUG-53766	<a href="#">STORM-1535</a>	Storm	Storm-HDFS - When using keytab, Kerberos ticket is not renewed with multiple bolts/executors.	2.4.2
BUG-55917	<a href="#">STORM-1711</a>	Storm	HiveUtils.authenticate reuses current UGI, mixes up principles in use.	2.3.6

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary	Source
BUG-57857	<a href="#">STORM-721</a>	Storm	Storm UI server should support SSL.	2.3.6

### Incorrect Results

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary	Source
BUG-50338		Hive	NULL table comparison within CASE does not work as previous Hive versions.	2.3.6
BUG-52194	<a href="#">RANGER-798</a>	Ranger	Ranger "Access > Audit" not showing anything with Source "Solr" due to Time filter Issue (GMT).	2.3.6
BUG-53374	<a href="#">HADOOP-12716</a>	HDFS	KerberosAuthenticator#265 use incorrect class to determine isKeyTab in JDK8.	2.3.6
BUG-53777	<a href="#">PIG-4690</a>	Pig	Pig on TEZ creates wrong result with replicated join.	2.4.2
BUG-53918	<a href="#">PHOENIX-2601</a>	Phoenix	Query result is incorrect when both index hint and limit are used.	2.4.2
BUG-53927	<a href="#">HIVE-13330</a>	Hive	Vectorization returns NULL for empty values for varchar/string data type.	2.4.2
BUG-54492	<a href="#">HIVE-12165</a>	Hive	Wrong result when hive.optimize.samplingorderby=true with some aggregate functions.	2.3.6
BUG-55552	<a href="#">HIVE-13200</a>	Hive	Aggregation functions returning empty rows on partitioned columns.	2.3.6
BUG-55811	<a href="#">HIVE-13242</a>	Hive	DISTINCT keyword is dropped by the parser for windowing.	2.3.6
BUG-56283	<a href="#">HIVE-12768</a>	Hive	Tez various wrong results with map <- (map , reduce) join on decimal query.	2.3.6
BUG-56434	<a href="#">MAPREDUCE-6680</a>	MapReduce	JHS directory scan issues on AzureFileSystem due to time inaccurate.	2.3.6
BUG-57030	<a href="#">PHOENIX-2876</a>	Phoenix	Phoenix ORDER BY clause does not work if the ORDER BY column is not present in the Select clause.	2.3.6
BUG-57038	<a href="#">RANGER-893</a>	Ranger	Ranger ugsync with LDAP is not able to fetch group information.	2.3.6

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary	Source
BUG-58985	<a href="#">HIVE-13602</a>	Hive	TPCH q16 return wrong result when CBO is on.	2.3.6
BUG-59049		Ranger	Ranger UI : policy status search is not working on policy listing page.	2.3.6

## Stability

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary	Source
BUG-18218	<a href="#">YARN-3846</a>	YARN	Scheduler queue filter failed to work because index of queue column changed.	2.3.6
BUG-41224		Hue	Workaround Hiveserver2 memory Leak.	2.4.0
BUG-41311	<a href="#">HDFS-9917</a>	HDFS	NameNode quickly runs out of memory at start due to queued incremental block reports.	2.3.6
BUG-43129		Hive	ACID update/delete query fails with ArrayIndexOutOfBoundsException when there is null struct.	2.3.6
BUG-46394		Hive	Hiveserver2 error seen after upgrade.	2.3.6
BUG-47626		HBase, Hive, Phoenix	Creating Hive table on HBase fails in HDP 2.3.x with Phoenix Dependencies.	2.3.6
BUG-48217	<a href="#">YARN-4325</a>	YARN	After restarting Yarn services, applications stays in ACCEPTED state.	2.3.6
BUG-49345		Hive	Tez queries stopped running - ApplicationNotFoundException.	2.3.6
BUG-49788	<a href="#">HBASE-14968</a>	Phoenix	Altering session consistency blocks SYSTEM.CATALOG table.	2.3.6
BUG-49898	<a href="#">HIVE-12684</a>	Hive	Autoquerygen tests failing in secure runs.	2.3.6
BUG-50148	<a href="#">HIVE-12660</a>	Hive	HS2 memory leak with .hiverc file use.	2.3.6
BUG-50215	<a href="#">FALCON-1595</a>	Falcon	In secure cluster, Falcon server loses ability to communicate with HDFS over time.	2.4.0
BUG-50410	<a href="#">HBASE-14987</a>	HBase	Compaction marker whose region name doesn't match current	2.3.6

