

Release Notes 3

Hortonworks Data Platform

Date of Publish: 2018-12-10



<http://docs.hortonworks.com>

Contents

HDP 3.1.0 Release Notes.....	3
Component Versions.....	3
Descriptions of New Features.....	3
Deprecation Notices.....	4
Terminology.....	4
Removed Components and Product Capabilities.....	5
Testing Unsupported Features.....	5
Descriptions of the Latest Technical Preview Features.....	5
Upgrading to HDP 3.1.0.....	5
Behavioral Changes.....	5
Apache Patch Information.....	6
Accumulo.....	6
Atlas.....	6
Calcite.....	8
DataFu.....	8
Hadoop.....	8
HBase.....	13
Hive.....	16
Kafka.....	24
Knox.....	24
Livy.....	24
Oozie.....	24
Phoenix.....	28
Pig.....	28
Ranger.....	28
Spark.....	30
Sqoop.....	31
Storm.....	31
Tez.....	32
Zeppelin.....	33
ZooKeeper.....	34
Fixed Common Vulnerabilities and Exposures.....	34
Fixed Issues.....	36
Known Issues.....	45
Legal Information.....	53

HDP 3.1.0 Release Notes

This document provides you with the latest information about the Hortonworks Data Platform (HDP) 3.1.0 release.

Component Versions

List of the official Apache component versions for Hortonworks Data Platform (HDP) 3.1.0. To ensure that you are working with the most recent stable software available, you must be familiar with the latest Apache component versions in HDP 3.1.0. You should also be aware of the available Technical Preview components and use them only in a testing environment.

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 install a patch, each of the HDP components should remain at the following package version levels to ensure a certified and supported copy of HDP 3.1.0.

Official Apache component versions for HDP 3.1.0:

- Apache Accumulo 1.7.0
- Apache Atlas 1.1.0
- Apache Calcite 1.16.0
- Apache DataFu 1.3.0
- Apache Druid 0.12.1 (incubating)
- Apache Hadoop 3.1.1
- Apache HBase 2.0.2
- Apache Hive 3.1.0
- Apache Kafka 2.0.0
- Apache Knox 1.0.0
- Apache Livy 0.5.0
- Apache Oozie 4.3.1
- Apache Phoenix 5.0.0
- Apache Pig 0.16.0
- Apache Ranger 1.2.0
- Apache Spark 2.3.2
- Apache Sqoop 1.4.7
- Apache Storm 1.2.1
- Apache TEZ 0.9.1
- Apache Zeppelin 0.8.0
- Apache ZooKeeper 3.4.6

You can use the following [Technical Previews](#) components, but only in a testing environment:

- Apache Superset 0.23.3



Note: For information about open source software licensing and notices, refer to the Licenses and Notices files included with the software installation package.

Descriptions of New Features

Hortonworks Data Platform (HDP) 3.1.0 provides you with several new capabilities. You must learn about the new features and their benefits to use the new features in HDP 3.1.0.

Apache component	Feature areas, benefits, and links to further information
Kafka	<p>Core Capabilities</p> <ul style="list-style-type: none"> • Kafka Streams is now officially supported. Kafka Streams is fully integrated with platform services like Ranger and Schema Registry, and supports secure/unsecure clusters. <p>Security</p> <ul style="list-style-type: none"> • Simplified access control management with support for prefixed ACLs for bulk management of entities. Access control for topic creation for specific topics/topic prefix. • Hostname verification to prevent SSL configuration man in the middle attacks. • Improved Encryption support with faster TLS and CRC32C implementations. Over-the-wire encryption is faster. • Simpler security configuration with SSL trust stores update without broker restart and security for Zookeeper listeners can be configured before starting brokers. <p>Performance</p> <ul style="list-style-type: none"> • Prevent indefinite consumer block with new configuration options. • Windowed aggregation performance in Kafka streams vastly improved. <p>Reliability</p> <ul style="list-style-type: none"> • Quota limit notifications to distinguish network errors from quota limits reached. • Better broker resiliency by reducing memory footprint of message down conversion. • Replication protocol improvements for fixing log divergence during fast leader failover.
Ranger	<p>Integration</p> <ul style="list-style-type: none"> • support Kafka 2.0 in Kafka Ranger Plugin. • Support `DelegationToken` resource type with `Describe` operation for Kafka. • Support for cluster resource type for Kafka. • Support for Create operation with Topic resource in Kafka. • Support for Knox Trusted Proxy. Ranger now supports doAs (Trusted proxy).

Deprecation Notices

When upgrading from one Hortonworks Data Platform release to another, you must be aware of any technology that is deprecated, moved, or removed from the latest release. Use the deprecation notices as a guide for your implementation plans.

Terminology

Items in this section are designated as follows:

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 through 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.

Removed Components and Product Capabilities

No components are deprecated or removed in this HDP release.

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

Testing Unsupported Features

Unsupported features are those features that are not officially supported in HDP 3.1.0. Although the features listed in the Technical Preview and Community Features sections exist within HDP, you must not use these features in production because Hortonworks does not support them.

Descriptions of the Latest Technical Preview Features

If you want to experiment with the Technical Preview features included in Hortonworks Data Platform 3.1.0, you should be aware that these features 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](#).

Apache Component	Technical Preview Feature
Spark	Native Atlas connector or hook for Spark to capture metadata and lineage
YARN	Introduced in a previous release: <ul style="list-style-type: none"> • Cloud Service Metrics for YARN (YARN-4488) • Cloud Service Health for YARN • In-Place upgrade of long-running applications in YARN

Upgrading to HDP 3.1.0

You must use Ambari to either do a major (HDP 2.6.x to HDP 3.1.0) or a minor (HDP 3.0.x to HDP 3.1.0) upgrade.

If you already have HDP 2.6.x or HDP 3.0.x installed, upgrading your cluster to HDP 3.1.0 means:

- Keeping the same configuration files you used for HDP 2.6.x or HDP 3.0.x.
- Keeping the same data and metadata in the same location you used for HDP 2.6.x or HDP 3.0.x.
- Installing any new components (added for the first time in HDP 2.6.x or HDP 3.0.x) side-by-side with existing components.

Before you begin

- Ensure that you know whether you want to do a major or a minor upgrade.
- Ensure that you know which HDP components you need to upgrade at your installation.
- Read the Prepare to Upgrade section in the *Ambari Upgrade Guide*.
- See the *Ambari Upgrade Guide* for upgrade information and instructions.

Related Information

[Ambari Major Upgrade Guide](#)

[Ambari Minor Upgrade Guide](#)

[Ambari Major Upgrade Guide for IBM Power Systems](#)

Behavioral Changes

Behavioral changes denote a marked change in behavior from the previously released version to this version of HDP. There are no behavioral changes in this version when compared to the previously released version.

Table 1: Behavioral Changes

Hortonworks Bug ID	Apache Component	Apache JIRA	Summary	Details
BUG-113252	Hive	N/A	Kerberos users connected to Hive using Hive Warehouse Connector no longer have to reconfigure for client and cluster modes	<p>Scenario: The Spark property <code>spark.security.credentials.hiveserver2.enabled</code> is configured based on whether you are running in client-mode or cluster-mode.</p> <p>Previous behavior:</p> <p>On Kerberized clusters, you have to toggle the Spark property <code>spark.security.credentials.hiveserver2.enabled</code> depending on whether you are running in client-mode or cluster-mode.</p> <p>New behavior:</p> <p>Hive automatically detects whether a Spark session is running on client-mode or cluster-mode, and automatically sets the <code>spark.security.credentials.hiveserver2.enabled</code> parameter.</p> <p>Expected Customer Action:</p> <p>When connected to Hive using the <code>HiveWarehouseConnector</code>, you no longer have to toggle the Spark property <code>spark.security.credentials.hiveserver2.enabled</code>.</p>

Apache Patch Information

The following sections list patches in each HDP 3.1.0 component beyond what was fixed in the base version of the Apache component.

Accumulo

This release provides Accumulo 1.7.0 and no additional Apache patches.

HDP 3.0.0 and HDP 3.0.1 provided Accumulo 1.7.0 and no additional Apache patches.

Atlas

This release provides Atlas 1.1.0 and the listed Apache patches.

