Hortonworks Data Platform

Release Notes

(April 3, 2017)

docs.hortonworks.com

Hortonworks Data Platform: Release Notes

Copyright © 2012-2017 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.

Table of Contents

List of Tables

1. HDP 2.6.0 Release Notes

This document provides you with the latest information about the HDP 2.6.0 release and its product documentation.

1.1. Component Versions

The official Apache versions of all HDP 2.6.0 components are listed below. All components listed here are official Apache releases of the most recent stable versions available.

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 2.6.0.

Official Apache versions for HDP 2.6.0:

- Apache Accumulo 1.7.0[1]

- Apache Atlas 0.8.0

- Apache DataFu 1.3.0

- Apache Falcon 0.10.0 [1]

- Apache Flume 1.5.2 [1]

- Apache Hadoop 2.7.3

- Apache HBase 1.1.2

- Apache Hive 1.2.1

- Apache Hive 2.1.0

- Apache Kafka 0.10.1.1 [1]

- Apache Knox 0.12.0

- Apache Mahout 0.9.0+ [1]

- Apache Oozie 4.2.0

- Apache Phoenix 4.7.0

- Apache Pig 0.16.0

- Apache Ranger 0.7.0

- Apache Slider 0.92.0 [1]

- Apache Spark 1.6.3

- Apache Spark 2.1.0

---

[1]Refer to Deprecation Notices for more information on deprecation plans for these componennts.

- Apache Sqoop 1.4.6

- Apache Storm 1.1.0 [1]

- Apache TEZ 0.7.0

- Apache Zeppelin 0.7.0

- Apache ZooKeeper 3.4.6

Later versions of a few Apache components are sometimes bundled in the HDP distribution in addition to the versions listed above. In this case, these later versions are listed in the Technical Previews table and should not substitute for the Apache component versions of the above list in a production environment.

Additional component versions:

- Cascading 3.0.0 [1]

- Hue 2.6.1 [1][2]

Note

For information on open source software licensing and notices, please refer to the Licenses and Notices files included with the software install package.

1.2. New Features

This section highlights new features in HDP 2.6.0.

Apache Components	Feature
Atlas	Enterprise Readiness • Atlas to Support KnoxSSO Authentication (AMBARI-18818, ATLAS-1244) Cloud • Shared Services HDC: Atlas (AMBARI-19007) • Install & configure Atlas hook components independently of Atlas admin components Core Capabilities • Manually create entities (ATLAS-1193) • Higher level API - Simplify and Streamline usecases • Integration of ATLAS with Hive2 (ATLAS-1280) Integration • SmartSense Data Capture for Atlas (ATLAS-1436)
HDFS	Enterprise Readiness • Detect slow disks and slow Datanodes.
Hive	HDP 2.6.0 includes two versions of Apache Hive, one based on the Hive 1 branch, meant for batch processing and one based on the Hive 2 branch, meant for interactive processing. Hortonworks applies stabilizations, bug fixes and features from the trunk code base as deemed appropriate, so neither of these is identical to any specific released build in Apache. Specific details on the differences are provided here.

---

[2]Hue is not supported on IBM PPC.

Apache Components	Feature
	HDP deployments driven by Apache Ambari can deploy both Hive 1 and Hive 2 side-by-side. In this setup, Hive 2 features are accessed through the Hive LLAP interface. • Hive LLAP (Hive 2 only) • ACID Merge (Hive 1 and 2)
Hue	Ease of Use • GA criteria for Hue migration tool
Knox	Security Features • Knox support for Hadoop User-group mapping (KNOX-237) • Support for Linux user authentication via PAM (KNOX-537) Integration • Add support for WebSockets to Knox
Phoenix	Reliability • Phoenix index stability improvements
Ranger	Security Features • Tag Based security expand for existing type common to Atlas + Ranger : Kafka, HDFS, Hbase (RANGER-1233) • Ranger Authorization for "show Databases/Tables/Columns/Views" or "describe database/schema/table" permissions on HiveServer2 (RANGER-1195, RANGER-1204) • Support for $username variable in Ranger policies (RANGER-698) Integration • SmartSense Rules: Product Analytics for Ranger (RANGER-1304) Enterprise Readiness • Support Governance and security features for Beacon - Ranger • Policy versions in use for each Ranger plugin (RANGER-1230) • UI hints for Audit search for Ranger Audit Logs & Policy Search on Report Manager or Policy Listing Page (RANGER-1227) • Change Ranger Admin start script to setup db for ranger (AMBARI-18065) • Export/Import of policies in Ranger (RANGER-1214) • Ranger UserSync improvement - Incremental, group filtered usersync with additional attributes (RANGER-1211)
Spark	Ease of Use • [Codev 2016-17] Support virtualenv in PySpark (SPARK-13587) Integration • Umbrella - Provide Spark 2 GA on HDP: Features across Spark, Livy & Zeppelin components • Spark HBase connector enhancements Enterprise readiness, security • Spark Thrift Server doAs support for identity propagation, for Spark SQL queries • Livy API (version 0.3) for REST-based access to Spark and Spark2 jobs, including support for Kerberos authentication and high availability

Apache Components	Feature
	Core capabilities • Support for user-defined functions in SparkR • Support in SparkR for additional large-scale machine learning algorithms
Storm	Integration • Connectivity Enhancements: Kinesis spout, openTSDB bolt, new Kafka Consumer spout
YARN	Enterprise Readiness • Queue Priorities to allow High Priority Services (e.g. LLAP) run in a shared cluster (YARN-5864) • Fault tolerance of Slider apps • Support port range for YARN apps
Zeppelin	Core capabilities • Zeppelin to work with Spark version 2.x Ease of Use • Improve JDBC interpreter • Improve Zeppelin Livy integration • Support multiple SQL statements in one notebook paragraph Enterprise readiness, security • Knox-based LDAP authentication (Zeppelin-1472) • Improvements to LDAP authentication (Zeppelin-1611) Integration • Umbrella - Provide Spark 2 GA on HDP: Features across Spark, Livy & Zeppelin components • Support for Phoenix interpreter

1.3. Deprecation Notices

This section points out any technology from previous releases that has been deprecated, moved, or removed from this release. Use this section as a guide for your implementation plans.

1.3.1. Terminology

Items in this section are designated in the following way:

Deprecated   Technology that Hortonworks is removing in a future HDP release. Marking an item as deprecated gives you time to plan for removal in a future HDP release.

Moving   Technology that Hortonworks is moving from a future HDP release and is making available via an alternative Hortonworks offering or subscription. Marking an item as moving gives you time to plan for removal in a future HDP release and plan for the alternative Hortonworks offering or subscription for the technology.

Removed   Technology that Hortonworks has removed from HDP and is no longer available or supported as of this release. Take note of technology marked as Removed since it can potentially affect your upgrade plans.

1.3.2. Deprecated Components and Product Capabilities

The following components are **marked deprecated** from HDP and will be **removed** in a future HDP release:

Component or Capability	Status	Marked Deprecated as of	Target Release for Removal	Comments
Apache Falcon	Deprecated	HDP 2.6.0	HDP 3.0.0	
Apache Flume	Deprecated	HDP 2.6.0	HDP 3.0.0	Consider Hortonworks DataFlow as an alternative for Flume use cases.
Apache Mahout	Deprecated	HDP 2.6.0	HDP 3.0.0	Consider Apache Spark as an alternative depending on the workload.
Apache Slider	Deprecated	HDP 2.6.0	HDP 3.0.0	Apache Slider functionality will be absorbed by Apache YARN.
Cascading	Deprecated	HDP 2.6.0	HDP 3.0.0	
Hue	Deprecated	HDP 2.6.0	HDP 3.0.0	Consider Ambari Views as the alternative.

The following components are **marked moving** from HDP and will be **moved** in a future HDP release to an alternative Hortonworks Subscription and Offering:

Component or Capability	Status	Marked Moving as of	Target Release for Move
Apache Accumulo	Moving	HDP 2.6.0	HDP 3.0.0
Apache Kafka	Moving	HDP 2.6.0	HDP 3.0.0
Apache Storm	Moving	HDP 2.6.0	HDP 3.0.0
Cloudbreak	Moving	HDP 2.6.0	HDP 3.0.0

1.3.3. Deprecated OSs, Java Versions, Databases, and Platforms

The following technologies are **deprecated** from HDP and will be **removed** in a future HDP release:

Technology	Status	Marked Deprecated as of	Target Release for Removal
Oracle JDK 1.7 OpenJDK 7	Deprecated	HDP 2.6.0	HDP 3.0.0
Ubuntu 12	Deprecated	HDP 2.6.0	HDP 3.0.0

Please contact Hortonworks Support or your Hortonworks Account Team if you have any questions.

1.4. Unsupported Features

Although the following features exist within HDP 2.6.0, Hortonworks does not currently support these specific capabilities:

- Technical Preview Features

- Community Features

1.4.1. Technical Preview Features

The following features are available within HDP 2.6.0 but are not ready for production deployment. Hortonworks encourages you to explore these technical preview features in non-production environments and provide feedback on your experiences through the Hortonworks Community Forums.

Table 1.1. Technical Previews

Component	Feature
Druid	New in this release: • Druid 0.9.2
HBase and Phoenix	New in this release: • HBase storage quota on HDFS • HBase backup/restore supports bulk loaded data • Phoenix repair tool - A repair tool for corrupted system tables (PHOENIX-3165) Introduced in a previous release: • Backup and restore • Phoenix operability with Microsoft .NET Driver • Phoenix Storage Handler for Hive integration • Cell-level ACLs • Column Family Encryption • Phoenix-Spark Integration • RPC Throttling • Support for `init.d` scripts
HDFS	New in this release: • G1 GC support for HDFS
Hive	Introduced in a previous release: • Stored Procedures Client-Side • LLAP / Spark security interface • Phoenix Storage Handler for Hive integration
Slider	Introduced in a previous release: • Support for Docker-based application packaging (SLIDER-780)

Component	Feature
Spark	New in this release: • Spark SQL Ranger integration for row and column security, for Spark 2.1 Introduced in a previous release: • DataSet API • GraphX
Storm	New in this release: • Storm SQL: run SQL queries over streaming data in Storm Introduced in a previous release: • Automatic back pressure • Pacemaker daemon • Resource-aware scheduling
YARN	Introduced in a previous release: • Add support for network I/O isolation/scheduling for containers (YARN-2140) • NodeManager: add cgroup support for disk I/O isolation (YARN-2619)

1.4.2. 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 non-production 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

Component	Feature
Falcon	Introduced in a previous release: • Prism Server • User Recipes
HBase	Introduced in a previous release: • HBase Column Family Encryption: use HDFS data at rest encryption instead • Use of memcached as block cache is unsupported (HBASE-13170) • ZooKeeper-less region assignment • Region size balancing (HBASE-13103)
HDFS	Introduced in a previous release: • HTTPFS (HDFS-2179)

Component	Feature
	HTTPFS is used with Hue, in combination with Knox, and standalone as an API gateway. We highly recommend that you replace HTTPFS with Knox where possible. • NameNode Federation (HDFS-1052) • Disk Hot Swap (HDFS-1362) • Block-volume device choosing (HDFS-1804) • viewFS (HADOOP-7257)
Knox	Introduced in a previous release: • Storm REST APIs
Slider	Introduced in a previous release: • Simplified Application Packaging
Spark	Introduced in a previous release: • Spark Standalone • Spark on Mesos • Jupyter/iPython Notebook
YARN	Introduced in a previous release: • Fair Scheduler • MapReduce Eclipse Plug-in • MapReduce Uber AM
Zeppelin	For R interpreter support (ZEPPELIN-156), only the following subset of the interpreters are enabled by default: • Spark • Hive • Shell • MarkDown • R • Livy • Angular The following interpreters are not enabled by default: • Phoenix

1.5. Upgrading from HDP 2.5.x to HDP 2.6.0

1.5.1. Before Upgrading

If you are managing your cluster manually (i.e. without Ambari) and you are running HDP 2.5.x with Ranger, to avoid timeouts for databases managing a large amount of users, manually execute one of the following SQL statements on Ranger db before starting the upgrade.

DB Flavor	SQL Statement
MySQL	`ALTER TABLE x_user MODIFY COLUMN user_name varchar(767) NOT NULL, ADD CONSTRAINT x_user_UK_user_name UNIQUE(user_name(767));`
Oracle	`ALTER TABLE x_user MODIFY(user_name VARCHAR(767)) ADD CONSTRAINT x_user_UK_user_name UNIQUE (user_name);`
Postgres	`ALTER TABLE x_user ALTER COLUMN user_name TYPE VARCHAR(767),ADD CONSTRAINT x_user_UK_user_name UNIQUE(user_name);`
SQLServer	`ALTER TABLE [dbo].[x_user] ALTER COLUMN [user_name] [varchar](767) NOT NULL; ALTER TABLE [dbo].[x_user] ADD CONSTRAINT [x_user$x_user_UK_user_name] UNIQUE ([user_name]);`
SQL Anywhere	`ALTER TABLE dbo.x_user ALTER user_name varchar(767) NOT NULL;ALTER TABLE dbo.x_user ADD CONSTRAINT x_user_UK_user_name UNIQUE NONCLUSTERED (user_name);`

Note

Execution of SQL statement on x_user table having one million users should not take more than ~30 seconds.

1.5.2. Upgrade Options

HDP 2.6 is a minor release that contains new features. If you are managing your cluster manually (i.e. without Ambari), follow instructions on how to perform a manual upgrade in the Command-Line Upgrade guide.

If you are managing your cluster with Ambari, this table summarize the various Upgrade Options.

Table 1.3. HDP Upgrade Options

Upgrade Path	Current HDP[a]	Target HDP [a]	Rolling Upgrade	Express Upgrade
Minor [b]	HDP 2.5	HDP 2.6	✅	✅
Minor [b]	HDP 2.4	HDP 2.5 or 2.6	✅	✅
Minor [b]	HDP 2.3	HDP 2.4, 2.5 or 2.6	✅	✅
Maintenance	HDP 2.5.x.y	HDP 2.5.x.y	✅	✅
Maintenance	HDP 2.4.x.y	HDP 2.4.x.y	✅	✅
Maintenance	HDP 2.3.x.y	HDP 2.3.x.y	✅	✅

[a]The instructions in this document sometimes refer to an HDP "version" using the HDP #.#.x.y convention. For example, you will see the HDP 2.3.x.y convention to refer to any HDP 2.3 "maintenance" release. Therefore, to use a specific HDP 2.3 maintenance

release, be sure to replace 2.3.x.y in the following instructions with the appropriate maintenance version, such as 2.3.0.0 for the HDP 2.3 GA release, or 2.3.2.0 for an HDP 2.3 maintenance release.

[b]A Minor upgrade can be performed from any current maintenance release to any target maintenance release in that minor release. For example, you can go from HDP 2.3.x.y to HDP 2.4.x.y as part of a Minor upgrade.

> ### Important
>
> Because HDP 2.3.6 contains features and fixes that are not applicable to HDP 2.4.0 or HDP 2.4.2, do not upgrade from HDP 2.3.6 to HDP 2.4.0 or HDP 2.4.2.

1.6. HDP 2.6.0 Repositories

HDP 2.6.0 for Linux supports the following:

- Debian 7

- RHEL/CentOS/Oracle LINUX 6

- RHEL/CentOS/Oracle LINUX 7

- SLES 12

- SLES 11 SP3/SP4

- Ubuntu 12

- Ubuntu 14

- Ubuntu 16

> ### Note
>
> The package identifier for HDP 2.6.0 components is **8**. For example:
>
> ```
> dfs -mkdir -p /hdp/apps/2.6.0.3-$BUILD/hive/
> ```
>
> would become:
>
> ```
> dfs -mkdir -p /hdp/apps/2.6.0.3-8/hive/
> ```

Table 1.4. Debian 7

Description	Link
Component metadata	http://public-repo-1.hortonworks.com/HDP/debian7/2.x/updates/2.6.0.3/build_metadata.txt
HDP Tars tarball	http://public-repo-1.hortonworks.com/HDP/debian7/2.x/updates/2.6.0.3/HDP-2.6.0.3-debian7-tars-tarball.tar.gz
HDP repo	http://public-repo-1.hortonworks.com/HDP/debian7/2.x/updates/2.6.0.3/hdp.list
HDP RPM tarball	http://public-repo-1.hortonworks.com/HDP/debian7/2.x/updates/2.6.0.3/HDP-2.6.0.3-debian7-deb.tar.gz
Slider app packages	http://public-repo-1.hortonworks.com/HDP/debian7/2.x/updates/2.6.0.3/slider-app-packages/pkg-list_public.txt

Table 1.5. RHEL/CentOS/Oracle LINUX 6

Description	Link
Component metadata	http://public-repo-1.hortonworks.com/HDP/centos6/2.x/updates/2.6.0.3/build_metadata.txt
HDP Tars tarball	http://public-repo-1.hortonworks.com/HDP/centos6/2.x/updates/2.6.0.3/HDP-2.6.0.3-centos6-tars-tarball.tar.gz
HDP repo	http://public-repo-1.hortonworks.com/HDP/centos6/2.x/updates/2.6.0.3/hdp.repo
HDP RPM tarball	http://public-repo-1.hortonworks.com/HDP/centos6/2.x/updates/2.6.0.3/HDP-2.6.0.3-centos6-rpm.tar.gz
Slider app packages	http://public-repo-1.hortonworks.com/HDP/centos6/2.x/updates/2.6.0.3/slider-app-packages/pkg-list.txt

Table 1.6. RHEL/CentOS/Oracle LINUX 7

Description	Link
Component metadata	http://public-repo-1.hortonworks.com/HDP/centos7/2.x/updates/2.6.0.3/build_metadata.txt
HDP Tars tarball	http://public-repo-1.hortonworks.com/HDP/centos7/2.x/updates/2.6.0.3/HDP-2.6.0.3-centos7-tars-tarball.tar.gz
HDP repo	http://public-repo-1.hortonworks.com/HDP/centos7/2.x/updates/2.6.0.3/hdp.repo
HDP RPM tarball	http://public-repo-1.hortonworks.com/HDP/centos7/2.x/updates/2.6.0.3/HDP-2.6.0.3-centos7-rpm.tar.gz
Slider app packages	http://public-repo-1.hortonworks.com/HDP/centos7/2.x/updates/2.6.0.3/slider-app-packages/pkg-list.txt

Table 1.7. SLES 11 SP3/SP4

Description	Link
Component metadata	http://public-repo-1.hortonworks.com/HDP/suse11sp3/2.x/updates/2.6.0.3/build_metadata.txt
HDP Tars tarball	http://public-repo-1.hortonworks.com/HDP/suse11sp3/2.x/updates/2.6.0.3/HDP-2.6.0.3-suse11sp3-tars-tarball.tar.gz
HDP repo	http://public-repo-1.hortonworks.com/HDP/suse11sp3/2.x/updates/2.6.0.3/hdp.repo
HDP RPM tarball	http://public-repo-1.hortonworks.com/HDP/suse11sp3/2.x/updates/2.6.0.3/HDP-2.6.0.3-suse11sp3-rpm.tar.gz
Slider app packages	http://public-repo-1.hortonworks.com/HDP/suse11sp3/2.x/updates/2.6.0.3/slider-app-packages/pkg-list.txt

Table 1.8. SLES 12

Description	Link
Component metadata	http://public-repo-1.hortonworks.com/HDP/sles12/2.x/updates/2.6.0.3/build_metadata.txt
HDP Tars tarball	http://public-repo-1.hortonworks.com/HDP/sles12/2.x/updates/2.6.0.3/HDP-2.6.0.3-sles12-tars-tarball.tar.gz
HDP repo	http://public-repo-1.hortonworks.com/HDP/sles12/2.x/updates/2.6.0.3/hdp.repo

Description	Link
HDP RPM tarball	http://public-repo-1.hortonworks.com/HDP/sles12/2.x/updates/2.6.0.3/HDP-2.6.0.3-sles12-rpm.tar.gz
Slider app packages	http://public-repo-1.hortonworks.com/HDP/sles12/2.x/updates/2.6.0.3/slider-app-packages/pkg-list.txt

Table 1.9. Ubuntu 12

Description	Link
Component metadata	http://public-repo-1.hortonworks.com/HDP/ubuntu12/2.x/updates/2.6.0.3/build_metadata.txt
HDP Tars tarball	http://public-repo-1.hortonworks.com/HDP/ubuntu12/2.x/updates/2.6.0.3/HDP-2.6.0.3-ubuntu12-tars-tarball.tar.gz
HDP repo	http://public-repo-1.hortonworks.com/HDP/ubuntu12/2.x/updates/2.6.0.3/hdp.list
HDP RPM tarball	http://public-repo-1.hortonworks.com/HDP/ubuntu12/2.x/updates/2.6.0.3/HDP-2.6.0.3-ubuntu12-deb.tar.gz
Slider app packages	http://public-repo-1.hortonworks.com/HDP/ubuntu12/2.x/updates/2.6.0.3/slider-app-packages/pkg-list.txt

Table 1.10. Ubuntu 14

Description	Link
Component metadata	http://public-repo-1.hortonworks.com/HDP/ubuntu14/2.x/updates/2.6.0.3/build_metadata.txt
HDP Tars tarball	http://public-repo-1.hortonworks.com/HDP/ubuntu14/2.x/updates/2.6.0.3/HDP-2.6.0.3-ubuntu14-tars-tarball.tar.gz
HDP repo	http://public-repo-1.hortonworks.com/HDP/ubuntu14/2.x/updates/2.6.0.3/hdp.list
HDP RPM tarball	http://public-repo-1.hortonworks.com/HDP/ubuntu14/2.x/updates/2.6.0.3/HDP-2.6.0.3-ubuntu14-deb.tar.gz
Slider app packages	http://public-repo-1.hortonworks.com/HDP/ubuntu14/2.x/updates/2.6.0.3/slider-app-packages/pkg-list.txt

Table 1.11. Ubuntu 16

Description	Link
Component metadata	http://public-repo-1.hortonworks.com/HDP/ubuntu16/2.x/updates/2.6.0.3/build_metadata.txt
HDP Tars tarball	http://public-repo-1.hortonworks.com/HDP/ubuntu16/2.x/updates/2.6.0.3/HDP-2.6.0.3-ubuntu16-tars-tarball.tar.gz
HDP repo	http://public-repo-1.hortonworks.com/HDP/ubuntu16/2.x/updates/2.6.0.3/hdp.list
HDP RPM tarball	http://public-repo-1.hortonworks.com/HDP/ubuntu16/2.x/updates/2.6.0.3/HDP-2.6.0.3-ubuntu16-deb.tar.gz
Slider app packages	http://public-repo-1.hortonworks.com/HDP/ubuntu16/2.x/updates/2.6.0.3/slider-app-packages/pkg-list.txt

Table 1.12. Helper Files

Description	Link
Manual install helper files TAR.GZ	http://public-repo-1.hortonworks.com/HDP/tools/2.6.0.3/hdp_manual_install_rpm_helper_files-2.6.0.3.8.tar.gz
Manual install helper files MD5	http://public-repo-1.hortonworks.com/HDP/tools/2.6.0.3/hdp_manual_install_rpm_helper_files-2.6.0.3.8.tar.gz.md5
Manual install helper files ASC	http://public-repo-1.hortonworks.com/HDP/tools/2.6.0.3/hdp_manual_install_rpm_helper_files-2.6.0.3.8.tar.gz.asc

1.7. Behavioral Changes

Behavioral changes denote a marked change in behavior from the previously released version to this version of software. In HDP 2.6.0, behavioral changes affect the following Hadoop components.

Table 1.13. Behavioral Changes

Hortonworks Bug ID	Apache Component	Apache JIRA	Summary	Details
BUG-68588	Spark	N/A	Livy Interpreter should run in Scoped mode	**Component Affected:** Livy Interpreter in Zeppelin **Scenario:** The Livy interpreter has been refactored in HDP 2.6.0 to change the mode from shared to per user scoped. Shared mode does not work for HDP 2.6.0. **Previous Behavior:** Use shared mode in HDP 2.5. **New Behavior:** Use per user scoped mode in HDP 2.6. **Workaround:** Change the mode to per user scoped mode when upgrading from HDP 2.5.x to HDP 2.6.0.
BUG-69438	Zeppelin	AMBARI-18975	Remove Zeppelin View from Ambar	**Component Affected:** Zeppelin, Ambari Views **Scenario:** Zeppelin view on Ambari is deprecated. **Previous Behavior:** Zeppelin view was available in the list of views on Ambari UI. **New Behavior:** There is no view for Zeppelin. Users can access Zeppelin via Quick links or by directly going to the configured Zeppelin hostname and port.
BUG-72046	Oozie Spark action	N/A	Yarn-client mode for the Spark action is deprecated	**Component Affected:** Spark action in Oozie **Scenario:** The yarn-client mode of the Spark action has been deprecated in HDP 2.6.0 and will be removed in HDP 3.0.0. **New Behavior:** No change in current behavior for HDP 2.6.0. For HDP 3.0.0, yarn-client mode for the Spark action in Oozie will no longer be available.
BUG-76651	Spark	N/A	Behavior change of Livy interpreter	**Component Affected:** Spark **Scenario:** Livy Interpreter in Zeppelin **Previous Behavior:** In HDP 2.5. 0, Livy would parse the code line by line, streaming the output to the frontend user as it is received.

Hortonworks Bug ID	Apache Component	Apache JIRA	Summary	Details
				New Behavior: In HDP 2.6.0, Livy recieves the code in blocks. This means the output is only displayed when the entire block is recieved.
RMP-8089	Zeppelin	N/A	Ambari service definition to change to work with Azure LLAP Clusters	**Component Affected:** Zeppelin **Scenario:** In the previous release of Zeppelin in HDP 2.5, Spark was a required service to install Zeppelin. **Previous Behavior:** On Ambari install wizard Spark was selected during the installation process. **New Behavior:** There is no Spark dependency while running Ambari install wizard. Spark/Spark 2 are optional and are required only if the user wish to run Spark Interpreter in Zeppelin.

1.8. Apache Patch Information

The following sections list patches in each HDP 2.6.0 component beyond what was fixed in the base version of the Apache component.

- Hadoop

- Accumulo

- Atlas

- DataFu

- Falcon

- Flume

- HBase

- Hive

- Kafka

- Knox

- Mahout

- Oozie

- Phoenix

- Pig

- Ranger

- Slider

- Spark

- Sqoop

- Storm

- Tez

- Zeppelin

- ZooKeeper

1.8.1. Hadoop

This release provides the following Apache patches:

- HADOOP-1023: Add more logging in WASB page blob read function.

- HADOOP-1026: Logging bug in WASB causing out of index exception.

- HADOOP-1540: breaks backward compatibilty.

- HADOOP-1540: Support file exclusion list in distcp.

- HADOOP-3733: "s3x:" URLs break when Secret Key contains a slash, even if encoded.

- HADOOP-5323: Trash documentation should describe its directory structure and configurations.

- HADOOP-5470: RunJar.unJar() should write the last modified time found in the jar entry to the uncompressed file.

- HADOOP-7139: Allow appending to existing SequenceFiles

- HADOOP-7256: Resource leak during failure scenario of closing of resources.

- HADOOP-7713: dfs -count -q should label output column

- HADOOP-7713: dfs -count -q should label output column .

- HADOOP-7817: RawLocalFileSystem.append() should give FSDataOutputStream with accurate .getPos()

- HADOOP-7930: Kerberos relogin interval in UserGroupInformation should be configurable

- HADOOP-8151: Error handling in snappy decompressor throws invalid exceptions.

- HADOOP-8436: NPE In getLocalPathForWrite

- HADOOP-8437: getLocalPathForWrite should throw IOException for invalid paths.

- HADOOP-8751: NPE in Token.toString() when Token is constructed using null identifier.

- HADOOP-8818: Use equals instead == in MD5MD5CRC32FileChecksum and TFileDumper.

- HADOOP-8934: Shell command ls should include sort options

- HADOOP-9121: InodeTree.java has redundant check for vName while throwing exception.

- HADOOP-9242: Duplicate surefire plugin config in hadoop-common.

- HADOOP-9477: Add posixGroups support for LDAP groups mapping service.

- HADOOP-10035: Cleanup TestFilterFileSystem.

- HADOOP-10048: LocalDirAllocator should avoid holding locks while accessing the filesystem.

- HADOOP-10300: Allowed deferred sending of call responses.

- HADOOP-10365: BufferedOutputStream in FileUtil#unpackEntries() should be closed in finally block.

- HADOOP-10406: TestIPC.testIpcWithReaderQueuing may fail.

- HADOOP-10597: RPC Server signals backoff to clients when all request queues are full.

- HADOOP-10597: RPC Server signals backoff to clients when all request queues are full.

- HADOOP-1062:added hadoop-azure-datalake and azure-data-lake-store-sdk to all the wasb locations.