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary	Source
			region's needs to be handled.	
BUG-51034	<a href="#">HIVE-11097</a>	Hive	MR mode query fails if one table path string starts with another's.	2.3.6
BUG-51454	<a href="#">HBASE-14471</a>	HBase	Thrift - HTTP Error 413 full HEAD if using Kerberos authentication.	2.3.6
BUG-51829	<a href="#">YARN-4820</a>	YARN	ResourceManager JMX Redirects in HA Mode Drops the Query Parameters.	2.4.2
BUG-52310	<a href="#">HBASE-14347</a>	HBase	Add a switch to DynamicClassLoader to disable it.	2.3.6
BUG-52418	<a href="#">YARN-4680</a>	YARN	TimerTasks leak in ATS V1.5 Writer.	2.3.6
BUG-52479	<a href="#">PHOENIX-1968</a> <a href="#">PHOENIX-2036</a> <a href="#">PHOENIX-2088</a> <a href="#">PHOENIX-2112</a> <a href="#">PHOENIX-2196</a> <a href="#">PHOENIX-2287</a> <a href="#">PHOENIX-2288</a> <a href="#">PHOENIX-2328</a> <a href="#">PHOENIX-2426</a> <a href="#">PHOENIX-2469</a> <a href="#">PHOENIX-2503</a> <a href="#">PHOENIX-2567</a> <a href="#">PHOENIX-2568</a> <a href="#">PHOENIX-2599</a> <a href="#">PHOENIX-2638</a>	Phoenix	Phoenix-Spark integration inoperative in HDP 2.3.4.	2.4.2
BUG-52780	<a href="#">PHOENIX-2667</a>	Phoenix	Race condition between IndexBuilder and Split for region lock.	2.3.6
BUG-52957	<a href="#">PHOENIX-2295</a>	Phoenix	NullPointerException when parsing a query with hints in UNION ALL.	2.4.2
BUG-53482	<a href="#">PIG-4814</a>	Pig	AvroStorage does not take NameNode HA as part of schema file URL.	2.4.2
BUG-53499	<a href="#">PIG-4790</a>	Pig	Cannot use values iterator on the previous K-V pair after moveToNext has been	2.4.2

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary	Source
			invoked to move to the next K-V.	
BUG-54199	<a href="#">PHOENIX-2616</a>	Phoenix	Indexes over immutable tables not marked as immutable.	2.4.2
BUG-54603	<a href="#">TEZ-3114</a>	Tez	.Shuffle OOM due to EventMetaData flood.	2.3.6
BUG-54650	<a href="#">KNOX-695</a>	Knox	Intermittent java.net.ConnectException: Connection timed out between Knox server and Hiveserver2.	2.4.2
BUG-54713	<a href="#">STORM-1689</a>	Storm	Large request header causes 413 ERROR when accessing Storm Logviewer.	2.4.2
BUG-55235	<a href="#">HIVE-10729</a>	Hive	Query failed when select complex columns from joined table (Tez map join only).	2.3.6
BUG-55895	<a href="#">PIG-4873</a>	Pig	Pig job fails with Backend error : java.lang.NullPointerException.	2.3.6
BUG-56083	<a href="#">HIVE-13394</a>	Hive	Analyze table fails in Tez on empty partitions/files/tables.	2.3.6
BUG-56423	<a href="#">HDFS-10312</a>	HDFS	NameNode is stuck at SafeMode due to not all data blocks validated.	2.3.6
BUG-56577	<a href="#">HADOOP-12559</a> <a href="#">HADOOP-12682</a> <a href="#">HADOOP-12716</a> <a href="#">HADOOP-12787</a> <a href="#">HADOOP-12878</a>	Accumulo, HDFS	KMS SPNEGO sequence does not work with WEBHDFS and KMS connection failures should trigger TGT renewal.	2.3.6
BUG-56919	<a href="#">YARN-2910</a>	YARN	FSLeafQueue can throw ConcurrentModificationException.	2.3.6
BUG-57460	<a href="#">HBASE-15837</a> <a href="#">PHOENIX-2883</a>	HBase, Phoenix	Invalid MemStore state during close of region for Phoenix table.	2.3.6
BUG-57513	<a href="#">HDFS-9198</a> <a href="#">HDFS-9958</a>	HDFS	BlockManager#createLocalBlocks can throw NPE for corruptBlocks on failed storages.	2.3.6
BUG-57653	<a href="#">HIVE-13051</a>	Hive, Hive2	Deadline class has numerous issues.	2.3.6
BUG-57732	<a href="#">HBASE-14712</a> <a href="#">HBASE-15100</a>	HBase	MasterProcWALS never clean up.	2.3.6
BUG-57814	<a href="#">HBASE-13318</a>	HBase	RpcServer.getListenerAddresses should handle when	2.3.6

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary	Source
			the accept channel is closed.	
BUG-57906	<a href="#">HBASE-14460</a>	HBase	Regression for CheckAndPut and Increment updates.	2.3.6
BUG-57907	<a href="#">PHOENIX-2892</a>	HBase	Scan for pre-warming the block cache for 2ndary index should be removed.	2.3.6
BUG-58254	<a href="#">HBASE-15811</a>	HBase	Batch Get after batch Put does not fetch all Cells.	2.3.6
BUG-59089	<a href="#">HIVE-13592</a>	Hive	Metastore calls map is not thread safe.	2.3.6

### Query Failure

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary	Source
BUG-54388 BUG-51701 BUG-53494	<a href="#">MAPREDUCE-6635</a>	MapReduce, Tez	java.lang.IndexOutOfBoundsException running select count(*) query on table with skip.header.line.count.	