- [ATLAS-2961](#): updated Hive hook to recognize configuration 'hdfs_path.convert_to_lowercase'
- [AATLAS-2950](#): UI - In newer versions of chrome i.e 70 onwards, if a new column is added the instead of right most it appears at the left most
- [ATLAS-2943](#): Export options null check added.
- [ATLAS-2900](#): Export connected addressed case where imported file is not usable.
- [BUG-113089](#): fix build issue for npm mime-db compatible version issue
- [ATLAS-2927](#): Update lineage query for Process entities
- [ATLAS-2926](#): ZipSink OOM
- [ATLAS-2921](#): Update import scripts exit status in case of java error
- [ATLAS-2924](#): Use encoded property name in `getEncodedProperty()` and `setEncodedProperty()` methods
- [ATLAS-1950](#): Import API: Improvement: Specify Supertypes in Import Transforms
- [ATLAS-2797](#): Atlas Cluster.
- [ATLAS-2798](#): Export & Import Audits.
- [ATLAS-2802](#): Atlas Client Update for Export and Import.
- [ATLAS-2804](#): Export & Import Detailed Audits.
- [ATLAS-2805](#): Applying classification updates modificationTimestamp
- [ATLAS-2806](#): Using replication attributes during export and import process.
- [ATLAS-2811](#): Skip Lineage Export option.
- [ATLAS-2738](#): Export Process: Support for incremental export.
- [ATLAS-2820](#): UI : Add replication audits tab for AtlasCluster entity to show Export & Import audit

- [ATLAS-2814](#): Cluster stores replication details.
- [ATLAS-2818](#): Entity tagging after import.
- [ATLAS-2843](#): AtlasClient updates for exportData and importData.
- [ATLAS-2845](#): Rename AtlasCluster to AtlasServer.
- [ATLAS-2856](#): added utility methods to RequestContext, to find number of active requests and earliest active request-time
- [ATLAS-2862](#): Incremental Export now uses request context to determine change marker.
- [ATLAS-2870](#): Improvement to AddClassification transform to use filters.
- [ATLAS-2869](#): Hdfs_path if requested are created and then proceeds with export.
- [ATLAS-2864](#): Improved incremental export queries.
- [ATLAS-2873](#): Atlas Import Transform Handler Implementation
- [ATLAS-2874](#): Include handling of Atlas Entity Transformers in current Import logic
- [ATLAS-2875](#): Implement clear attribute value transformer for Atlas Entity Transformer
- [ATLAS-2882](#): AddClassification transform for new transforms
- [ATLAS-2886](#): Support for fully qualified server name
- [ATLAS-2892](#): Delete by name REST endpoint.
- [ATLAS-2895](#): Server full name processing
- [ATLAS-2888](#): Change marker fix for server name. Unit test fix.
- [ATLAS-2906](#): Allow transforms to be applied when entity-level transforms are present.
- [ATLAS-2897](#): Better handling of empty zip files. Unit test fix.
- [ATLAS-2909](#): ChangeMarker updated during initialization.
- [ATLAS-2923](#): Options copied to AttributeDefinition
- [ATLAS-2922](#): Multiplicity computation updated for SET/optional
- [ATLAS-2724](#): UI enhancement for Avro schemas and other JSON-valued attributes
- [ATLAS-2786](#): updated HBase client to use all zookeeper entries
- [ATLAS-2827](#): Fix to handle failure in saving indexable string property of large size
- [ATLAS-2844](#): added eventTime field in EntityNotificationV2
- [ATLAS-2847](#): UI - Add grunt scripts in package.json
- [ATLAS-2848](#): Fix property name in ApplicationProperties
- [ATLAS-2849](#): Basic search result list for classification= '_NOT_CLASSIFIED' returns type vertices in addition to matching entity vertices
- [ATLAS-2852](#): Export Import simultaneous requests updated error code.
- [ATLAS-2853](#): updated to send entity-notifications after successful graph transaction commit
- [ATLAS-2857](#): Fix Atlas release version in /api/atlas/version API.
- [ATLAS-2860](#): Remove dependency on atlas-application.properties for AtlasClient.
- [ATLAS-2872](#): updated HiveHook to purge name cache periodically
- [ATLAS-2877](#): updated notification processing to wait only before retry
- [ATLAS-2878](#): avoid retrieval of entityWithExtInfo when extInfo is not needed
- [ATLAS-2891](#): updated hook notification processing with option to ignore potentially incorrect hive_column_lineage
- [ATLAS-2905](#): Generate lineage information for process entities
- [ATLAS-2907](#): UI - show lineage information for process entities
- [ATLAS-2917](#): Filter added for Lineage to hide Process entity.
- [ATLAS-2918](#): Update JanusGraph version to 0.3.1
- [ATLAS-2920](#): Update JanusGraph Solr clients to use all zookeeper entries

HDP 3.0.0 and HDP 3.0.1 provided Atlas 1.0.0 and the following Apache patches:

- [ATLAS-2740](#): Updated import-hbase.sh script to populate values for missed HBase column-family attributes.
- [ATLAS-2751](#): Atlas is not consuming messages from ATLAS_HOOK topic after recovering from zookeeper connection timeout.

Calcite

This release provides Calcite 1.16.0 and no additional Apache patches.

HDP 3.0.0 and HDP 3.0.1 provided Calcite 1.16.0 and no additional Apache patches.

DataFu

This release provides DataFu 1.3.0 and no additional Apache patches.

HDP 3.0.0 and HDP 3.0.1 provided DataFu 1.3.0 and no additional Apache patches.

Hadoop

This release provides Hadoop Common 3.1.1 and no additional Apache patches.

HDP 3.0.1 provided Hadoop Common 3.1.1 and no additional Apache patches.

HDP 3.0.0 provided Hadoop Common 3.1.0 and the following Apache patches:

- [HADOOP-15571](#): Multiple FileContexts created with the same configuration object should be allowed to have different umask. Contributed by Vinod Kumar Vavilapalli.
- [HDFS-13712](#): BlockReaderRemote.read() logging improvement. Contributed by Gergo Repas.
- [HDFS-13715](#): diskbalancer does not work if one of the blockpools are empty on a Federated cluster. Contributed by Bharat Viswanadham
- [HDFS-13715](#): Fix TestNameNodePrunesMissingStorages test failures on Windows. Contributed by Lukas Majercak.
- [HDFS-13635](#): Incorrect message when block is not found. Contributed by Gabor Bota.
- [HADOOP-15554](#): Improve JIT performance for Configuration parsing. Contributed by Todd Lipcon.
- [HDFS-13702](#): Remove HTrace hooks from DFSClient to reduce CPU usage. Contributed by Todd Lipcon.
- [HDFS-13703](#): Avoid allocation of CorruptedBlocks hashmap when no corrupted blocks are hit. Contributed by Todd Lipcon.
- [HADOOP-15574](#): Suppress build error if there are no docs after excluding private annotations. Contributed by Takano Asanuma.
- [HADOOP-15548](#): Randomize local dirs. Contributed by Jim Brennan.
- [HDFS-13705](#): The native ISA-L library loading failure should be made warning rather than an error message. Contributed by Shashikant Banerjee.
- [HADOOP-15458](#): TestLocalFileSystem#testFSOutputStreamBuilder fails on Windows. Contributed by Xiao Liang.
- [HDFS-13692](#): StorageInfoDefragmenter floods log when compacting StorageInfo TreeSet. Contributed by Bharat Viswanadham.
- [HDFS-13682](#): Cannot create encryption zone after KMS auth token expires.
- [HADOOP-15527](#): Improve delay check for stopping processes. Fixed script location. Contributed by Vinod Kumar Vavilapalli
- [HADOOP-15527](#): Improve delay check for stopping processes. Contributed by Vinod Kumar Vavilapalli
- [HADOOP-15533](#): Make WASB listStatus messages consistent. Contributed by Esfandiar Manii
- [HADOOP-15523](#): Shell command timeout given is in seconds whereas it is taken as millisec while scheduling. Contributed by Bilwa S T.
- [HDFS-13681](#): Fix TestStartup.testNNFailToStartOnReadOnlyNNDir test failure on Windows. Contributed by Xiao Liang.
- [HDFS-13686](#): Add overall metrics for FSNamesystemLock. Contributed by Lukas Majercak.
- [HDFS-13174](#): hdfs mover -p /path times out after 20 min. Contributed by Istvan Fajth.
- [HDFS-13676](#): TestEditLogRace fails on Windows. Contributed by Zuoming Zhang.
- [HDFS-13673](#): TestNameNodeMetrics fails on Windows. Contributed by Zuoming Zhang.
- [HDFS-13675](#): Speed up TestDFSAdminWithHA. Contributed by Lukas Majercak.
- [HDFS-13563](#): TestDFSAdminWithHA times out on Windows. Contributed by Lukas Majercak.
- [HDFS-13641](#): Add metrics for edit log tailing. Contributed by Chao Sun.