- HADOOP-10775: Shell operations to fail with meaningful errors on windows if winutils.exe not found.

- HADOOP-10823: TestReloadingX509TrustManager is flaky.

- HADOOP-10965: Print fully qualified path in CommandWithDestination error messages.

- HADOOP-10971: Add -C flag to make `hadoop fs -ls` print filenames only.

- HADOOP-11031: Design Document for Credential Provider API.

- HADOOP-11098: [JDK8] Max Non Heap Memory default changed between JDK7 and 8.

- HADOOP-11212: NetUtils.wrapException to handle SocketException explicitly.

- HADOOP-11218: Add TLSv1.1,TLSv1.2 to KMS, HttpFS, SSLFactory.

- HADOOP-11252: RPC client does not time out by default

- HADOOP-11262: Enable YARN to use S3A.

- HADOOP-11361: Fix a race condition in MetricsSourceAdapter.updateJmxCache.

- HADOOP-11404: Clarify the "expected client Kerberos principal is null" authorization message.

- HADOOP-11491: HarFs incorrectly declared as requiring an authority.

- HADOOP-11581: Multithreaded correctness Warnings #org.apache.hadoop.fs.shell.Ls

- HADOOP-11613: Remove commons-httpclient dependency from hadoop-azure.

- HADOOP-11628: SPNEGO auth does not work with CNAMEs in JDK8.

- HADOOP-11685: StorageException complaining "no lease ID" during HBase distributed log splitting.

- HADOOP-11687: Ignore x-* and response headers when copying an Amazon S3 object.

- HADOOP-11692: Improve authentication failure WARN message to avoid user confusion.

- HADOOP-11713: ViewFileSystem should support snapshot methods.

- HADOOP-11717: Add Redirecting WebSSO behavior with JWT Token in Hadoop Auth

- HADOOP-11772: RPC Invoker relies on static ClientCache which has synchronized(this) blocks.

- HADOOP-11780: Prevent IPC reader thread death.

- HADOOP-11785: Reduce the number of listStatus operation in distcp buildListing

- HADOOP-11802: DomainSocketWatcher thread terminates sometimes after there is an I/O error during requestShortCircuitShm

- HADOOP-11812: Implement listLocatedStatus for ViewFileSystem to speed up split calculation

- HADOOP-11827: Speed-up distcp buildListing() using threadpool

- HADOOP-11859: PseudoAuthenticationHandler fails with httpcomponents v4.4.

- HADOOP-11864: JWTRedirectAuthenticationHandler breaks java8 javadocs.

- 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-11918: Listing an empty s3a root directory throws FileNotFound.

- HADOOP-11924: Tolerate JDK-8047340-related exceptions in Shell#isSetSidAvailable preventing class init.

- HADOOP-11932: MetricsSinkAdapter may hang when being stopped.

- HADOOP-11960: Enable Azure-Storage Client Side logging.

- HADOOP-12001: Fixed LdapGroupsMapping to include configurable Posix UID and GID attributes during the search.

- HADOOP-12006: Remove unimplemented option for from document in branch-2.7.

- HADOOP-12009: Clarify FileSystem.listStatus() sorting order & fix FileSystemContractBaseTest:testListStatus. backport to 2.8.

- HADOOP-12017: Hadoop archives command should use configurable replication factor when closing

- HADOOP-12043: Display warning if defaultFs is not set when running fs commands.

- HADOOP-12045: Enable LocalFileSystem#setTimes to change atime.

- HADOOP-12049: Control http authentication cookie persistence via configuration.

- HADOOP-12052:IPC client downgrades all exception types to IOE, breaks callers trying to use them.

- HADOOP-12054: RPC client should not retry for InvalidToken exceptions.

- HADOOP-12056: Use DirectoryStream in DiskChecker#checkDirs to detect errors when listing a directory.

- HADOOP-12058: Fix dead links to DistCp and Hadoop Archives pages.

- HADOOP-12059: S3Credentials should support use of CredentialProvider.

- HADOOP-12061: Incorrect command in single cluster setup document.

- HADOOP-12073: Azure FileSystem PageBlobInputStream does not return -1 on EOF.

- HADOOP-12074: in Shell.java#runCommand() rethrow InterruptedException as InterruptedIOException

- HADOOP-12076: Incomplete Cache Mechanism in CredentialProvider API.

- HADOOP-12089: StorageException complaining " no lease ID" when updating FolderLastModifiedTime in WASB.

- HADOOP-12095: org.apache.hadoop.fs.shell.TestCount fails.

- HADOOP-12097: Allow port range to be specified while starting webapp.

- 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-12119: hadoop fs -expunge does not work for federated namespace

- HADOOP-12158: Improve error message in TestCryptoStreamsWithOpensslAesCtrCryptoCodec when OpenSSL is not installed.

- HADOOP-12159: Move DistCpUtils#compareFs() to org.apache.hadoop.fs.FileUtil and fix for HA namespaces

- HADOOP-12161: Add getStoragePolicy API to the FileSystem interface

- HADOOP-12169:ListStatus on empty dir in S3A lists itself instead of returning an empty list. author: Pieter Reuse.

- HADOOP-12185: NetworkTopology is not efficient adding/getting/removing nodes.

- HADOOP-12186: ActiveStandbyElector shouldn't call monitorLockNodeAsync multiple times

- HADOOP-12189: Improve CallQueueManager#swapQueue to make queue elements drop nearly impossible.

- HADOOP-12191: Bzip2Factory is not thread safe.

- HADOOP-12213: Interrupted exception can occur when Client#stop is called.

- HADOOP-12229:Fix inconsistent subsection titles in filesystem.md.

- HADOOP-12230: hadoop-project declares duplicate, conflicting curator dependencies.

- HADOOP-12232: Upgrade Tomcat dependency to 6.0.44.

- HADOOP-12239: StorageException complaining " no lease ID" when updating FolderLastModifiedTime in WASB.

- HADOOP-12242: Add in-page TOC to filesystem specification pages.

- HADOOP-12252: LocalDirAllocator should not throw NPE with empty string configuration.

- HADOOP-12258: Need translate java.nio.file.NoSuchFileException to FileNotFoundException to avoid regression.

- HADOOP-12259: Utility to Dynamic port allocation

- HADOOP-12269: Update aws-sdk dependency to 1.10.6

- HADOOP-12280: Skip unit tests based on maven profile rather than NativeCodeLoader.isNativeCodeLoaded

- HADOOP-12284: UserGroupInformation doAs can throw misleading exception

- HADOOP-12291: Add support for nested groups in LdapGroupsMapping.

- HADOOP-12292: Make use of DeleteObjects optional.

- HADOOP-12296: when setnetgrent returns 0 in linux, exception should be thrown.

- HADOOP-12304: Applications using FileContext fail with the default file system configured to be wasb/s3/etc.

- HADOOP-12317: Applications fail on NM restart on some linux distro because NM container recovery declares AM container as LOST

- HADOOP-12324: Better exception reporting in SaslPlainServer.

- HADOOP-12325: RPC Metrics : Add the ability track and log slow RPCs.

- HADOOP-12334: Change Mode Of Copy Operation of HBase WAL Archiving to bypass Azure Storage Throttling after retries.

- HADOOP-12345: Pad hostname correctly in CredentialsSys.java.

- HADOOP-12346: Increase some default timeouts / retries for S3a connector.

- HADOOP-12348: MetricsSystemImpl creates MetricsSourceAdapter with wrong time unit parameter.

- HADOOP-12350: WASB Logging: Improve WASB Logging around deletes, reads and writes.

- HADOOP-12358: Add -safely flag to rm to prompt when deleting many files.

- HADOOP-12359: hadoop fs -getmerge doc is wrong.

- HADOOP-12366: expose calculated paths.

- HADOOP-12374: Description of hdfs expunge command is confusing.

- HADOOP-12386: RetryPolicies.RETRY_FOREVER should be able to specify a retry interval.

- HADOOP-12404: Disable caching for JarURLConnection to avoid sharing JarFile with other users when loading resource from URL in Configuration class.

- HADOOP-12406: Fixed AbstractMapWritable.readFields to use the thread's ClassLoader to load class instead of System ClassLoader.

- HADOOP-12407: Test failing: hadoop.ipc.TestSaslRPC.

- HADOOP-12413: AccessControlList should avoid calling getGroupNames in isUserInList with empty groups.

- HADOOP-12415: Add io.netty dependency to hadoop-nfs and to hadoop-hdfs, needed to build Bigtop successfuly, see BIGTOP-2049.

- HADOOP-12416: Trash messages should be handled by Logger instead of being delivered on System.out.

- HADOOP-12416: Trash messages should be handled by Logger instead of being delivered on System.out.

- HADOOP-12423: Handle failure of registering shutdownhook by ShutdownHookManager in static block

- HADOOP-12426: Add Entry point for Kerberos health check.

- 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-12442: Display help if the command option to 'hdfs dfs' is not valid

- HADOOP-12444: Support lazy seek in S3AInputStream. Rajesh Balamohan via stevel.

- HADOOP-12449: TestDNS and TestNetUtils failing if no network.

- HADOOP-12453: Support decoding KMS Delegation Token with its own Identifier.

- HADOOP-12457: [JDK8] Fix a failure of compiling common by javadoc.

- HADOOP-12463: Fix TestShell.testGetSignalKillCommand failure on windows.

- HADOOP-12464: Interrupted client may try to fail-over and retry.

- HADOOP-12465: Incorrect javadoc in WritableUtils.java.

- HADOOP-12469: distcp should not ignore the ignoreFailures option.

- HADOOP-12472: Make GenericTestUtils.assertExceptionContains robust.

- HADOOP-12478: Shell.getWinUtilsPath() has been renamed Shell.getWinutilsPath().

- HADOOP-12481: JWTRedirectAuthenticationHandler doesn't Retain Original Query String.

- HADOOP-12482: Race condition in JMX cache update.

- HADOOP-12483: Maintain wrapped SASL ordering for postponed IPC responses.

- 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-12520: Use XInclude in hadoop-azure test configuration to isolate Azure Storage account keys for service integration tests.

- HADOOP-12526: there are duplicate dependency definitions in pom's

- HADOOP-12533: Introduce FileNotFoundException in WASB for read and seek API.

- HADOOP-12535: Run FileSystem contract tests with hadoop-azure.

- HADOOP-12537: S3A to support Amazon STS temporary credentials.

- 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-12545: Hadoop javadoc has broken links for AccessControlList, ImpersonationProvider, DefaultImpersonationProvider, and DistCp.

- HADOOP-12548: Read s3a creds from a Credential Provider.

- HADOOP-12551: Introduce FileNotFoundException for WASB FileSystem API

- HADOOP-12555: WASB to read credentials from a credential provider.

- HADOOP-12559: KMS connection failures should trigger TGT renewal.

- HADOOP-12564: Upgrade JUnit3 TestCase to JUnit 4 in org.apache.hadoop.io package.

- HADOOP-12565: Replace DSA with RSA for SSH key type in SingleCluster.md.

- HADOOP-12568: Update core-default.xml to describe posixGroups support.

- HADOOP-12570: HDFS Secure Mode Documentation updates.

- HADOOP-12573: TestRPC.testClientBackOff failing.

- HADOOP-12577: Bump up commons-collections version to 3.2.2 to address a security flaw.

- HADOOP-12588: Fix intermittent test failure of TestGangliaMetrics.

- HADOOP-12589: Fix intermittent test failure of TestCopyPreserveFlag

- HADOOP-12590: TestCompressorDecompressor failing without stack traces

- HADOOP-12602: TestMetricsSystemImpl#testQSize occasionally fails.

- HADOOP-12603: TestSymlinkLocalFSFileContext#testSetTimesSymlinkToDir occasionally fail.

- HADOOP-12604: Exception may be swallowed in KMSClientProvider.

- HADOOP-12605: Fix intermittent failure of TestIPC.testIpcWithReaderQueuingi.

- HADOOP-12608: Fix exception message in WASB when connecting with anonymous credential.

- HADOOP-12609: Fix intermittent failure of TestDecayRpcScheduler.

- HADOOP-12613: TestFind.processArguments occasionally fails.

- HADOOP-12622: Improve the loggings in RetryPolicies and RetryInvocationHandler.

- HADOOP-12634: Change Lazy Rename Pending Operation Completion of WASB to address case of potential data loss due to partial copy.

- HADOOP-12635: Adding Append API support for WASB.

- HADOOP-12636: Prevent ServiceLoader failure init for unused FileSystems.

- HADOOP-12653: Use SO_REUSEADDR to avoid getting "Address already in use" when using kerberos and attempting to bind to any port on the local IP address.

- HADOOP-12659: Incorrect usage of config parameters in token manager of KMS.

- HADOOP-12663: Remove Hard-Coded Values From FileSystem.java.

- HADOOP-12666: Support Microsoft Azure Data Lake - as a file system in Hadoop.

- HADOOP-12667:s3a to support createNonRecursive API.

- HADOOP-12668: Support excluding weak Ciphers in HttpServer2 through ssl- server.conf.

- HADOOP-12672: RPC timeout should not override IPC ping interval

- HADOOP-12675: Fix description about retention period in usage of expunge command.

- HADOOP-12678: Handle empty rename pending metadata file during atomic rename in redo path.

- HADOOP-12682: Fix TestKMS#testKMSRestart* failure.

- HADOOP-12688: Fix deadlinks in Compatibility.md.

- HADOOP-12689: S3 filesystem operations stopped working correctly

- HADOOP-12691: Add CSRF Filter for REST APIs to Hadoop Common.

- HADOOP-12691: Move files to correct location.

- HADOOP-12696: Add tests for S3Filesystem Contract

- HADOOP-12699: TestKMS#testKMSProvider intermittently fails during 'test rollover draining'.

- HADOOP-12700: Remove unused import in TestCompressorDecompressor.java.

- HADOOP-12706: TestLocalFsFCStatistics#testStatisticsThreadLocalDataCleanUp times out occasionally.

- HADOOP-12711: Remove dependency on commons-httpclient for ServletUtil

- HADOOP-12715: TestValueQueue#testgetAtMostPolicyALL fails intermittently.

- HADOOP-12716: KerberosAuthenticator#doSpnegoSequence use incorrect class to determine isKeyTab in JDK8.

- HADOOP-12717: NPE when trying to rename a directory in Windows Azure Storage FileSystem.

- HADOOP-12723:S3A: Add ability to plug in any AWSCredentialsProvider.

- HADOOP-12731: Remove useless boxing/unboxing code.

- HADOOP-12736: TestTimedOutTestsListener#testThreadDumpAndDeadlocks sometimes times out.

- HADOOP-12738: Create unit test to automatically compare Common related classes and core-default.xml.

- HADOOP-12751: While using kerberos Hadoop incorrectly assumes names with '@' to be non-simple.

- HADOOP-12752: Improve diagnostics/use of envvar/sysprop credential propagation.

- HADOOP-12758: Extend CSRF Filter with UserAgent Checks.

- HADOOP-12761: incremental maven build is not really incremental

- HADOOP-12765: HttpServer2 should switch to using the non-blocking SslSelectChannelConnector to prevent performance degradation when handling SSL connections.

- HADOOP-12767: Update apache httpclient version to 4.5.2; httpcore to 4.4.4.

- HADOOP-12772: NetworkTopologyWithNodeGroup.getNodeGroup() can loop infinitely for invalid 'loc' values.

- HADOOP-12773: HBase classes fail to load with client/job classloader enabled

- HADOOP-12774: s3a should use UGI.getCurrentUser.getShortname() for username.

- HADOOP-12776: Remove getaclstatus call for non-acl commands in getfacl.

- HADOOP-12780: During WASB atomic rename handle crash when one directory has been renamed but not file under it.

- HADOOP-12782: Faster LDAP group name resolution with ActiveDirectory.

- HADOOP-12786: "hadoop key" command usage is not documented.

- HADOOP-12787: KMS SPNEGO sequence does not work with WEBHDFS.

- HADOOP-12788: OpensslAesCtrCryptoCodec should log which random number generator is used.

- HADOOP-12789: log classpath of ApplicationClassLoader at INFO level.

- HADOOP-12792: TestUserGroupInformation#testGetServerSideGroups fails in chroot.

- HADOOP-12793: Write a new group mapping service guide

- HADOOP-12794: Support additional compression levels for GzipCodec.

- HADOOP-12795: KMS does not log detailed stack trace for unexpected errors.

- HADOOP-12800: Copy docker directory from 2.8 to 2.7/2.6 repos to enable pre-commit Jenkins runs.

- HADOOP-12801: Suppress obsolete S3FileSystem tests.

- HADOOP-12804: Read Proxy Password from Credential Providers in S3 FileSystem.

- HADOOP-12805: Annotate CanUnbuffer with @InterfaceAudience.Public

- HADOOP-12807: S3AFileSystem should read AWS credentials from environment variables.

- HADOOP-12810: FileSystem#listLocatedStatus causes unnecessary RPC calls

- HADOOP-12817: Enable TLS v1.1 and 1.2

- HADOOP-12825: Log slow name resolutions.

- HADOOP-12828: Print user when services are started.

- HADOOP-12829: StatisticsDataReferenceCleaner swallows interrupt exceptions

- HADOOP-12831: LocalFS/FSOutputSummer NPEs in constructor if bytes per checksum set to 0

- HADOOP-12844:Recover when S3A fails on IOException in read().

- HADOOP-12846: Credential Provider Recursive Dependencies.

- HADOOP-12847: hadoop daemonlog should support https and SPNEGO for Kerberized cluster.

- HADOOP-12849: TestSymlinkLocalFSFileSystem fails intermittently.

- HADOOP-12851: S3AFileSystem Uptake of ProviderUtils.excludeIncompatibleCredentialProviders.

- HADOOP-12853: Change WASB documentation regarding page blob support.

- HADOOP-12870: Fix typo admininistration in CommandsManual.md.

- HADOOP-12871: Fix dead link to NativeLibraries.html in CommandsManual.md.

- HADOOP-12872: Fix formatting in ServiceLevelAuth.md.

- HADOOP-12875: [Azure Data Lake] Support for contract test and unit test cases.

- HADOOP-12878: KMS SPNEGO sequence does not work with WEBHDFS.

- HADOOP-12886: Exclude weak ciphers in SSLFactory through ssl-server.xml.

- HADOOP-12891: S3AFileSystem should configure Multipart Copy threshold and chunk size.

- HADOOP-12893: Update LICENSE.txt and NOTICE.txt.

- HADOOP-12893: Update LICENSE.txt and NOTICE.txt.

- HADOOP-12893: Update LICENSE.txt and NOTICE.txt".

- HADOOP-12895: SSLFactory#createSSLSocketFactory exception message is wrong.

- HADOOP-12901: Add warning log when KMSClientProvider cannot create a connection to the KMS server.

- HADOOP-12902: JavaDocs for SignerSecretProvider are out-of-date in AuthenticationFilter.

- HADOOP-12903: IPC Server should allow suppressing exception logging by type,.

- HADOOP-12903: IPC Server should allow suppressing exception logging by type, not log 'server too busy' messages.

- HADOOP-12906: AuthenticatedURL should convert a 404/Not Found into an FileNotFoundException.

- HADOOP-12916: Allow RPC scheduler/callqueue backoff using response times.

- HADOOP-12926: lz4.c does not detect 64-bit mode properly.

- HADOOP-12929: JWTRedirectAuthenticationHandler must accommodate null expiration time.

- HADOOP-12943: Add -w -r options in dfs -test command.

- HADOOP-12947: Update documentation Hadoop Groups Mapping to add static group mapping, negative cache.

- HADOOP-12950: ShutdownHookManager should have a timeout for each of the Registered shutdown hook.

- HADOOP-12954: Add a way to change hadoop.security.token.service.use_ip

- HADOOP-12958: PhantomReference for filesystem statistics can trigger OOM.

- HADOOP-12962: KMS key names are incorrectly encoded when creating key.

- HADOOP-12963:Allow using path style addressing for accessing the s3 endpoint.

- HADOOP-12964: Http server vulnerable to clickjacking.

- HADOOP-12973: Make DU pluggable.

- HADOOP-12975: Add jitter to CachingGetSpaceUsed's thread

- HADOOP-12977:s3a to handle delete("/", true) robustly.

- HADOOP-12982:Document missing S3A and S3 properties.

- HADOOP-12984: Add GenericTestUtils.getTestDir method and use it for emporary directory in tests.

- HADOOP-12985: Support MetricsSource interface for DecayRpcScheduler Metrics.

- HADOOP-12989: Some tests in org.apache.hadoop.fs.shell.find occasionally time out.

- HADOOP-12993: Change ShutdownHookManger complete shutdown log from INFO to DEBUG.

- HADOOP-12994: Specify PositionedReadable, add contract tests, fix problems.

- HADOOP-13008: Add XFS Filter for UIs to Hadoop Common.

- HADOOP-13011: Clearly Document the Password Details for Keystore-based Credential Providers

- HADOOP-13018: Make Kdiag check whether hadoop.token.files points to existent and valid files.

- HADOOP-13024: Distcp with -delete feature on raw data not implemented.

- HADOOP-13026: Should not wrap IOExceptions into a AuthenticationException in KerberosAuthenticator.

- HADOOP-13028:add low level counter metrics for S3A; use in read performance tests.

- HADOOP-13030: Handle special characters in passwords in KMS startup script.

- HADOOP-13037: Refactor Azure Data Lake Store as an independent FileSystem.

- HADOOP-13039: Add documentation for configuration property ipc.maximum.data.length

- HADOOP-13042: Restore lost leveldbjni LICENSE and NOTICE changes.

- HADOOP-13047:S3a Forward seek in stream length to be configurable.

- HADOOP-13052: ChecksumFileSystem mishandles crc file permissions.

- HADOOP-13058:S3A FS fails during init against a read-only FS if multipart purge.

- HADOOP-13065: Add a new interface for retrieving FS and FC Statistics

- HADOOP-13072: WindowsGetSpaceUsed constructor should be public

- HADOOP-13079: Add -q option to Ls to print ? instead of non-printable characters.

- HADOOP-13084: Fix ASF License warnings in branch-2.7.

- 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-13105: Support timeouts in LDAP queries in LdapGroupsMapping.

- HADOOP-13113:Enable parallel test execution for hadoop-aws. Chris Nauroth via stevel.

- HADOOP-13116:Jets3tNativeS3FileSystemContractTest does not run.

- HADOOP-13122:Customize User-Agent header sent in HTTP requests by S3A. Chris Nauroth via stevel.

- HADOOP-13130: s3a failures can surface as RTEs, not IOEs.

- HADOOP-13131: Add tests to verify that S3A supports SSE-S3 encryption.

- HADOOP-13138: Unable to append to a SequenceFile with Compression.NONE.

- HADOOP-13139:Branch-2: S3a to use thread pool that blocks clients.

- HADOOP-13140: FileSystem#initialize must not attempt to create StorageStatistics objects with null or empty schemes.

- HADOOP-13145:In DistCp, prevent unnecessary getFileStatus call when not preserving metadata.

- HADOOP-13146: Refactor RetryInvocationHandler.

- HADOOP-13155: Implement TokenRenewer to renew and cancel delegation tokens in KMS.

- HADOOP-13158:S3AFileSystem#toString might throw NullPointerException due to null cannedACL.

- HADOOP-13159: Fix potential NPE in Metrics2 source for DecayRpcScheduler.

- HADOOP-13162: Consider reducing number of getFileStatus calls in S3AFileSystem.mkdirs.

- HADOOP-13163:Reuse pre-computed filestatus in Distcp-CopyMapper

- HADOOP-13164:Optimize S3AFileSystem::deleteUnnecessaryFakeDirectories.

- HADOOP-13168: Support Future.get with timeout in ipc async calls

- HADOOP-13169: Randomize file list in SimpleCopyListing.

- HADOOP-13171: Add StorageStatistics to S3A; instrument some more operations.

- HADOOP-13179: GenericOptionsParser is not thread-safe because commons-cli OptionBuilder is not thread-safe.

- HADOOP-13183: S3A proxy tests fail after httpclient/httpcore upgrade.

- HADOOP-13188:S3A file-create should throw error rather than overwrite directories.

- HADOOP-13192: org.apache.hadoop.util.LineReader cannot handle multibyte delimiters correctly.

- HADOOP-13197: Add non-decayed call metrics for DecayRpcScheduler.

- HADOOP-13202: Avoid possible overflow in org.apache.hadoop.util.bloom.BloomFilter#getNBytes.

- HADOOP-13203: S3a: Consider reducing the number of connection aborts by setting correct length in s3 request.

- HADOOP-13207: Specify FileSystem listStatus, listFiles and RemoteIterator.

- HADOOP-13208: S3A listFiles(recursive=true) to do a bulk listObjects instead of walking the pseudo-tree of directories.

- HADOOP-13212:Provide an option to set the socket buffers in S3AFileSystem

- HADOOP-13213: Small Documentation bug with AuthenticatedURL in hadoop-auth.

- HADOOP-13237: s3a initialization against public bucket fails if caller lacks any credentials.

- HADOOP-13239:Deprecate s3:// in branch-2.

- HADOOP-13241: document s3a better.

- HADOOP-13242: Authenticate to Azure Data Lake using client ID and keys.

- HADOOP-13244: o.a.h.ipc.Server#Server should honor handlerCount when queueSizePerHandler is specified in consturctor.

- HADOOP-13249: RetryInvocationHandler need wrap InterruptedException in IOException when call Thread.sleep.

- HADOOP-13251: Authenticate with Kerberos credentials when renewing KMS delegation token.

- HADOOP-13252: Tune S3A provider plugin mechanism.

- HADOOP-13254: Create framework for configurable disk checkers

- HADOOP-13255: KMSClientProvider should check and renew tgt when doing delegation token operations.

- HADOOP-13257: Improve Azure Data Lake contract tests.

- HADOOP-13263: Reload cached groups in background after expiry.

- HADOOP-13270: BZip2CompressionInputStream finds the same compression marker twice in corner case, causing duplicate data blocks.

- HADOOP-13280: FileSystemStorageStatistics#getLong("readOps") should return readOps + largeReadOps

- HADOOP-13283: Support reset operation for new global storage statistics and per FS storage stats.

- HADOOP-13285: DecayRpcScheduler MXBean should only report decayed CallVolumeSummary.

- HADOOP-13287: TestS3ACredentials#testInstantiateFromURL fails if AWS secret key contains +.

- HADOOP-13288: Guard null stats key in FileSystemStorageStatistics

- HADOOP-13290: Appropriate use of generics in FairCallQueue.

- HADOOP-13291: Probing stats in DFSOpsCountStatistics/S3AStorageStatistics should be correctly implemented.

- HADOOP-13299: JMXJsonServlet is vulnerable to TRACE.

- HADOOP-13305: Define common statistics names across schemes.

- HADOOP-13309: Document S3A known limitations in file ownership and permission model.

- HADOOP-13315: FileContext#umask is not initialized properly.

- HADOOP-13316: Enforce Kerberos authentication for required ops in DelegationTokenAuthenticator.

- HADOOP-13320: Fix arguments check in documentation for WordCount v2.0.

- HADOOP-13323: Downgrade stack trace on FS load from Warn to debug.

- HADOOP-13324: s3a tests don't authenticate with S3 frankfurt

- HADOOP-13336:S3A to support per-bucket configuration.

- HADOOP-13350: Additional fix to LICENSE and NOTICE.

- HADOOP-13351: TestDFSClientSocketSize buffer size tests are flaky

- HADOOP-13352: Make X-FRAME-OPTIONS configurable in HttpServer2.

- HADOOP-13353: LdapGroupsMapping getPassward shouldn't return null when IOException throws.

- HADOOP-13354: Update WASB driver to use the latest version

- HADOOP-13368: DFSOpsCountStatistics$OpType#fromSymbol and s3a.Statistic#fromSymbol should be O(1) operation.

- HADOOP-13381: KMS clients should use KMS Delegation Tokens from current UGI.

- HADOOP-13382: Remove unneeded commons-httpclient dependencies from POM files in Hadoop and sub-projects, while adding it in to the only place it is still needed, HADOOP-openstack.

- HADOOP-13387:users always get told off for using S3 —even when not using it.

- HADOOP-13389:TestS3ATemporaryCredentials.testSTS error when using IAM credentials.

- HADOOP-13392: [Azure Data Lake] OAuth2 configuration should be default set to true when AdlFileSystem is configured.

- HADOOP-13393: Omit unsupported fs.defaultFS setting in ADLS documentation.

- HADOOP-13403: AzureNativeFileSystem rename/delete performance improvements.

- HADOOP-13405:doc for fs.s3a.acl.default indicates incorrect values.

- HADOOP-13406:S3AFileSystem: Consider reusing filestatus in delete() and mkdirs().

- HADOOP-13422: ZKDelegationTokenSecretManager JaasConfig does not work well with other ZK users in process.

- HADOOP-13433:Race in UGI.reloginFromKeytab.