### Upgrade

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary	Source
BUG-53878		Kafka	Running bin/kafka-console-consumer.sh outputs "No brokers found in ZK". Script modified to address this issue.	2.4.2
BUG-54155	<a href="#">RANGER-882</a>	Ranger	Ranger policy engine initialization failure caused NameNode startup to fail.	2.4.2
BUG-55263		Storm	Storm Nimbus error: "Could not find leader nimbus from seed hosts".	2.3.6

### Usability

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary	Source
BUG-47662	<a href="#">RANGER-218</a>	Ranger	UserSync - group sync is marked as INTERNAL even if the group is synced from LDAP.	2.4.2
BUG-48573	<a href="#">OOZIE-2322</a>	Oozie	Oozie Web UI doesn't work with Kerberos in Internet Explorer 10 or 11 and curl.	2.4.2
BUG-49428	<a href="#">HBASE-13153</a>	HBase	Bulk Loaded HFile Replication.	2.3.6

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary	Source
BUG-50576	<a href="#">SQOOP-2779</a>	Sqoop	Sqoop Metastore doesn't seem to recognize –schema option.	2.3.6
BUG-52257	<a href="#">FALCON-1315</a>	Falcon	Falcon UI Hive mirroring : Unable to provide security credentials.	2.4.2
BUG-52260	<a href="#">SQOOP-2737</a>	Sqoop	Cannot import table from Oracle with column with spaces in name.	2.4.2
BUG-52702	<a href="#">SQOOP-2779</a>	Sqoop	Sqoop Metastore doesn't seem to recognize –schema option.	2.4.2
BUG-52951	<a href="#">FALCON-1647</a>	Falcon	Unable to create feed : FilePermission error under cluster staging directory.	2.4.2
BUG-54681	<a href="#">RANGER-889</a>	Ranger	Create new method for partners to use for integration.	2.4.2
BUG-55216	<a href="#">FALCON-748</a>	Falcon	Falcon webUI return 413 error when SSL is enabled.	2.4.2
BUG-58351	<a href="#">OOZIE-2112</a>	Hive, Oozie, Tez	Empty 'ExternalChildID' returned when using Hive Action in Oozie.	2.3.6
BUG-58450	<a href="#">PHOENIX-2568</a>	Flume, Hive, Phoenix	Flume dependency of Phoenix is bringing in an old version of Thrift (0.7.0).	2.3.6

## Performance

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary	Source
BUG-42479	<a href="#">HBASE-14575</a>	HBase	Relax region read lock for compactions.	2.4.0
BUG-49426	<a href="#">HBASE-14606</a>	HBase	TestSecureLoadIncrementalFiles tests timed out in trunk build on Apache.	2.3.6
BUG-49427	<a href="#">HBASE-14604</a>	HBase	Improve MoveCostFunction in StochasticLoadBalancer.	2.3.6
BUG-50147	<a href="#">HBASE-15014</a>	HBase	Fix slow filterCellByStore in WALsplitter.	2.3.6
BUG-52708	<a href="#">RANGER-809</a>	Ranger	Remove reference to getLocalHost() call from Audit code to avoid performance impact.	2.4.2
BUG-52729	<a href="#">RANGER-821</a>	Ranger	NameNode shutdown hangs due to Ranger HDFS plugin.	2.4.2



Hortonworks Bug ID	Apache JIRA	Apache Component	Summary	Source
BUG-53028	<a href="#">RANGER-794</a> <a href="#">RANGER-836</a> <a href="#">RANGER-844</a>	Ranger	Policy download optimizations.	2.4.2
BUG-53657	<a href="#">HDFS-9710</a>	HDFS	Change DN to send block receipt IBRs in batches.	2.3.6
BUG-53757	<a href="#">HIVE-11424</a> <a href="#">HIVE-13233</a>	Hive, Hive2	Stats are not properly accounted for when multiple COUNT(DISTINCT) are in the query.	2.3.6
BUG-55291	<a href="#">HDFS-10223</a>	HDFS	peerFromSocketAndKey performs SASL exchange before setting connection timeouts.	2.3.6
BUG-56776	<a href="#">MAPREDUCE-6514</a> <a href="#">MAPREDUCE-6689</a>	MapReduce	MR job can infinitely increase number of reducer resource requests and job hangs as ask is not updated after ramping down of all reducers.	2.3.6
BUG-56869	<a href="#">HDFS-9902</a>	HDFS	Support different values of dfs.datanode.du.reserved per storage-type.	2.3.6