- [BUG-103209](#):Dependency upgrades related to blackduck scan.
- [HADOOP-15527](#): Improve delay check for stopping processes. Contributed by Vinod Kumar Vavilapalli
- [HADOOP-15532](#): TestBasicDiskValidator fails with NoSuchFileException. Contributed by Giovanni Matteo Fumarola.
- [HDFS-13511](#): Provide specialized exception when block length cannot be obtained. Contributed by Gabor Bota.
- [HADOOP-15529](#): ContainerLaunch#testInvalidEnvVariableSubstitutionType is not supported in Windows. Contributed by Giovanni Matteo Fumarola.
- [HADOOP-15506](#): Upgrade Azure Storage Sdk version to 7.0.0 and update corresponding code blocks. Contributed by Esfandiar Manii.
- [HDFS-13653](#): Make dfs.client.failover.random.order a per nameservice configuration. Contributed by Ekanth Sethuramalingam.
- [HADOOP-15499](#): Performance severe drops when running RawErasureCoderBenchmark with NativeRSRawErasureCoder. Contributed by Sammi Chen.
- [HDFS-13667](#): Typo: Marking all datadoes as stale. Contributed by Nanda Kumar
- [HDFS-12670](#): can't renew HDFS tokens with only the hdfs client jar. Contributed by Arpit Agarwal.
- [HDFS-13664](#): Refactor ConfiguredFailoverProxyProvider to make inheritance easier. Contributed by Chao Sun.
- [HADOOP-15516](#): Add test cases to cover FileUtil#readLink. Contributed by Giovanni Matteo Fumarola.
- [HDFS-13659](#): Add more test coverage for contentSummary for snapshottable path. Contributed by Wei-Chiu Chuang.
- [HDFS-12950](#): [oiv] ls will fail in secure cluster. Contributed by Wei-Chiu Chuang.
- [HADOOP-15217](#): FsURLConnection does not handle paths with spaces. Contributed by Joseph Fourny and Zsolt Venczel.
- [HDFS-13545](#): "guarded" is misspelled as "gaurded" in FSPermissionChecker.java. Contributed by Jianchao Jia.
- [HDFS-13650](#): Randomize baseDir for MiniDFSCluster in TestDFSStripedInputStream and TestDFSStripedInputStreamWithRandomECPolicy. Contributed by Anbang Hu.
- [HDFS-13649](#): Randomize baseDir for MiniDFSCluster in TestReconstructStripedFile and TestReconstructStripedFileWithRandomECPolicy. Contributed by Anbang Hu.
- [HDFS-13652](#): Randomize baseDir for MiniDFSCluster in TestBlockScanner. Contributed by Anbang Hu.
- [HADOOP-15137](#): ClassNotFoundException: org.apache.hadoop.yarn.server.api.DistributedSchedulingAMProtocol when using hadoop-client-minicluster. Contributed by Bharat Viswanadham
- [HDFS-13281](#) Namenode#createFile should be /.reserved/raw/ aware.. Contributed by Rushabh S Shah
- [HDFS-13155](#): BlockPlacementPolicyDefault.chooseTargetInOrder Not Checking Return Value for NULL. Contributed by Zsolt Venczel.
- [HDFS-13339](#): Volume reference can't be released and may lead to deadlock when DataXceiver does a check volume. Contributed by liaoyuxiangqin and Zsolt Venczel.
- [HDFS-13648](#): Fix TestGetConf#testGetJournalNodes on Windows due to a mismatch line separator. Contributed by Giovanni Matteo Fumarola.
- [HDFS-13651](#): TestReencryptionHandler fails on Windows. Contributed by Anbang Hu.
- MapReduce JHS should honor yarn.webapp.filter-entity-list-by-user(addendum). Contributed by Sunil Govindan.
- [HDFS-13637](#): RBF: Router fails when threadIndex (in ConnectionPool) wraps around Integer.MIN_VALUE. Contributed by CR Hota.
- [HDFS-13636](#): Cross-Site Scripting vulnerability in HttpServer2 (Contributed by Haibo Yan via Daniel Templeton)
- [HDFS-13582](#): Improve backward compatibility for HDFS-13176 (WebHdfs file path gets truncated when having semicolon (;) inside). Contributed by Zsolt Venczel.
- [HDFS-13602](#): Add checkOperation(WRITE) checks in FSNamesystem. Contributed by Chao Sun.
- [HDFS-12978](#): Fine-grained locking while consuming journal stream. Contributed by Konstantin Shvachko. (cherry picked from commit 442dd87dcd3e9913a1d97793fde879ce51d701e7)
- [HDFS-13631](#): TestDFSAdmin#testCheckNumOfBlocksInReportCommand should use a separate MiniDFSCluster path. Contributed by Anbang Hu.
- [BUG-104243](#):Dependency jackson-databind dependency related to blackduck scan.

- [HADOOP-15406](#): hadoop-nfs dependencies for mockito and junit are not test scope
- [HDFS-13611](#): Unsafe use of Text as a ConcurrentHashMap key in PBHelperClient.
- [HDFS-13629](#): Some tests in TestDiskBalancerCommand fail on Windows due to MiniDFSCluster path conflict and improper path usage. Contributed by Anbang Hu.
- [HDFS-13632](#): Randomize baseDir for MiniJournalCluster in MiniQJMHACluster for TestDFSAdminWithHA. Contributed by Anbang Hu.
- Additional check when unpacking archives. Contributed by Wilfred Spiegelenburg.
- [HADOOP-15497](#): TestTrash should use proper test path to avoid failing on Windows. Contributed by Anbang Hu.
- [HADOOP-15498](#): TestHadoopArchiveLogs (#testGenerateScript, #testPrepareWorkingDir) fails on Windows. Contributed by Anbang Hu.
- [HDFS-13591](#): TestDFSShell#testSetrepLow fails on Windows. Contributed by Anbang Hu.
- [HDFS-13627](#): TestErasureCodingExerciseAPIs fails on Windows. Contributed by Anbang Hu.
- [HADOOP-15449](#): Increase default timeout of ZK session to avoid frequent NameNode failover
- [HDFS-13628](#): Update Archival Storage doc for Provided Storage
- [HDFS-13620](#): Randomize the test directory path for TestHDFSFileSystemContract. Contributed by Anbang Hu.
- [HDFS-13619](#): TestAuditLoggerWithCommands fails on Windows. Contributed by Anbang Hu.
- [HADOOP-15473](#): Configure serialFilter in KeyProvider to avoid UnrecoverableKeyException caused by JDK-8189997. Contributed by Gabor Bota.
- [HDFS-13618](#): Fix TestDataNodeFaultInjector test failures on Windows. Contributed by Xiao Liang.
- [HADOOP-15486](#): Make NetworkTopology#netLock fair. Contributed by Nanda kumar.
- Skip the proxy user check if the ugi has not been initialized. Contributed by Daryn Sharp
- [HDFS-13540](#): DFSStripedInputStream should only allocate new buffers when reading. Contributed by Xiao Chen.
- [HADOOP-15478](#): WASB: hflush() and hsync() regression. Contributed by Thomas Marquardt.
- [HDFS-13493](#): Reduce the HttpServer2 thread count on DataNodes. Contributed by Erik Krogen.
- Additional check when unpacking archives. Contributed by Jason Lowe and Akira Ajisaka.
- [HDFS-13587](#): TestQuorumJournalManager fails on Windows. Contributed by Anbang Hu.
- [HDFS-13588](#): Fix TestFsDatasetImpl test failures on Windows. Contributed by Xiao Liang.
- [HADOOP-15450](#): Avoid fsync storm triggered by DiskChecker and handle disk full situation. Contributed by Arpit Agarwal.
- [HDFS-13601](#): Optimize ByteString conversions in PBHelper.
- [HDFS-13589](#): Add dfsAdmin command to query if upgrade is finalized. Contributed by Hanisha Koneru
- [HADOOP-15299](#): Bump Jackson 2 version to Jackson 2.9.x.
- [HDFS-13593](#): TestBlockReaderLocalLegacy#testBlockReaderLocalLegacyWithAppend fails on Windows. Contributed by Anbang Hu.
- [HDFS-13592](#): TestNameNodePrunesMissingStorages#testNameNodePrunesUnreportedStorages does not shut down cluster properly. Contributed by Anbang Hu.
- [HDFS-13586](#): Fsync fails on directories on Windows. Contributed by Lukas Majercack.
- [HDFS-13560](#): Insufficient system resources exist to complete the requested service for some tests on Windows. Contributed by Anbang Hu.
- [HDFS-13556](#): TestNestedEncryptionZones does not shut down cluster. Contributed by Anbang Hu.
- [HDFS-13554](#): TestDatanodeRegistration#testForcedRegistration does not shut down cluster. Contributed by Anbang Hu.
- [HDFS-13558](#): TestDatanodeHttpXFrame does not shut down cluster. Contributed by Anbang Hu.
- [BUG-103397](#). [HDFS-13589](#). Add dfsAdmin command to query if upgrade is finalized.
- [HADOOP-15469](#): S3A directory committer commit job fails if _temporary directory created under dest. Contributed by Steve Loughran.
- [HADOOP-15250](#): Split-DNS MultiHomed Server Network Cluster Network IPC Client Bind Addr Wrong Contributed by Ajay Kumar
- [HDFS-13570](#): TestQuotaByStorageType,TestQuota,TestDFSOutputStream fail on Windows. Contributed by Anbang Hu.
- [HDFS-13581](#): DN UI logs link is broken when https is enabled. Contributed by Shashikant Banerjee.