- HADOOP-13434: Add quoting to Shell class.

- HADOOP-13443: KMS should check the type of underlying keyprovider of KeyProviderExtension before falling back to default.

- HADOOP-13446: Support running isolated unit tests separate from AWS integration tests.

- HADOOP-13447: Refactor S3AFileSystem to support introduction of separate metadata repository and tests.

- HADOOP-13458: LoadBalancingKMSClientProvider#doOp should log IOException stacktrace.

- HADOOP-13459: hadoop-azure runs several test cases repeatedly, causing unnecessarily long running time.

- HADOOP-13461: NPE in KeyProvider.rollNewVersion.

- HADOOP-13466: Add an AutoCloseableLock class.

- HADOOP-13467: Shell#getSignalKillCommand should use the bash builtin on Linux.

- HADOOP-13470: GenericTestUtils$LogCapturer is flaky.

- HADOOP-13494: ReconfigurableBase can log sensitive information.

- HADOOP-13502: Split fs.contract.is-blobstore flag into more descriptive flags for use by contract tests.

- HADOOP-13512: ReloadingX509TrustManager should keep reloading in case of exception.

- HADOOP-13513: Java 1.7 support for org.apache.hadoop.fs.azure testcases.

- HADOOP-13535: Add jetty6 acceptor startup issue workaround to branch-2.

- HADOOP-13540:improve section on troubleshooting s3a auth problems.

- HADOOP-13541:explicitly declare the Joda time version S3A depends on.

- HADOOP-13546: Override equals and hashCode to avoid connection leakage

- HADOOP-13552: RetryInvocationHandler logs all remote exceptions.

- HADOOP-13552: RetryInvocationHandler logs all remote exceptions.

- HADOOP-13558: UserGroupInformation created from a Subject incorrectly tries to renew the Kerberos ticket.

- HADOOP-13560: S3ABlockOutputStream to support huge

- HADOOP-13565: KerberosAuthenticationHandler#authenticate should not rebuild SPN based on client request.

- HADOOP-13590: Retry until TGT expires even if the UGI renewal thread encountered exception.

- HADOOP-13599: s3a close() to be non-synchronized, so avoid risk of deadlock on shutdown.

- HADOOP-13614: Purge some superfluous/obsolete S3 FS tests that are slowing test runs down.

- HADOOP-13638: KMS should set UGI's Configuration object properly.

- HADOOP-13674: S3A can provide a more detailed error message when accessing a bucket through an incorrect S3 endpoint.

- HADOOP-13675: Bug in return value for delete() calls in WASB.

- HADOOP-13680: fs.s3a.readahead.range to use getLongBytes.

- HADOOP-13686: Adding additional unit test for Trash

- HADOOP-13692: hadoop-aws should declare explicit dependency on Jackson 2 jars to prevent classpath conflicts.

- HADOOP-13707: If kerberos is enabled while HTTP SPNEGO is not configured, some links cannot be accessed.

- HADOOP-13710: Supress CachingGetSpaceUsed from logging interrupted exception stacktrace.

- HADOOP-13716: Add LambdaTestUtils class for tests; fix eventual consistency problem in contract test setup.

- HADOOP-13727: S3A: Reduce high number of connections to EC2 Instance Metadata Service caused by InstanceProfileCredentialsProvider.

- HADOOP-13735:ITestS3AFileContextStatistics.testStatistics() failing.

- HADOOP-13737: Cleanup DiskChecker interface.

- HADOOP-13738: DiskChecker should perform some disk IO.

- HADOOP-13742: Expose NumOpenConnectionsPerUser as a metric. Brahma Reddy Battula.

- HADOOP-13749: KMSClientProvider combined with KeyProviderCache can result in wrong UGI being used.

- HADOOP-13782: Make MutableRates metrics thread-local write, aggregate-on-read.

- HADOOP-13792: Stackoverflow for schemeless defaultFS with trailing slash.

- HADOOP-13801:regression: ITestS3AMiniYarnCluster failing.

- HADOOP-13804: MutableStat mean loses accuracy if add(long, long) is used.

- HADOOP-13822: Use GlobalStorageStatistics.INSTANCE.reset() at FileSystem#clearStatistics(). Contribute by Brahma Reddy Battula .

- HADOOP-13823: s3a rename: fail if dest file exists.

- HADOOP-13824: FsShell can suppress the real error if no error message is present.

- HADOOP-13826: S3A Deadlock in multipart copy due to thread pool limits.

- HADOOP-13830:Intermittent failure of ITestS3NContractRootDir.testRecursiveRootListing "Can not create a Path from an empty string".

- HADOOP-13831: Correct check for error code to detect Azure Storage Throttling and provide retries.

- HADOOP-13838: KMSTokenRenewer should close providers

- HADOOP-13847: KMSWebApp should close KeyProviderCryptoExtension.

- HADOOP-13855: Fix a couple of the s3a statistic names to be consistent with the rest.

- HADOOP-13857: S3AUtils.translateException to map

- HADOOP-13863: Azure: Add a new SAS key mode for WASB.

- HADOOP-13871: ITestS3AInputStreamPerformance.testTimeToOpenAndReadWholeFileBlocks performance awful.

- HADOOP-13890: Maintain HTTP/host as SPNEGO SPN support and fix KerberosName parsing.

- HADOOP-13930: Azure: Add Authorization support to WASB.

- HADOOP-13945: Azure: Add Kerberos and Delegation token support to WASB client.

- HADOOP-13956: Read ADLS credentials from Credential Provider.

- HADOOP-13962: Update ADLS SDK to 2.1.4.

- HADOOP-13988: KMSClientProvider does not work with WebHDFS and Apache Knox w/ ProxyUser.

- HADOOP-14017: User friendly name for ADLS user and group.

- HADOOP-14019: Fix some typos in the s3a docs.

- HADOOP-14028: S3A BlockOutputStreams doesn't delete temporary files in multipart uploads or handle part upload failures.

- HADOOP-14029: Fix KMSClientProvider for non-secure proxyuser use case.

- HADOOP-14048: REDO operation of WASB#AtomicRename should create placeholder blob for destination folder.

- HADOOP-14049: Honour AclBit flag associated to file/folder permission for Azure datalake account.

- HADOOP-14055: SwiftRestClient includes pass length in exception if auth fails.

- HADOOP-14081:S3A: Consider avoiding array copy in S3ABlockOutputStream

- HADOOP-14091: AbstractFileSystem implementaion for 'wasbs' scheme. Contributed Varada Hemeswari.

- HADOOP-14100: Upgrade Jsch jar to latest version to fix vulnerability in old versions.

- HADOOP-14138:Remove S3A ref from META-INF service discovery, rely on existing core-default entry.

- HADOOP-14150: Implement getHomeDirectory() method in NativeAzureFileSystem.

- HADOOP-14214: DomainSocketWatcher::add()/delete() should not self interrupt while looping await().

- HDFS-2043: TestHFlush failing intermittently.

- HDFS-2580: NameNode#main(...) can make use of GenericOptionsParser.

- HDFS-27: HDFS CLI with –config set to default config complains log file not found error.

- HDFS-3677: dfs.namenode.edits.dir.required is missing from hdfs- default.xml.

- HDFS-3918: EditLogTailer shouldn't log WARN when other node is in standby mode.

- HDFS-4015: Safemode should count and report orphaned blocks.

- HDFS-4176: EditLogTailer should call rollEdits with a timeout.

- HDFS-4185: Add a metric for number of active leases

- HDFS-4366: Block Replication Policy Implementation May Skip Higher-Priority Blocks for Lower-Priority Blocks.

- HDFS-4396: Add START_MSG/SHUTDOWN_MSG for ZKFC

- HDFS-4937: ReplicationMonitor can infinite-loop in BlockPlacementPolicyDefault#chooseRandom.

- HDFS-4946: Allow preferLocalNode in BlockPlacementPolicyDefault to be configurable.

- HDFS-5640: Add snapshot methods to FileContext.

- HDFS-6054: MiniQJMHACluster should not use static port to avoid binding failure in unit test.

- HDFS-6101: TestReplaceDatanodeOnFailure fails occasionally.

- HDFS-6184: Capture NN's thread dump when it fails over.

- HDFS-6407: Add sorting and pagination in the datanode tab of the NN Web UI.

- HDFS-6481: DatanodeManager#getDatanodeStorageInfos() should check the length of storageIDs.

- HDFS-6533: TestBPOfferService#testBasicFunctionalitytest fails intermittently.

- HDFS-6581: Support for writing to single replica in RAM. Applied HDFS-6581.merge.15.patch to Champlain. This is a consolidated patch containing all feature changes and was applied to trunk.

- HDFS-6666: Abort NameNode and DataNode startup if security is enabled but block access token is not enabled.

- HDFS-6763: Initialize file system-wide quota once on transitioning to active

- HDFS-6832: Fix the usage of 'hdfs namenode' command.

- HDFS-6860: BlockStateChange logs are too noisy.

- HDFS-6945: BlockManager should remove a block from excessReplicateMap and decrement ExcessBlocks metric when the block is removed.

- HDFS-7163: WebHdfsFileSystem should retry reads according to the configured retry policy.

- HDFS-7166: SbNN Web UI shows #Under replicated blocks and #pending deletion blocks.

- HDFS-7284: Add more debug info to BlockInfoUnderConstruction#setGenerationStampAndVerifyReplicas.

- HDFS-7314: When the DFSClient lease cannot be renewed, abort open-for- write files rather than the entire DFSClient.

- HDFS-7390: Provide JMX metrics per storage type.

- HDFS-7452: skip StandbyException log for getCorruptFiles()

- HDFS-7483: Display information per tier on the Namenode UI.

- HDFS-7597: DelegationTokenIdentifier should cache the TokenIdentifier to UGI mapping.

- HDFS-7608: hdfs dfsclient newConnectedPeer has no write timeout

- HDFS-7609: Avoid retry cache collision when Standby NameNode loading edits.

- 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-7725: Incorrect "nodes in service" metrics caused all writes to fail.

- HDFS-7758: Retire FsDatasetSpi#getVolumes() and use FsDatasetSpi#getVolumeRefs() instead

- HDFS-7890: Improve information on Top users for metrics in RollingWindowsManager and lower log level.

- HDFS-7923: The DataNodes should rate-limit their full block reports by asking the NN on heartbeat messages

- HDFS-7928: Scanning blocks from disk during rolling upgrade startup takes a lot of time if disks are busy.

- HDFS-7933: fsck should also report decommissioning replicas.

- HDFS-7964: Add support for async edit logging.

- HDFS-7978: Add LOG.isDebugEnabled() guard for some LOG.debug(..).

- HDFS-7990: IBR delete ack should not be delayed.

- HDFS-8008: Support client-side back off when the datanodes are congested.

- HDFS-8008: Support client-side back off when the datanodes are congested.

- HDFS-8009: Signal congestion on the DataNode.

- HDFS-8009: Signal congestion on the DataNode.

- HDFS-8046: Allow better control of getContentSummary

- HDFS-8055: NullPointerException when topology script is missing.

- HDFS-8071:(Contributed by Tsz Wo Nicholas Sze).

- HDFS-8099: Change "DFSInputStream has been closed already" message to debug log level.

- HDFS-8101: DFSClient use of non-constant DFSConfigKeys pulls in WebHDFS classes at runtime.

- HDFS-8113: Add check for null BlockCollection pointers in BlockInfoContiguous structures.

- HDFS-8131: Implement a space balanced block placement policy.

- HDFS-8144: Split TestLazyPersistFiles into multiple tests.

- HDFS-8152: Refactoring of lazy persist storage cases.

- HDFS-8155: Support OAuth2 in WebHDFS.

- HDFS-8180: AbstractFileSystem Implementation for WebHdfs.

- HDFS-8205: CommandFormat#parse() should not parse option as value of option.

- HDFS-8209: Support different number of datanode directories in MiniDFSCluster.

- HDFS-8211: DataNode UUID is always null in the JMX counter.

- HDFS-8219: setStoragePolicy with folder behavior is different after cluster restart.

- HDFS-8224: Schedule a block for scanning if its metadata file is corrupt.

- HDFS-8229: LAZY_PERSIST file gets deleted after NameNode restart.

- HDFS-8232: Missing datanode counters when using Metrics2 sink interface.

- HDFS-8247: TestDiskspaceQuotaUpdate#testAppendOverTypeQuota is failing.

- HDFS-8276: LazyPersistFileScrubber should be disabled if scrubber interval configured zero.

- HDFS-8278: When computing max-size-to-move in Balancer, count only the storage with remaining >= default block size.

- HDFS-8309: Skip unit test using DataNodeTestUtils#injectDataDirFailure() on Windows.

- HDFS-8310: Fix TestCLI.testAll "help: help for find" on Windows.

- HDFS-8311: DataStreamer.transfer() should timeout the socket InputStream.

- HDFS-8337: Accessing httpfs via webhdfs doesn't work from a jar with kerberos.

- HDFS-8345: Storage policy APIs must be exposed via the FileSystem interface.

- HDFS-8384: Allow NN to startup if there are files having a lease but are not under construction.

- HDFS-8431: hdfs crypto class not found in Windows.

- HDFS-8432: Introduce a minimum compatible layout version to allow downgrade in more rolling upgrade use cases.

- HDFS-8435: Support CreateFlag in WebHDFS.

- HDFS-8492: DN should notify NN when client requests a missing block

- HDFS-8539: Hdfs doesnt have class 'debug' in windows.

- HDFS-8542: WebHDFS getHomeDirectory behavior does not match specification.

- HDFS-8546: Use try with resources in DataStorage and Storage.

- HDFS-8548: Minicluster throws NPE on shutdown.

- HDFS-8549: Abort the balancer if an upgrade is in progress.

- HDFS-8554: TestDatanodeLayoutUpgrade fails on Windows.

- HDFS-8573: Move creation of restartMeta file logic from BlockReceiver to ReplicaInPipeline.

- HDFS-8578: On upgrade, Datanode should process all storage/data dirs in parallel.

- HDFS-8581: ContentSummary on / skips further counts on yielding lock

- HDFS-8593: Calculation of effective layout version mishandles comparison to current layout version in storage.

- HDFS-8615: Correct HTTP method in WebHDFS document.

- HDFS-8633: Fix setting of dfs.datanode.readahead.bytes in hdfs-default.xml to match DFSConfigKeys.

- HDFS-8642: Make TestFileTruncate more reliable.

- HDFS-8647: Abstract BlockManager's rack policy into BlockPlacementPolicy.

- HDFS-8656: Preserve compatibility of ClientProtocol#rollingUpgrade after finalization.

- HDFS-8659: Block scanner INFO message is spamming logs.

- HDFS-8674: Improve performance of postponed block scans.

- HDFS-8676: Delayed rolling upgrade finalization can cause heartbeat expiration.

- HDFS-8696: Make the lower and higher watermark in the DN Netty server configurable.

- HDFS-8709: Clarify automatic sync in FSEditLog#logEdit.

- HDFS-8722: Optimize datanode writes for small writes and flushes.

- HDFS-8729: Fix TestFileTruncate#testTruncateWithDataNodesRestartImmediately which occasionally failed.

- HDFS-8767: RawLocalFileSystem.listStatus() returns null for UNIX pipefile.

- HDFS-8772: Fix TestStandbyIsHot#testDatanodeRestarts which occasionally fails.

- HDFS-8778: TestBlockReportRateLimiting#testLeaseExpiration can deadlock.

- HDFS-8785: TestDistributedFileSystem is failing in trunk.

- HDFS-8792: BlockManager#postponedMisreplicatedBlocks should use a LightWeightHashSet to save memory.

- HDFS-8797: WebHdfsFileSystem creates too many connections for pread.

- HDFS-8806: Inconsistent metrics: number of missing blocks with replication factor 1 not properly cleared

- HDFS-8809: HDFS fsck reports under construction blocks as CORRUPT.

- HDFS-8815: DFS getStoragePolicy implementation using single RPC call

- HDFS-8816: Improve visualization for the Datanode tab in the NN UI.

- 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-8828: Utilize Snapshot diff report to build diff copy list in distcp.

- HDFS-8829: Make SO_RCVBUF and SO_SNDBUF size configurable for DataTransferProtocol sockets and allow configuring auto-tuning.

- HDFS-8831: Trash Support for deletion in HDFS encryption zone.

- HDFS-8844: TestHDFSCLI does not cleanup the test directory

- **HDFS-8845**: DiskChecker should not traverse the entire tree

- **HDFS-8846**: Add a unit test for INotify functionality across a layout version upgrade

- **HDFS-8850**: VolumeScanner thread exits with exception if there is no block pool to be scanned but there are suspicious blocks.

- **HDFS-8852**: HDFS architecture documentation of version 2.x is outdated about append write support.

- **HDFS-8855**: Webhdfs client leaks active NameNode connections.

- **HDFS-8856**: Make LeaseManager#countPath O(1).

- **HDFS-8859**: Improve DataNode ReplicaMap memory footprint to save about 45%.

- **HDFS-8860**: Remove unused Replica copyOnWrite code

- **HDFS-8863**: The remaining space check in BlockPlacementPolicyDefault is flawed.

- **HDFS-8867**: Enable optimized block reports.

- **HDFS-8879**: Quota by storage type usage incorrectly initialized upon namenode restart.

- **HDFS-8880**: NameNode metrics logging.

- **HDFS-8883**: NameNode Metrics : Add FSNameSystem lock Queue Length.

- **HDFS-8884**: Fail-fast check in BlockPlacementPolicyDefault#chooseTarget.

- **HDFS-8885**: ByteRangeInputStream used in webhdfs does not override available().

- **HDFS-8887**: Expose storage type and storage ID in BlockLocation.

- **HDFS-8891**: HDFS concat should keep srcs order.

- **HDFS-8896**: DataNode object isn't GCed when shutdown, because it has GC root in ShutdownHookManager.

- **HDFS-8898**: Create API and command-line argument to get quota and quota usage without detailed content summary.

- **HDFS-8911**: NameNode Metric : Add Editlog counters as a JMX metric.

- **HDFS-8914**: Document HA support in the HDFS HdfsDesign.md.

- **HDFS-8923**: Add -source flag to balancer usage message.

- **HDFS-8930**: Block report lease may leak if the 2nd full block report comes when NN is still in safemode

- **HDFS-8932**: NPE thrown in NameNode when try to get TotalSyncCount metric before editLogStream initialization.

- **HDFS-8939**: Test(S)WebHdfsFileContextMainOperations failing on branch-2.

- HDFS-8946: Improve choosing datanode storage for block placement.

- HDFS-8950: NameNode refresh doesn't remove DataNodes that are no longer in the allowed list.

- HDFS-8964: When validating the edit log, do not read at or beyond the file offset that is being written

- 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-8969: Clean up findbugs warnings for HDFS-8823 and HDFS-8932.

- HDFS-8983: NameNode support for protected directories.

- HDFS-8983: NameNode support for protected directories.

- HDFS-8995: Flaw in registration bookeeping can make DN die on reconnect.

- HDFS-9008: Balancer#Parameters class could use a builder pattern.

- HDFS-9008:was not backported correctly.

- HDFS-9009: Send metrics logs to NullAppender by default.

- HDFS-9019: Adding informative message to sticky bit permission denied exception.

- HDFS-9033: dfsadmin -metasave prints "NaN" for cache used%.

- HDFS-9038: DFS reserved space is erroneously counted towards non-DFS used.

- HDFS-9042: Update document for the Storage policy name

- HDFS-9043: Doc updation for commands in HDFS Federation

- HDFS-9047: Retire libwebhdfs.

- HDFS-9048: DistCp documentation is out-of-dated

- HDFS-9063: Correctly handle snapshot path for getContentSummary.

- HDFS-9072: Fix random failures in TestJMXGet.

- HDFS-9082: Change the log level in WebHdfsFileSystem.initialize() from INFO to DEBUG.

- HDFS-9083: Replication violates block placement policy.

- HDFS-9106: Transfer failure during pipeline recovery causes permanent write failures.

- 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-9128: TestWebHdfsFileContextMainOperations and TestSWebHdfsFileContextMainOperations fail due to invalid HDFS path on Windows.

- HDFS-9130: Use GenericTestUtils#setLogLevel to the logging level.

- HFDS-9137:should be backported correctly.

- HDFS-9142: Separating Configuration object for namenode(s) in MiniDFSCluster.

- HDFS-9145: Tracking methods that hold FSNamesytemLock for too long.

- HDFS-9148: Incorrect assert message in TestWriteToReplica#testWriteToTemporary

- 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-9198: Coalesce IBR processing in the NN.

- HDFS-9205: Do not schedule corrupt blocks for replication.

- HDFS-9210: Fix some misuse of %n in VolumeScanner#printStats.

- HDFS-9220: Reading small file

- HDFS-9221: HdfsServerConstants#ReplicaState#getState should avoid calling values() since it creates a temporary array.

- HDFS-9226: Unit Tests failing due to mockito version issue.

- HDFS-9239: DataNode Lifeline Protocol: an alternative protocol for reporting DataNode liveness.

- HDFS-9251: Refactor TestWriteToReplica and TestFsDatasetImpl to avoid explicitly creating Files in the tests code.

- HDFS-9259: Make SO_SNDBUF size configurable at DFSClient side for hdfs write scenario.

- HDFS-9264: Minor cleanup of operations on FsVolumeList#volumes.

- HDFS-9273: ACLs on root directory may be lost after NN restart.

- HDFS-9276: Failed to Update HDFS Delegation Token for long running application in HA mode.

- 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-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-9313: Possible NullPointerException in BlockManager if no excess replica can be chosen.

- HDFS-9314: Improve BlockPlacementPolicyDefault's picking of excess replicas.

- HDFS-9317: Document fsck -blockId and -storagepolicy options in branch-2.7.

- HDFS-9343: Empty caller context considered invalid.

- HDFS-9347: Invariant assumption in TestQuorumJournalManager.shutdown() is wrong

- HDFS-9350: Avoid creating temprorary strings in Block.toString() and getBlockName()

- HDFS-9354: Fix TestBalancer#testBalancerWithZeroThreadsForMove on Windows.

- HDFS-9358: TestNodeCount#testNodeCount timed out.

- HDFS-9362: TestAuditLogger#testAuditLoggerWithCallContext assumes Unix line endings, fails on Windows.

- HDFS-9364: Unnecessary DNS resolution attempts when creating NameNodeProxies.

- HDFS-9383: TestByteArrayManager#testByteArrayManager fails.

- HDFS-9384: TestWebHdfsContentLength intermittently hangs and fails due to TCP conversation mismatch between client and server.

- HDFS-9395: Make HDFS audit logging consistant.

- HDFS-9397: Fix typo for readChecksum() LOG.warn in BlockSender.java.

- HDFS-9402: Switch DataNode.LOG to use slf4j.

- HDFS-9406: FSImage may get corrupted after deleting snapshot.

- HDFS-9412: getBlocks occupies FSLock and takes too long to complete.

- HDFS-9413: getContentSummary() on standby should throw StandbyException.

- HDFS-9415: Document dfs.cluster.administrators and dfs.permissions.superusergroup.

- HDFS-9426: Rollingupgrade finalization is not backward compatible.

- HDFS-9428: Fix intermittent failure of TestDNFencing.testQueueingWithAppend.

- 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 messags.

- HDFS-9445: Datanode may deadlock while handling a bad volume.

- HDFS-9466: TestShortCircuitCache#testDataXceiverCleansUpSlotsOnFailure is flaky

- HDFS-9467: Fix data race accessing writeLockHeldTimeStamp in FSNamesystem.

- HDFS-9470: Encryption zone on root not loaded from fsimage after NN restart. Xiao Chen via wang.

- HDFS-9476: TestDFSUpgradeFromImage#testUpgradeFromRel1BBWImage occasionally fail

- HDFS-9478: Reason for failing ipc.FairCallQueue contruction should be thrown.

- HDFS-9484: NNThroughputBenchmark$BlockReportStats should not send empty block reports.

- HDFS-9493: Test o.a.h.hdfs.server.namenode.TestMetaSave fails in trunk.

- HDFS-9500: Fix software version counts for DataNodes during rolling upgrade.

- HDFS-9505: HDFS Architecture documentation needs to be refreshed.

- HDFS-9516: Truncate file fails with data dirs on multiple disks.

- HDFS-9517: Fix missing @Test annotation on TestDistCpUtils.testUnpackAttributes

- HDFS-9521: TransferFsImage.receiveFile should account and log separate times for image download and fsync to disk.

- HDFS-9525: hadoop utilities need to support provided delegation tokens.

- HDFS-9530: ReservedSpace is not cleared for abandoned Blocks

- HDFS-9533: seen_txid in the shared edits directory is modified during bootstrapping.

- HDFS-9534: Add CLI command to clear storage policy from a path.

- HDFS-9549: TestCacheDirectives#testExceedsCapacity is flaky

- HDFS-9555: LazyPersistFileScrubber should still sleep if there are errors in the clear progress.

- HDFS-9557: Reduce object allocation in PB conversion.

- HDFS-9566: Remove expensive 'BlocksMap#getStorages(Block b, final DatanodeStorage.State state)' method

- HDFS-9569: Log the name of the fsimage being loaded for better supportability.

- 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-9584: NPE in distcp when ssl configuration file does not exist in class path.

- HDFS-9589: Block files which have been hardlinked should be duplicated before the DataNode appends to the them

- HDFS-9600: Do not check replication if the block is under construction.

- HDFS-9605: Add links to failed volumes to explorer.html in HDFS Web UI.

- HDFS-9608: Disk IO imbalance in HDFS with heterogeneous storages.

- HDFS-9612: DistCp worker threads are not terminated after jobs are done.

- HDFS-9618: Fix mismatch between log level and guard in BlockManager#computeRecoveryWorkForBlocks

- HDFS-9619: SimulatedFSDataset sometimes can not find blockpool for the correct namenode

- HDFS-9623: Update example configuration of block state change log in log4j.properties.

- HDFS-9624: DataNode start slowly due to the initial DU command operations.

- HDFS-9625: set replication for empty file failed when set storage policy

- HDFS-9626: TestBlockReplacement#testBlockReplacement fails occasionally.

- HDFS-9629: Update the footer of Web UI to show year 2016

- HDFS-9630: DistCp minor refactoring and clean up.

- HDFS-9634: webhdfs client side exceptions don't provide enough details.

- HDFS-9638: to branch-2.7.

- HDFS-9644: Update encryption documentation to reflect nested EZs.

- HDFS-9655: NN should start JVM pause monitor before loading fsimage.

- HDFS-9661: Deadlock in DN.FsDatasetImpl between moveBlockAcrossStorage moveBlockAcrossStorage and createRbw.

- HDFS-9669: TcpPeerServer should respect ipc.server.listen.queue.size

- HDFS-9670: DistCp throws NPE when source is root.

- HDFS-9672: o.a.h.hdfs.TestLeaseRecovery2 fails intermittently

- HDFS-9688: Test the effect of nested encryption zones in HDFS downgrade.

- HDFS-9696: Garbage snapshot records linger forever.

- HDFS-9701: DN may deadlock when hot-swapping under load.

- HDFS-9706: Log more details in debug logs in BlockReceiver's constructor.

- HDFS-9710: DN can be configured to send block receipt IBRs in batches.

- HDFS-9711: Integrate CSRF prevention filter in WebHDFS.

- HDFS-9713: DataXceiver#copyBlock should return if block is pinned.

- HDFS-9715: Check storage ID uniqueness on datanode startup

- HDFS-9721: Allow Delimited PB OIV tool to run upon fsimage that contains INodeReference

- HDFS-9724: Degraded performance in WebHDFS listing as it does not reuse ObjectMapper.

- HDFS-9726: Refactor IBR code to a new class.

- HDFS-9730: Storage ID update does not happen when there is a layout change.

- HDFS-9739: DatanodeStorage.isValidStorageId() is broken.

- HDFS-9740: Use a reasonable limit in DFSTestUtil.waitForMetric().

- HDFS-9743: Fix TestLazyPersistFiles#testFallbackToDiskFull in branch-2.7.

- HDFS-9748: Avoid duplication in pendingReplications when addExpectedReplicasToPending is called twice.

- HDFS-9752: Permanent write failures may happen to slow writers during datanode rolling upgrades

- HDFS-9760: WebHDFS AuthFilter cannot be configured with custom AltKerberos auth handler

- HDFS-9764: DistCp doesn't print value for several arguments including -numListstatusThreads.

- HDFS-9765: TestBlockScanner#testVolumeIteratorWithCaching fails intermittently.

- HDFS-9766: TestDataNodeMetrics#testDataNodeTimeSpend fails intermittently.

- HDFS-9768: Reuse ObjectMapper instance in HDFS to improve the performance.

- HDFS-9772: TestBlockReplacement#testThrottler doesn't work as expected.

- HDFS-9777: Fix typos in DFSAdmin command line and documentation.(Wei-Chiu Chuang via umamahesh).