## Other

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary	Source
BUG-50101		Hadoop Common	Remove ivy jar from hadoop-lzo source code.	2.4.0
BUG-51738	<a href="#">HBASE-15147</a>	HBase	Backport HBASE-15147: Shell should use Admin.listTableNames() instead of Admin.listTables().	2.4.2
BUG-53095	<a href="#">HADOOP-12825</a>	Hadoop Common	Log slow name resolutions.	2.4.2
BUG-53390	<a href="#">HADOOP-12794</a>	Hadoop Common	Support additional compression levels for GzipCodec	2.3.6
BUG-53871	<a href="#">HBASE-15128</a>	HBase	Disable region splits and merges switch in master.	2.4.2
BUG-54764	<a href="#">MAPREDUCE-6436</a>	MapReduce	JobHistoryServer cache issues.	2.4.2
BUG-56772	<a href="#">HADOOP-12950</a> <a href="#">HADOOP-12993</a> <a href="#">RANGER-891</a>	HDFS, Ranger	NameNode went down on the production cluster.	2.3.6
BUG-57450	<a href="#">HUE-2880</a>	Hue	Failure when uploading large	2.3.6

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary	Source
			files to a kerberized HTTPFS.	

## 1.9. Known Issues

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary
BUG-27132	<a href="#">KNOX-556</a>	Knox	Knox needs better diagnosability for keystore failures.
BUG-34480	<a href="#">KNOX-603</a>	Knox	Potential resource leak in BaseKeystoreService.createKeystore.
BUG-34843	<a href="#">KNOX-539</a>	Knox	Identity mapping auditing should record mapped identity.
BUG-36435	<a href="#">KNOX-554</a>	Knox	gateway.path not taken into account for Knox URL.
BUG-39514		Oozie	Version Disclosure Vulnerabilities in Oozie Web UI.
BUG-41406		YARN	YARN UI messages not clear about why log aggregation not happened when meeting maximum HDFS subdir limit.
BUG-41606		YARN	AM Container link in YARN RM UI links to a 404 page after application completion.
BUG-46128		Storm	Kafka spout showing errors in the log streaming workflow.
BUG-49715	<a href="#">KNOX-639</a>	Knox	Using empty master secret for Knox causes NullPointerException.
BUG-49952	<a href="#">YARN-3604</a> <a href="#">YARN-3987</a>	YARN	removeApplication in ZKRMStateStore should also disable watch.  am container complete msg ack to NM once RM receive it.
BUG-49955	<a href="#">PHOENIX-2531</a>	Phoenix	Phoenix thin-client jar has incorrect META-INF/services/java.sql.Driver file  <b>Issue:</b> The Phoenix Thin Client Driver, org.apache.phoenix.queryserver.client.Driver, is not automatically registered in the JDBC DriverManager.  <b>Workaround:</b>  1. Add the following to your Java application before trying to obtain the driver:

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary
			<pre>Properties props = new Properties(); String URL = "jdbc:phoenix:thin:URL= http:// localhost:8765"; try { Class.forName("org. apache.phoenix. queryserver.client. Driver"); } catch (Exception e) { throw new RuntimeException(e); } DriverManager. getConnection(URL, props);</pre> <p>2. Manually register the driver.</p>
BUG-50692	<a href="#">YARN-4576</a>	YARN	Extend blacklist mechanism to protect AM failed multiple times on failure nodes.
BUG-53388		Kafka	After non-secure cluster upgrade to 2.3.4, running bin/kafka-console-consumer.sh outputs "No brokers found in ZK".
BUG-55124		HBase Phoenix	Deadlock on region opening when secondary indexes are recovering.
BUG-59387	<a href="#">HADOOP-13105</a>	Hadoop Common	Support timeouts in LDAP queries in LdapGroupsMapping.
BUG-59397	<a href="#">HDFS-10485</a>	HDFS	Fix findbugs warning in FSEditLog.java.
BUG-60205	<a href="#">SQOOP-2470</a>	Sqoop	Incremental Hive import with append not working after validation check for –hive-import and –import.
BUG-60233	<a href="#">TEZ-3294</a>	Tez	Re-submitted dag that fails java opts checks in first try fails due to illegal argument error on next submission.
BUG-60770		HDFS	<p>HDFS distcp SaslException: Checksum failed</p> <p><b>Applies to:</b> MIT Secured / Kerberized HDP Install</p> <p><b>Summary:</b> MapReduce job posting or job status retrieval might fail with new installations with the following exception:</p> <pre>"main" java.io. IOException: org. apache.hadoop.ipc. RemoteException(org. apache.hadoop. yarn.exceptions. YarnRuntimeException):</pre>