- [HDFS-13559](#): TestBlockScanner does not close TestContext properly. Contributed by Anbang Hu.
- [HDFS-13512](#): WebHdfs getFileStatus doesn't return ecPolicy. Contributed by Ajay Kumar.
- [HDFS-13550](#): TestDebugAdmin#testComputeMetaCommand fails on Windows. Contributed by Anbang Hu.
- [HDFS-13557](#): TestDFSAdmin#testListOpenFiles fails on Windows. Contributed by Anbang Hu.
- [BUG-103055](#) Remove aiyun, openstack from hadoop-cloud-project
- [HDFS-13567](#): TestNameNodeMetrics#testGenerateEDEKTime, TestNameNodeMetrics#testResourceCheck should use a different cluster basedir. Contributed by Anbang Hu.
- [HDFS-13548](#): TestResolveHdfsSymlink#testFcResolveAfs fails on Windows. Contributed by Anbang Hu.
- [HDFS-11700](#): TestHDFSServerPorts#testBackupNodePorts doesn't pass on Windows. Contributed by Anbang Hu.
- [HDFS-13551](#): TestMiniDFSCluster#testClusterSetStorageCapacity does not shut down cluster. Contributed by Anbang Hu.
- [HADOOP-15442](#): ITestS3AMetrics.testMetricsRegister can't know metrics source's name. Contributed by Sean Mackrory.
- [HDFS-13544](#): Improve logging for JournalNode in federated cluster.
- [HDFS-13539](#) DFSStripedInputStream NPE when reportChecksumFailure.
- Add 2.9.1 release notes and changes documents
- [BUG-102169](#) Include Google Connector in Hadoop
- [HADOOP-15441](#): Log kms url and token service at debug level. Contributed by Gabor Bota
- [HDFS-13542](#): TestBlockManager#testNeededReplicationWhileAppending fails due to improper cluster shutdown in TestBlockManager#testBlockManagerMachinesArray on Windows. Contributed by Anbang Hu.
- [HDFS-13346](#): RBF: Fix synchronization of router quota and nameservice quota.
- [HADOOP-15454](#): TestRollingFileSystemSinkWithLocal fails on Windows. Contributed by Xiao Liang.
- [HDFS-13537](#): TestHdfsHelper does not generate jceks path properly for relative path in Windows. Contributed by Xiao Liang.
- [HDFS-12981](#): renameSnapshot a Non-Existent snapshot to itself should throw error. Contributed by Kitti Nanasi.
- [HADOOP-15446](#): WASB: PageBlobInputStream.skip breaks HBASE replication. Contributed by Thomas Marquardt
- [HADOOP-14841](#) Kms client should disconnect if unable to get output stream from connection. Contributed by Rushabh S Shah
- [HDFS-13525](#): RBF: Add unit test TestStateStoreDisabledNameservice. Contributed by Yiquan Lin.
- [HDFS-13481](#): TestRollingFileSystemSinkWithHdfs#testFlushThread: test failed intermittently (Contributed by Gabor Bota via Daniel Templeton)
- [HDFS-13512](#) . WebHdfs getFileStatus doesn't return ecPolicy. Contributed by Ajay Kumar.
- [HDFS-13488](#): RBF: Reject requests when a Router is overloaded. Contributed by Inigo Goiri.
- [HDFS-13503](#): Fix TestFsck test failures on Windows. Contributed by Xiao Liang.
- [HADOOP-15390](#): Yarn RM logs flooded by DelegationTokenRenewer trying to renew KMS tokens (xiaochen via rkanter)
- [HDFS-13283](#): Percentage based Reserved Space Calculation for DataNode. Contributed by Lukas Majercak.
- [HDFS-13509](#): Bug fix for breakHardlinks() of ReplicaInfo/LocalReplica, and fix TestFileAppend failures on Windows. Contributed by Xiao Liang.
- [HDFS-13508](#): RBF: Normalize paths (automatically) when adding, updating, removing or listing mount table entries. Contributed by Ekanth S.
- [HDFS-13484](#): RBF: Disable Nameservices from the federation. Contributed by Inigo Goiri.
- [HDFS-13434](#): RBF: Fix dead links in RBF document. Contributed by Chetna Chaudhari.
- [HDFS-13490](#): RBF: Fix setSafeMode in the Router. Contributed by Inigo Goiri.
- [HDFS-13499](#): RBF: Show disabled name services in the UI. Contributed by Inigo Goiri.
- [HDFS-13326](#): RBF: Improve the interfaces to modify and view mount tables. Contributed by Gang Li.
- [HADOOP-15411](#): AuthenticationFilter should use Configuration.getPropsWithPrefix instead of iterator. (Suma Shivaprasad via wangda)
- [HDFS-13336](#): Test cases of TestWriteToReplica failed in windows. Contributed by Xiao Liang.

- [HDFS-13356](#): Balancer:Set default value of minBlockSize to 10mb. Contributed by Bharat Viswanadham
- [BUG-101422](#): Disable HDFS and Hadoop unit tests that are flaky in Mawo.
- [HDFS-13408](#): MiniDFSCluster to support being built on randomized base directory. Contributed by Xiao Liang
- [HDFS-13433](#): webhdfs requests can be routed incorrectly in federated cluster. Contributed by Arpit Agarwal.
- [HDFS-13478](#): RBF: Disabled Nameservice store API. Contributed by Inigo Goiri.
- [HDFS-10183](#): Prevent race condition during class initialization. Contributed by Pavel Avgustinov.
- [HDFS-13453](#): RBF: getMountPointDates should fetch latest subdir time/date when parent dir is not present but / parent/child dirs are present in mount table. Contributed by Dibyendu Karmakar.
- [HADOOP-15396](#): Some java source files are executable
- [HDFS-13464](#): Fix javadoc in FsVolumeList#handleVolumeFailures. Contributed by Shashikant Banerjee
- [HADOOP-15366](#): Add a helper shutdown routine in HadoopExecutor to ensure clean shutdown. Contributed by Shashikant Banerjee.
- [HDFS-13426](#): Fix javadoc in FsDatasetAsyncDiskService#removeVolume. Contributed by Shashikant Banerjee.
- [HDFS-13463](#): Fix javadoc in FsDatasetImpl#checkAndUpdate. Contributed by Shashikant Banerjee.
- [HADOOP-15391](#): Add missing css file in hadoop-aws, hadoop-aliyun, hadoop-azure and hadoop-azure-datalake modules.
- [HDFS-12828](#): OIV ReverseXML Processor fails with escaped characters
- [BUG-100895](#): Ignore UTs failing due to environmental issues.
- [HDFS-13439](#): Add test case for read block operation when it is moved. Contributed by Ajay Kumar.
- [HDFS-13435](#): RBF: Improve the error loggings for printing the stack trace.
- [HDFS-13330](#): ShortCircuitCache#fetchOrCreate never retries. Contributed by Gabor Bota.
- [HDFS-13311](#): RBF: TestRouterAdminCLI#testCreateInvalidEntry fails on Windows. Contributed by Inigo Goiri
- [HDFS-13056](#): Expose file-level composite CRCs in HDFS which are comparable across different instances/ layouts. Contributed by Dennis Huo.
- [HDFS-13466](#): RBF: Add more router-related information to the UI.
- [HDFS-13462](#): Add BIND_HOST configuration for JournalNode's HTTP and RPC Servers. Contributed by Lukas Majercak.
- [HDFS-13129](#): Add a test for DfsAdmin refreshSuperUserGroupsConfiguration. Contributed by Mukul Kumar Singh
- [HDFS-13433](#): webhdfs requests can be routed incorrectly in federated cluster. Contributed by Arpit Agarwal.
- [HDFS-11900](#): Hedged reads thread pool creation not synchronized. Contributed by John Zhuge.
- [HDFS-11043](#): TestWebHdfsTimeouts fails. Contributed by Xiaoyu Yao and Chao Sun.
- [HDFS-13315](#): Add a test for the issue reported in HDFS-11481 which is fixed by HDFS-10997. Contributed by Yongjun Zhang.
- [HADOOP-15332](#): Fix typos in hadoop-aws markdown docs. Contributed by Gabor Bota.
- [HADOOP-15331](#): Fix a race condition causing parsing error of java.io.BufferedInputStream in class org.apache.hadoop.conf.Configuration. Contributed by Miklos Szegedi.
- [HADOOP-15330](#): Remove jdk1.7 profile from hadoop-annotations module
- [HADOOP-15062](#): TestCryptoStreamsWithOpensslAesCtrCryptoCodec fails on Debian 9. Contributed by Miklos Szegedi.
- [HADOOP-14667](#): Flexible Visual Studio support. Contributed by Allen Wittenauer
- [HDFS-13438](#): Fix javadoc in FsVolumeList#removeVolume. Contributed by Shashikant Banerjee.
- [HDFS-13418](#): NetworkTopology should be configurable when enable DFSNetworkTopology. Contributed by Tao Jie.
- [HADOOP-14970](#): MiniHadoopClusterManager doesn't respect lack of format option. Contributed by Erik Krogen
- [HADOOP-15379](#): Make IrqHandler.bind() public. Contributed by Ajay Kumar
- [HDFS-13436](#): Fix javadoc of package-info.java
- [HDFS-7101](#): Potential null dereference in DFSck#doWork(). Contributed by skrho and Ted Yu.
- [HDFS-13427](#): Fix the section titles of transparent encryption document.
- [HDFS-13386](#): RBF: Wrong date information in list file(-ls) result. Contributed by Dibyendu Karmakar.
- [HDFS-13428](#): RBF: Remove LinkedList From StateStoreFileImpl.java. Contributed by BELUGA BEHR.

- [HDFS-13045](#): RBF: Improve error message returned from subcluster. Contributed by Inigo Goiri.
- [HADOOP-15346](#) S3ARetryPolicy for 400/BadArgument to be "fail". Contributed by Steve Loughran.
- [HADOOP-15313](#): TestKMS should close providers.
- [HADOOP-15357](#): Configuration.getPropsWithPrefix no longer does variable substitution. Contributed by Jim Brennan
- [HDFS-13328](#): Abstract ReencryptionHandler recursive logic in separate class. Contributed by Surendra Singh Lilhore.
- [HADOOP-15376](#): Remove double semi colons on imports that make Clover fall over.
- [HDFS-13420](#): License header is displayed in ArchivalStorage/MemoryStorage html pages. Contributed by Akira Ajisaka.
- [HADOOP-15374](#): Add links of the new features of 3.1.0 to the top page
- [HDFS-13384](#): RBF: Improve timeout RPC call mechanism. Contributed by Inigo Goiri.
- [HDFS-13410](#): RBF: Support federation with no subclusters. Contributed by Inigo Goiri.
- [HDFS-13380](#): RBF: mv/rm fail after the directory exceeded the quota limit. Contributed by Yiqun Lin.
- [HDFS-13388](#): RequestHedgingProxyProvider calls multiple configured NNs all the time. Contributed by Jinglun.
- [HDFS-13402](#): RBF: Fix java doc for StateStoreFileSystemImpl. Contributed by Yiran Wu.
- Added CHANGES/RELEASES/Jdiff for 3.1.0 release
- [HDFS-13353](#): RBF: TestRouterWebHDFSContractCreate failed. Contributed by Takanobu Asanuma.
- [HDFS-13350](#): Negative legacy block ID will confuse Erasure Coding to be considered as striped block. (Contributed by Lei (Eddy) Xu).
- [HDFS-13364](#): RBF: Support NamenodeProtocol in the Router. Contributed by Inigo Goiri.
- [HADOOP-15355](#): TestCommonConfigurationFields is broken by HADOOP-15312. Contributed by LiXin Ge.
- [HADOOP-15317](#): Improve NetworkTopology chooseRandom's loop.
- [HDFS-13289](#): RBF: TestConnectionManager#testCleanup() test case need correction. Contributed by Dibyendu Karmakar.
- [HADOOP-15253](#): Should update maxQueueSize when refresh call queue. Contributed by Tao Jie.
- [HADOOP-14999](#): AliyunOSS: provide one asynchronous multi-part based uploading mechanism. Contributed by Genmao Yu.
- [HADOOP-12862](#): LDAP Group Mapping over SSL can not specify trust store. Contributed by Wei-Chiu Chuang and Konstantin Shvachko.
- [HDFS-13087](#): Snapshotted encryption zone information should be immutable. Contributed by LiXin Ge.
- [HADOOP-15342](#): Updating ADLS connector to use the current SDK version (2.2.7). Contributed by Atul Sikaria.