- HDFS-9779: . TestReplicationPolicyWithNodeGroup NODE variable picks wrong rack value.

- HDFS-9781: FsDatasetImpl#getBlockReports can occasionally throw NullPointerException.

- HDFS-9784: Example usage is not correct in Transparent Encryption document.

- HDFS-9790: HDFS Balancer should exit with a proper message if upgrade is not finalized

- HDFS-9797: Log Standby exceptions thrown by RequestHedgingProxyProvider at DEBUG Level

- HDFS-9799: Reimplement getCurrentTrashDir to remove incompatibility.

- HDFS-9812: Streamer threads leak if failure happens when closing DFSOutputStream.

- HDFS-9839: Reduce verbosity of processReport logging.

- HDFS-9842: dfs.datanode.balance.bandwidthPerSec should accept friendly size units.

- HDFS-9844: Correct path creation in getTrashRoot to handle root dir.

- 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-9864: Correct reference for RENEWDELEGATIONTOKEN and CANCELDELEGATIONTOKEN in webhdfs doc.

- HDFS-9865: TestBlockReplacement fails intermittently in trunk

- HDFS-9871: "Bytes Being Moved" -ve(-1 B) when cluster was already balanced.

- HDFS-9874: Long living DataXceiver threads cause volume shutdown to block.

- HDFS-9880: TestDatanodeRegistration fails occasionally.

- HDFS-9881: DistributedFileSystem#getTrashRoot returns incorrect path for encryption zones.

- HDFS-9882: Add heartbeatsTotal in Datanode metrics.

- HDFS-9902: Support different values of dfs.datanode.du.reserved per storage type.

- HDFS-9904: testCheckpointCancellationDuringUpload occasionally fails.

- HDFS-9905: WebHdfsFileSystem#runWithRetry should display original stack trace on error.

- HDFS-9906: Remove spammy log spew when a datanode is restarted.

- HDFS-9917: IBR accumulate more objects when SNN was down for sometime.

- HDFS-9926: MiniDFSCluster leaks dependency Mockito via DataNodeTestUtils.

- HDFS-9941: Do not log StandbyException on NN, other minor logging fixes.

- HDFS-9947: Block#toString should not output information from derived classes

- HDFS-9949: Add a test case to ensure that the DataNode does not regenerate its UUID when a storage directory is cleared

- HDFS-9958: BlockManager#createLocatedBlocks can throw NPE for corruptBlocks on failed storages.

- HDFS-10178: Permanent write failures can happen if pipeline recoveries occur for the first packet.

- HDFS-10182: Hedged read might overwrite user's buf.

- HDFS-10186: DirectoryScanner: Improve logs by adding full path of both actual and expected block directories.

- HDFS-10189: PacketResponder#toString should include the downstreams for PacketResponderType.HAS_DOWNSTREAM_IN_PIPELINE

- HDFS-10199: Unit tests TestCopyFiles, TestDistCh, TestLogalyzer under org.apache.hadoop.tools are failing.

- HDFS-10216: Distcp -diff throws exception when handling relative path.

- HDFS-10217: show 'blockScheduled' tooltip in datanodes table.

- HDFS-10223: peerFromSocketAndKey performs SASL exchange before setting connection timeouts.

- HDFS-10225: DataNode hot swap drives should disallow storage type changes.

- HDFS-10228: TestHDFSCLI fails.

- HDFS-10235: Last contact for Live Nodes should be relative time.

- HDFS-10239: Fsshell mv fails if port usage doesn't match in src and destination paths.

- HDFS-10245: Fix the findbugs warnings in branch-2.7.

- HDFS-10253: Fix TestRefreshCallQueue failure

- HDFS-10261: TestBookKeeperHACheckpoints doesn't handle ephemeral HTTP ports.

- HDFS-10264: Logging improvements in FSImageFormatProtobuf.Saver.

- HDFS-10270: TestJMXGet:testNameNode() fails.

- HDFS-10271: Extra bytes are getting released from reservedSpace for append

- HDFS-10275: TestDataNodeMetrics failing intermittently due to TotalWriteTime counted incorrectly.

- HDFS-10277: PositionedReadable test testReadFullyZeroByteFile failing in HDFS.

- HDFS-10279: Improve validation of the configured number of tolerated failed volumes.

- HDFS-10281: TestPendingCorruptDnMessages fails intermittently.

- HDFS-10291: TestShortCircuitLocalRead failing

- HDFS-10301: BlockReport retransmissions may lead to storages falsely being declared zombie if storage report processing happens out of order .

- HDFS-10301: Remove FBR tracking state to fix false zombie storage detection for interleaving block reports.

- HDFS-10309: Balancer doesn't honor dfs.blocksize value defined with suffix k(kilo), m(mega), g(giga).

- 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-10319: Balancer should not try to pair storages with different types.

- HDFS-10320: Rack failures may result in NN terminate.

- HDFS-10324: Trash directory in an encryption zone should be pre-created with correct permissions.

- HDFS-10329: Bad initialisation of StringBuffer in RequestHedgingProxyProvider.

- HDFS-10330: Add Corrupt Blocks Information in Metasave output.

- 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-10342: BlockManager#createLocatedBlocks should not check corrupt replicas if none are corrupt.

- HDFS-10343: BlockManager#createLocatedBlocks may return blocks on failed storages.

- HDFS-10344: DistributedFileSystem#getTrashRoots should skip encryption zone that does not have .Trash.

- HDFS-10347: Namenode report bad block method doesn't log the bad block or datanode.

- HDFS-10360: DataNode may format directory and lose blocks if current/VERSION is missing.

- HDFS-10367: TestDFSShell.testMoveWithTargetPortEmpty fails with Address bind exception.

- HDFS-10372: Fix for failing TestFsDatasetImpl#testCleanShutdownOfVolume.

- HDFS-10377: CacheReplicationMonitor shutdown log message should use INFO level

- HDFS-10390: Implement asynchronous setAcl/getAclStatus for DistributedFileSystem

- HDFS-10396: Using -diff option with DistCp may get "Comparison method violates its general contract" exception.

- HDFS-10397: Distcp should ignore -delete option if -diff option is provided instead of exiting.

- HDFS-10410: RedundantEditLogInputStream.LOG is set to wrong class.

- HDFS-10415: TestDistributedFileSystem#MyDistributedFileSystem attempts to set up statistics before initialize() is called

- HDFS-10417: Improve error message from checkBlockLocalPathAccess.

- HDFS-10423: Increase default value of httpfs maxHttpHeaderSize.

- HDFS-10424: DatanodeLifelineProtocol not able to use under security cluster

- HDFS-10431: Refactor and speedup TestAsyncDFSRename

- HDFS-10437: ReconfigurationProtocol not covered by HDFSPolicyProvider.

- HDFS-10438: When NameNode HA is configured to use the lifeline RPC server, it should log the address of that server.

- HDFS-10440: Improve DataNode web UI

- HDFS-10448: CacheManager#addInternal tracks bytesNeeded incorrectly when dealing with replication factors other than 1

- HDFS-10455: Logging the username when deny the setOwner operation.

- HDFS-10457: DataNode should not auto-format block pool directory if VERSION is missing.

- HDFS-10458: getFileEncryptionInfo should return quickly for non-encrypted cluster.

- HDFS-10468: HDFS read ends up ignoring an interrupt.

- HDFS-10469: Add number of active xceivers to datanode metrics.

- HDFS-10471: DFSAdmin#SetQuotaCommand's help msg is not correct.

- HDFS-10474: hftp copy fails when file name with Chinese+special char in branch-2

- HDFS-10481: HTTPFS server should correctly impersonate as end user to open file.

- HDFS-10485: Fix findbugs warning in FSEditLog.java

- HDFS-10488: Update WebHDFS documentation regarding CREATE and MKDIR default permissions.

- HDFS-10493: Add links to datanode web UI in namenode datanodes page.

- HDFS-10499: TestNameNodeMetadataConsistency#testGenerationStampInFuture Fails Intermittently.

- HDFS-10508: DFSInputStream should set thread's interrupt status after catching InterruptException from sleep.

- HDFS-10512: VolumeScanner may terminate due to NPE in DataNode.reportBadBlocks.

- HDFS-10525: Fix NPE in CacheReplicationMonitor#rescanCachedBlockMap

- HDFS-10544: Balancer doesn't work with IPFailoverProxyProvider.

- HDFS-10556: DistCpOptions should be validated automatically.

- HDFS-10569: A bug causes OutOfIndex error in BlockListAsLongs.

- HDFS-10579: HDFS web interfaces lack configs for X-FRAME-OPTIONS protection.

- HDFS-10589: Javadoc for HAState#HAState and HAState#setStateInternal contains non-existent params.

- HDFS-10609: Uncaught InvalidEncryptionKeyException during pipeline recovery may abort downstream applications.

- HDFS-10617: PendingReconstructionBlocks.size() should be synchronized.

- HDFS-10623: Remove unused import of httpclient.HttpConnection from TestWebHdfsTokens.

- HDFS-10625: VolumeScanner to report why a block is found bad.

- HDFS-10627: Volume Scanner marks a block as "suspect" even if the exception is network-related.

- HDFS-10642: TestLazyPersistReplicaRecovery#testDnRestartWithSavedReplicas fails intermittently.

- HDFS-10643: [Unit Test] Namenode should use loginUser(hdfs) to generateEncryptedKey.

- HDFS-10643: HDFS namenode should always use service user

- HDFS-10653: Optimize conversion from path string to components.

- HDFS-10655: Fix path related byte array conversion bugs.

- HDFS-10660: Expose storage policy apis via HDFSAdmin interface.

- HDFS-10684: WebHDFS DataNode calls fail without parameter createparent.

- HDFS-10688: BPServiceActor may run into a tight loop for sending block report when hitting IOException.

- HDFS-10703: HA NameNode Web UI should show last checkpoint time.

- HDFS-10710: In BlockManager#rescanPostponedMisreplicatedBlocks(), postponed misreplicated block counts should be retrieved within the NN lock protection.Contributed by GAO Rui.

- HDFS-10713: Throttle FsNameSystem lock warnings.

- HDFS-10715: NPE when applying AvailableSpaceBlockPlacementPolicy.

- HDFS-10716: In Balancer, the target task should be removed when its size < 0.

- HDFS-10735:Distcp using webhdfs on secure HA clusters fails with StandbyException .

- HDFS-10752: Several log refactoring/improvement suggestion in HDFS.

- HDFS-10757: KMSClientProvider combined with KeyProviderCache can result in wrong UGI being used.

- HDFS-10763: Open files can leak permanently due to inconsistent lease update.

- HDFS-10763: Open files can leak permanently due to inconsistent lease update TestTrashWithSecureEncryptionZones diverged from HDFS-10906.

- HDFS-10764: Fix INodeFile#getBlocks to not return null.

- HDFS-10793: Fix HdfsAuditLogger binary incompatibility introduced by HDFS-9184. HDFS-9700. DFSClient and DFSOutputStream should set TCP_NODELAY on sockets for DataTransferProtocol.

- HDFS-10798: Make the threshold of reporting FSNamesystem lock contention configurable.

- HDFS-10817: Add Logging for Long-held NN Read Locks.

- HDFS-10822: Log DataNodes in the write pipeline. John Zhuge via Lei Xu.

- HDFS-10913: Introduce fault injectors to simulate slow mirrors.

- HDFS-10917: Collect peer performance statistics on DataNode.

- HDFS-10922: Adding additional unit tests for Trash

- HDFS-10930: Refactor: Wrap Datanode IO related operations.

- HDFS-10934: TestDFSShell#testStat fails intermittently.

- HDFS-10956: Remove rename/delete performance penalty when not using snapshots.

- HDFS-10959: Update DataNode to use DatasetVolumeChecker.

- HDFS-10962: TestRequestHedgingProxyProvider is flaky.

- HDFS-10963: Reduce log level when network topology cannot find enough datanodes.

- HDFS-10965: Add unit test for HDFS command 'dfsadmin -printTopology'.

- HDFS-10966: Enhance Dispatcher logic on deciding when to give up a source DataNode.

- HDFS-10984: Expose nntop output as metrics.

- HDFS-10987: Make Decommission less expensive when lot of blocks present.

- HDFS-11003: Expose XmitsInProgress through DataNodeMXBean.

- HDFS-11008: Change unit test for testing parsing "-source" parameter in Balancer CLI.

- HDFS-11011: Add unit tests for HDFS command 'dfsadmin -set/clrSpaceQuota'.

- HDFS-11015: Enforce timeout in balancer.

- HDFS-11018: Incorrect check and message in FsDatasetImpl#invalidate.

- HDFS-11031: Add additional unit test for DataNode startup behavior when volumes fail.

- HDFS-11047: Remove deep copies of FinalizedReplica to alleviate heap consumption on DataNode.

- HDFS-11054: Suppress verbose log message in BlockPlacementPolicyDefault.

- HDFS-11056: Addendum fix for self-healing branch.

- HDFS-11056: Concurrent append and read operations lead to checksum error.

- HDFS-11069: Tighten the authorization of datanode RPC.

- HDFS-11076: Add unit test for extended Acls.

- HDFS-11083: Add unit test for DFSAdmin -report command.

- HDFS-11087: NamenodeFsck should check if the output writer is still writable.

- HDFS-11112: Journal Nodes should refuse to format non-empty directories.

- HDFS-11114: Support for running async disk checks in DataNode.

- HDFS-11119: Support for parallel checking of StorageLocations on DataNode startup.

- HDFS-11122: TestDFSAdmin#testReportCommand fails due to timed out.

- HDFS-11128: CreateEditsLog throws NullPointerException.

- HDFS-11148: Update DataNode to use StorageLocationChecker at startup.

- HDFS-11149: Fix build break.

- HDFS-11149: Support for parallel checking of FsVolumes.

- HDFS-11160: VolumeScanner reports write-in-progress replicas as corrupt incorrectly.

- HDFS-11180: Intermittent deadlock in NameNode when failover happens.

- HDFS-11182: Addendum fix for Fenton to fix a unit test.

- HDFS-11182: Update DataNode to use DatasetVolumeChecker.

- HDFS-11194: Maintain aggregated peer performance metrics on NameNode.

- HDFS-11195: Return error when appending files by webhdfs rest api fails.

- HDFS-11197: Listing encryption zones fails when deleting a EZ that is on a snapshotted directory.

- HDFS-11209: SNN can't checkpoint when rolling upgrade is not finalized.

- HDFS-11224: Lifeline message should be ignored for dead nodes

- HDFS-11229: HDFS-11056 failed to close meta file.

- HDFS-11240: Remove snapshot version of SDK dependency from Azure Data Lake Store File System.

- HDFS-11251: ConcurrentModificationException during DataNode#refreshVolumes .

- HDFS-11253: FileInputStream leak on failure path in BlockSender.

- HDFS-11258: File mtime change could not save to editlog.

- HDFS-11263: ClassCastException when we use Bzipcodec for Fsimage compression.

- HDFS-11267: Avoid redefinition of storageDirs in NNStorage and cleanup its accessors in Storage.

- HDFS-11274: Datanode should only check the failed volume upon IO errors.

- HDFS-11279: Cleanup unused DataNode#checkDiskErrorAsync().

- HDFS-11280: Allow WebHDFS to reuse HTTP connections to NN.

- HDFS-11282: Document the missing metrics of DataNode Volume IO operations.

- HDFS-11292: log lastWrittenTxId etc info in logSyncAll.

- HDFS-11299: Support multiple Datanode File IO hooks.

- HDFS-11307: The rpc to portmap service for NFS has hardcoded timeout.

- HDFS-11337:

- HDFS-11339: Support File IO sampling for Datanode IO profiling hooks.

- HDFS-11369: Change exception message in StorageLocationChecker.

- HDFS-11379: DFSInputStream may infinite loop requesting block locations.

- HDFS-11387: Socket reuse address option is not honored in PrivilegedNfsGatewayStarter.

- HDFS-11391: Numeric usernames do no work with WebHDFS FS write access.

- HDFS-11395: RequestHedgingProxyProvider#RequestHedgingInvocationHandler hides the Exception thrown from NameNode.

- HDFS-11403: Zookeper ACLs on NN HA enabled clusters should be handled consistently.

- HDFS-11461: DataNode Disk Outlier Detection.

- HDFS-11462: Fix occasional BindException in TestNameNodeMetricsLogger.

- HDFS-11476: Fix NPE in FsDatasetImpl#checkAndUpdate

- HDFS-11477: Simplify file IO profiling configuration.

- HDFS-11479: Socket re-use address option should be used in SimpleUdpServer.

- HDFS-11508: Fix bind failure in SimpleTCPServer & Portmap where bind fails because socket is in TIME_WAIT state.

- HDFS-11517: Expose slow disks via DataNode JMX.

- HDFS-11533: reuseAddress option should be used for child channels in Portmap and SimpleTcpServer.

- MAPREDUCE-5485: Allow repeating job commit by extending OutputCommitter API

- MAPREDUCE-6197: Cache MapOutputLocations in ShuffleHandler.

- MAPREDUCE-6213: NullPointerException caused by job history server addr not resolvable.

- MAPREDUCE-6213: NullPointerException caused by job history server addr not resolvable.

- MAPREDUCE-6273: HistoryFileManager should check whether summaryFile exists to avoid FileNotFoundException causing HistoryFileInfo into MOVE_FAILED state.

- MAPREDUCE-6302:

- MAPREDUCE-6338: MR AppMaster does not honor ephemeral port range.

- MAPREDUCE-6350: JobHistory doesn't support fully-functional search.

- MAPREDUCE-6404: Allow AM to specify a port range for starting its webapp.

- MAPREDUCE-6436: JobHistory cache issue.

- MAPREDUCE-6436: JobHistory cache issue.

- MAPREDUCE-6460: TestRMContainerAllocator.testAttemptNotFoundCausesRMCommunicatorException fails.

- MAPREDUCE-6485: Create a new task attempt with failed map task priority if in-progress attempts are unassigned.

- MAPREDUCE-6492: AsyncDispatcher exit with NPE on TaskAttemptImpl#sendJHStartEventForAssignedFailTask.

- MAPREDUCE-6492: AsyncDispatcher exit with NPE on TaskAttemptImpl#sendJHStartEventForAssignedFailTask.

- MAPREDUCE-6514: Job hangs as ask is not updated after ramping down of all reducers.

- MAPREDUCE-6543: Migrate MR Client test cases part 2.

- MAPREDUCE-6549: multibyte delimiters with LineRecordReader cause duplicate records

- 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-6577: MR AM unable to load native library without MR_AM_ADMIN_USER_ENV set.

- MAPREDUCE-6579: TestNetworkedJob.testNetworkedJob failed due to YARN-3946.

- MAPREDUCE-6593: TestJobHistoryEventHandler.testTimelineEventHandling fails because of NPE.

- MAPREDUCE-6616: Fail to create jobhistory file if there are some multibyte characters in the job name.

- MAPREDUCE-6618: YarnClientProtocolProvider leaking the YarnClient thread.(xgong).

- MAPREDUCE-6621: Memory Leak in JobClient#submitJobInternal().

- MAPREDUCE-6625: TestCLI#testGetJob fails occasionally

- MAPREDUCE-6635: Unsafe long to int conversion in UncompressedSplitLineReader and IndexOutOfBoundsException.

- MAPREDUCE-6635: Unsafe long to int conversion in UncompressedSplitLineReader and IndexOutOfBoundsException.

- MAPREDUCE-6639: Process hangs in LocatedFileStatusFetcher if FileSystem.get throws.

- MAPREDUCE-6672: TestTeraSort fails on Windows.

- MAPREDUCE-6684: High contention on scanning of user directory under immediate_done in Job History Server.

- MAPREDUCE-6689: MapReduce job can infinitely increase number of reducer resource requests.

- MAPREDUCE-6693: ArrayIndexOutOfBoundsException occurs when the length of the job name is equal to mapreduce.jobhistory.jobname.limit.

- MAPREDUCE-6701: application master log can not be available when clicking jobhistory's am logs link.

- MAPREDUCE-6711: After enabling preemption on cluster, jobs failed without making any attempts made after being preempted.

- MAPREDUCE-6724: Single shuffle to memory must not exceed Integer#MAX_VALUE.

- MAPREDUCE-6762: ControlledJob#toString failed with NPE when job status is not successfully updated.

- MAPREDUCE-6789: Fix TestAMWebApp failure.

- MAPREDUCE-6792: Allow user's full principal name as owner of MapReduce staging directory in JobSubmissionFiles#JobStagingDir().

- MAPREDUCE-6797: Job history server scans can become blocked on a single, slow entry.

- MAPREDUCE-6852: Job#updateStatus() failed with NPE due to race condition.

- YARN-857: Localization failures should be available in container diagnostics.

- YARN-998: Keep NM resource updated through dynamic resource config for RM/NM restart.

- 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-1462: AHS API and other AHS changes to handle tags for completed MR jobs.

- YARN-1556: NPE getting application report with a null appId.

- YARN-1880: Cleanup TestApplicationClientProtocolOnHA

- YARN-1993: Cross-site scripting vulnerability in TextView.java.

- YARN-2019: Retrospect on decision of making RM crashed if any exception throw in ZKRMStateStore.

- YARN-2046: Out of band heartbeats are sent only on container kill and possibly too early.

- YARN-2046: Out of band heartbeats are sent only on container kill and possibly too early.

- YARN-2194: Fix bug causing CGroups functionality to fail on RHEL7.

- YARN-2392: Add more diags about app retry limits on AM failures.

- YARN-2495: Allow admin specify labels from each NM

- 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-2666: TestFairScheduler.testContinuousScheduling fails Intermittently.

- YARN-2696: Queue sorting in CapacityScheduler should consider node label.

- 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-2821: Fixed a problem that DistributedShell AM may hang if restarted.

- 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-3021: YARN's delegation-token handling disallows certain trust setups to operate properly over DistCp.

- YARN-3026: Move application-specific container allocation logic from LeafQueue to FiCaSchedulerApp.

- YARN-3102: Decommisioned Nodes not listed in Web UI.

- YARN-3102: Decommisioned Nodes not listed in Web UI.

- **YARN-3110**: Few issues in ApplicationHistory web ui.

- **YARN-3136**: Fixed a synchronization problem of AbstractYarnScheduler#getTransferredContainers.

- **YARN-3148**: Allow CORS related headers to passthrough in WebAppProxyServlet.

- **YARN-3215**: Respect labels in CapacityScheduler when computing headroom.

- **YARN-3216**: Max-AM-Resource-Percentage should respect node labels.

- **YARN-3225**: New parameter of CLI for decommissioning node gracefully in RMAdmin CLI.

- **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-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-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-3305**: Normalize AM resource request on app submission.

- **YARN-3318**: Create Initial OrderingPolicy Framework and FifoOrderingPolicy.

- **YARN-3319**: Implement a FairOrderingPolicy.

- **YARN-3326**: Support RESTful API for getLabelsToNodes.

- **YARN-3343**: Increased TestCapacitySchedulerNodeLabelUpdate#testNodeUpdate timeout.

- **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-3356**: Capacity Scheduler FiCaSchedulerApp should use ResourceUsage to track used-resources-by-label.

- **YARN-3357**: Move TestFifoScheduler to FIFO package.

- **YARN-3360**: Add JMX metrics to TimelineDataManager .

- YARN-3361: CapacityScheduler side changes to support non-exclusive node labels.

- YARN-3362: Add node label usage in RM CapacityScheduler web UI.

- 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-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-3387: Previous AM's container completed status couldn't pass to current AM if AM and RM restarted during the same time.

- YARN-3394: Enrich WebApplication proxy documentation.

- YARN-3397: yarn rmadmin 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-3410: YARN admin should be able to remove individual application records from RMStateStore.

- YARN-3413: Changed Nodelabel attributes

- YARN-3425: NPE from RMNodeLabelsManager.serviceStop when NodeLabelsManager.serviceInit failed.

- YARN-3433: Jersey tests failing with Port in Use -again.

- YARN-3435: AM container to be allocated Appattempt AM container shown as null.

- 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-3451: Display attempt start time and elapsed time on the web UI.

- YARN-3453: Ensure preemption logic in FairScheduler uses DominantResourceCalculator in DRF queues to prevent unnecessary thrashing.

- YARN-3459: Fix failiure of TestLog4jWarningErrorMetricsAppender.

- YARN-3463: Integrate OrderingPolicy Framework with CapacityScheduler.

- YARN-3467: Expose allocatedMB, allocatedVCores, and runningContainers metrics on running Applications in RM Web UI.

- **YARN-3480**: Remove attempts that are beyond max-attempt limit from state store.

- **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-3505**: Node's Log Aggregation Report with SUCCEED should not cached in RMApps.

- **YARN-3511**: Add errors and warnings page to ATS.

- **YARN-3517**: RM web ui for dumping scheduler logs should be for admins only

- **YARN-3521**: Support return structured NodeLabel objects in REST API.

- **YARN-3530**: ATS throws exception on trying to filter results without otherinfo.

- **YARN-3541**: Add version info on timeline service / generic history web UI and REST API.

- **YARN-3543**: ApplicationReport should be able to tell whether the Application is AM managed or not.

- **YARN-3552**: RM Web UI shows -1 running containers for completed apps

- **YARN-3565**: NodeHeartbeatRequest/RegisterNodeManagerRequest should use NodeLabel object instead of String.

- **YARN-3579**: CommonNodeLabelsManager should support NodeLabel instead of string label name when getting node-to-label/label-to-label mappings.

- **YARN-3580**: [JDK8] TestClientRMService.testGetLabelsToNodes fails.

- **YARN-3581**: Deprecate -directlyAccessNodeLabelStore in RMAdminCLI.

- **YARN-3583**: Support of NodeLabel object instead of plain String in YarnClient side.

- **YARN-3589**: RM and AH web UI display DOCTYPE wrongly.

- **YARN-3591**: Resource localization on a bad disk causes subsequent containers failure.

- **YARN-3593**: Add label-type and Improve "DEFAULT_PARTITION" in Node Labels Page.

- **YARN-3600**: AM container link is broken

- **YARN-3604**: removeApplication in ZKRMStateStore should also disable watch.

- **YARN-3632**: Ordering policy should be allowed to reorder an application when demand changes.

- **YARN-3635**: Refactored current queue mapping implementation in CapacityScheduler to use a generic PlacementManager framework.

- **YARN-3654**: ContainerLogsPage web UI should not have meta-refresh.

- **YARN-3695**: ServerProxy

- YARN-3695: ServerProxy

- YARN-3700: Made generic history service load a number of latest applications according to the parameter or the configuration.

- YARN-3707: RM Web UI queue filter doesn't work.

- YARN-3717: Expose app/am/queue's node-label-expression to RM web UI / CLI / REST-API.

- YARN-3727: For better error recovery, check if the directory exists before using it for localization.

- YARN-3740: Fixed the typo in the configuration name: APPLICATION_HISTORY_PREFIX_MAX_APPS.

- 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-3769: Consider user limit when calculating total pending resource for preemption policy in Capacity Scheduler.

- YARN-3785: Support for Resource as an argument during submitApp call in MockRM test class.

- YARN-3846: RM Web UI queue filter is not working for sub queue.

- YARN-3849:(Too much of preemption activity causing continuos killing of containers across queues) to 2.3-maint.

- YARN-3873: PendingApplications in LeafQueue should also use OrderingPolicy.

- 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 id has not been reset synchronously.

- YARN-3920: FairScheduler container reservation on a node should be configurable to limit it to large containers

- YARN-3932: SchedulerApplicationAttempt#getResourceUsageReport and UserInfo should based on total-used-resources.

- YARN-3941: Proportional Preemption policy should try to avoid sending duplicate PREEMPT_CONTAINER event to scheduler.

- YARN-3946: Update exact reason as to why a submitted app is in ACCEPTED state to app's diagnostic message.

- YARN-3948: Display Application Priority in RM Web UI.(Sunil G via rohithsharmaks).

- YARN-3961: Expose pending, running and reserved containers of a queue in REST api and yarn top

- YARN-3966: Fix excessive loggings in CapacityScheduler.

- YARN-3971: Skip RMNodeLabelsManager#checkRemoveFromClusterNodeLabelsOfQueue on nodelabel recovery.

- YARN-3978: Configurably turn off the saving of container info in Generic AHS

- YARN-3982: container-executor parsing of container-executor.cfg broken in trunk and branch-2.

- YARN-3983: Refactored CapacitySchedulеri#FiCaSchedulerApp to easier extend container allocation logic.

- YARN-3986: getTransferredContainers in AbstractYarnScheduler should be present in YarnScheduler interface.

- YARN-3987: am container complete msg ack to NM once RM receive it.

- YARN-4000: RM crashes with NPE if leaf queue becomes parent queue during restart.

- YARN-4023: Publish Application Priority to TimelineServer.

- YARN-4026: Refactored ContainerAllocator to accept a list of priorites rather than a single priority.

- YARN-4082: Container shouldn't be killed when node's label updated.