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary
			<pre> java.io.IOException: Failed on local exception: java.io. IOException: Couldn't setup connection for jhs/fqdn@REALM to fqdn/ ip.ip.ip.ip:8020; Host Details : local host is: "fqdn/ip.ip.ip.ip"; destination host is: "fqdn":8020; at org.apache.hadoop. mapreduce.v2.hs. CachedHistoryStorage. getFullJob(CachedHistoryStorage. java:147) at org.apache. hadoop.mapreduce. v2.hs.JobHistory. getJob(JobHistory. java:217) at org.apache.hadoop. mapreduce.v2.hs. HistoryClientService \$HSClientProtocolHandler \$1. run(HistoryClientService. java:226) at org.apache.hadoop. mapreduce.v2.hs. HistoryClientService \$HSClientProtocolHandler \$1. run(HistoryClientService. java:222) at java.security. AccessController. doPrivileged(Native Method) ....                     </pre> <p><b>Solution:</b> Restart the NameNode in the cluster.</p>
BUG-67482	<a href="#">RANGER-1136</a>	Ranger	<p><b>Description of Problem:</b> Ranger audit to HDFS fails with TGT errors</p> <p><b>Workaround:</b> Currently, there is no known workaround for this issue.</p>

Technical Service Bulletin	Apache JIRA	Apache Component	Summary
TSB-405	N/A	N/A	<p><b>Impact of LDAP Channel Binding and LDAP signing changes in Microsoft Active Directory</b></p> <p>Microsoft has introduced changes in LDAP Signing and LDAP Channel Binding to increase the security for communications between LDAP clients and Active Directory domain controllers. These optional changes will have an impact on how 3rd party products integrate with Active Directory using the LDAP protocol.</p> <p><b>Workaround</b></p> <p>Disable LDAP Signing and LDAP Channel Binding features in Microsoft Active Directory if they are enabled</p> <p>For more information on this issue, see the corresponding Knowledge article: <a href="#">TSB-2021 405: Impact of LDAP Channel</a></p>

Technical Service Bulletin	Apache JIRA	Apache Component	Summary
			<a href="#">Binding and LDAP signing changes in Microsoft Active Directory</a>
TSB-406	N/A	HDFS	<p><b>CVE-2020-9492 Hadoop filesystem bindings (ie: webhdfs) allows credential stealing</b></p> <p>WebHDFS clients might send SPNEGO authorization header to remote URL without proper verification. A maliciously crafted request can trigger services to send server credentials to a webhdfs path (ie: webhdfs://...) for capturing the service principal</p> <p>For more information on this issue, see the corresponding Knowledge article: <a href="#">TSB-2021-406: CVE-2020-9492 Hadoop filesystem bindings (ie: webhdfs) allows credential stealing</a></p>
TSB-434	<a href="#">HADOOP-17208</a> , <a href="#">HADOOP-17304</a>	Hadoop	<p><b>KMS Load Balancing Provider Fails to invalidate Cache on Key Delete</b></p> <p>For more information on this issue, see the corresponding Knowledge article: <a href="#">TSB 2020-434: KMS Load Balancing Provider Fails to invalidate Cache on Key Delete</a></p>
TSB-465	N/A	HBase	<p><b>Corruption of HBase data stored with MOB feature</b></p> <p>For more information on this issue, see the corresponding Knowledge article: <a href="#">TSB 2021-465: Corruption of HBase data stored with MOB feature on upgrade from CDH 5 and HDP 2</a></p>
TSB-497	N/A	Solr	<p><b>CVE-2021-27905: Apache Solr SSRF vulnerability with the Replication handler</b></p> <p>The Apache Solr ReplicationHandler (normally registered at "/replication" under a Solr core) has a "masterUrl" (also "leaderUrl" alias) parameter. The "masterUrl" parameter is used to designate another ReplicationHandler on another Solr core to replicate index data into the local core. To help prevent the CVE-2021-27905 SSRF vulnerability, Solr should check these parameters against a similar configuration used for the "shards" parameter.</p> <p>For more information on this issue, see the corresponding Knowledge article: <a href="#">TSB 2021-497: CVE-2021-27905: Apache Solr SSRF vulnerability with the Replication handler</a></p>
TSB-512	N/A	HBase	<p><b>HBase MOB data loss</b></p> <p>HBase tables with the MOB feature enabled may encounter problems which result in data loss.</p> <p>For more information on this issue, see the corresponding Knowledge article: <a href="#">TSB 2021-512: HBase MOB data loss</a></p>

## 1.10. Documentation Errata

The following section contains late additions or corrections to the product documentation.

- [Flume: Kafka Sink \[106\]](#)
- [Hive Sink \[107\]](#)
- [Configuring Pig Scripts to Use HCatalog in Oozie Workflows \[110\]](#)
- [Configuring a Sqoop Action to Use Tez to Load Data into a Hive Table \[112\]](#)

## 1.10.1. Flume: Kafka Sink

This is a Flume Sink implementation that can publish data to a Kafka topic. One of the objectives is to integrate Flume with Kafka so that pull-based processing systems can process the data coming through various Flume sources. This currently supports Kafka 0.8.x series of releases.