HBase

This release provides HBase 2.0.2 and the listed Apache patches.

- [HBASE-20723](#): Custom hbase.wal.dir results in data loss because we write recovered edits into a different place than where the recovering region server looks for them
- [HBASE-20722](#): Make RegionServerTracker only depend on children changed event
- [HBASE-20668](#): Avoid permission change if ExportSnapshot's copy fails
- [HBASE-20657](#): Retrying RPC call for ModifyTableProcedure may get stuck
- [HBASE-20635](#): Add missing shaded-pb conversion method for UserPermission
- [HBASE-20627](#): Relocate RS Group pre/post hooks from RSGroupAdminServer to RSGroupAdminEndpoint
- [HBASE-20579](#): Include original exception in wrapped exception
- [HBASE-20579](#): Improve snapshot manifest copy in ExportSnapshot
- [HBASE-20547](#): Restore from backup will fail if done from a different file system
- [HBASE-20531](#): RS may throw NPE when close meta regions in shutdown procedure.
- [HBASE-20495](#): REST unit test fails with NoClassDefFoundError against hadoop3
- [HBASE-20196](#): Maintain all regions with same size in memstore flusher
- [HBASE-20136](#): TestKeyValue misses ClassRule and Category annotations
- [HBASE-19735](#): Create a client-tarball assembly

- [HBASE-19478](#): Utilize multi-get to speed up WAL file checking in BackupLogCleaner (Toshihiro Suzuki)
- [HBASE-18135](#): Implement mechanism for RegionServers to report file archival for space quotas
- [HBASE-18133](#): Decrease quota reaction latency by HBase
- [HBASE-17549](#): HBase-Spark Module: Corrected - Incorrect log at println and unwanted comment code
- [HBASE-21440](#): Assign procedure on the crashed server is not properly interrupted
- [HBASE-21344](#): Fix TestMetaShutdownHandler (addendum)
- [HBASE-21384](#): Procedure with holdlock=false should not be restored lock when restarts
- [HBASE-21344](#): hbase:meta location in ZooKeeper set to OPENING by the procedure which eventually failed but precludes Master from assigning it forever
- [HBASE-21215](#): Figure how to invoke hbck2; make it easy to find
- [HBASE-21372](#): Set hbase.assignment.maximum.attempts to Long.MAX
- [HBASE-21363](#): Rewrite the buildingHoldCleanupTracker method in WALProcedureStore
- [HBASE-21364](#): Procedure holds the lock should put to front of the queue after restart
- [HBASE-21342](#): FileSystem in use may get closed by other bulk load call in secure bulkLoad
- [HBASE-21349](#): Do not run CatalogJanitor or Normalizer when cluster is shutting down
- [HBASE-21338](#): Warn if balancer is an ill-fit for cluster size
- [HBASE-21073](#): Redo concept of maintenance mode
- [HBASE-20937](#): ArrayIndexOutOfBoundsException when rolling back procedure
- [HBASE-21321](#): Backport HBASE-21278 to branch-2.1 and branch-2.0 ("Do not rollback successful sub procedures when rolling back a procedure")
- [HBASE-21354](#): Procedure may be deleted improperly during master restarts resulting in 'Corrupt'(addendum)
- [HBASE-21356](#): bulkLoadHFile API should ensure that rs has the source hfile's write permissions
- [HBASE-21336](#): Addendum remove unused code in HBTU
- [HBASE-21355](#): (addendum) replace the expensive reload storefiles with reading the merge result of compacted storefiles and current storefiles
- [HBASE-21336](#): Simplify the implementation of WALProcedureMap
- [HBASE-21355](#): HStore's storeSize is calculated repeatedly which causing the confusing region split
- [HBASE-21348](#): Fix failing TestRegionBypass, broke by HBASE-21291
- [HBASE-21345](#): [hbck2] Allow version check to proceed even though master is 'initializing'.
- [HBASE-21200](#): Memstore flush doesn't finish because of seekToPreviousRow() in memstore scanner.
- [HBASE-21323](#): Should not skip force updating for a sub procedure even if it has been finished
- [HBASE-21075](#): Confirm that we can (rolling) upgrade from 2.0.x and 2.1.x to 2.2.x after HBASE-20881
- [HBASE-21292](#): IdLock.getLockEntry() may hang if interrupted
- [HBASE-21291](#): Add a test for bypassing stuck state-machine procedures
- [HBASE-21323](#): Should not skip force updating for a sub procedure even if it has been finished; ADDENDUM
- [HBASE-21288](#): HostingServer in UnassignProcedure is not accurate
- [HBASE-21323](#): Should not skip force updating for a sub procedure even if it has been finished
- [HBASE-21055](#): NullPointerException when balanceOverall() but server balance info is null
- [HBASE-21198](#): Exclude dependency on net.minidev:json-smart
- [HBASE-21310](#) and [HBASE-21311](#): Addendum fix failed UTs, some UTs are not present on branch-2.1 and some are a bit different in the implementation
- [HBASE-21327](#): Fix minor logging issue where we don't report servername if no associated SCP
- [HBASE-12320](#): [canary] Cleanup of usage and add commentary
- [HBASE-21311](#): Split TestRestoreSnapshotFromClient
- [HBASE-21263](#): Mention compression algorithm along with other storefile details
- [HBASE-21242](#): [amv2] Miscellaneous minor log and assign procedure create improvements; ADDENDUM Fix TestHRegionInfo AND TestRegionInfoDisplay
- [HBASE-21315](#): The getActiveMinProcId and getActiveMaxProcId of BitSetNode are incorrect if there are no active procedure
- [HBASE-21310](#): Split TestCloneSnapshotFromClient
- [HBASE-21266](#): Not running balancer because processing dead regionservers, but empty dead rs list

- [HBASE-21290](#): No need to instantiate BlockCache for master which not carry table
- [HBASE-21260](#): The whole balancer plans might be aborted if there are more than one plans to move a same region
- [HBASE-21309](#): Increase the waiting timeout for TestProcedurePriority
- [HBASE-21271](#): [amv2] Don't throw UnsupportedOperationException when rollback called on Assign/Unassign; spiral of death
- [HBASE-21251](#): Refactor RegionMover
- [HBASE-21259](#): [amv2] Revived deadservers; recreated serverstatenode
- [HBASE-21303](#): [shell] clear_deadservers with no args fails
- [HBASE-21289](#): Remove the log "'hbase.regionserver.maxlogs' was deprecated." in AbstractFSWAL
- [HBASE-21242](#): [amv2] Miscellaneous minor log and assign procedure create improvements
- [HBASE-21254](#): Need to find a way to limit the number of proc wal files
- [HBASE-21103](#): nightly job should make sure cached yetus will run.
- [HBASE-21287](#): Allow configuring test master initialization wait time.
- [HBASE-20764](#): build broken when latest commit is gpg signed
- [HBASE-21271](#): [amv2] Don't throw UnsupportedOperationException when rollback called on Assign/Unassign; spiral of death
- [HBASE-21250](#): Addendum remove unused modification in hbase-server module
- [HBASE-21250](#): Refactor WALProcedureStore and add more comments for better understanding the implementation
- [HBASE-21223](#): [amv2] Remove abort_procedure from shell
- [HBASE-20506](#): Add doc and test for unused RetryCounter, useful-looking utility
- [HBASE-21185](#): - WALPrettyPrinter: Additional useful info to be printed by wal printer tool, for debugability purposes
- [HBASE-21281](#): Upgrade bouncycastle to latest - addendum adds test dependency
- [HBASE-21178](#): [BC break] : Get and Scan operation with a custom converter_class not working
- [HBASE-21282](#): Upgrade to latest jetty 9.2 and 9.3 versions
- [HBASE-21281](#): Upgrade bouncycastle to latest
- [HBASE-21261](#): Add log4j.properties for hbase-rsgroup tests
- [HBASE-19275](#): TestSnapshotFileCache never worked properly
- [HBASE-21196](#): HTableMultiplexer clears the meta cache after every put operation
- [HBASE-21249](#): Add jitter for ProcedureUtil.getBackoffTimeMs
- [HBASE-21233](#): Allow the procedure implementation to skip persistence of the state after a execution
- [HBASE-21237](#): Use CompatRemoteProcedureResolver to dispatch open/close region requests to RS
- [HBASE-21228](#): Memory leak since AbstractFSWAL caches Thread object and never clean later
- Revert "[HBASE-20657](#) Retrying RPC call for ModifyTableProcedure may get stuck"
- [HBASE-21212](#): Wrong flush time when update flush metric
- [HBASE-21208](#): Bytes#toShort doesn't work without unsafe
- Amend [HBASE-20704](#): Sometimes some compacted storefiles are not archived on region close
- [HBASE-21203](#): TestZKMainServer#testCommandLineWorks won't pass with default 4lw whitelist
- [HBASE-21206](#): Scan with batch size may return incomplete cells
- [HBASE-20704](#): Sometimes some compacted storefiles are not archived on region close
- [HBASE-21097](#): Flush pressure assertion may fail in testFlushThroughputTuning
- [HBASE-21189](#): flaky job should gather machine stats
- [HBASE-21190](#): Log files and count of entries in each as we load from the MasterProcWAL store
- [HBASE-21171](#): Reimplement the retry backoff logic for ReopenTableRegionsProcedure
- [HBASE-21179](#): Fix the number of actions in responseTooSlow log
- [HBASE-21174](#): [REST] Failed to parse empty qualifier in TableResource#getScanResource
- [HBASE-21181](#): Use the same filesystem for wal archive directory and wal directory
- [HBASE-21158](#): Empty qualifier cell is always returned when using QualifierFilter