- 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.

- YARN-4101: RM should print alert messages if ZooKeeper and Resourcemanager gets connection issue.

- YARN-4108: CapacityScheduler: Improve preemption to only kill containers that would satisfy the incoming request.

- YARN-4113: RM should respect retry-interval when uses RetryPolicies.RETRY_FOREVER.

- YARN-4140: RM container allocation delayed incase of app submitted to Nodelabel partition.

- YARN-4149: yarn logs -am should provide an option to fetch all the log files.

- YARN-4155: TestLogAggregationService.testLogAggregationServiceWithInterval failing.

- YARN-4162: CapacityScheduler: Add resource usage by partition and queue capacity by partition to REST API.

- YARN-4215: RMNodeLabels Manager Need to verify and replace node labels for the only modified Node Label Mappings in the request.

- YARN-4225: Add preemption status to yarn queue -status for capacity scheduler.

- YARN-4233: YARN Timeline Service plugin: ATS v1.5.

- YARN-4245: Generalize config file handling in container-executor.

- YARN-4270: Limit application resource reservation on nodes for non-node/rack specific requests

- YARN-4285: Display resource usage as percentage of queue and cluster in the RM UI

- YARN-4287: Capacity Scheduler: Rack Locality improvement

- YARN-4288: Fixed RMProxy to retry on IOException from local host.

- YARN-4290: Add -showDetails option to YARN Nodes CLI to print all nodes reports information.

- YARN-4304: AM max resource configuration per partition to be displayed/updated correctly in UI and in various partition related metrics.

- YARN-4309: Add container launch related debug information to container logs when a container fails.

- YARN-4313: Race condition in MiniMRYarnCluster when getting history server address.

- YARN-4315: NaN in Queue percentage for cluster apps page.

- YARN-4329: [YARN-5437] Allow fetching exact reason as to why a submitted app is in ACCEPTED state in Fair Scheduler

- YARN-4345: yarn rmadmin -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-4365: FileSystemNodeLabelStore should check for root dir existence on startup.

- YARN-4384: updateNodeResource CLI should not accept negative values for resource.

- YARN-4390: Do surgical preemption based on reserved container in CapacityScheduler.

- YARN-4392: ApplicationCreatedEvent event time resets after RM restart/failover.

- YARN-4393: TestResourceLocalizationService#testFailedDirsResourceRelease fails intermittently.

- YARN-4403:

- YARN-4405: Support node label store in non-appendable file system.

- YARN-4414: Nodemanager connection errors are retried at multiple levels.

- YARN-4414: Nodemanager connection errors are retried at multiple levels.

- YARN-4416: Deadlock due to synchronised get Methods in AbstractCSQueue.

- YARN-4417: Make RM and Timeline-server REST APIs more consistent.

- YARN-4418: AM Resource Limit per partition can be updated to ResourceUsage as well.

- YARN-4422: Generic AHS sometimes doesn't show started, node, or logs on App page.

- 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-4428: Redirect RM page to AHS page when AHS turned on and RM page is not available.

- YARN-4439: Clarify NMContainerStatus#toString method.

- YARN-4454: NM to nodelabel mapping going wrong after RM restart.

- YARN-4477: FairScheduler: Handle condition which can result in an infinite loop in attemptScheduling.

- YARN-4479: Change CS LeafQueue pendingOrderingPolicy to hornor recovered apps.

- YARN-4497: RM might fail to restart when recovering apps whose attempts are missing.

- YARN-4502: Fix two AM containers get allocated when AM restart.

- YARN-4522: Queue acl can be checked at app submission.

- YARN-4535: Fix checkstyle error in CapacityScheduler.java

- YARN-4546: YARN-4546. ResourceManager crash due to scheduling opportunity overflow.

- YARN-4546: YARN-4546. ResourceManager crash due to scheduling opportunity overflow.

- YARN-4556: TestFifoScheduler.testResourceOverCommit fails.

- YARN-4557: Fix improper Queues sorting in PartitionedQueueComparator when accessible-node-labels=*.

- YARN-4565:

- YARN-4573: Fix test failure in TestRMAppTransitions#testAppRunningKill and testAppKilledKilled.

- YARN-4584: RM startup failure when AM attempts greater than max-attempts.

- YARN-4591: YARN Web UIs should provide a robots.txt. Junping Du.

- YARN-4598: Invalid event: RESOURCE_FAILED at CONTAINER_CLEANEDUP_AFTER_KILL.

- YARN-4598: Invalid event: RESOURCE_FAILED at CONTAINER_CLEANEDUP_AFTER_KILL.

- YARN-4607: Pagination support for AppAttempt page TotalOutstandingResource Requests table.

- YARN-4610: Reservations continue looking for one app causes other apps to starve.

- YARN-4617: LeafQueue#pendingOrderingPolicy should always use fixed ordering policy instead of using same as active applications ordering policy.

- YARN-4623: TestSystemMetricsPublisher#testPublishAppAttemptMetricsForUnmanagedAM fails with NPE on branch-2.7.

- YARN-4625: Make ApplicationSubmissionContext and ApplicationSubmissionContextInfo more consistent.

- YARN-4633: TestRMRestart.testRMRestartAfterPreemption fails intermittently.

- YARN-4634: Scheduler UI/Metrics need to consider cases like non-queue label mappings.

- YARN-4667: RM Admin CLI for refreshNodesResources throws NPE when nothing is configured.

- YARN-4680: Fix TimerTasks leak in ATS V1.5 Writer.

- YARN-4696: TimelineClient to add flush operation for deterministic writes

- YARN-4697: NM aggregation thread pool is not bound by limits

- YARN-4699: Scheduler UI and REST o/p is not in sync when -replaceLabelsOnNode is used to change label of a node.

- YARN-4709: NMWebServices produces incorrect JSON for containers.

- YARN-4709: NMWebServices produces incorrect JSON for containers.

- YARN-4723: NodesListManager$UnknownNodeId ClassCastException.

- YARN-4737: Add CSRF filter support in YARN

- YARN-4737: Add CSRF filter support in YARN

- YARN-4746: yarn web services should convert parse failures of appId, appAttemptId and containerId to 400.

- YARN-4767: Network issues can cause persistent RM UI outage.

- 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-4785: Inconsistent value type of the "type" field for LeafQueueInfo in response of RM REST API.

- YARN-4811: Generate histograms in ContainerMetrics for actual container resource usage.

- YARN-4814: ATS 1.5 timelineclient impl call flush after every event write.

- YARN-4815: ATS 1.5 timelineclinet impl try to create attempt directory for every event call.

- YARN-4817: TimelineClient ATSv1.5 logging is very noisy.

- YARN-4822: Refactor existing Preemption Policy of CS for easier adding new approach to select preemption candidates.

- YARN-4832: NM side resource value should get updated if change applied in RM side.

- YARN-4842: Fixed "yarn logs" command to guess

- YARN-4844: Add getMemorySize/getVirtualCoresSize to o.a.h.y.api.records.Resource.

- YARN-4846: Fix random failures for TestCapacitySchedulerPreemption #testPreemptionPolicyShouldRespectAlreadyMarkedKillableContainers.

- YARN-4865: Track Reserved resources in ResourceUsage and QueueCapacities.

- YARN-4890: Unit test intermittent failure: TestNodeLabelContainerAllocation#testQueueUsedCapacitiesUpdate.

- YARN-4893: Fix intermittent failures in TestRMAdminService.

- YARN-4905: Improved "yarn logs" command-line to optionally show log metadata also.

- YARN-4913: Yarn logs should take a -out option to write to a directory.

- YARN-4916: TestNMProxy.tesNMProxyRPCRetry fails.

- YARN-4920: ATS/NM should support a link to dowload/get the logs in text format.

- YARN-4934: Reserved Resource for QueueMetrics needs to be handled correctly in few cases.

- YARN-4955: Add retry for SocketTimeoutException in TimelineClient.

- YARN-4966: Improve yarn logs to fetch container logs without specifying nodeId.

- YARN-4984: LogAggregationService shouldn't swallow exception in handling createAppDir() which cause thread leak.

- YARN-4990: Re-direction of a particular log file within in a container in NM UI does not redirect properly.

- YARN-5023: TestAMRestart#testShouldNotCountFailureToMaxAttemptRetry random failure.

- YARN-5048: DelegationTokenRenewer#skipTokenRenewal may throw NPE.

- YARN-5069: TestFifoScheduler.testResourceOverCommit race condition.

- YARN-5076: YARN web interfaces lack XFS

- YARN-5080: Cannot obtain logs using YARN CLI -am for either KILLED or RUNNING AM.

- YARN-5083: YARN CLI for AM logs does not give any error message if entered invalid am value.

- YARN-5088: Improve "yarn log" command-line to read the last K bytes for the log files.

- YARN-5089: Improve "yarn log" command-line "logFiles" option to support regex.

- YARN-5112: Excessive log warnings for directory permission issue on NM recovery.

- YARN-5115: Avoid setting CONTENT-DISPOSITION header in the container-logs web-service.

- YARN-5121: fix some container-executor portability issues.

- YARN-5122: "yarn logs" for running containers should print an explicit footer saying that the log may be incomplete.

- YARN-5131: Distributed shell AM fails when extra container arrives during finishing. Wangda Tan via junping_du.

- YARN-5141: Get Container logs for the Running application from Yarn Logs CommandLine.

- YARN-5190: Registering/unregistering container metrics triggered by ContainerEvent and ContainersMonitorEvent are conflict which cause uncaught exception in ContainerMonitorImpl.

- YARN-5191: Renamed the newly added "download=true" option for getting logs via NMWebServices and AHSWebServices to be a better "format" option.

- YARN-5199: Close LogReader in in AHSWebServices#getStreamingOutput and FileInputStream in NMWebServices#getLogs.

- YARN-5200: Enhanced "yarn logs" to be able to get a list of containers whose logs are aggregated via a "show_container_log_info" option.

- YARN-5206: RegistrySecurity includes id:pass in exception text if considered invalid.

- YARN-5208: Run TestAMRMClient TestNMClient TestYarnClient TestClientRMTokens TestAMAuthorization tests with hadoop.security.token.service.use_ip enabled.

- YARN-5213: Fix a bug in LogCLIHelpers which cause TestLogsCLI#testFetchApplictionLogs fails intermittently.

- YARN-5214: Fixed locking in DirectoryCollection to avoid hanging NMs when various code-paths hit slow disks.

- YARN-5223: Container line in yarn logs output for a live application should include the hostname for the container.

- YARN-5224: Added new web-services /containers/{containerid}/logs & /containers/{containerid}/logs/{filename} and using them in "yarn logs" CLI to get logs of finished containers of a running application.

- YARN-5226: Remove AHS enable check from LogsCLI#fetchAMContainerLogs.

- YARN-5227: Yarn logs command: no need to specify applicationId when specifying containerId.

- YARN-5251: Yarn CLI to obtain App logs for last 'n' bytes fails.

- YARN-5266: Wrong exit code while trying to get app logs using regex via CLI.

- YARN-5270: Solve miscellaneous issues caused by YARN-4844.

- YARN-5294: Pass remote ip address down to YarnAuthorizationProvider.

- YARN-5296: NMs going OutOfMemory because ContainerMetrics leak in ContainerMonitorImpl.

- YARN-5297: Avoid printing a stack trace when recovering an app after the RM restarts.

- YARN-5337: Fix Dshell OOM issues.

- YARN-5339: Passing file to -out for YARN log CLI doesnt give warning or error code.

- YARN-5353: ResourceManager can leak delegation tokens when they are shared across apps.

- YARN-5363: For AM containers, or for containers of running-apps, "yarn logs" incorrectly only

- YARN-5456: container-executor support for FreeBSD, NetBSD, and others if conf path is absolute.

- YARN-5470: Differentiate exactly match with regex in yarn log CLI.

- YARN-5476: Not existed application reported as ACCEPTED state by YarnClientImpl.

- YARN-5523: Yarn running container log fetching causes OutOfMemoryError

- YARN-5524: Yarn live log aggregation does not throw if command line arg is wrong.

- YARN-5539: TimelineClient failed to retry on java.net.SocketTimeoutException: Read timed out.

- YARN-5549: AMLauncher#createAMContainerLaunchContext() should not log the command to be launched indiscriminately.

- YARN-5551: Ignore file backed pages from memory computation when smaps is enabled.

- YARN-5659: getPathFromYarnURL should use standard methods.

- YARN-5677: RM should transition to standby when connection is lost for an extended period.

- YARN-5704: Provide config knobs to control enabling/disabling new/work in progress features in container-executor.

- YARN-5711: Propogate exceptions back to client when using hedging RM failover provider.

- YARN-5718: TimelineClient

- YARN-5773: RM recovery too slow due to LeafQueue#activateApplications.

- YARN-5864: Capacity Scheduler queue priority.

- YARN-5873: RM crashes with NPE if generic application history is enabled.

- YARN-5915: ATS 1.5 FileSystemTimelineWriter causes flush() to be called after every event write.

- YARN-5923: Unable to access logs for a running application if YARN_ACL_ENABLE is enabled.

- YARN-5999: AMRMClientAsync will stop if any exceptions thrown on allocate call.

- YARN-6011: Add a new web service to list the files on a container in AHSWebService.

- YARN-6029: CapacityScheduler deadlock when ParentQueue#getQueueUserAclInfo is called by one thread and LeafQueue#assignContainers is releasing excessive reserved container by another thread.

- YARN-6068: Log aggregation get failed when NM restart even with recovery.

- YARN-6081:LeafQueue#getTotalPendingResourcesConsideringUserLimit should deduct reserved from pending to avoid unnecessary preemption of reserved container.

- YARN-6099: Improve webservice to list aggregated log files.

- YARN-6100: Improve YARN webservice to output aggregated container logs.

- YARN-6108: Improve AHS webservice to accept NM address as a parameter to get container logs.

- YARN-6113: Re-direct NM Web Service to get container logs for finished applications.

- YARN-6123: Add a test to make sure queues of orderingPolicy will be updated when childQueues is added or removed.

- YARN-6126: Obtaining app logs for Running application fails with json parse error.

- YARN-6137: Yarn client implicitly invoke ATS client which accesses HDFS.

- YARN-6145: Improve log message on fail over.

- YARN-6174: Log files pattern should be same for both running and finished container.

- YARN-6189: ApplicationNotFoundException due to RM failed over should throw more user friendly exception.

- YARN-6228: EntityGroupFSTimelineStore should allow configurable cache stores.

- YARN-6270: WebUtils.getRMWebAppURLWithScheme() needs to honor RM HA setting.

- YARN-6313: yarn logs cli should provide logs for a completed container even when application is still running.

- YARN-6314: Potential infinite redirection on YARN log redirection web service.

- YARN-6332: Make RegistrySecurity use short user names for ZK ACLs.

- YARN-6367: YARN logs CLI needs alway check containerLogsInfo/containerLogInfo before parse the JSON object from NMWebService.

1.8.2. Accumulo

This release provides Accumulo 1.7.1 with no additional Apache patches.

1.8.3. Atlas

This release provides Atlas 0.8.0 and the following Apache patches:

- ATLAS-1333: fix for incorrect json4s library version in Atlas Storm hook.

- ATLAS-1469: User provided values for tag attributes should be evaluated for its data type..

- ATLAS-1470: Tag is not associated to an entity when incorrect value for byte type is provided.

- ATLAS-1503: optimization of export implementation.

- ATLAS-1611: incorrect error code for negative tests (#2).

- ATLAS-1630: basic search performance improvement (#3).

- ATLAS-1659: Update full-text search string when traits/tags are added/removed from entity.

- ATLAS-1664: Able to add already added tag to an entity through REST API.

- ATLAS-1665: export optimization to reduce file-size and export-time.

- ATLAS-1666: fix for IT failures caused by earlier commit.

- ATLAS-1666: Updated exception handling to avoid use of generic exceptions.

- ATLAS-1667: Atlas UI cleanup for lineage.

- ATLAS-1669: Removed Unused class.

- ATLAS-1670: fix for IT failures introduced due to ATLAS-1664.

- ATLAS-1671: fix for missing client IP in Ranger audit log for Atlas authorizations.

- ATLAS-1675: type-registry re-initialization failure when instance goes from passive to active state.

- ATLAS-1676: update Hive hook to set hdfs_path.clusterName attribute.

- ATLAS-1682: UI update to add new overlay loader to tables; fix incorrect state while switching between basic and Advance search.

- ATLAS-1687: hbase_column_family and hbase_column are not loaded after upgrade.

1.8.4. DataFu

This release provides DataFu 1.3.0 with no additional Apache patches.

1.8.5. Falcon

This release provides Falcon 0.10.0 and the following Apache patches:

- FALCON-1944: Ability to provide additional DistCP options.

- FALCON-2066: Extend support for Import and Export of various database types.

- FALCON-2072: Hive2 URLs in Falcon should allow additional configuration elements in the URL.

- FALCON-2082: Add CSRF filter for REST APIs.

- FALCON-2083: Handle connector specific direct mode and verbose options for Database Import and Export.

- FALCON-2093: Database import and export to support TD connector.

- FALCON-2095: Hive Replication jobs are failing with UnknownHostException in NN HA.

- FALCON-2117: Implement X-Frame-Options header for Falcon UI.

- FALCON-2118: Proposal for new UI changes for FALCON.

- FALCON-2273: Disallow external entity injection and clean up some log messages.

- FALCON-2280: Unable to create mirror on WASB target.

- FALCON-2281: HiveDRTest tests are getting permissions denied.

1.8.6. Flume

This release provides Flume 1.5.2 and the following Apache patch:

- FLUME-1334: Flume startup script for Windows.

- FLUME-1734: Hive Sink based on the new Hive Streaming support.

- FLUME-2058: TestFlumeEventQueue in FileChannel fails on Windows.

- FLUME-2068: File Channel issue - recovering from BadCheckpoint exception on Windows.

- FLUME-2095: JMS source with TIBCO (patch-1).

- FLUME-2122: Minor cleanups of User guide.

- FLUME-2123: Morphline Solr sink missing short type name.

- FLUME-2136: Fix intermittent test failure in TestMonitoredCounterGroup on Windows.

- FLUME-2137: Fix StagedInstall.java to invoke the correct startup script on Windows.

- FLUME-2145: TestCheckpointRebuilder.testFastReplay fails on Windows.

- FLUME-2146: Windows: Tmp file creation in TestBodyTextEventSerializer.java needs fixing.

- FLUME-2150: Fix TestFileChannelEncryption failure on Windows.

- FLUME-2151: Windows: Update TestExecSource to use native commands on Windows.

- FLUME-2162: TestHDFSEventSinkOnMiniCluster.maxUnderReplicationTest fails on hadoop2.

- FLUME-2175: Update Developer Guide with notes on how to upgrade Protocol Buffer version.

- FLUME-2218: TestFileChannelIntegrityTool tests failing on Windows.

- FLUME-2219: Windows: Flume tests need to know location of Hadoop native libraries (hadoop.dll).

- FLUME-2224: Disable File channel dual checkpointing on Windows.

- 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-2358: File Channel needs to close BackingStore & EventQueue before deleting files in checkpoint directory.

- FLUME-2359: TestFileChannelIntegrityTool throws exception on class teardown on Windows.

- 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-2442: Need an alternative to providing clear text passwords in flume config.

- FLUME-2450: Improve replay index insertion speed.

- FLUME-2451: HDFS Sink Cannot Reconnect After NameNode Restart.

- FLUME-2501: Updating HttpClient lib version to ensure compat with Solr.

- FLUME-2508: LineDeserializer causes incorrect behavior in SpoolDir Source on Windows.

- FLUME-2511: Allow configuration of enabled protocols in Avro source and Rpc client.

- FLUME-2520: HTTP Source should be able to block a prefixed set of protocols..

- FLUME-2530: Resource leaks found by Coverity tool.

- FLUME-2532: Windows : TestReliableSpoolingFileEventReader - need to close filereader.

- FLUME-2533: HTTPS tests fail on Java 6.

- FLUME-2534: Windows : Windows : TestFlumeEventQueue has many failures.

- FLUME-2541: Bug in TestBucketWriter.testSequenceFileCloseRetries.

- 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-2632: High CPU on KafkaSink.

- FLUME-2655: Update documentation for hdfs.closeTries based on FLUME-2586.

- FLUME-2662: Upgrade to Commons-IO 2.4.

- FLUME-2663: Address Build warnings of duplicate dependencies listed.

- FLUME-2722: Windows : TestKafkaSourceUtil Unit Tests fail on Windows.

- FLUME-2729: Allow pollableSource backoff times to be configurable.

- FLUME-2754: Hive Sink skipping first transaction in each Batch of Hive Transactions.

- FLUME-2761: Move Hive sink out of preview mode.

- FLUME-2804: Hive sink should try to clean up transactions when flume exits.

- FLUME-2812: Fix semaphore leak causing java.lang.Error: Maximum permit count exceeded in MemoryChannel.

- FLUME-2841: Upgrade commons-collections to 3.2.2.

- FLUME-2854: parameterizing jetty version..

- FLUME-2865: Upgrade thrift version (0.9.2).

- FLUME-2947: Upgrade Hive and thrift dependencies.

- FLUME-2991: ExecSource command execution starts before starting the sourceCounter.

- FLUME-3031: Sequence source should reset its counter for event body on channel exception.

- FULME-2544: Windows: Incorrect Path Separator used in HDFS path (HDFS Sink).

1.8.7. HBase

This release provides HBase 1.1.2 and the following Apache patches:

- BUG-45027:/BUG-45361 Fix hostname used with Kerberos on Dual-Homed clusters..

- HBASE-13318: RpcServer.getListenerAddress should handle when the accept channel is closed (Andrew Purtell).

- HBASE-13829: Add more ThrottleType (Guanghao Zhang).

- HBASE-14123: HBase Backup/Restore Phase 2.

- HBASE-14140: HBase Backup Admin API.

- HBASE-14148: Web UI Framable Page.

- HBASE-14172: Upgrade existing thrift binding using thrift 0.9.3 compiler (Josh Elser).

- 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-14283: Reverse scan doesn't work with HFile inline index/bloom blocks.

- HBASE-14302: TableSnapshotInputFormat should not create back references when restoring snapshot.

- HBASE-14307: Incorrect use of positional read api in HFileBlock (Chris Nauroth).

- HBASE-14313: After a Connection sees ConnectionClosingException it never recovers (Elliott Clark).

- HBASE-14407: NotServingRegion: hbase region closed forever (Shuaifeng Zhou).

- HBASE-14449: Rewrite deadlock prevention for concurrent connection close.

- HBASE-14535: Integration test for rpc connection concurrency / deadlock testing (Enis Soztutar).

- HBASE-14536: Balancer & SSH interfering with each other leading to unavailability.

- HBASE-14598: ByteBufferOutputStream grows its HeapByteBuffer beyond JVM limitations (Ian Friedman).

- HBASE-14635: Fix flaky test TestSnapshotCloneIndependence.

- HBASE-14655: Narrow the scope of doAs() calls to region observer notifications for compaction.

- HBASE-14667: HBaseFsck constructors have diverged (Matteo).

- HBASE-14680: Two configs for snapshot timeout and better defaults.

- HBASE-14712: Increase MasterProcWALs clean up granularity (Matteo Bertozzi).

- HBASE-14805: status should show the master in shell (Enis).

- HBASE-14818: user_permission does not list namespace permissions (li xiang).

- HBASE-14843: TestWALProcedureStore.testLoad is flakey (Matteo Bertozz).

- HBASE-14850: C++ client implementation (TECH PREVIEW. DO NOT DEPLOY IN PROD).

- HBASE-14906: Improvements on FlushLargeStoresPolicy (Yu Li).

- HBASE-14947: WALProcedureStore improvements (Matteo Bertozzi).

- HBASE-14963: Remove Guava dependency from HBase client code (ddas).

- HBASE-15019: Replication stuck when HDFS is restarted (Matteo).

- HBASE-15100: Master WALProcs still never clean up (Matteo Bertozzi).

- HBASE-15125: HBaseFsck's adoptHdfsOrphan function creates region with wrong end key boundary (chenrongwei).

- HBASE-15136: Explore different queuing behaviors while busy.

- HBASE-15142: Procedure v2 - Basic WebUI listing the procedures (Matteo).

- HBASE-15144: Procedure v2 - Web UI displaying Store state.

- HBASE-15146: Don't block on Reader threads.

- HBASE-15221: Reload the cache on re-tried puts in HTableMultiplexer and adds a close() method to HTableMultiplexer.

- HBASE-15232: Handle region location cache mgmt in AsyncProcess for multi()'s.

- HBASE-15292: Refined ZooKeeperWatcher to prevent ZooKeeper's callback while construction (Hiroshi Ikeda).

- HBASE-15306: Make RPC call queue length dynamically configurable.

- HBASE-15377: Per-RS Get metric is time based, per-region metric is size-based (Heng Chen).

- HBASE-15378: Scanner cannot handle heartbeat message with no results (Phil Yang).

- HBASE-15465: userPermission returned by getUserPermission() for the selected namespace does not have namespace set (li xiang).

- HBASE-15518: Add per-table metrics.

- HBASE-15519: Add per-user metrics.

- HBASE-15636: hard coded wait time out value in HBaseTestingUtility#waitUntilAllRegionsAssigned might cause test failure (Stephen Yuan Jiang).

- HBASE-15752: ClassNotFoundException is encountered when custom WAL codec is not found in WALPlayer job.

- HBASE-15767: upgrade httpclient to 4.3.6.

- HBASE-15808: Reduce potential bulk load intermediate space usage and waste (Jerry He).

- HBASE-15817: Backup history should mention the type (full or incremental) of the backup.

- HBASE-15842: SnapshotInfo should display ownership information.

- HBASE-15850: Localize the configuration change in testCheckTableLocks to reduce flakiness of TestHBaseFsck test suite (Stephen Yuan Jiang).

- HBASE-15858: Some region server group shell commands don't work.

- HBASE-15861: Add support for table sets in restore operation .

- HBASE-15862: Backup - Delete- Restore does not restore deleted data.

- HBASE-15873: ACL for snapshot restore / clone is not enforced.

- HBASE-15880: RpcClientImpl#tracedWriteRequest incorrectly closes HTrace span (Mikhail).

- HBASE-15884: NPE in StoreFileScanner#skipKVsNewerThanReadpoint during reverse scan (Sergey Soldatov).

- HBASE-15899: HBase incremental restore should handle namespaces properly.

- HBASE-15904: Use comma as separator for list of tables in BackupInfo.

- HBASE-15925: compat-module maven variable not evaluated (Sean Busbey).

- HBASE-15928: hbase backup delete command does not remove backup root dir from hdfs.

- HBASE-15931: Add log for long-running tasks in AsyncProcess (Yu Li).

- HBASE-15933: NullPointerException may be thrown from SimpleRegionNormalizer#getRegionSize().

- HBASE-15953: HBase backup set list command does not list any existing sets.

- HBASE-15971: Regression: Random Read/WorkloadC slower in 1.x than 0.98 (Stack).

- HBASE-15972: hbase backup set command should not accept non-existing table.

- HBASE-16006: FileSystem should be obtained from specified path in WALInputFormat#getSplits().

- HBASE-16007: Job's Configuration should be passed to TableMapReduceUtil#addDependencyJars() in WALPlayer.

- HBASE-16012: Major compaction can't work due to obsolete scanner read point in RegionServer (Guanghao Zhang).

- HBASE-16016: Fix overflow deadline value in AssignmentManager#waitForAssignment (Stephen Yuan Jiang).

- HBASE-16023: Fastpath for the FIFO rpcscheduler.

- HBASE-16037: Backup/Restore: Make automatic mode default one (Vladimir Rodionov).

- HBASE-16045: endtime argument for VerifyReplication was incorrectly specified in usage.

- HBASE-16048: Tag InternalScanner with LimitedPrivate(HBaseInterfaceAudience.COPROC).

- HBASE-16052: Improve HBaseFsck Scalability (Ben Lau).

- HBASE-16056: Procedure v2 - fix master crash for FileNotFound (Matteo Bertozzi).

- HBASE-16059: Region normalizer fails to trigger merge action where one of the regions is empty.

- HBASE-16062: Improper error handling in WAL Reader/Writer creation (Vladimir Rodionov).

- HBASE-16064: delete backup command shows HDFS permission error when deleting the intended backup.

- HBASE-16065: hbase backup set describe command does not inform if the set does not exist (Vladimir Rodionov).

- HBASE-16088: hbase restore gives a misleading message if the table doesn't exist.

- HBASE-16132: Scan does not return all the result when regionserver is busy (binlijin).

- HBASE-16237: Blocks for hbase:meta table are not cached in L1 cache.

- HBASE-16267: Remove commons-httpclient dependency from hbase-rest module.

- HBASE-16284: Unauthorized client can shutdown the cluster.

- HBASE-16288: HFile intermediate block level indexes might recurse forever creating multi TB files.