Property Name	Default	Description
<b>type</b>	-	Must be set to <code>org.apache.flume.sink.kafka.KafkaSink</code> .
<b>brokerList</b>	-	List of brokers Kafka-Sink will connect to, to get the list of topic partitions. This can be a partial list of brokers, but we recommend at least two for HA. The format is a comma separated list of <code>hostname:port</code> .
<b>topic</b>	<code>default-flume-topic</code>	The topic in Kafka to which the messages will be published. If this parameter is configured, messages will be published to this topic. If the event header contains a "topic" field, the event will be published to that topic overriding the topic configured here.
<b>batchSize</b>	100	How many messages to process in one batch. Larger batches improve throughput while adding latency.
<b>requiredAcks</b>	1	How many replicas must acknowledge a message before it is considered successfully written. Accepted values are 0 (Never wait for acknowledgement), 1 (wait for leader only), -1 (wait for all replicas) Set this to -1 to avoid data loss in some cases of leader failure.
<b>Other Kafka Producer Properties</b>	-	These properties are used to configure the Kafka Producer. Any producer property supported by Kafka can be used. The only requirement is to prepend the property name with the prefix "Kafka.". For example: <code>kafka.producer.type</code> .

Note: Kafka Sink uses the topic and key properties from the FlumeEvent headers to send events to Kafka. If the topic exists in the headers, the event will be sent to that specific topic, overriding the topic configured for the Sink. If key exists in the headers, the key will be used by Kafka to partition the data between the topic partitions. Events with same key will be sent to the same partition. If the key is null, events will be sent to random partitions.

An example configuration of a Kafka sink is given below. Properties starting with the prefix Kafka (the last 3 properties) are used when instantiating the Kafka producer. The properties that are passed when creating the Kafka producer are not limited to the properties given in this example. It is also possible include your custom properties here and access them inside the preprocessor through the Flume Context object passed in as a method argument.

```

a1.sinks.k1.type = org.apache.flume.sink.kafka.KafkaSink a1.sinks.k1.topic =
mytopic
a1.sinks.k1.brokerList = localhost:9092
a1.sinks.k1.requiredAcks = 1
a1.sinks.k1.batchSize = 20
a1.sinks.k1.channel = c1

```

## 1.10.2. Hive Sink

This sink streams events containing delimited text or JSON data directly into a Hive table or partition. Events are written using Hive transactions. As soon as a set of events are committed to Hive, they become immediately visible to Hive queries. Partitions to which flume will stream to can either be pre-created or, optionally, Flume can create them if they are missing. Fields from incoming event data are mapped to corresponding columns in the Hive table.

Property Name	Default	Description
channel	-	
type	-	The component type name, needs to be hive.
hive.metastore	-	Hive metastore URI (e.g. thrift://a.b.com:9083).
hive.database	-	Hive database name
hive.table	-	Hive table name.
hive.partition	-	Comma separated list of partition values identifying the partition to write to. May contain escape sequences. E.g.: If the table is partitioned by (continent:string, country:string, time:string) then 'Asia,India,2014-02-26-01-21' will indicate continent=Asia,country=India,time=2014-02-26-01-21.
hive.txnsPerBatchAsk	100	Hive grants a batch of transactions instead of single transactions to streaming clients like Flume. This setting configures the number of desired transactions per Transaction Batch. Data from all transactions in a single batch end up in a single file. Flume will write a maximum of batchSize events in each transaction in the batch. This setting in conjunction with batchSize provides control over the size of each file. Note that eventually Hive will transparently compact these files into larger files.
heartBeatInterval	240	(In seconds) Interval between consecutive heartbeats sent to Hive to keep unused transactions from expiring. Set this value to 0 to disable heartbeats.
autoCreatePartitions	true	Flume will automatically create the necessary Hive partitions to stream to.
batchSize	15000	Max number of events written to Hive in a single Hive transaction.
maxOpenConnections	500	Allow only this number of open connections. If this number is

Property Name	Default	Description
		exceeded, the least recently used connection is closed.
callTimeout	10000	(In milliseconds) Timeout for Hive & HDFS I/O operations, such as openTxn, write, commit, abort.
serializer	-	Serializer is responsible for parsing out field from the event and mapping them to columns in the hive table. Choice of serializer depends upon the format of the data in the event. Supported serializers: DELIMITED and JSON.
roundUnit	minute	The unit of the round down value - second, minute or hour.
roundValue	1	Rounded down to the highest multiple of this (in the unit configured using hive.roundUnit), less than current time.
timeZone	Local	Name of the timezone that should be used for resolving the escape sequences in partition, e.g. Time America/Los_Angeles.
useLocalTimeStamp	false	Use the local time (instead of the timestamp from the event header) while replacing the escape sequences.