- [HBASE-21144](#): AssignmentManager.waitForAssignment is not stable
- [HBASE-21143](#): Update findbugs-maven-plugin to 3.0.4
- [HBASE-21001](#): ReplicationObserver fails to load in HBase 2.0.0
- [HBASE-20307](#): LoadTestTool prints too much zookeeper logging (Colin Garcia)
- [HBASE-21157](#): Split TableInputFormatScan to individual tests
- [HBASE-21127](#): TableRecordReader need to handle cursor result too
- [HBASE-20892](#): [UI] Start / End keys are empty on table.jsp
- [HBASE-21132](#): return wrong result in rest multiget

HDP 3.0.1 provided HBase HBase 2.0.0 and the following Apache patches:

- [HBASE-20723](#): Custom hbase.wal.dir results in data loss because we write recovered edits into a different place than where the recovering region server looks for them
- [HBASE-20722](#): Make RegionServerTracker only depend on children changed event
- [HBASE-20668](#): Avoid permission change if ExportSnapshot's copy fails
- [HBASE-20657](#): Retrying RPC call for ModifyTableProcedure may get stuck
- [HBASE-20635](#): Add missing shaded-pb conversion method for UserPermission
- [HBASE-20627](#): Relocate RS Group pre/post hooks from RSGroupAdminServer to RSGroupAdminEndpoint
- [HBASE-20579](#): Include original exception in wrapped exception
- [HBASE-20579](#): Improve snapshot manifest copy in ExportSnapshot
- [HBASE-20547](#): Restore from backup will fail if done from a different file system
- [HBASE-20531](#): RS may throw NPE when close meta regions in shutdown procedure.
- [HBASE-20495](#): REST unit test fails with NoClassDefFoundError against hadoop3
- [HBASE-20196](#): Maintain all regions with same size in memstore flusher
- [HBASE-20136](#): TestKeyValue misses ClassRule and Category annotations
- [HBASE-19735](#): Create a client-tarball assembly
- [HBASE-19478](#): Utilize multi-get to speed up WAL file checking in BackupLogCleaner (Toshihiro Suzuki)
- [HBASE-18135](#): Implement mechanism for RegionServers to report file archival for space quotas
- [HBASE-18133](#): Decrease quota reaction latency by HBase
- [HBASE-17549](#): HBase-Spark Module: Corrected - Incorrect log at println and unwanted comment cod

Hive

This release provides Hive 3.1.0 and the listed Apache patches.

- [HIVE-20827](#): Inconsistent results for empty arrays
- [HIVE-20998](#): HiveStrictManagedMigration utility should update DB/Table location as last migration steps
- [HIVE-20542](#): Incremental REPL DUMP progress information log message is incorrect
- [HIVE-20953](#): Remove a function from function registry when it can not be added to the metastore when creating it.
- [HIVE-20884](#): Support bootstrap of tables from HDP 2.6.5 to HDP 3.0. Added managed to external testcase.
- [HIVE-20926](#): Semi join reduction hint fails when bloom filter entries are high or when there are no stats
- [HIVE-20979](#): Fix memory leak in hive streaming
- [HIVE-20838](#): Timestamps with timezone are set to null when using the streaming API
- [HIVE-20988](#): Wrong results for group by queries with primary key on multiple columns
- [HIVE-20985](#): If select operator inputs are temporary columns vectorization may reuse some of them as output
- [HIVE-20842](#): Fix logic introduced in HIVE-20660 to estimate statistics for group by
- [HIVE-20951](#): LLAP: Set Xms to 50% always
- [HIVE-20822](#): Hive JDBC Storage Handler (Postgres): "java.lang.RuntimeException: While invoking method 'public org.apache.calcite.rel.rel2sql.SqlImplementor\$Result" while executing TPCDS query GRP_2_85
- [HIVE-20884](#): Support bootstrap of tables from HDP 2.6.5 to HDP 3.0.
- [HIVE-20926](#): Semi join reduction hint fails when bloom filter entries are high or when there are no stats
- [HIVE-20976](#): Hive JDBC Storage Handler: TPCDS_GRP_3_88 doesn't output proper results

- [HIVE-20626](#): Backport Log more details when druid metastore transaction fails
- [HIVE-20954](#): Vector RS operator is not using uniform hash function for TPC-DS query 95
- [HIVE-20691](#): Fix org.apache.hadoop.hive.cli.TestMiniLlapCliDriver.testCliDriver[cttl]
- [HIVE-10296](#): Cast exception observed when hive runs a multi join query on metastore , since postgres pushes the filter into the join, and ignores the condition before applying cast
- [HIVE-20940](#): Bridge cases in which Calcite's type resolution is more stricter than Hive.
- [HIVE-20944](#): Not validate stats during query compilation
- [HIVE-20949](#): Improve PKFK cardinality estimation in physical planning
- [HIVE-20940](#): Bridge cases in which Calcite's type resolution is more stricter than Hive.
- [HIVE-20937](#): Postgres jdbc query fail with "LIMIT must not be negative"
- [HIVE-20775](#): Factor cost of each SJ reduction when costing a follow-up reduction
- [HIVE-19882](#): Fix QTestUtil session lifecycle addendum
- [HIVE-20920](#): Use SQL constraints to improve join reordering algorithm
- [HIVE-20918](#): Flag to enable/disable pushdown of computation from Calcite into JDBC connection
- [HIVE-20775](#): Factor cost of each SJ reduction when costing a follow-up reduction
- [HIVE-20910](#): Insert in bucketed table fails due to dynamic partition sort optimization
- [HIVE-20682](#): Async query execution can potentially fail if shared sessionHive is closed by master thread
- [HIVE-19552](#): Backport fix Druid kafka mini tests
- [HIVE-19882](#): Fix QTestUtil session lifecycle
- [HIVE-20822](#): Hive JDBC Storage Handler: TPCDS_GRP_3_88 doesn't output proper results
- [HIVE-20101](#): BloomKFilter: Avoid using the local byte[] arrays entirely
- [HIVE-20899](#): Keytab URI for LLAP YARN Service is restrictive to support HDFS only
- [HIVE-20744](#): Use SQL constraints to improve join reordering algorithm
- [HIVE-20898](#): For time related functions arguments may not be casted to a non nullable type
- [HIVE-20746](#): HiveProtoHookLogger does not close file at end of day.
- [HIVE-20873](#): Use Murmur hash for VectorHashKeyWrapperTwoLong to reduce hash collision
- [HIVE-20881](#): Constant propagation oversimplifies projections
- [HIVE-20853](#): Expose ShuffleHandler.registerDag in the llap daemon API
- [HIVE-20804](#): Further improvements to group by optimization with constraints
- [HIVE-20768](#): Adding Tumbling Window UDF
- [HIVE-20813](#): udf to_epoch_milli need to support timestamp without time zone as well.
- [HIVE-20868](#): SMB Join fails intermittently when TezDummyOperator has child op in getFinalOp in MapRecordProcessor
- [HIVE-20617](#): Fix type of constants in IN expressions to have correct type
- [HIVE-20486](#): Vectorization support for Kafka Storage Handler
- [HIVE-20767](#): Multiple project between join operators may affect join reordering using constraints
- [HIVE-20822](#): Improvements to push computation to JDBC from Calcite
- [HIVE-20835](#): Interaction between constraints and MV rewriting may create loop in Calcite planner
- [HIVE-20821](#): Rewrite SUM0 into SUM + COALESCE combination
- [HIVE-20820](#): MV partition on clause position
- [HIVE-20788](#): Extended SJ reduction may backtrack columns incorrectly when creating filters
- [HIVE-20477](#): OptimizedSql is not shown if the expression contains INs
- [HIVE-20834](#): Hive QueryResultCache entries keeping reference to SemanticAnalyzer from cached query
- [HIVE-20829](#): JdbcStorageHandler range split throws NPE
- [HIVE-20830](#): JdbcStorageHandler range query assertion failure in some cases
- [HIVE-20815](#): JdbcRecordReader.next shall not eat exception
- [HIVE-20792](#): Inserting timestamp with zones truncates the data
- [HIVE-20638](#): Upgrade version of Jetty to 9.3.25.v20180904
- [HIVE-20765](#):HIVE- : fetch partitions for txn stats validation in get_aggr_stats with one call
- [HIVE-20718](#): Add perf cli driver with constraints