- HBASE-16314: Retry on table snapshot failure.

- HBASE-16315: RegionSizeCalculator prints region names as binary without escapes.

- HBASE-16319: Fix TestCacheOnWrite after HBASE-16288 (Andrew Purtell).

- HBASE-16350: Undo server abort from HBASE-14968.

- HBASE-16358: Backup create command: log success/failure with backup id.

- HBASE-16359: NullPointerException in RSRpcServices.openRegion().

- HBASE-16367: Race between master and region server initialization may lead to premature server abort.

- HBASE-16377: ServerName check is ineffective in region_mover.rb.

- HBASE-16429: FSHLog: deadlock if rollWriter called when ring buffer filled with appends.

- HBASE-16604: Scanner retries on IOException can cause the scans to miss data.

- HBASE-16675: Average region size may be incorrect.

- HBASE-16678: MapReduce jobs do not update counters from ScanMetrics.

- HBASE-16721: Concurrency issue in WAL unflushed seqId tracking.

- HBASE-16750: Updating asciidoctor plugin version.

- HBASE-16824: Writer.flush() can be called on already closed streams in WAL roll.

- HBASE-16931: Setting cell's seqId to zero in compaction flow might cause RS down.

- HBASE-16985: TestClusterId failed due to wrong hbase rootdir.

- HBASE-17292: Add observer notification before bulk loaded hfile is moved to region directory.

- HBASE-5291: Add Kerberos HTTP SPNEGO authentication support to HBase web consoles.

1.8.8. Hive

This release provides Hive 1.2.1 and Hive 2.1.0 in addition to the following patches:

Hive 1.2.1 Apache patches:

- HIVE-7224: Set incremental printing to true by default in Beeline.

- HIVE-7239: Fix bug in HiveIndexedInputFormat implementation that causes incorrect query result when input backed by Sequence/RC files.

- HIVE-9941: sql std authorization on partitioned table: truncate and inser.

- HIVE-10562: Add versioning/format mechanism to NOTIFICATION_LOG entries, expand MESSAGE siz.

- HIVE-10924: add support for MERGE statemen.

- HIVE-11030: Enhance storage layer to create one delta file per writ.

- HIVE-11293: HiveConnection.setAutoCommit(true) throws exception .

- HIVE-11594: Analyze Table for column names with embedded space.

- HIVE-11616: DelegationTokenSecretManager reuses the same objectstore, which has concurrency issue.

- HIVE-11935: Race condition in HiveMetaStoreClient: isCompatibleWith and clos.

- HIVE-12077: MSCK Repair table should fix partitions in batche.

- HIVE-12594: X lock on partition should not conflict with S lock on DB.

- HIVE-12664: Bug in reduce deduplication optimization causing ArrayOutOfBoundException.

- HIVE-12968: genNotNullFilterForJoinSourcePlan: needs to merge predicates into the multi-AND.

- HIVE-13014: RetryingMetaStoreClient is retrying too aggressiveley .

- HIVE-13083: Writing HiveDecimal to ORC can wrongly suppress present strea.

- HIVE-13185: orc.ReaderImp.ensureOrcFooter() method fails on small text files with IndexOutOfBoundsException.

- HIVE-13423: Handle the overflow case for decimal datatype for sum().

- HIVE-13527: Using deprecated APIs in HBase client causes zookeeper connection leaks.

- HIVE-13539: HiveHFileOutputFormat searching the wrong directory for HFiles .

- HIVE-13756: Map failure attempts to delete reducer _temporary dir on pig multi-quer.

- HIVE-13836: DbNotifications giving an error = Invalid state. Transaction has already started.

- HIVE-13872: Queries failing with java.lang.ClassCastException when vectorization is enable.

- HIVE-13936: Add streaming support for row_numbe.

- HIVE-13966: DbNotificationListener: can loose DDL operation notification.

- HIVE-14037: java.lang.ClassNotFoundException for the jar in hive.reloadable.aux.jars.path in mapreduc.

- HIVE-14170: Beeline IncrementalRows should buffer rows and incrementally re-calculate width if TableOutputFormat is used .

- HIVE-14229: the jars in hive.aux.jar.paths are not added to session classpath.

- HIVE-14229: the jars in hive.aux.jar.paths are not added to session classpath .

- HIVE-14251: Union All of different types resolves to incorrect data.

- HIVE-14278: Migrate TestHadoop20SAuthBridge.java from Unit3 to Unit.

- HIVE-14279: fix mvn test TestHiveMetaStore.testTransactionalValidatio.

- HIVE-14290: Refactor HIVE-14054 to use Collections#newSetFromMap.

- HIVE-14375: hcatalog-pig-adaptor pom.xml uses joda-time 2.2 instead of ${joda.version} that uses 2.8.1.

- HIVE-14399: Fix test flakiness of org.apache.hive.hcatalog.listener.TestDbNotificationListener.cleanupNotifs.

- HIVE-14436: Hive 1.2.1/Hitting "ql.Driver: FAILED: IllegalArgumentException Error: , expected at the end of 'decimal(9'" after enabling hive.optimize.skewjoin and with MR engine .

- HIVE-14445: upgrade maven surefire to 2.19.1.

- HIVE-14457: Partitions in encryption zone are still trashed though an exception is returned.

- HIVE-14519: Multi insert query bug .

- HIVE-14520: We should set a timeout for the blocking calls in TestMsgBusConnection.

- HIVE-14591: HS2 is shut down unexpectedly during the startup time .

- HIVE-14607: ORC split generation failed with exception: java.lang.ArrayIndexOutOfBoundsException: 1.

- HIVE-14659: OutputStream won't close if caught exception in funtion unparseExprForValuesClause in SemanticAnalyzer.java .

- HIVE-14690: Query fail when hive.exec.parallel=true, with conflicting session di.

- HIVE-14693: Some paritions will be left out when partition number is the multiple of the option hive.msck.repair.batch.size.

- HIVE-14715: Hive throws NumberFormatException with query with Null value.

- HIVE-14762: Add logging while removing scratch spac.

- HIVE-14773: NPE aggregating column statistics for date column in partitioned table .

- HIVE-14774: Canceling query using Ctrl-C in beeline might lead to stale locks.

- HIVE-14805: Subquery inside a view will have the object in the subquery as the direct input.

- HIVE-14837: JDBC: standalone jar is missing hadoop core dependencie.

- HIVE-14865: Fix comments after HIVE-14350.

- HIVE-14922: Add perf logging for post job completion step.

- HIVE-14924: MSCK REPAIR table with single threaded is throwing null pointer exception.

- HIVE-14928: Analyze table no scan mess up schema.

- HIVE-14929: Adding JDBC test for query cancellation scenari.

- HIVE-14935: Add tests for beeline force optio.

- HIVE-14943: Base Implementation (merge statement.

- HIVE-14948: properly handle special characters in identifier.

- HIVE-14959: Fix DISTINCT with windowing when CBO is enabled/disable.

- HIVE-14966: Backport: JDBC: HiveConnction never saves HTTP cookies.

- HIVE-14992: Relocate several common libraries in hive jdbc uber ja.

- HIVE-14993: make WriteEntity distinguish writeTyp.

- HIVE-15002: HiveSessionImpl#executeStatementInternal may leave locks in an inconsistent state.

- HIVE-15010: Make LockComponent aware if it's part of dynamic partition operatio.

- HIVE-15060: Remove the autoCommit warning from beeline .

- HIVE-15099: PTFOperator.PTFInvocation didn't properly reset the input partition.

- HIVE-15124: Fix OrcInputFormat to use reader's schema for include boolean arra.

- HIVE-15137: metastore add partitions background thread should use current username.

- HIVE-15151: Bootstrap support for replv2.

- HIVE-15178: ORC stripe merge may produce many MR jobs and no merge if split size is small .

- HIVE-15180: Extend JSONMessageFactory to store additional information about metadata objects on different table events.

- HIVE-15231: query on view with CTE and alias fails with table not found error .

- HIVE-15232: Add notification events for functions and indexes.

- HIVE-15284: Add junit test to test replication scenarios.

- HIVE-15291: Comparison of timestamp fails if only date part is provided.

- HIVE-15294: Capture additional metadata to replicate a simple insert at destination.

- HIVE-15307: Hive MERGE: "when matched then update" allows invalid column names.

- HIVE-15322: Skipping "hbase mapredcp" in hive script for certain service.

- HIVE-15327: Outerjoin might produce wrong result depending on joinEmitInterval value .

- HIVE-15332: REPL LOAD & DUMP support for incremental CREATE_TABLE/ADD_PTN.

- HIVE-15333: Add a FetchTask to REPL DUMP plan for reading dump uri, last repl id as ResultSet.

- HIVE-15355: Concurrency issues during parallel moveFile due to HDFSUtils.setFullFileStatu.

- HIVE-15365: Add new methods to MessageFactory API (corresponding to the ones added in JSONMessageFactory).

- HIVE-15366: REPL LOAD & DUMP support for incremental INSERT events.

- HIVE-15426: Fix order guarantee of event executions for REPL LOAD.

- HIVE-15437: avro tables join fails when - tbl join tbl_postfix .

- HIVE-15448: ChangeManager.

- HIVE-15466: REPL LOAD & DUMP support for incremental DROP_TABLE/DROP_PTN.

- HIVE-15469: Fix REPL DUMP/LOAD DROP_PTN so it works on non-string-ptn-key tables.

- HIVE-15472: JDBC: Standalone jar is missing ZK dependencie.

- HIVE-15473: Progress Bar on Beeline clien.

- HIVE-15478: Add file + checksum list for create table/partition during notification creation (whenever relevant.

- HIVE-15522: REPL LOAD & DUMP support for incremental ALTER_TABLE/ALTER_PTN including renames.

- HIVE-15525: Hooking ChangeManager to "drop table", "drop partition.

- HIVE-15534: Update db/table repl.last.id at the end of REPL LOAD of a batch of events.

- HIVE-15542: NPE in StatsUtils::getColStatistics when all values in DATE column are NULL.

- HIVE-15550: fix arglist logging in schematool .

- HIVE-15551: memory leak in directsql for mysql+bonecp specific initialization.

- HIVE-15551: memory leak in directsql for mysql+bonecp specific initialization .

- HIVE-15569: failures in RetryingHMSHandler.

- HIVE-15579: Support HADOOP_PROXY_USER for secure impersonation in hive metastore client.

- HIVE-15588: Vectorization: Fix deallocation of scratch columns in VectorUDFCoalesce, etc to prevent wrong reus.

- HIVE-15589: Flaky org.apache.hadoop.hive.ql.lockmgr.TestDbTxnManager.testHeartbeater .

- HIVE-15668: change REPL DUMP syntax to use "LIMIT" instead of "BATCH" keyword.

- HIVE-15684: Wrong posBigTable used in VectorMapJoinOuterFilteredOperato.

- HIVE-15714: backport HIVE-11985 (and HIVE-12601) to branch-1 .

- HIVE-15717: JDBC: Implement rowDeleted, rowInserted and rowUpdated to return false.

- HIVE-15752: MSCK should add output WriteEntity for table in semantic analysis .

- HIVE-15755: NullPointerException on invalid table name in ON clause of Merge statemen.

- HIVE-15774: Ensure DbLockManager backward compatibility for non-ACID resources .

- HIVE-15803: msck can hang when nested partitions are present.

- HIVE-15830: Allow additional ACLs for tez jobs.

- HIVE-15839: Don't force cardinality check if only WHEN NOT MATCHED is specifie.

- HIVE-15840: Webhcat test TestPig_5 failing with Pig on Tez at check for percent complete of jo.

- HIVE-15846: CNF error without hadoop jars, Relocate more dependencies (e.g. org.apache.zookeeper) for JDBC uber ja.

- HIVE-15846: Relocate more dependencies (e.g. org.apache.zookeeper) for JDBC uber ja.

- HIVE-15847: In Progress update refreshes seem slo.

- HIVE-15848: count or sum distinct incorrect when hive.optimize.reducededuplication set to true.

- HIVE-15851: SHOW COMPACTIONS doesn't show JobI.

- HIVE-15871: Add cross join check in SQL MERGE stm.

- HIVE-15871: enable cardinality check by defaul.

- HIVE-15872: The PERCENTILE_APPROX UDAF does not work with empty se.

- HIVE-15879: Fix HiveMetaStoreChecker.checkPartitionDirs metho.

- HIVE-15889: Some tasks still run after hive cli is shutdow.

- HIVE-15891: Detect query rewrite scenario for UPDATE/DELETE/MERGE and fail fast.

- HIVE-15917: incorrect error handling from BackgroundWork can cause beeline query to hang.

- HIVE-15935: ACL is not set in ATS dat.

- HIVE-15936: ConcurrentModificationException in ATSHoo.

- HIVE-15941: Fix o.a.h.hive.ql.exec.tez.TezTask compilation issue with tez maste.

- HIVE-15950: Make DbTxnManager use Metastore client consistently with caller.

- HIVE-15970: Merge statement implementation clashes with AST rewrite.

- HIVE-15999: Fix flakiness in TestDbTxnManager2.

- HIVE-16014: HiveMetastoreChecker should use hive.metastore.fshandler.threads instead of hive.mv.files.thread for pool siz.

- HIVE-16028: Fail UPDATE/DELETE/MERGE queries when Ranger authorization manager is use.

- HIVE-16045: Print progress bar along with operation lo.

- HIVE-16050: Regression: Union of null with non-nul.

- HIVE-16070: fix nonReserved list in IdentifiersParser..

- HIVE-16086: Fix HiveMetaStoreChecker.checkPartitionDirsSingleThreaded metho.

- HIVE-16090: Addendum to HIVE-1601.

- HIVE-16102: Grouping sets do not conform to SQL standar.

- HIVE-16114: NullPointerException in TezSessionPoolManager when getting the sessio.

- HIVE-16160: OutOfMemoryError: GC overhead limit exceeded on Hs2 longevity test.

- HIVE-16170: Exclude relocation of org.apache.hadoop.security.* in the JDBC standalone ja.

- HIVE-16172: Switch to a fairness lock to synchronize HS2 thrift clien.

- HIVE-16175: Possible race condition in InstanceCache.

- HIVE-16181: Make logic for hdfs directory location extraction more generic, in webhcat test drive.

Hive 2.1.0 Apache Patches:

- HIVE-9941: sql std authorization on partitioned table: truncate and insert.

- HIVE-12492: Inefficient join ordering in TPCDS query19 causing 50-70% slowdown.

- HIVE-14278: Migrate TestHadoop23SAuthBridge.java from Unit3 to Unit4.

- HIVE-14360: Starting BeeLine after using !save, there is an error logged: "Error setting configuration: conf".

- HIVE-14362: Support explain analyze in Hive.

- HIVE-14367: Estimated size for constant nulls is 0.

- HIVE-14405: Have tests log to the console along with hive.log.

- HIVE-14432: LLAP signing unit test may be timing-dependent.

- HIVE-14445: upgrade maven surefire to 2.19.1.

- HIVE-14612: org.apache.hive.service.cli.operation.TestOperationLoggingLayout.testSwitchLogLayout failure .

- HIVE-14655: LLAP input format should escape the query string being passed to getSplits().

- HIVE-14929: Adding JDBC test for query cancellation scenario.

- HIVE-14935: Add tests for beeline force option.

- HIVE-14959: Fix DISTINCT with windowing when CBO is enabled/disabled.

- HIVE-14959: Fix DISTINCT with windowing when CBO is enabled/disabled.

- HIVE-15002: HiveSessionImpl#executeStatementInternal may leave locks in an inconsistent state.

- HIVE-15069: Optimize MetaStoreDirectSql:: aggrColStatsForPartitions during query compilation.

- HIVE-15084: Flaky test: TestMiniTezCliDriver:explainanalyze_1, 2, 3, 4, 5.

- HIVE-15099: PTFOperator.PTFInvocation didn't properly reset the input partition.

- HIVE-15570: LLAP: Exception in HostAffinitySplitLocationProvider when running in container mode.

- HIVE-15668: change REPL DUMP syntax to use "LIMIT" instead of "BATCH" keyword.

- HIVE-15789: Vectorization: limit reduce vectorization to 32Mb chunks.

- HIVE-15799: LLAP: rename VertorDeserializeOrcWriter.

- HIVE-15809: Typo in the PostgreSQL database name for druid service.

- HIVE-15830: Allow additional ACLs for tez jobs.

- HIVE-15847: In Progress update refreshes seem slow.

- HIVE-15848: count or sum distinct incorrect when hive.optimize.reducededuplication set to true.

- HIVE-15851: SHOW COMPACTIONS doesn't show JobID.

- HIVE-15872: The PERCENTILE_APPROX UDAF does not work with empty set.

- HIVE-15874: Invalid position alias in Group By when CBO failed.

- HIVE-15877: Upload dependency jars for druid storage handler.

- HIVE-15879: Fix HiveMetaStoreChecker.checkPartitionDirs method.

- HIVE-15884: Optimize not between for vectorization.

- HIVE-15903: Compute table stats when user computes column stats.

- HIVE-15928: Druid/Hive integration: Parallelization of Select queries in Druid handler.

- HIVE-15935: ACL is not set in ATS data.

- HIVE-15938: position alias in order by fails for union queries.

- HIVE-15941: Fix o.a.h.hive.ql.exec.tez.TezTask compilation issue with tez master.

- [HIVE-15948](): Failing test: TestCliDriver, TestSparkCliDriver join31.

- [HIVE-15951](): Make sure base persist directory is unique and deleted.

- [HIVE-15955](): Provide additional explain plan info to facilitate display of runtime filtering and lateral joins.

- [HIVE-15958](): LLAP: Need to check why 1000s of ipc threads are created.

- [HIVE-15959](): LLAP: fix headroom calculation and move it to daemon.

- [HIVE-15969](): Failures in TestRemoteHiveMetaStore, TestSetUGIOnOnlyServer.

- [HIVE-15971](): LLAP: logs urls should use daemon container id instead of fake container id.

- [HIVE-15991](): Flaky Test: TestEncryptedHDFSCliDriver encryption_join_with_different_encryption_keys.

- [HIVE-15994](): Grouping function error when grouping sets are not specified.

- [HIVE-15999](): Fix flakiness in TestDbTxnManager2 .

- [HIVE-16002](): Correlated IN subquery with aggregate asserts in sq_count_check UDF.

- [HIVE-16005](): miscellaneous small fixes to help with llap debuggability.

- [HIVE-16010](): incorrect conf.set in TezSessionPoolManager.

- [HIVE-16012](): BytesBytes hash table - better capacity exhaustion handling.

- [HIVE-16013](): Fragments without locality can stack up on nodes.

- [HIVE-16014](): HiveMetastoreChecker should use hive.metastore.fshandler.threads instead of hive.mv.files.thread for pool size.

- [HIVE-16015](): LLAP: some Tez INFO logs are too noisy II.

- [HIVE-16015](): Modify Hive log settings to integrate with tez reduced logging.

- [HIVE-16018](): Add more information for DynamicPartitionPruningOptimization.

- [HIVE-16020](): LLAP: Reduce IPC connection misses.

- [HIVE-16022](): BloomFilter check not showing up in MERGE statement queries.

- [HIVE-16023](): Bad stats estimation in TPCH Query 12.

- [HIVE-16028](): Fail UPDATE/DELETE/MERGE queries when Ranger authorization manager is used.

- [HIVE-16033](): LLAP: Use PrintGCDateStamps for gc logging.

- [HIVE-16034](): Hive/Druid integration: Fix type inference for Decimal DruidOutputFormat.

- [HIVE-16040](): union column expansion should take aliases from the leftmost branch.

- [HIVE-16045](): Print progress bar along with operation log.

- HIVE-16050: Regression: Union of null with non-null.

- HIVE-16054: AMReporter should use application token instead of ugi.getCurrentUser.

- HIVE-16065: Vectorization: Wrong Key/Value information used by Vectorizer.

- HIVE-16067: LLAP: send out container complete messages after a fragment completes.

- HIVE-16068: BloomFilter expectedEntries not always using NDV when it's available during runtime filtering.

- HIVE-16070: fix nonReserved list in IdentifiersParser.g.

- HIVE-16072: LLAP: Add some additional jvm metrics for hadoop-metrics2.

- HIVE-16078: improve abort checking in Tez/LLAP.

- HIVE-16082: Allow user to change number of listener thread in LlapTaskCommunicator.

- HIVE-16086: Fix HiveMetaStoreChecker.checkPartitionDirsSingleThreaded method.

- HIVE-16090: Addendum to HIVE-16014.

- HIVE-16094: queued containers may timeout if they don't get to run for a long time.

- HIVE-16097: minor fixes to metrics and logs in LlapTaskScheduler.

- HIVE-16098: Describe table doesn't show stats for partitioned tables.

- HIVE-16102: Grouping sets do not conform to SQL standard.

- HIVE-16103: LLAP: Scheduler timeout monitor never stops with slot nodes.

- HIVE-16104: LLAP: preemption may be too aggressive if the pre-empted task doesn't die immediately.

- HIVE-16114: NullPointerException in TezSessionPoolManager when getting the session.

- HIVE-16115: Stop printing progress info from operation logs with beeline progress bar.

- HIVE-16122: NPE Hive Druid split introduced by HIVE-15928.

- HIVE-16132: DataSize stats don't seem correct in semijoin opt branch.

- HIVE-16133: Footer cache in Tez AM can take too much memory.

- HIVE-16135: Vectorization: unhandled constant type for scalar argument.

- HIVE-16137: Default value of hive config hive.auto.convert.join.hashtable.max.entries should be set to 40m instead of 4m.

- HIVE-16140: Stabilize few randomly failing tests.

- HIVE-16142: ATSHook NPE via LLAP.

- HIVE-16150: LLAP: HiveInputFormat:getRecordReader: Fix log statements to reduce memory pressure.

- HIVE-16154: Determine when dynamic runtime filtering should be disabled.

- HIVE-16160: OutOfMemoryError: GC overhead limit exceeded on Hs2 longevity tests.

- HIVE-16161: Standalone hive jdbc jar throws ClassNotFoundException.

- HIVE-16167: Remove transitive dependency on mysql connector jar.

- HIVE-16168: log links should use the NM nodeId port instead of web port.

- HIVE-16170: Exclude relocation of org.apache.hadoop.security.* in the JDBC standalone jar.

- HIVE-16172: Switch to a fairness lock to synchronize HS2 thrift client.

- HIVE-16175: Possible race condition in InstanceCache.

- HIVE-16180: LLAP: Native memory leak in EncodedReader.

- HIVE-16190: Support expression in merge statement.

- HIVE-16211: MERGE statement failing with ClassCastException.

- HIVE-16215: counter recording for text cache may not fully work.

- HIVE-16229: Wrong result for correlated scalar subquery with aggregate.

- HIVE-16236: BuddyAllocator fragmentation - short-term fix.

- HIVE-16238: LLAP: reset/end has to be invoked for o.a.h.hive.q.io.orc.encoded.EncodedReaderImpl.

- HIVE-16245: Vectorization: Does not handle non-column key expressions in MERGEPARTIAL mode.

- HIVE-16260: Remove parallel edges of semijoin with map joins.

- HIVE-16274: Support tuning of NDV of columns using lower/upper bounds.

- HIVE-16278: LLAP: metadata cache may incorrectly decrease memory usage in mem manager.

- HIVE-16282: Semijoin: Disable slow-start for the bloom filter aggregate task.

- HIVE-16298: Add config to specify multi-column joins have correlated columns.

- HIVE-16305: Additional Datanucleus ClassLoaderResolverImpl leaks causing HS2 OOM.

- HIVE-16310: Get the output operators of Reducesink when vectorization is on.

- HIVE-16318: LLAP cache: address some issues in 2.2/2.3.

- HIVE-16319: Fix NPE in ShortestJobFirstComparator.

- HIVE-16323: HS2 JDOPersistenceManagerFactory.pmCache leaks after HIVE-14204.

• HIVE-16325: Sessions are not restarted properly after the configured interval.

1.8.9. Kafka

This release provides Kafka 0.10.1.1 with no additional Apache patches.

1.8.10. Knox

This release provides Knox 0.9.0 with no additional Apache patches.

1.8.11. Mahout

In HDP-2.3.x and 2.4.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 2.4.x is from the "mahout-0.10.x" branch of Apache Mahout, as of 19 December 2014, revision 0f037cb03e77c096 in GitHub.

In HDP-2.5.x and 2.6.x, we removed the "commons-httpclient" library from Mahout because we view it as an obsolete library with possible security issues., and upgraded the Hadoop-Client in Mahout to version 2.7.3, the same version used in HDP-2.5. As a result:

• Previously compiled Mahout jobs will need to be recompiled in the HDP-2.5 or 2.6 environment.

• There is a small possibility that some Mahout jobs may encounter "ClassNotFoundException" or "could not load class" errors related to "org.apache.commons.httpclient", "net.java.dev.jets3t", or related class name prefixes. If these errors happen, you may consider whether to manually install the needed jars in your classpath for the job, if the risk of security issues in the obsolete library is acceptable in your environment.

• There is an even smaller possibility that some Mahout jobs may encounter crashes in Mahout's hbase-client code calls to the hadoop-common libraries, due to binary compatibility problems. Regrettably, there is no way to resolve this issue except revert to the HDP-2.4.2 version of Mahout, which may have security issues. Again, this should be very unusual, and is unlikely to occur in any given Mahout job suite.

1.8.12. Oozie

This release provides Oozie 4.2.0 and the following Apache patches:

• OOZIE-1402: Increase retry interval for non-progressing coordinator action with fix value.

• OOZIE-1837: LauncherMainHadoopUtils sensitive to clock skew.

• OOZIE-1922: MemoryLocksService fails if lock is acquired multiple times in same thread and released.

- OOZIE-1976: Specifying coordinator input datasets in more logical ways.

- OOZIE-2030: Configuration properties from global section is not getting set in Hadoop job conf when using sub-workflow action in Oozie workflow.xml.

- OOZIE-2159: 'oozie validate' command should be moved server-side.

- OOZIE-2165: Job log fetching can fail in Oozie HA mode when using doAs impersonation.

- OOZIE-2168: Oozie flow and action names have 50 char limit.

- OOZIE-2173: DISCLAIMER.txt still refers to Apache Incubator.

- OOZIE-2178: fix javadoc to compile on JDK8.

- OOZIE-2185: Make oozie cli source conf/oozie-client-env.sh.

- OOZIE-2187: Add a way to specify a default JT/RM and NN.

- OOZIE-2215: Support glob in FS EL function.

- OOZIE-2222: Oozie UI parent job should be clickable.

- OOZIE-2228: Statustransit service doesn't pick bundle with suspend status.

- OOZIE-2245: Service to periodically check database schema.

- OOZIE-2246: CoordinatorInputCheckCommand does not behave properly when har file is one of data dependency and doesn't exist.

- OOZIE-2250: Show log for WAITING and TIMEDOUT coord actions.

- OOZIE-2251: Expose instrumental matrices in Realtime Graphing tool.

- OOZIE-2257: BundleStatusTransitXCommand doesn't update last modified time.

- OOZIE-2260: Only set "java.io.tmpdir" to "./tmp" for the AM.

- OOZIE-2262: Fix log streaming from other server with start/end filter.

- OOZIE-2264: Fix coord:offset(n,"DAY") to resolve correct data set.

- OOZIE-2266: Fix 'total' actions returned in coordinator job info.

- OOZIE-2268: Update ActiveMQ version for security and other fixes.

- OOZIE-2270: Audit log doesn't get loaded when "audit log" tab is clicked for coord job.

- OOZIE-2271: Upgrade Tomcat to 6.0.44.

- OOZIE-2272: Use Hadoop's CredentialProvider for passwords in oozie-site.

- OOZIE-2277: Honor oozie.action.sharelib.for.spark in Spark jobs.

- OOZIE-2283: Documentation should not say that System.exit is not allowed in Java Action.

- OOZIE-2284: HBaseCredentials should only add hbase-default.xml and hbase-site.xml to actionConf.

- OOZIE-2285: Change in concurrency should trigger coord action ready command.

- OOZIE-2286: Update Log4j and Log4j-extras to latest 1.2.x release.

- OOZIE-2287: Add support for deleting hcat partitions in fs action delete.

- OOZIE-2289: hive-jdbc dependency in core/pom.xml should be compile.

- OOZIE-2291: Hive2 workflow.xml.security should have "cred" in action tag instead of "hive2" tag.

- OOZIE-2293: Oozie 4.0.1 build failed while building Catalog.

- OOZIE-2300: TestAuthFilterAuthOozieClient.testClientAuthTokenCache fails with Hadoop 2.7.0 and later.

- OOZIE-2303: Typo in documentation.

- OOZIE-2305: Compile Oozie with Hive-1.2.0.

- OOZIE-2308: Add support for bundle:conf() function.

- OOZIE-2309: Enable the coord:dateOffset() function in /coordinator-app/datasets/dataset/@initial-instance.