Following serializers are provided for Hive sink:

- **JSON:** Handles UTF8 encoded Json (strict syntax) events and requires no configuration. Object names in the JSON are mapped directly to columns with the same name in the Hive table. Internally uses org.apache.hive.hcatalog.data.JsonSerDe but is independent of the Serde of the Hive table. This serializer requires HCatalog to be installed.
- **DELIMITED:** Handles simple delimited textual events. Internally uses LazySimpleSerde but is independent of the Serde of the Hive table.

Property Name	Default	Description
serializer.delimiter	,	(Type: string) The field delimiter in the incoming data. To use special characters, surround them with double quotes like "\t".
serializer.fieldnames	-	The mapping from input fields to columns in hive table. Specified as a comma separated list (no spaces) of hive table columns names, identifying the input fields in order of their occurrence. To skip fields leave the column name unspecified. E.g.. 'time,,IP,message' indicates the 1st, 3rd and 4th fields in input map to time, IP and message columns in the hive table.
serializer.serdeSeparator	Ctrl-A	(Type: character) Customizes the separator used by underlying serde. There can be a gain in efficiency if the fields in serializer.fieldnames are in same order as table columns, the serializer.delimiter is same as the serializer.serdeSeparator and number of fields in serializer.fieldnames is less than or equal to number



Property Name	Default	Description
		of table columns, as the fields in incoming event body do not need to be reordered to match order of table columns. Use single quotes for special characters like '\t'. Ensure input fields do not contain this character. <b>Note:</b> If serializer.delimiter is a single character, preferably set this to the same character.

The following are the escape sequences supported:

Alias	Description
%{host}	Substitute value of event header named "host". Arbitrary header names are supported.
%t	Unix time in milliseconds .
%a	Locale's short weekday name (Mon, Tue, ...)
%A	Locale's full weekday name (Monday, Tuesday, ...)
%b	Locale's short month name (Jan, Feb, ...)
%B	Locale's long month name (January, February, ...)
%c	Locale's date and time (Thu Mar 3 23:05:25 2005)
%d	Day of month (01)
%D	Date; same as %m/%d/%y
%H	Hour (00..23)
%I	Hour (01..12)
%j	Day of year (001..366)
%k	Hour ( 0..23)
%m	Month (01..12)
%M	Minute (00..59)
%p	Locale's equivalent of am or pm
%s	Seconds since 1970-01-01 00:00:00 UTC
%S	Second (00..59) %y last two digits of year (00..99)
%Y	Year (2015)
%z	+hhmm numeric timezone (for example, -0400)

Example Hive table:

```
create table weblogs ( id int , msg string )
partitioned by (continent string, country string, time string)
clustered by (id) into 5 buckets
stored as orc;
```

Example for agent named a1:

```
a1.channels = c1
a1.channels.c1.type = memory
a1.sinks = k1
```

```
al.sinks.k1.type = hive

al.sinks.k1.channel = c1

al.sinks.k1.hive.metastore = thrift://127.0.0.1:9083
al.sinks.k1.hive.database = logsdb

al.sinks.k1.hive.table = weblogs
al.sinks.k1.hive.partition = asia,${country},%y-%m-%d-%H-%M
al.sinks.k1.useLocalTimeStamp = false

al.sinks.k1.round = true

al.sinks.k1.roundValue = 10

al.sinks.k1.roundUnit = minute

al.sinks.k1.serializer = DELIMITED
al.sinks.k1.serializer.delimiter = "\t"
al.sinks.k1.serializer.serdeSeparator = '\t'
al.sinks.k1.serializer.fieldnames =id,,msg
```

**Note:** For all of the time related escape sequences, a header with the key “timestamp” must exist among the headers of the event (unless useLocalTimeStamp is set to true). One way to add this automatically is to use the TimestampInterceptor.

The above configuration will round down the timestamp to the last 10th minute. For example, an event with timestamp header set to 11:54:34 AM, June 12, 2012 and ‘country’ header set to ‘india’ will evaluate to the partition (continent=‘asia’,country=‘india’,time=‘2012-06-12-11-50’. The serializer is configured to accept tab separated input containing three fields and to skip the second field.

## 1.10.3. Configuring Pig Scripts to Use HCatalog in Oozie Workflows

To access HCatalog with a Pig action in an Oozie workflow, you need to modify configuration information to point to the Hive metastore URIs.

There are two methods for providing this configuration information. Which method you use depends upon how often your Pig scripts access the HCatalog.

### 1.10.3.1. Configuring Individual Pig Actions to Access HCatalog

If only a few individual Pig actions access HCatalog, do the following:

1. Identify the URI (host and port) for the Thrift Metastore server.
  - a. In Ambari, click **Hive > Configs > Advanced**.
  - b. Make note of the URI in the **hive.metastore.uris** field in the General section.