- [HIVE-20763](#): Add google cloud storage to the exim uri schema whitelist
- [HIVE-20703](#): Put dynamic sort partition optimization under cost based decision
- [HIVE-20762](#): NOTIFICATION_LOG cleanup interval is hardcoded as 60s and is too small
- [HIVE-6980](#): Drop table by using direct sql
- [HIVE-20716](#): Set default value for hive.cbo.stats.correlated.multi.key.joins to true
- [HIVE-18986](#): Table rename will run java.lang.StackOverflowError in dataNucleus if the table contains large number of columns
- [HIVE-20720](#): Add partition column option to JDBC handler
- [HIVE-20735](#): Adding Support for Kerberos Auth, Removed start/end offset columns, remove the best effort mode and made 2pc default for EOS
- [HIVE-20754](#): Hive: java.lang.NoClassDefFoundError: org/apache/hadoop/security/SaslRpcServer exception occurs while using Hive JDBC standalone jar
- [HIVE-20761](#): [Intermittent]Hive DDL's failing with Error- Couldn't acquire the DB log notification lock because we reached the maximum # of retries: 5 retries.
- [HIVE-20644](#):HIVE- : Avoid exposing sensitive information through a Hive Runtime exception
- [HIVE-20731](#): keystore file in JdbcStorageHandler should be authorized
- [HIVE-20649](#): LLAP aware memory manager for Orc writers
- [HIVE-20648](#): LLAP: Vector group by operator should use memory per executor
- [HIVE-20696](#): msck_*.q tests are broken
- [HIVE-20537](#): Multi-column joins estimates with uncorrelated columns different in CBO and Hive
- [HIVE-20639](#): Add ability to Write Data from Hive Table/Query to Kafka Topic
- [HIVE-20561](#): Use the position of the Kafka Consumer to track progress instead of Consumer Records offsets
- [HIVE-20481](#): Add the Kafka Key record as part of the row
- [HIVE-20377](#): Hive Kafka Storage Handler
- [HIVE-20710](#): Constant folding may not create null constants without types
- [HIVE-20719](#): UPDATE with hive.optimize.sort.dynamic.partition corrupts table
- [HIVE-20660](#): Group by statistics estimation could be improved by bounding the total number of rows to source table
- [HIVE-17043](#): Remove non unique columns from group by keys if not referenced later
- [HIVE-20692](#): Enable folding of NOT x IS (NOT) [TRUE|FALSE] expressions
- [HIVE-14431](#): Recognize COALESCE as CASE
- [HIVE-20618](#): During join selection BucketMapJoin might be chosen for non bucketed tables
- [HIVE-20712](#): HivePointLookupOptimizer should extract deep cases
- [HIVE-20498](#):HIVE- Support date type for column stats autogather
- [HIVE-20657](#): pre-allocate LLAP cache at init time
- [HIVE-18871](#): hive on tez execution error due to set hive.aux.jars.path to hdfs://
- [HIVE-20583](#): Use canonical hostname only for kerberos auth in HiveConnection
- [HIVE-19993](#): Using a table alias which also appears as a column name is not possible
- [HIVE-20183](#): Inserting from bucketed table can cause data loss, if the source table contains empty bucket
- [HIVE-19850](#): Dynamic partition pruning in Tez is leading to 'No work found for tablescan' error
- [HIVE-20349](#): Implement Retry Logic in HiveDruidSplit for Scan Queries
- [HIVE-20700](#): Add config to disable rollup for druid tables
- [HIVE-20698](#): Add better message for NPE when inserting rows with null timestamp to druid
- [HIVE-20686](#): Sync query IDs between druid and Hive
- [HIVE-20684](#): Make compute stats work for Druid tables
- [HIVE-20582](#): Avoid hflush in hive proto logger
- [HIVE-20702](#): Account for overhead from datastructure aware estimations during mapjoin selection
- [HIVE-20704](#): Extend HivePreFilteringRule to support other functions
- [HIVE-20651](#): JdbcStorageHandler password should be encrypted

- [HIVE-20677](#): JDBC handler ordering problem - single split flag [HIVE-20652](#): JdbcStorageHandler push join of two different datasource to jdbc driver [HIVE-20669](#): JdbcStorageHandler push union of two different datasource to jdbc driver
- [HIVE-20455](#): Log spew from security.authorization.PrivilegeSynchronizer.run
- [HIVE-20676](#): HiveServer2: PrivilegeSynchronizer is not set to daemon status
- [HIVE-17040](#): Join elimination in the presence of FK relationship
- [HIVE-20636](#): Improve number of null values estimation after outer join
- [HIVE-20326](#): Create constraints with RELY as default instead of NO RELY
- [HIVE-20646](#): Partition condition not being pushed into metastore query if it has IS NOT NULL.
- [HIVE-18778](#): Needs to capture input/output entities in explain
- [HIVE-20623](#): Shared work: Extend sharing of map-join cache entries in LLAP
- [HIVE-20095](#): Fix feature to push computation to jdbc external tables
- [HIVE-20540](#):HIVE- : Vectorization : Support loading bucketed tables using sorted dynamic partition optimizer - II
- [HIVE-20603](#): "Wrong FS" error when inserting to partition after changing table location filesystem
- [HIVE-20620](#): manifest collisions when inserting into bucketed sorted MM tables with dynamic partitioning
- [HIVE-20612](#): Create new join multi-key correlation flag for CBO
- [HIVE-20625](#): Regex patterns not working in SHOW MATERIALIZED VIEWS '<pattern>'
- [HIVE-20593](#): Load Data for partitioned ACID tables fails with bucketId out of range: -1
- [HIVE-20522](#): HiveFilterSetOpTransposeRule may throw assertion error due to nullability of fields
- [HIVE-20517](#): Creation of staging directory and Move operation is taking time in S3
- [HIVE-20499](#): GetTablesOperation pull all the tables meta irrespective of auth.
- [HIVE-19432](#): GetTablesOperation is too slow if the hive has too many databases and tables
- [HIVE-18873](#): Skipping predicate pushdown for MR silently at HiveInputFormat can cause storage handlers to produce erroneous result
- [HIVE-20267](#): Concurrent async queries intermittently fails with LockException and cause memory leak.
- [HIVE-20632](#): Query with get_splits UDF fails if materialized view is created on queried table.
- [HIVE-20521](#): HS2 doAs=true has permission issue with hadoop.tmp.dir, with MR and S3A filesystem .
- [HIVE-1118](#): background updates fix
- [HIVE-20399](#): CTAS w/a custom table location that is not fully qualified fails for MM tables
- [HIVE-20311](#): add txn stats checks to some more paths
- [HIVE-20507](#): Beeline: Add a utility command to retrieve all uris from beeline-site.xml
- [HIVE-20503](#): Use datastructure aware estimations during mapjoin selection
- [HIVE-20513](#): Vectorization: Improve Fast Vector MapJoin Bytes Hash Tables
- [HIVE-20296](#): Improve HivePointLookupOptimizerRule to be able to extract from more sophisticated contexts
- [HIVE-20541](#): NoSuchElementException in REPL dump.
- [HIVE-19985](#): Benchmark - LLAP performance issue for fully compacted ACID tables
- [HIVE-20409](#): Hive ACID: Update/delete/merge leave behind the staging directory
- [HIVE-20491](#): Fix mapjoin size estimations for Fast implementation
- [HIVE-19097](#): related equals and in operators may cause inaccurate stats estimations
- [HIVE-20439](#): Use the inflated memory limit during join selection for llap
- [HIVE-20260](#): NDV of a column shouldn't be scaled when row count is changed by filter on another column
- [HIVE-20418](#): LLAP IO may not handle ORC files that have row index disabled correctly for queries with no columns selected
- [HIVE-20508](#):HIVE- : Hive does not support user names of type "user@realm"
- [HIVE-20193](#): cboInfo is not present in the explain plan json
- [HIVE-20510](#): Vectorization : Support loading bucketed tables using sorted dynamic partition optimizer
- [HIVE-20515](#): Empty query results when using results cache and query temp dir, results cache dir in different filesystems
- [HIVE-20476](#): CopyUtils used by REPL LOAD and EXPORT/IMPORT operations ignore distcp error
- [HIVE-20450](#): Add replication test for LOAD command on ACID table.