- OOZIE-2312: oozie doesn't purge audit and error log.

- OOZIE-2314: Unable to kill old instance child job by workflow or coord rerun by Launcher.

- OOZIE-2315: TestOozieCLI.testshareLibUpdate_withSecurity fails with Hadoop 2.

- OOZIE-2318: Provide better solution for specifying SSL truststore to Oozie Client.

- OOZIE-2320: TestZKXLogStreamingService.testStreamingWithMultipleOozieServers_coordActionList is failing.

- OOZIE-2322: Oozie Web UI doesn't work with Kerberos in Internet Explorer 10 or 11 and curl.

- OOZIE-2324: A syntax error in the kill node causes the workflow to get stuck and other problems.

- OOZIE-2325: Shell action fails if user overrides oozie.launcher.mapreduce.map.env.

- OOZIE-2328: Coordinator endtime change should check if the last action is in database.

- OOZIE-2329: Make handling yarn restarts configurable.

- OOZIE-2330: Spark action should take the global jobTracker and nameNode configs by default and allow file and archive elements.

- OOZIE-2332: Add ability to provide Hive and Hive 2 Action queries inline in workflows.

- OOZIE-2343: Shell Action should take Oozie Action config and setup HADOOP_CONF_DIR.

- OOZIE-2344: Enabling 'oozie.action.jobinfo.enable' doesn't inject the job information into the map/reduce job's configuration..

- OOZIE-2345: Parallel job submission for forked actions.

- OOZIE-2346: Add sub-workflow information like the super-parent id and workflow depth into the 'oozie.job.info' property.

- OOZIE-2347: Remove unnecessary new Configuration()/new jobConf() calls from oozie.

- OOZIE-2348: Recovery service keeps on recovering coord action of suspended jobs.

- OOZIE-2350: Package changes for release.

- OOZIE-2355: Hive2 Action doesn't pass along oozie configs to jobconf.

- OOZIE-2356: Add a way to enable/disable credentials in a workflow.

- OOZIE-2358: Coord rerun cleanup should reuse hcat connections.

- OOZIE-2360: Spark Action fails due to missing mesos jar.

- OOZIE-2362: SQL injection in BulkJPAExecutor.

- OOZIE-2365: oozie fail to start when smtp password not set.

- OOZIE-2367: fs delete should support skipTrash option.

- OOZIE-2368: coord:dateOffset and coord:dateTzOffset can't be used for coord initial-instance.

- OOZIE-2369: coord:offset doesn't resolve correctly.

- OOZIE-2370: Docs for Coordinator Action Status Notification has wrong property name.

- OOZIE-2376: Default action configs not honored if no <configuration> section in workflow.

- OOZIE-2377: Hive2 Action should not propagate oozie.hive2.* properties to Beeline.

- OOZIE-2378: org.apache.oozie.client.TestOozieCLI.testAdminInstrumentation is flakey.

- OOZIE-2379: org.apache.oozie.command.coord.TestCoordPushDependencyCheckXCommand.testRequeueOnExceptio is flakey.

- OOZIE-2380: Oozie Hive action failed with wrong tmp path.

- OOZIE-2382: org.apache.oozie.action.hadoop.TestPigMain.testPig_withNullExternalID is flakey.

- OOZIE-2385: org.apache.oozie.TestCoordinatorEngineStreamLog.testCoordLogStreaming is flakey.

- OOZIE-2386: org.apache.oozie.util.TestXConfiguration.testSubstituteVar is flakey.

- OOZIE-2391: spark-opts value in workflow.xml is not parsed properly.

- OOZIE-2394: Oozie can execute command without holding lock.

- OOZIE-2397: LAST_ONLY and NONE don't properly handle READY actions.

- OOZIE-2400: Workflow xml configuration parser cannot deal with namespace prefix.

- OOZIE-2401: Typo in twiki docs with FilesAchives instead of FilesArchives.

- OOZIE-2402: oozie-setup.sh sharelib create takes a long time on large clusters.

- OOZIE-2407: AbandonedService should not send mail if there is no abandoned coord.

- OOZIE-2411: Add BCC to oozie email action.

- OOZIE-2413: Kerberos credentials can expire if the KDC is slow to respond.

- OOZIE-2419: HBase credentials are not correctly proxied.

- OOZIE-2428: TestSLAService, TestSLAEventGeneration flakey tests.

- OOZIE-2429: TestEventGeneration test is flakey.

- OOZIE-2430: Add root logger for hive,sqoop action.

- OOZIE-2431: Support hive.metastore.uris as alias for hcat.metastore.uri.

- OOZIE-2432: TestPurgeXCommand fails.

- OOZIE-2434: inconsistent coord action status and workflow job status.

- OOZIE-2435: TestCoordChangeXCommand is flakey.

- OOZIE-2436: Fork/join workflow fails with "oozie.action.yarn.tag must not be null".

- OOZIE-2437: Exclude additional Hadoop dependencies from Spark Sharelib.

- OOZIE-2438: Oozie client "jobs -filter" diagnostic message clarification.

- OOZIE-2439: FS Action no longer uses name-node from global section or default NN.

- OOZIE-2441: SubWorkflow action with propagate-configuration but no global section throws NPE on submit.

- OOZIE-2444: Need conditional logic in bundles.

- OOZIE-2445: Doc for - Specifying coordinator input datasets in more logical ways.

- OOZIE-2446: Job does not fail during submission if non existent credential is specified.

- OOZIE-2447: Illegal character 0x0 oozie client.

- OOZIE-2452: Coordinator Functional Specification - EL Constants Typo.

- OOZIE-2466: Repeated failure of TestMetricsInstrumentation.testSamplers.

- OOZIE-2467: Oozie can shutdown itself on long GC pause.

- OOZIE-2470: Remove infinite socket timeouts in the Oozie email action.

- OOZIE-2471: Show child job url tab for distcp.

- OOZIE-2474: <job-xml> is not being applied to the launcher job.

- OOZIE-2475: Oozie does not cleanup action dir of killed actions.

- OOZIE-2476: When one of the action from fork fails with transient error, WF never joins.

- OOZIE-2481: Add YARN_CONF_DIR in the Shell action.

- OOZIE-2482: Pyspark job fails with Oozie.

- OOZIE-2485: Oozie client keeps trying to use expired auth token.

- OOZIE-2486: TestSLAEventsGetForFilterJPAExecutor is flakey.

- OOZIE-2489: XML parsing is vulnerable.

- OOZIE-2490: Oozie can't set hadoop.security.token.service.use_ip.

- OOZIE-2492: JSON security issue in js code.

- OOZIE-2503: show ChildJobURLs to spark action.

- OOZIE-2504: Create a log4j.properties under HADOOP_CONF_DIR in Shell Action.

- OOZIE-2509: SLA job status can stuck in running state.

- OOZIE-2510: TestLogStreamer.testStreamLog() fails in other timezones.

- OOZIE-2511: SubWorkflow missing variable set from option if config-default is present in parent workflow.

- OOZIE-2512: ShareLibservice returns incorrect path for jar.

- OOZIE-2522: There can be multiple coord submit from bundle in case of ZK glitch.

- OOZIE-2529: Support adding secret keys to Credentials of Launcher.

- OOZIE-2531: Prevent Spark trying for token which is already available.

- OOZIE-2532: patch apply does not handle binary files.

- OOZIE-2535: User can't disable uber mode.

- OOZIE-2537: SqoopMain does not set up log4j properly.

- OOZIE-2541: Possible resource leak in Hive2Credentials.

- OOZIE-2542: Option to disable OpenJPA BrokerImpl finalization.

- OOZIE-2546: Improperly closed resources in OozieDBCLI.

- OOZIE-2547: Add mapreduce.job.cache.files to spark action.

- OOZIE-2548: Flaky test TestZKLocksService.testLockReaper.

- OOZIE-2550: Flaky tests in TestZKUUIDService.java.

- OOZIE-2551: Feature request: epoch timestamp generation.

- OOZIE-2553: Cred tag is required for all actions in the workflow even if an action does not require it.

- OOZIE-2556: TestAbandonedCoordChecker.testCatchupJob is flaky.

- OOZIE-2563: Pass spark-defaults.conf to spark action.

- OOZIE-2567: HCat connection is not closed while getting hcat cred.

- OOZIE-2570: remove -PtestPatchCompile from patch testing as there is no such profile.

- OOZIE-2577: Flaky tests TestCoordActionInputCheckXCommand.testTimeout and testTimeoutWithException.

- OOZIE-2578: Oozie example distcp job fails to run within an encrypted zone with checksum match error.

1.8.13. Phoenix

This release provides Phoenix 4.7.0 and the following Apache patches:

- PHOENIX-86: UDF testing in phoenix is failing with NoSuchMethodError.

- PHOENIX-117: Phoenix is failing to establish connection on secure envs with NoClassDefFoundError.

- PHOENIX-118: Remove very chatty logging and add notification around cache evictions.

- PHOENIX-930: duplicated columns cause query exception and drop table exception (Junegunn Choi, Kalyan Hadoop).

- PHOENIX-1523: Make it easy to provide a tab literal as separator for CSV imports.

- PHOENIX-1754: Try to handle keytab paths on Windows filesystems.

- PHOENIX-2276: Addednum2 to fix test failures.

- PHOENIX-2276: addendum for fixing test failures.

- PHOENIX-2276: Creating index on a global view on a multi-tenant table fails with NPE.

- PHOENIX-2276: Fix test failure.

- PHOENIX-2743: HivePhoenixHandler for big-big join with predicate push down.

- PHOENIX-2748: Disable auto-commit during bulk load.

- PHOENIX-2758: Ordered GROUP BY not occurring with leading PK equality expression.

- PHOENIX-2784: phoenix-spark: Allow coercion of DATE fields to TIMESTAMP when loading DataFrames(Josh Mahonin).

- PHOENIX-2803: Initializing ToCharFunction with the List <Expression> constructor fails.

- PHOENIX-2817: Phoenix-Spark plugin doesn't work in secured env.

- PHOENIX-2820: Investigate why SortMergeJoinIT has a sort in the explain plan.

- PHOENIX-2828: Ordinality should be 1-based in UNNEST WITH ORDINALITY.

- PHOENIX-2845: Timestamp ignores nanos in literal expressions.

- PHOENIX-2856: HBaseManagedTimeTableReuseTest category isn't getting run.

- PHOENIX-2871: Workaround sqlline issue.

- PHOENIX-2878: CURRENT_TIME fails to provide correct time when projected table is using.

- PHOENIX-2886: Union ALL with Char column not present in the table in Query 1 but in Query 2 throw exception (Alicia Ying Shu).

- PHOENIX-2890: Extend IndexTool to allow incremental index rebuilds.

- PHOENIX-2894: Sort-merge join works incorrectly with DESC columns.

- PHOENIX-2898: HTable not closed in ConnectionQueryServicesImpl (Alex Araujo).

- PHOENIX-2900: Unable to find hash cache once a salted table 's first region has split (chenglei).

- PHOENIX-2905: hadoop-2.5.1 artifacts are in the dependency tree.

- PHOENIX-2908: phoenix-core depends on both antlr 3.5 and antlr 2.7.7 (William Yang).

- PHOENIX-2912: Broken IT tests after PHOENIX-2905.

- PHOENIX-2919: PreparedStatement Returns Incorrect Number of Deleted Records.

- PHOENIX-2920: Incorrect Queries on Multi-tenant tables with WHERE clause containing Row Value Constructor.

- PHOENIX-2930: Cannot resolve columns aliased to its own name.

- PHOENIX-2934: Checking a coerce expression at top level should not be necessary for Union All query (Alicia Ying Shu).

- PHOENIX-2935: IndexMetaData cache can expire when a delete and or query running on server.

- PHOENIX-2936: Missing antlr runtime on server side after PHOENIX-2908.

- PHOENIX-2940: Query the stats table and cache stats in the client.

- PHOENIX-2942: Order by incorrect for RVC.

- PHOENIX-2952: Addendum to fix test case failure (Ram).

- PHOENIX-2952: array_length return negative value (Joseph Sun).

- PHOENIX-2961: select on a composite primary key fails when DESC is set in the table.

- PHOENIX-2983: ClassCastException on auto coerce of BIGINT to DECIMAL.

- PHOENIX-2994: Coalesce function may return null.

- PHOENIX-3004: Allow configuration in hbase-site to define realms other than the server's.

- PHOENIX-3006: Fix all ScalarFunctions to implement clone(List) or <init> (List). (James Taylor & Lars Hofhansl).

- PHOENIX-3008: Prevent upgrade of existing multi-tenant table to map to namespace until we support it correctly.

- PHOENIX-3011: Fix missing apache licenses.

- PHOENIX-3013: TO_CHAR fails to handle indexed null value (Junegunn Choi).

- PHOENIX-3014: SELECT DISTINCT pk ORDER BY pk DESC gives the wrong results with salted tables..

- PHOENIX-3015: Any metadata changes may cause unpredictable result when local indexes are using.

- PHOENIX-3021: Using local index during compaction is producing NPE.

- PHOENIX-3057: Set incremental=false for sqlline-thin.

- PHOENIX-3078:. Hive storage handler does not work if phoenix.table.name contains the schema name.

- PHOENIX-3126: Tie a driver instance to a specific user (Prabhjyot Singh).

- PHOENIX-3134: Ingestion of VARBINARY data should fail when data is not base64-encoded varbinary fields bulk load difference between MR/psql and upserts.

- PHOENIX-3136: Do not relocate org.apache.calcite in queryserver/thin-client.

- PHOENIX-3149: Local index got corrupted if intermediate compaction happen during the split..

- PHOENIX-3152: Incorrect comparator in QueryOptimizer may cause IllegalArgumentException.

- PHOENIX-3161: Improve performance of automatic index rebuilding.

- PHOENIX-3164: Cache UGI instances for remote users in PQS.

- PHOENIX-3189: Perform Kerberos login before ConnectionInfo is constructed.

- PHOENIX-3199: ServerCacheClient sends cache to all regions unnecessarily (chenglei).

- PHOENIX-3223: Add `hadoop classpath` to PQS classpath when available.

- PHOENIX-3236: Problem with shading apache commons on Azure..

- PHOENIX-3241: Support both DATE and TIMESTAMP in CONVERT_TZ().

- PHOENIX-32451: Support both DATE and TIMESTAMP in CONVERT_TZ().

- PHOENIX-3246: Treat U+2002 as whitespace in parser.

- PHOENIX-3271: Distribute UPSERT SELECT across cluster.

- PHOENIX-3280: Automatic attempt to rebuild all disabled index(James Taylor).

- PHOENIX-3346: Hive PhoenixStorageHandler doesn't work well with column.

- PHOENIX-3386: PhoenixStorageHandler throws NPE if local tasks executed via child.

- PHOENIX-3387: Hive PhoenixStorageHandler fails with join on numeric fields.

- PHOENIX-3408: arithmetic/mathematical operations with Decimal columns failed in Hive with PheonixStorageHandler..

- PHOENIX-3416: Memory leak in PhoenixStorageHandler.

- PHOENIX-3422: PhoenixQueryBuilder doesn't make value string correctly for char(/varchar) column type..

- PHOENIX-3423: PhoenixObjectInspector doesn't have information on length of the column..

- PHOENIX-3486: RoundRobinResultIterator doesn't work correctly because of setting Scan's cache size inappropriately in PhoenixInputForamt.

- PHOENIX-3490: PhoenixStorageHandler doesn't handler hbase configuration properly and don't work in secured env.

- PHOENIX-3499: Enable null value for quote character for CSVBulkLoad tool.

- PHOENIX-3503: PhoenixStorageHandler doesn't work properly when execution engine of Hive is Tez..

- PHOENIX-3504: Spark integration doesn't work with decimal columns that are using default precision.

- PHOENIX-3505: Avoid NPE on close() in OrderedResultIterator.

- PHOENIX-3512: PhoenixStorageHandler makes erroneous query string when handling between clauses with date constants..

- PHOENIX-3521: Scan over local index may return incorrect result after flush & compaction.

- PHOENIX-3540: BUG-70567 Fix Time data type in Phoenix Spark integration.

- PHOENIX-3544: Fix UNSIGNED_DATE and UNSIGNED_TIME type in phoenix-spark.

- PHOENIX-3565: Set failIfNoTests=false on surefire and failsafe.

- PHOENIX-3608: KeyRange interset should return EMPTY_RANGE when one of it is NULL_RANGE(Rajeshbabu).

- PHOENIX-3609: Detect and fix corrupted local index region during compaction.

- PHOENIX-3610: Fix tableName used to get the index maintainers while creating HalfStoreFileReader for local index store.

- PHOENIX-3652: User requires all Create privilege on 'default' namespace (Ankit Singhal).

- PHOENIX-3662: PhoenixStorageHandler throws ClassCastException (by Jeongdae Kim).

- PHOENIX-3684: Handle _HOST in principal specified in JDBC url.

- PHOENIX-3686: Allow client-authentication to be disabled for PQS.

- PHOENIX-3694: Drop schema does not invalidate schema from the server cache.

- PHOENIX-3698: No-args constructor for IndexedWALEditCodec.

- PHOENIX-3708: Tests introduced in PHOENIX-3346 doesn't work well with failsafe plugin.

1.8.14. Pig

This release provides Pig 0.16.0 and the following Apache patches:

- PIG-5159: Fix Pig not saving grunt history.

- PIG-5175: Upgrade jruby to 1.7.26.

1.8.15. Ranger

This release provides Ranger 0.7.0 and the following Apache patches:

- RANGER-1378: Update MySQL Schema to fix issues related to only_full_group_by restriction of MySQL 5.7 version..

- RANGER-1383: Use resource matchers for filtering service policies.

- RANGER-1392: Hive test connection is failing even if jdbc.url configured is correct to 2.6-maint.

- RANGER-1392: Revert "RANGER-1392: Hive test connection is failing even if jdbc.url configured is correct to 2.6-maint".

- RANGER-1401: Add consolidated db schema script for SQLServer DB flavor.

- RANGER-1405: groups are not shown if exact user name is passed in search filter.

- RANGER-1406: Audit spoolfile not getting created when ranger service user didn't have permission to log into Solr.

- RANGER-1407: Service update transaction log is not generated in some cases.

- RANGER-1409: User role get deleted from table when he tries to update his role to a restricted role.

- RANGER-1413: Fix issues uncovered by static code analysis.

- RANGER-1413: Good coding practice in Ranger recommended by static code analysis.

- RANGER-1417: Ranger Upgrade is failing for Oracle DB flavor.

- RANGER-1422: Ranger Knox Plugin audit doesn't have the access type populated.

- RANGER-1428: In certain scenario user data contains junk email-id.

- RANGER-1434: Enable Group Search First causes issues when Enable Group Sync is disabled - 2.6-maint branch.

- RANGER-1435: Allow different files to be specified for unix based usersync - 2.6-maint.

- RANGER-1435: fixed minor issue of resource filenames from previous commit.

- RANGER-1440: Improve install script to retry failing statements.

- RANGER-1448: Change of import / export icons on Ranger UI.

- RANGER-1453: Ranger KMS failed to start with Exception] : More than one Master Key exists.

- RANGER-1459: Ranger update policy API is failing on Postgres / Oracle for case sensitive ACLs.

- RANGER-1477: 'show databases' fails with access-denied when user doesn't have access to some of the databases.

1.8.16. Slider

This release provides Slider 0.92.0 and the following Apache patches:

- SLIDER-1220: Fix funtest failures when sasl security is configured for registry

1.8.17. Spark

This release provides Spark 1.6.3 and the following Apache patches:

- SPARK-6717: Clear shuffle files after checkpointing in ALS.

- SPARK-6735: Add window based executor failure tracking mechanism for long running service.

- SPARK-6847: Stack overflow on updateStateByKey which followed by a stream with checkpoint set.

- SPARK-7481: Add spark-cloud module to pull in aws+azure object store FS accessors; test integration.

- SPARK-7889: Jobs progress of apps on complete page of HistoryServer shows uncompleted.

- SPARK-10582: using dynamic-executor-allocation, if AM failed, the new AM will be started. But the new AM does not allocate executors to driver.

- SPARK-11137: Make StreamingContext.stop() exception-safe.

- SPARK-11314: Add service API and test service for Yarn Cluster schedulers.

- SPARK-11315: Add YARN extension service to publish Spark events to YARN timeline service (part of SPARK-1537).

- SPARK-11323: Add History Service Provider to service application histories from YARN timeline server (part of SPARK-1537).

- SPARK-11627: Spark Streaming backpressure mechanism has no initial rate limit, receivers receive data at the maximum speed , it might cause OOM exception.

- SPARK-12001: StreamingContext cannot be completely stopped if the stop() is interrupted.

- SPARK-12009: Avoid re-allocate yarn container while driver want to stop all executors.

- SPARK-12142: Can't request executor when container allocator us bit ready.

- SPARK-12241: Improve failure reporting in Yarn client obtainTokenForHBase().

- SPARK-12353: wrong output for countByValue and countByValueAndWIndow.

- SPARK-12513: SocketReceiver hang in Netcat example.

- SPARK-12523: Support long-running of the Spark on HBase and hive metastore.

- SPARK-12920: Fix high CPU usage in Spark thrift server with concurrent users..

- SPARK-12948: OrcRelation uses HadoopRDD which can broadcast conf objects frequently..

- SPARK-12967: NettyRPC races with SparkContext.stop() and throws exception.

- SPARK-12998: Enable OrcRelation even when connecting via spark thrift server..

- SPARK-13021: Fail fast when custom RDD's violate RDD.partition's API contract.

- SPARK-13117: WebUI should use the local ip not 0.0.0.0.

- SPARK-13278: Launcher fails to start with JDK 9 EA.

- SPARK-13308: ManagedBuffers passed to OneToOneStreamManager need to be freed in non error cases.

- SPARK-13360: pyspark related enviroment variable is not propagated to driver in yarn-cluster mode.

- SPARK-13468: Fix a corner case where the page UI should show DAG but it doesn't show.

- SPARK-13478: Use real user when fetching delegation tokens.

- SPARK-13885: Fix attempt id regression for Spark running on Yarn.

- SPARK-13902: Make DAGScheduler not to create duplicate stage.

- SPARK-14062: Fix log4j and upload metrics.properties automatically with distributed cache.

- SPARK-14091: Consider improving performance of SparkContext.getCallSite()..

- SPARK-15067: YARN executors are launched with fixed perm gen size.

- SPARK-1537: Add integration with Yarn's Application Timeline Server.

- SPARK-15705: Change the default value of spark.sql.hive.convertMetastoreOrc to false.

- SPARK-15844: HistoryServer doesn't come up if spark.authenticate = true.

- SPARK-15990: Add rolling log aggregation support for Spark on yarn.

- SPARK-16110: Can't set Python via spark-submit for YARN cluster mode when PYSPARK_PYTHON & PYSPARK_DRIVER_PYTHON are set.

- SPARK-19033: HistoryServer still uses old ACLs even if ACLs are updated.

- SPARK-19306: Fix inconsistent state in DiskBlockObjectWriter when exception occurred.

- SPARK-19970: Table owner should be USER instead of PRINCIPAL in kerberized clusters.

1.8.18. Sqoop

This release provides Sqoop 1.4.6 with no additional Apache patches.

1.8.19. Storm

This release provides Storm 1.1.0 with no additional Apache patches.

1.8.20. Tez

This release provides Tez 0.7.0 and the following Apache patches:

- TEZ-2712: Tez UI: Display the vertex description in the tooltip of vertex in DAG view UI.

- TEZ-2852: TestVertexImpl fails due to race in AsyncDispatcher.

- TEZ-3000: Fix TestContainerReuse.

- TEZ-3115: Shuffle string handling adds significant memory overhead.

- TEZ-3257: Add more unit test coverage for container reuse.

- TEZ-3267: Publish queue name to ATS as part of dag summary.

- TEZ-3404: Move blocking call for YARN Timeline domain creation from client side to AM.

- TEZ-3419: Tez UI: Applications page shows error, for users with only DAG level ACL permission.

- TEZ-3434: Add unit tests for flushing of recovery events.

- TEZ-3470: Tez UI: Make the build work in IBM PPC.

- TEZ-3484: Tez UI: Remove.travis.yml from webapp folder.

- TEZ-3494: Support relative url for tez-ui.history-url.base config.

- TEZ-3496: Tez UI: Optimize display of all tasks table.

- TEZ-3502: Tez UI: Search in All DAGs page doesn't work with numeric values.

- TEZ-3504: Tez UI: Duration is displaying invalid values when start or end time is invalid.

- TEZ-3529: Tez UI: Add 'All Queries' table in the landing page along 'All DAGs' page.

- TEZ-3530: Tez UI: Add query details page, and link the page from All Queries table.

- TEZ-3531: Tez UI: All Queries table: Improve searchability.

- TEZ-3546: Tez UI: On sorting asc - Not Available must be at the top.

- TEZ-3554: Add a link to get to all logs from Tez UI while job is running.

- TEZ-3554: Add a link to get to all logs from Tez UI while job is running.

- TEZ-3555: Tez UI: Build is failing in RHEL6.

- TEZ-3556: Tez UI: Display query configurations.

- TEZ-3566: Avoid caching fs isntances in TokenCache after a point.

- TEZ-3568: Update SecurityUtils configuration to pick user provided configuration.

- TEZ-3570: Tez UI: Wait for sometime before tooltips are displayed.

- TEZ-3571: Tez UI: Display a Total Timeline View for Hive Queries.

- TEZ-3574: Container reuse won't pickup extra dag level local resource.

- TEZ-3575: RM have started forwarding origin. Use that in AMWebController for CORS support.

- TEZ-3580: Tez UI: Pagination broken on queries page.

- TEZ-3582: Exception swallowed in PipelinedSorter causing incorrect results.

- TEZ-3583: Tez UI: UTs are flaky because of a dependency issue.

- TEZ-3591: Tez UI: Logs url is broken if not open in a new tab.

- TEZ-3592: Tez UI: Search issues.

- TEZ-3593: Tez UI: Issues in timeline page.

- TEZ-3594: Tez UI: Graphical view tooltip issues.

- TEZ-3594: Tez UI: Graphical view tooltip issues.

- TEZ-3598: Tez UI: Text formatting changes.

- TEZ-3602: Tez UI: Query Name field is not required.

- TEZ-3609: Improve ATSv15 performance for DAG entities read calls.

- TEZ-3610: TEZ UI 0.7 0.9 compatibility for url query params and tez-app sub-routes.

- TEZ-3615: Tez UI: Table changes.

- TEZ-3615: Tez UI: Table changes.

- TEZ-3619: Tez UI: Improve DAG Data download.

- TEZ-3619: Tez UI: Improve DAG Data download.

- TEZ-3626: Tez UI: First Task Start Time and Last Task Finish Time values are showing up incorrectly on Tez UI.

- TEZ-3627: Use queue name available in RegisterApplicationMasterResponse for publishing to ATS.

- TEZ-3629: Tez UI: Enable the UI to display log links from LLAP.

- TEZ-3629: Tez UI: Enable the UI to display log links from LLAP.

- TEZ-3630: Tez UI: Use DAG status for controlling auto-refresh polling.

- TEZ-3639: Tez UI: Footer pagination is improper in landing page.

- TEZ-3640: Tez UI: Add associated llap application id to queries page.

- TEZ-3640: Tez UI: Add associated llap application id to queries page.

- TEZ-3642: Tez UI: Auto-refresh is not stopping when DAG is the main entity.

- TEZ-3653: Tez UI: Swimlane tooltip is not proper for running DAGs.

1.8.21. Zeppelin

This release provides Zeppelin 0.7.0 and the following Apache patches:

- ZEPPELIN-1588: bumping nvd3 to 1.8.5.

- ZEPPELIN-1725: replace cleanDirectory with forceDelete.

- ZEPPELIN-1964: Layout info is lost after refresh.

- ZEPPELIN-2045: Pass interpreter properties with "spark." as prefix to SparkConf.

- ZEPPELIN-2062: Hive proxy user doesn't work if interpreter name is not JDBC.

- ZEPPELIN-2063: Hive Jdbc interpreter does not relogin if kerberos ticket expired when hive.server2.transport.mode is htt.

- ZEPPELIN-2068: Change interpreter.json & credentials.json permission to 600.

- ZEPPELIN-2075: Can't stop infinite `while` statement in pyspark Interpreter.

- ZEPPELIN-2083: default interpreter list should honour order of zeppelin.interpreters property in zeppelin-site.xml.

- ZEPPELIN-2084: z.show doesn't show dataframe.

- ZEPPELIN-2094: Decrease npm install retry time (for branch-0.7).

- ZEPPELIN-2105: yarn installed by frontend-maven-plugin causes build failure (BUG).

- ZEPPELIN-2133: All interpreters sometimes throw random Connection refuse errors.

- ZEPPELIN-2147: zeppelin should redirect to login page after browser ... ...session expires.