This information is also stored in the `hive.default.xml` file.

2. Add the following two properties to the <configuration> elements in each Pig action.



## Note

Replace `[host:port(default:9083)]` in the example below with the host and port for the Thrift metastore server.

```
<configuration>
  <property>
    <name>hive.metastore.uris</name>
    <value>thrift://[host:port(default:9083)]</value>
    <description>A comma separated list of Metastore uris the client can
use to contact the
Metastore server.</description>
  </property>
  <property>
    <name>oozie.action.sharelib.for.pig</name>
    <value>pig,hive,hcatalog</value>
    <description>A comma separated list of libraries to be used by the
Pig action.</description>
  </property>
</configuration>
```

### 1.10.3.2. Configuring All Pig Actions to Access HCatalog

If all of your Pig actions access HCatalog, do the following:

1. Add the following line to the `job.properties` files, located in your working directory:

```
oozie.action.sharelib.for.pig=pig,hive,hcatalog
<!-- A comma separated list of libraries to be used by the Pig action.-->
```

2. Identify the URI (host and port) for the Thrift Metastore server.

- a. In Ambari, click **Hive > Configs > Advanced**.
- b. Make note of the URI in the `hive.metastore.uris` field in the General section.

This information is also stored in the `hive.default.xml` file.

3. Add the following property to the `<configuration>` elements in each Pig action.



## Note

Replace `[host:port(default:9083)]` in the example below with the host and port for the Thrift Metastore server.

```
<configuration>
  <property>
    <name>hive.metastore.uris</name>
    <value>thrift://[host:port(default:9083)]</value>
    <description>A comma separated list of Metastore uris the client can
use to contact the
Metastore server.</description>
  </property>
</configuration>
```

## 1.10.4. Configuring a Sqoop Action to Use Tez to Load Data into a Hive Table

You can use the Tez execution engine to load data into a Hive table using the `--hive-import` option,

In the code example in each step, replace the sample text in [square brackets] with the appropriate information for your configuration.

1. Create a workflow directory.

```
hdfs dfs -mkdir -p [/user/dummy/app]
```

2. Create a lib directory in the workflow directory.

```
hdfs dfs -mkdir -p [/user/dummy/app/lib]
```

3. Copy the database JDBC driver jar file to the lib directory.

```
hadoop fs -copyFromLocal [/usr/share/java/mysql-connector-java.jar]
[/user/dummy/app/lib]
```

4. Copy the `hive-site.xml` and `tez-site.xml` files to a location accessible by the workflow. For example:

```
hadoop fs -copyFromLocal [/etc/oozie/conf/action-conf/hive/hive-site.xml /
user/dummy/app]
hadoop fs -copyFromLocal [/etc/oozie/conf/action-conf/hive/tez-site.xml /
user/dummy/app]
```

5. In the Sqoop action of the workflow, do the following:

- Add `hive-site` and `tez-site` resources in the `<file>` element of the Sqoop action in the workflow.

```
<file>/user/dummy/app/hive-site.xml#hive-site.xml</file>
<file>/user/dummy/app/tez-site.xml#tez-site.xml</file>
```

- Include the `--hive-import` option in the `<command>` element.

```
<command>import --connect [jdbc:mysql://db_host:port/database] --
username [user]
--password [pwd] --driver c[om.mysql.jdbc.Driver] --table [table_name]
--hive-import -m 1 </command>
```

6. Add the following into the `job.properties` file.

```
oozie.use.system.libpath=true
oozie.action.sharelib.for.sqoop=sqoop,hive
```

More information regarding the Sqoop parameters can be found in the Apache documentation at [https://sqoop.apache.org/docs/1.4.6/SqoopUserGuide.html#\\_importing\\_data\\_into\\_hive](https://sqoop.apache.org/docs/1.4.6/SqoopUserGuide.html#_importing_data_into_hive)

### Example Workflow Action

Replace all sample text in [square brackets] in the example below with the appropriate workflow name, URI, paths, file names, etc. for your configuration.

```
<action name="sqoop-node">
  <sqoop xmlns="uri:oozie:sqoop-action:0.2">
    <job-tracker>${jobTracker}</job-tracker>
    <name-node>${nameNode}</name-node>
    <configuration>
      <property>
        <name>mapred.job.queue.name</name>
        <value>${queueName}</value>
      </property>
    </configuration>
    <command>import --connect [jdbc:mysql://db_host:port/database] --
username [user]
--password [pwd] --driver [com.mysql.jdbc.Driver] --table [table_name] --hive-
import -m 1</command>
    <file>[/user/dummy/app/hive-site.xml#hive-site.xml]</file>
    <file>[/user/dummy/app/tez-site.xml#tez-site.xml]</file>
  </sqoop>
  <ok to="end"/>
  <error to="killJob"/>
</action>
```