- ZEPPELIN-2148: On creation of Bar graph zeppelin UI shows it as minigraph .

- ZEPPELIN-2149: Each interpreter should have a seperate log file..

- ZEPPELIN-2161: Nested Group Support in LdapRealm for AD.

- ZEPPELIN-2164: Typo in Insufficient Privileges popup of Zeppelin.

- ZEPPELIN-2166: HeliumBundleFactoty can't transfile imported es6+.

- ZEPPELIN-2167: User with insufficient privileges can still restore files by renaming files in/ ou....

- ZEPPELIN-2172: Redirect to home if notebook authentication fails in realtime.

- ZEPPELIN-2173: Duplicate user names populated in the Note permission box.

- ZEPPELIN-2175: Jdbc interpreter sometime doesn't show detailed error message.

- ZEPPELIN-2178: Prevent from cleaning output in "Personalized Mode".

- ZEPPELIN-2199: Fix lapply issue in sparkR.

- ZEPPELIN-2202: Disable personalized mode btn when note is running (branch-0.7).

1.8.22. ZooKeeper

This release provides ZooKeeper 3.4.6 and the following Apache patches:

- ZOOKEEPER-2227: stmk four-letter word fails execution at server while reading trace mask argument.

1.9. Common Vulnerabilities and Exposures

1.9.1. CVE-2016-8746

Summary: Apache Ranger path matching issue in policy evaluation
Severity: Normal
Vendor: Hortonworks
Versions Affected: All HDP 2.5 versions including Apache Ranger versions 0.6.0/0.6.1/0.6.2
Users affected: All users of the ranger policy admin tool.
Impact: Ranger policy engine incorrectly matches paths in certain conditions when a policy contains wildcards and recursive flags.
Fix detail: Fixed policy evaluation logic
Recommended Action: Users should upgrade to HDP 2.5.4+ (with Apache Ranger 0.6.3+).

1.9.2. CVE-2016-8751

Summary: Apache Ranger stored cross site scripting issue
Severity: Normal
Vendor: Hortonworks
Versions Affected: All HDP 2.3/2.4/2.5 versions including Apache Ranger versions 0.5.x/0.6.0/0.6.1/0.6.2
Users affected: All users of the ranger policy admin tool.
Impact: Apache Ranger is vulnerable to a Stored Cross-Site Scripting when entering custom policy conditions. Admin users can store some arbitrary javascript code execute when normal users login and access policies.
Fix detail: Added logic to sanitize the user input.
Recommended Action: Users should upgrade to HDP 2.5.4+ (with Apache Ranger 0.6.3+).

1.10. Fixed Issues

Fixed issues represents selected issues that were previously logged via Hortonworks Support, but are now addressed in the current release. These issues may have been reported in previous versions within the Known Issues section; meaning they were reported by customers or identified by Hortonworks Quality Engineering team.

Incorrect Results

Hortonworks Bug ID	Apache JIRA	Summary
BUG-62638	PHOENIX-3608	Phoenix query dropping/ignoring additional query conditions
BUG-67879	HIVE-15331	Decimal: most multiplication returns incorrect results (NULL)
BUG-68526	HIVE-15493	Output of outer join query is different between HDP 2.3 and HDP 2.5
BUG-69200	HIVE-13936	Incorrect result when a query having left outer join with subquery and hive.optimize.ppd enabled

Hortonworks Bug ID	Apache JIRA	Summary
BUG-69892	HIVE-15338	Wrong result from vectorized DATEDIFF with scalar parameter of type DATE/TIMESTAMP
BUG-73355	SQOOP-3014	Sqoop with HCatalog import loose precision for large numbers that does not fit into double
BUG-76290	HIVE-16050	Union of null with non-null

Other

Hortonworks Bug ID	Apache JIRA	Summary
BUG-65326	RANGER-1355	RangerAuthorizer for Hive MSCK command
BUG-69874	PHOENIX-3505	(fenton) NPE on phoenix jdbc statement close
BUG-70505	HADOOP-12666, HADOOP-12738, HADOOP-12875, HADOOP-13037, HADOOP-13242, HADOOP-13257, HADOOP-13392, HADOOP-13393, HADOOP-13956, HADOOP-13962, HADOOP-14017, HADOOP-14049, HDFS-11240	Support Microsoft Azure Data Lake - as a file system in Hadoop
BUG-70961	N/A	.Hive-staging folders created from Spark hiveContext are not getting cleaned up
BUG-76643	HADOOP-14048	REDO operation of WASB#AtomicRename should create placeholder blob for destination folder

Performance

Hortonworks Bug ID	Apache JIRA	Summary
BUG-58370	HIVE-13878	Vectorized text about 2x slower on this simple query
BUG-62071	HBASE-16675	HBase normalizer default heuristic result in too may splits
BUG-65712	HDFS-10569, HDFS-9484	(fenton) NNThroughputBenchmark $BlockReportStats should not send empty block reports
BUG-67729	HBASE-16773	Integrate HBASE-16773 AccessController should access local region if possible
BUG-68699	HDFS-11047	Remove deep copies of FinalizedReplica to alleviate heap consumption on DataNode
BUG-69510	STORM-2197	NimbusCLient connections leak due to leakage in ThriftCLient
BUG-70717	HADOOP-12185	NetworkTopology is not efficient adding/getting/removing nodes
BUG-71848	MAPREDUCE-6724	Single shuffle to memory must not exceed Integer#MAX_VALUE

Potential Data Loss

Hortonworks Bug ID	Apache JIRA	Summary
BUG-66878	SPARK-16664	Include SPARK-16664 in HDP

Hortonworks Bug ID	Apache JIRA	Summary
BUG-68473	HIVE-15202	compactions can step on each other causing data loss
BUG-73796	HBASE-16649	Port HBASE-16649 - Truncate table with splits preserved can cause both data loss and truncated data appeared again

Query Failure

Hortonworks Bug ID	Apache JIRA	Summary
BUG-45149	HIVE-11985, HIVE-15714	CREATE EXTERNAL TABLE with Hive metastore in Oracle fails with ORA-01461
BUG-67793	HIVE-15488	Vectorizer should avoid vectorizing non-primitive types
BUG-72178	HIVE-16065	[LLAP]: Vectorized tpc-ds query11 with UNION causes ClassCastException ...DecimalColumnVector cannot be cast to ...LongColumnVector
BUG-74311	HIVE-13872	Queries failing with java.lang.ClassCastException when vectorization is enabled
BUG-77383	HIVE-16245	Vectorization: Does not handle non-column key expressions in MERGEPARTIAL mode – tpcds query11 fails with a ClassCastException - DoubleColumnVector to LongColumnVector

Security

Hortonworks Bug ID	Apache JIRA	Summary
BUG-65326	RANGER-1355	RangerAuthorizer for Hive MSCK command
BUG-68038	N/A	Integrate ZK patch for ZK C cli
BUG-68973	N/A	Kafkaspout (old API) failing with secure kafka in HDP 2.5
BUG-69197	PHOENIX-3004	Phoenix Query Server only accepts SPNEGO-authenticated clients from server's realm
BUG-70551	ACCUMULO-4534	(2.5-maint) SCA : XML External Entity Injection (Input Validation and Representation, Data Flow)
BUG-71078	ATLAS-1402	Sanitize user inputs taken in atlas UI
BUG-74097	ATLAS-1508	Atlas AD authentication Active Directory Authentication should authenticate on sAMAccountName attribute
BUG-74129	RANGER-1338	Ranger Plugin failed to download policy when JaasConfig alone is used to set the UGI instead of Principal/Keytab
BUG-74172	PHOENIX-1754	Phoenix jdbc url can not support keytab PATH in window file system
BUG-76127	PHOENIX-3652	User requires all Create privilege on 'default' namespace to connect using sqlline.py - Fenton

Hortonworks Bug ID	Apache JIRA	Summary
BUG-76645	HBASE-14148	Web UI Framable Page

Stability

Hortonworks Bug ID	Apache JIRA	Summary
BUG-43023	HDFS-9038	dfs.datanode.du.reserved is not considered while computing available space (HDFS-9038)
BUG-51037	ZOOKEEPER-1936	Regionservers exit for FATAL exception during initialization
BUG-53884	HIVE-14170, HIVE-7224	beeline OOM on large results with default params
BUG-58185	HIVE-15208	Drilldown link on HS2 UI for a running query results in an NPE
BUG-64503	HBASE-14798	Integrate HBASE-14798 NPE reporting server load causes regionserver abort; causes TestAcidGuarantee to fail
BUG-64573	HBASE-16429	Integrate HBASE-16429 FSHLog: deadlock if rollWriter called when ring buffer filled with appends
BUG-64977	YARN-5476	Pig-HA stuck on nonexistent application id
BUG-65709	HDFS-10569	A bug causes OutOfIndex error in BlockListAsLongs
BUG-65799	RANGER-1173	Improve Ranger database schema import logic and make it more robust
BUG-65984	N/A	(fenton) - Increase the kite sdk version to 1.1 for fixing Sqoop parquet support issues
BUG-66078	HIVE-15181	TxnHandler::performTimeout gets stuck in an infinite loop
BUG-66387	STORM-2042	Nimbus client connections not closed properly causing connection leaks (Storm-2042)
BUG-68046	HADOOP-11780	Prevent IPC reader thread death
BUG-68047	HDFS-10609	Uncaught InvalidEncryptionKeyException during pipeline recovery may abort downstream applications
BUG-68060	STORM-2158	OutOfMemoryError in Nimbus' SimpleTransportPlugin
BUG-68077	HDFS-10301	BlockReport retransmissions may lead to storages falsely being declared zombie
BUG-68526	HIVE-15493	Output of outer join query is different between HDP 2.3 and HDP 2.5
BUG-68737	YARN-5718	TimelineClient and other YARN places shouldn't overwrite hdfs client retry setting which affect NN failed-over case.
BUG-68946	TEZ-3566	HiveServer2 OOM - lots of DFSClient instances retained
BUG-68960	MAPREDUCE-6684, MAPREDUCE-6797	High contention on scanning of user directory under immediate_done in Job History Server

Hortonworks Bug ID	Apache JIRA	Summary
BUG-69235	AMBARI-18936	DataNode JVM heap settings should include CMSInitiatingOccupancy
BUG-69387	YARN-4109	[Fenton] Excessive logging happening while scheduler UI page is visited
BUG-69393	N/A	Falcon feed fails to aws s3 with hdp 2.5; Oozie joda time version does not meet required jar version 2.2 or later
BUG-69396	OOZIE-2705	Oozie Spark action ignores spark.executor.extraJavaOptions and spark.driver.extraJavaOptions
BUG-69659	N/A	Oozie job fails with NPE at LauncherMapper#setRecoveryId
BUG-69892	HIVE-15338	Wrong result from vectorized DATEDIFF with scalar parameter of type DATE/TIMESTAMP
BUG-70331	HADOOP-13863	Hadoop - Azure: Add a new SAS key mode for WASB
BUG-70423	HBASE-17023	Region left unassigned due to AM and SSH each thinking others would do the assignment work
BUG-70674	HADOOP-13565, HADOOP-13890	Access NameNode WebUI and got "GSSException: No valid credentials provided" error when using JDK 1.8.0_60. However, no problem on same cluster if uses jdk-1.7.0_65
BUG-70718	HADOOP-13831	HADOOP-13831 - Correct check for error code to detect Azure Storage Throttling and provide retries
BUG-70957	N/A	Spark Thrift Server failing after 24hrs due to delegation token renewal failure
BUG-71261	HIVE-15645	LLAP jobs won't run because AMs are launched in the wrong queue
BUG-71771	YARN-5555	Yarn RM UI % of Queue at application level is wrong
BUG-71825	HADOOP-12017	Hadoop archives command should use configurable replication factor when closing
BUG-71848	MAPREDUCE-6724	Single shuffle to memory must not exceed Integer#MAX_VALUE
BUG-72167	TEZ-3575	RM have started forwarding origin. Use that in AMWebController for CORS support
BUG-73010	N/A	FullTableBackupProcedure#toStringClassDetails() should protect against empty table list
BUG-73376	TEZ-3494	Support relative url for tez-ui.history-url.base config
BUG-73728	STORM-2360	Storm hive bolt throws NoSuchMethodError and NullPointerException
BUG-74157	HIVE-15579	Port HADOOP_PROXY_USER to fenton
BUG-74311	HIVE-13872	Queries failing with java.lang.ClassCastException when vectorization is enabled

Hortonworks Bug ID	Apache JIRA	Summary
BUG-74312	HIVE-15542	Getting NPE running a simple hive join in Tez engine after HDP upgrade
BUG-74953	RANGER-1379	Seeing HTTP 404 in the logs for PolicyRefresher intermittently for different services
BUG-76495	HIVE-16014	HiveMetastoreChecker should use hive.metastore.fshandler.threads instead of hive.mv.files.thread for pool size
BUG-76739	HADOOP-14150	Implement getHomeDirectory() method in NativeAzureFileSystem
BUG-76956	HDFS-11461	DataNode Disk Outlier Detection
BUG-76957	HDFS-11477	Simplify file IO profiling configuration

Supportability

Hortonworks Bug ID	Apache JIRA	Summary
BUG-66650	HDFS-6184	Capture NN's thread dump when it fails over
BUG-74475	YARN-6145	YARN-6145 Improve log message on fail over
BUG-74594	YARN-4496, YARN-5711, YARN-6145	Connection refused exception call stack creating panic among end users
BUG-74953	RANGER-1379	Seeing HTTP 404 in the logs for PolicyRefresher intermittently for different services
BUG-77461	HDFS-11292	Log lastWrittenTxId info in logSyncAll

Usability

Hortonworks Bug ID	Apache JIRA	Summary
BUG-54415	PHOENIX-2783, PHOENIX-2790, PHOENIX-930	Creating secondary index with duplicate name causes table inaccessible
BUG-58753	PHOENIX-3360	Our default secondary index configuration is wrong for RPC controller
BUG-61383	HIVE-14600	LLAP daemon does not shutdown if registration fails
BUG-64005	HADOOP-13552	Hdfs prints WARN messages during file creation causing issues for Storm
BUG-64511	HDFS-9618	Fix mismatch between log level and guard in BlockManager#computeRecoveryWorkForBlocks
BUG-64573	HBASE-16429	Integrate HBASE-16429 FSHLog: deadlock if rollWriter called when ring buffer filled with appends
BUG-64977	YARN-5476	Pig-HA stuck on nonexistent application id
BUG-65680	TEZ-3433	Tez UI: Searching using wrong ID causes error in all DAGs page
BUG-66066	ACCUMULO-4424, ACCUMULO-4456	(fenton) Do not wait to start Thrift servers until lock is acquired
BUG-66253	SPARK-17612	Spark2-Hive: ParseException for `DESC ~ PARTITION`

Hortonworks Bug ID	Apache JIRA	Summary
BUG-66359	SLIDER-1169	Slider not honoring zookeeper quorum values passed
BUG-66649	HIVE-14933	LLAP may fail to start if Python dependencies like argparse are not there
BUG-67340	PHOENIX-2930	Cannot resolve columns aliased to its own name
BUG-68080	HDFS-11054	Suppress verbose log message in BlockPlacementPolicyDefault
BUG-68146	SPARK-18160	spark.files & spark.jars should not be passed to driver in yarn mode
BUG-68690	HIVE-15120	Hive is not using Storage Based Authorization while dropping external tables.
BUG-68895	PHOENIX-3241	[PHOENIX-3241] Convert_tz doesn't allow timestamp data type
BUG-69590	N/A	Oozie allowing workflow names up to 40 characters
BUG-70061	N/A	Pig job statuses remain 'Running' for succeeded jobs in query history
BUG-70717	HADOOP-12185	NetworkTopology is not efficient adding/getting/removing nodes
BUG-70753	RANGER-698	Ranger policy should support variables like $user
BUG-72376	N/A	Kafka cli command fails when JMX port is added
BUG-72548	HADOOP-13988	WebHDFS through Knox does not work with TDE
BUG-73351	SQOOP-2408, SQOOP-2582	Parquet file support for Sqoop for HDP 2.4.0
BUG-73352	SQOOP-3003	Sqoop import fails to query with split-by/boundary-query using Oracle Date/Timestamp
BUG-73357	SQOOP-2995	backward incompatibility introduced by Custom Tool options
BUG-73358	SQOOP-3033	Sqoop option –skip-dist-cache is not saved as a parameter when saving Sqoop Job
BUG-73359	SQOOP-2999	Sqoop ClassNotFoundException (org.apache.commons.lang3.StringUtils) is thrown when executing Oracle direct import map task
BUG-74129	RANGER-1338	Ranger Plugin failed to download policy when JaasConfig alone is used to set the UGI instead of Principal/Keytab
BUG-74700	HBASE-17460	enable_table_replication can not perform cyclic replication of a table
BUG-75000	HDFS-11391	Cannot write to WebHDFS with username starting with numbers
BUG-75116	HBASE-17648	Integrate HBASE-17648 HBase Table-level synchronization fails between two secured(kerberized) cluster
BUG-75586	PHOENIX-3686	De-couple PQS's use of Kerberos to talk to HBase and client authentication

Hortonworks Bug ID	Apache JIRA	Summary
BUG-75764	SQOOP-2349	Unable read from a locked DB2 AS400 table with "transaction.isolation=TRANSACTION_READ_UNCOMMITTED
BUG-75952	N/A	hive variables being set as conf variables from connection url

1.11. Known Issues

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary
BUG-60904	KNOX-823	Knox	**Component Affected:** Knox **Description of Problem:** When Ambari is being proxied by Apache Knox the QuickLinks are not rewritten to go back through the gateway. **Workaround:** If all access to Ambari is through Knox in the deployment the new Ambari quicklink profile may be used to hide and/or change URLs to go through Knox permanently. A future release will make these reflect the gateway appropriately.
BUG-61512	FALCON-1965	Falcon	**Component Affected:** Falcon **Description of Problem:** When upgrading to HDP 2.6 or later with Falcon, you might encounter the following error when starting the ActiveMQ server: ERROR - [main:] ~ Failed to start ActiveMQ JMS Message Broker. Reason: java.lang.NegativeArraySizeException (BrokerService:528) You might also encounter this error if downgrading Falcon from HDP 2.6 to an earlier release. **Workaround:** If you encounter this error, delete the ActiveMQ history and then restart Falcon. If you want to retain the history, be sure to back up the ActiveMQ history prior to deleting it. ``` cd <ACTIVEMQ_DATA_DIR> rm -rf ./localhost cd /usr/hdp/current/falcon-server su -l <FALCON_USER> ./bin/falcon-stop ./bin/falcon-start ```
BUG-65977	SPARK-14922	Spark	**Component Affected:** Spark **Description of Problem:** Since Spark 2.0.0, `DROP PARTITION BY RANGE` is not supported grammatically. In other words, only '=' is supported while `<', '>', '<=', '>='` aren't. **Associated Error Message:** ``` scala> sql("alter table t drop partition (b<1) ").show org.apache.spark.sql.catalyst.parser.ParseException: mismatched input '<' expecting {')', ','}(line 1, pos 31) == SQL == alter table t drop partition (b<1) -------------------------------^^^ ```

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary
			Workaround: To drop partition, use the exact match with '='. `scala> sql("alter table t drop partition (b=0) ").show`
BUG-68628	SPARK-16605, SPARK-16628	Spark	**Component Affected:** Spark **Description of Problem:** Column names are not getting created for Spark DataFrame **Workaround:** There are two workaround options for this issue: • Use Spark to create the tables instead of Hive. • Set: sqlContext.setConf("spark.sql.hive.convertMetastoreOrc", "false")
BUG-68632	SPARK-18355	Spark	**Component Affected:** Spark **Description of Problem:** The property spark.sql.hive.convertMetastoreOrc is set to "true" by default. This may impact performance. **Workaround:** You can set this property as "false" to gain some performance improvement. However, you cannot use this property for ORC tables with new columns made by `ALTER TABLE`. The default is false for safety.
BUG-70956	N/A	Zeppelin	**Component Affected:** Zeppelin **Description of Problem:** A Hive query submitted to the %jdbc interpreter returns a proxy validation error. **Associated error messages:** • HiveSQLException: Failed to validate proxy privilege of zeppelin for <user> • The hiveserver2.log file lists a permission denied exception: org.apache.hadoop.ipc.RemoteException(org.apache.hadoop.security.AccessCo Permission denied: user=<user>, access=WRITE, inode="/user/<user>":hdfs:hdfs:drwxr-xr-x at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.check(FSPermi ... org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformatio **Workaround:** 1. Create the user account on all worker nodes. For example, if the account is user3, issue the following command: $ adduser user3 2. Restart the %jdbc interpreter.
BUG-70956	N/A	Zeppelin	**Component Affected:** Zeppelin **Description of Problem:** When used with Hive, the %jdbc interpreter might require Hadoop common jar files that need to be added manually. **Workaround:** 1. On the Interpreters page, add a new repository. 2. Set the repository ID to "hortonworks". 3. Set the URL to "http://nexus-private.hortonworks.com/nexus/content/groups/public/".

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary
			4. Identify the version of HDP running on the cluster. You can find this in the Zeppelin UI by clicking on your user ID and choosing "About Zeppelin." The number after the Zeppelin version is the HDP version number. 5. In the Dependencies section on the Interpreters page, remove existing jar files and add the following three jar files, with the correct HDP version for your cluster. The following example uses version 2.6.0.0-484: • org.apache.hive:hive-jdbc::standalone:1.2.1000.2.6.0.0-484 Xerces:xerces:2.4.0 • org.apache.hadoop:hadoop-common:2.7.3.2.6.0.0-484
BUG-74152	PHOENIX-3688	Phoenix	**Component Affected:** Phoenix **Description of Problem:** Rebuild(ALTER INDEX IDX ON TABLE REBUILD) of indexes created on the table having row_timestamp column will result in no data visible to the User for that Index. **Workaround:** Drop the index and recreate the same index. There will not be any extra overhead of recreating index when compared with rebuild Index.
BUG-75179	ZEPPELIN-2170	Zeppelin	**Component Affected:** Zeppelin **Description of Problem:** Zeppelin does not show all WARN messages thrown by spark-shell. At the Zeppelin notebook level, we cannot change the log level that comes as an output. **Associated Error Message:** There is no error message. This error is only related to increasing or decreasing logging in notebook output. **Workaround:** Currently, there is no workaround for this.
BUG-76417	N/A	YARN	**Component Affected:** YARN **Description of problem:** The default heap size (Xmx) for YARN Timeline Service may not be enough depending on your workloads. Please look at yarn.timeline-service.entity-group-fs-store.app-cache-size in conjunction. **Workaround:** Increase the Timeline Service heap size to be a minimum of 8GB, increasing further as the total load increases.
BUG-76996	N/A	Spark	**Component Affected:** Spark (Livy) **Description of Problem:** When upgrading from HDP-2.5.x to HDP-2.6.0, the Livy interpreter is configured with a scope of 'Global'. **Workaround:** After upgrading from HDP 2.5 to HDP 2.6, set the scope for the Livy interpreter.
BUG-76996	N/A	Spark 2 (Livy)	**Description of Problem:** When upgrading from HDP-2.5.x to HDP-2.6.0 and using Spark2, the Livy interpreter is configured with a scope of 'global', and should be changed to 'scoped'. **Workaround:** After upgrading from HDP 2.5 to HDP 2.6, set the interpreter mode for %livy (Spark 2) to "scoped" using the pulldown menu in the %livy section of the Interpreters page.

Hortonworks Bug ID	Apache JIRA	Apache Component	Summary
BUG-77238	ZEPPELIN-2261	Spark	**Component Affected:** Spark (Livy), Zeppelin **Description of Problem:** Encrypted calls to Livy REST APIs are not supported, and Zeppelin calls to Livy over an encrypted connection are not supported.
BUG-78035	ZEPPELIN-1263	Spark	**Component Affected:** Zeppelin **Description of Problem:** spark.driver.memory will not take effect, the driver memory is always 1G. **Workaround:** To change the driver memory, specify it in the SPARK_DRIVER_MEMORY property on the interpreter setting page for your spark interpreter.
BUG-78244	RANGER-1484	Ranger	See RangerUI: Escape of policy condition text entered in the policy form for a complete description and workaround options.
HOPS-35	N/A	Ambari, Falcon	**Description of Problem:** Falcon check service is failing in PPC **Workaround:** You will need to install the Berkeley db jar file into the ambari-server host. You can do this by following these steps: 1. **wget -O je-5.0.73.jar http://search.maven.org/ remotecontent?filepath=com/sleepycat/je/5.0.73/ je-5.0.73.jar** 2. **cp je-5.0.73.jar /usr/share/** 3. **chmod 644 /usr/share/je-5.0.73.jar** 4. **ambari-server setup –jdbc-db=bdb –jdbc-driver=/usr/ share/je-5.0.73.jar** 5. **ambari-server restart** 6. Restart Falcon service

1.12. Documentation Errata

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

1.12.1. RangerUI: Escape of policy condition text entered in the policy form

Component Affected: Ranger

1.12.1.1. Description of Problem

If a user wants to create policy with custom policy conditions and the expression or text contains special characters, then policy enforcement will not work. Special characters are converted into ASCII before saving the policy into the database.

Special Characters: & < > " ` '

For example, the condition `tags.attributes['type']='abc'` would get converted to the following once the policy is saved:User can see the policy condition with these chars by opening policy in edit mode.

```
tags.attds[&#x27;dsds&#x27;]=&#x27;cssdfs&#x27;
```

You can see the policy condition with these characters by opening the policy in edit mode.

1.12.1.2. Workaround

1.12.1.2.1. Option #1: Create/Update policy via Ranger Rest API

REST URL: `http://<host>:6080/service/plugins/policies`

Creating policy with policy condition:

The below example will create policy with tags as `tags-test` and assign it to `public` group with policy condition as `tags.attr['type']=='abc'` by selecting all hive component permission like `select,update,create,drop,alter,index,lock,all`.

Example:

```
curl -H "Content-Type: application/json" -X POST http://
localhost:6080/service/plugins/policies -u admin:admin -d
'{"policyType":"0","name":"P100","isEnabled":true,"isAuditEnabled":true,"descr
{"tag":{"values":["tags-
test"],"isRecursive":"","isExcludes":false}},"policyItems":
[{"groups":["public"],"conditions":[{"type":"accessed-after-
expiry","values":[]},{"type":"tag-expression","values":
["tags.attr['type']=='abc'"]}],"accesses":
[{"type":"hive:select","isAllowed":true},
{"type":"hive:update","isAllowed":true},
{"type":"hive:create","isAllowed":true},
{"type":"hive:drop","isAllowed":true},
{"type":"hive:alter","isAllowed":true},
{"type":"hive:index","isAllowed":true},
{"type":"hive:lock","isAllowed":true},
{"type":"hive:all","isAllowed":true}]}],"denyPolicyItems":
[],"allowExceptions":[],"denyExceptions":[],"service":"tagdev"}'
```

Update existing policy with policy condition:

The below example will update policy with tags as `tags-test` and assign it to `public` group with policy condition as `tags.attr['type']=='abc'` by selecting all hive component permission like `select,update,create,drop,alter,index,lock,all`.

REST URL: `http://<host-name>:6080/service/plugins/policies/<policy-id>`

Example:

```
curl -H "Content-Type: application/json" -X PUT
http://localhost:6080/service/plugins/policies/18 -u
admin:admin -d '{"id":18,"guid":"ea78a5ed-07a5-447a-978d-
e636b0490a54","isEnabled":true,"createdBy":"Admin","updatedBy":"Admin","create
{"tag":{"values":
["tags"],"isExcludes":false,"isRecursive":false}},"policyItems":
```

```
[{"accesses":[{"type":"hive:select","isAllowed":true},
{"type":"hive:update","isAllowed":true},
{"type":"hive:create","isAllowed":true},
{"type":"hive:drop","isAllowed":true},
{"type":"hive:alter","isAllowed":true},
{"type":"hive:index","isAllowed":true},
{"type":"hive:lock","isAllowed":true},
{"type":"hive:all","isAllowed":true}],"users":[],"groups":
["public"],"conditions":[{"type":"ip-range","values":
["tags.attributes['type']=abc"]}],"delegateAdmin":false}],"denyPolicyItems":
[],"allowExceptions":[],"denyExceptions":[],"dataMaskPolicyItems":
[],"rowFilterPolicyItems":[]}'
```

1.12.1.2.2. Option #2: Apply Javascript changes

Steps to update JS file :

1. 1. Find out PermissionList.js file under /usr/hdp/current/ranger-admin

2. 2. Find out definition of renderPolicyCondtion function (line no:404).

3. 3. Remove below line from that function i.e under display function(line no:434)

```
val = _.escape(val);//Line No:460
```

After removing the above line, the Ranger UI will allow you to create policies with policy condition that can contain special characters and policy evaluation will be successful for the same policy.