- [HIVE-20433](#): Implicit String to Timestamp conversion is slow
- [HIVE-20446](#): CachedStore: bug UT fixes when CachedStore is enabled
- [HIVE-19748](#): CREATE, INSERT commands fails with MetaException on HDI azure cluster due to NPE in MetaStoreUtils.decimalToDouble
- [HIVE-18624](#): Parsing time is extremely high (~10 min) for queries with complex select expressions
- [HIVE-20187](#): Incorrect query results in hive when hive.convert.join.bucket.mapjoin.tez is set to true
- [HIVE-20443](#): txn stats cleanup in compaction txn handler is unneeded
- [HIVE-20444](#): Parameter is not properly quoted in DbNotificationListener.addWriteNotificationLog
- [HIVE-20431](#): txn stats write ID check triggers on set location
- [HIVE-20329](#): Long running repl load (incr/bootstrap) causing OOM error
- [HIVE-19924](#): Tag distcp jobs run by Repl Load
- [HIVE-20411](#): Hive.loadPartition doesn't support catalogs
- [HIVE-20412](#): Hive upgrade: HBasestoragehandler tables fail to upgrade at HiveMetaStoreClient.alter_table with Null Pointer Exception - committing HIVE-20412.1.patch directly to HDP
- [HIVE-20410](#): aborted Insert Overwrite on transactional table causes "Not enough history available for..." error
- [HIVE-20407](#): Hive upgrade: Alter table(orc-bucket) owned by hive fails at seedWriteIdOnAcidConversion with Duplicate entry for key 'NEXT_WRITE_ID_IDX'
- [HIVE-20413](#): "cannot insert NULL" for TXN_WRITE_NOTIFICATION_LOG in Oracle
- [HIVE-20393](#): Semijoin Reduction : markSemiJoinForDPP behaves inconsistently
- [HIVE-20407](#): add txn stats support to upgrade tool
- [HIVE-14898](#): HS2 shouldn't log callstack for an empty auth header error
- [HIVE-20378](#): don't update stats during alter for txn table conversion
- [HIVE-20357](#): Introduce initOrUpgradeSchema option to schema tool
- [HIVE-19250](#): Schema column definitions inconsistencies in MySQL
- [HIVE-20391](#): HiveAggregateReduceFunctionsRule may infer wrong return type when decomposing aggregate function
- [HIVE-20383](#): Invalid queue name and synchronisation issues in hive proto events hook.
- [HIVE-20389](#): NPE in SessionStateUserAuthenticator when authenticator=SessionStateUserAuthenticator
- [HIVE-20345](#): Drop database may hang if the tables get deleted from a different call
- [HIVE-20033](#): PowerBI unable to connect to LLAP cluster with large Databases/Tables
- [HIVE-20098](#): Statistics: NPE when getting Date column partition statistics
- [HIVE-19860](#): HiveServer2 ObjectInspectorFactory memory leak with cachedUnionStructObjectInspector
- [HIVE-20382](#): Materialized views: Introduce heuristic to favour incremental rebuild
- [HIVE-20379](#): Rewriting with partitioned materialized views may reference wrong column
- [HIVE-18620](#): Improve error message while dropping a table that is part of a materialized view
- [HIVE-20067](#): mm_all test failing.
- [HIVE-19316](#): StatsTask fails due to ClassCastException
- [HIVE-20149](#): TestHiveCli failing/timing out
- [HIVE-20340](#): Druid Needs Explicit CASTs from Timestamp to STRING when the output of timestamp function is used as String
- [HIVE-20344](#): PrivilegeSynchronizer for SBA might hit AccessControlException
- [HIVE-20320](#): Turn on hive.optimize.remove.sq_count_check flag
- [HIVE-20354](#):HIVE- : Semijoin hints dont work with merge statements
- [HIVE-20264](#): Bootstrap repl dump with concurrent write and drop of ACID table makes target inconsistent
- [HIVE-20361](#): ReplDumpTaskTest is failing
- [HIVE-20316](#): Skip external table file listing for create table event
- [HIVE-19800](#): TestUpgradeTool.testPostUpgrade failing
- [HIVE-20337](#): Queries returning incorrect results using CachedStore. ObjectStore works fine
- [HIVE-20347](#): hive.optimize.sort.dynamic.partition should work with partitioned CTAS and MV
- [HIVE-20336](#): Masking and filtering policies for materialized views

- [HIVE-20301](#): Enable vectorization for materialized view rewriting tests
- [HIVE-20314](#): Include partition pruning in materialized view rewriting
- [HIVE-20302](#): LLAP: non-vectorized execution in IO ignores virtual columns, including ROW__ID
- [HIVE-19408](#): Improve show materialized views statement to show more information about invalidation
- [HIVE-20335](#): Add tests for materialized view rewriting with composite aggregation functions
- [HIVE-20165](#): Enable ZLIB for streaming ingest
- [HIVE-20147](#): Hive streaming ingest is contented on synchronized logging
- [HIVE-19800](#): HIVE- PreUpgrade Tool: Support upgrade from >= 2.6.0
- [HIVE-20315](#): Vectorization: Fix more NULL Wrong Results issues and avoid unnecessary casts/conversions : Vectorization: Backport HIVE-20315 to Atlantic-M05: Vectorization: Fix more NULL Wrong Results issues and avoid unnecessary casts/conversions
- [HIVE-19812](#): Disable external table replication by default via a configuration property
- [HIVE-20277](#): Vectorization: Case expressions that return BOOLEAN are not supported for FILTER
- [HIVE-20130](#): Better logging for information schema synchronizer
- [HIVE-20118](#): SessionStateUserAuthenticator.getGroupNames is always empty
- [HIVE-20292](#): Adding constraints on TPCDS schema, changes query plan for query93 and degrades its perf by 4x times
- [HIVE-19970](#): Replication dump has a NPE when table is empty
- [HIVE-20120](#): Incremental repl load DAG generation is causing OOM error
- [HIVE-20322](#): FlakyTest: TestMiniDruidCliDriver
- [HIVE-19829](#): Incremental replication load should create tasks in execution phase rather than semantic phase
- [HIVE-20322](#): FlakyTest: TestMiniDruidCliDriver
- [HIVE-19927](#): Last Repl ID set by bootstrap dump is not proper and may cause loss of data if have ACID tables.
- [HIVE-17683](#): Add explain locks <sql> command
- [HIVE-20209](#): Metastore connection fails for first attempt in repl dump
- [HIVE-20294](#): Vectorization: Fix NULL Wrong Results issues in COALESCE ELT
- [HIVE-20245](#): Vectorization: Fix NULL Wrong Results issues in BETWEEN IN
- [HIVE-20207](#): Vectorization: Fix NULL Wrong Results issues in Filter Compare
- [HIVE-20281](#): SharedWorkOptimizer fails with 'operator cache contents and actual plan differ'
- [HIVE-20252](#): Semijoin Reduction : Cycles due to semi join branch may remain undetected if small table side has a map join upstream.
- [HIVE-14493](#): Partitioning support for materialized views
- [HIVE-20241](#): Support partitioning spec in CTAS statements
- [HIVE-20212](#): Hiveserver2 in http mode emitting metric default.General.open_connections incorrectly
- [HIVE-20105](#): Druid-Hive: tpcds query on timestamp throws java.lang.IllegalArgumentException: Cannot create timestamp, parsing error
- [HIVE-20263](#): Typo in HiveReduceExpressionsWithStatsRule variable
- [HIVE-20090](#): Extend creation of semijoin reduction filters to be able to discover new opportunities
- [HIVE-19568](#): Active/Passive HS2 HA: Disallow direct connection to passive HS2 instance
- [HIVE-20082](#): HiveDecimal to string conversion doesn't format the decimal correctly
- [HIVE-20204](#): Type conversion during IN () comparisons is using different rules from other comparison operations
- [HIVE-20240](#): Semijoin Reduction : Use local variable to check for external table condition
- [HIVE-19313](#): TestJdbcWithDBTokenStoreNoDoAs tests are failing
- [HIVE-20228](#): configure repl configuration directories based on user running hiveserver2
- [HIVE-19935](#): Hive WM session killed: Failed to update LLAP tasks count
- [HIVE-19800](#): TestUpgradeTool.testPostUpgrade failing
- [HIVE-19891](#): inserting into external tables with custom partition directories may cause data loss
- [HIVE-4367](#): enhance TRUNCATE syntax to drop data of external table
- [HIVE-19992](#): Vectorization: Follow-on to HIVE-19951 --> add call to SchemaEvolution.isOnlyImplicitConversion to disable encoded LLAP I/O for ORC only when data type conversion is not implicit : Vectorization: Backport HIVE-19992 to Atlantic-M05: Follow-on to HIVE-19951 --

> add call to SchemaEvolution.isOnlyImplicitConversion to disable encoded LLAP I/O for ORC only when data type conversion is not implicit

- [HIVE-20069](#): Fix reoptimization in case of DPP and Semijoin optimization
- [HIVE-20152](#): reset db state, when repl dump fails, so rename table can be done
- [HIVE-20197](#): Vectorization: Add DECIMAL_64 testing, add Date/Interval/Timestamp arithmetic, and add more GROUP BY Aggregation tests : Vectorization: Backport HIVE-20197 to Atlantic-M05: Add DECIMAL_64 testing, add Date/Interval/Timestamp arithmetic, and add more GROUP BY Aggregation tests
- [HIVE-20174](#): Vectorization: Fix NULL Wrong Results issues in GROUP BY Aggregation Functions
- [HIVE-20192](#): HiveServer2 with embedded metastore is leaking JDOPersistenceManager objects
- [HIVE-19104](#): When test MetaStore is started with retry the instances should be independent
- [HIVE-20164](#): Murmur Hash : Make sure CTAS and IAS use correct bucketing version
- [HIVE-20116](#):HIVE- - TestOperationLoggingAPIWithTez failing
- [HIVE-20088](#): Beeline config location path is assembled incorrectly
- [HIVE-20011](#): Move away from append mode in proto logging hook
- [HIVE-19360](#): CBO: Add an "optimizedSQL" to QueryPlan object
- [HIVE-20127](#):HIVE- : fix some issues with LLAP Parquet cache
- [HIVE-19387](#):HIVE- Upgrade: Post the second truncate call, select query fails to fetch new results if query.results.cache is enabled
- [HIVE-20129](#): Revert to position based schema evolution for orc tables
- [HIVE-20126](#): OrcInputFormat does not pass conf to orc reader options
- [HIVE-20123](#): Fix masking tests after HIVE-19617
- [HIVE-19628](#):HIVE- : possible NPE in LLAP testSigning
- [HIVE-19759](#): Flaky test: TestRpc#testServerPort
- [HIVE-20111](#): Making storage handler tables external by default breaks HBase
- [HIVE-20112](#): Making storage handler tables external by default breaks Accumulo
- [HIVE-20102](#): Ranger row-level filter policies don't work with CTAS
- [HIVE-20001](#): With doas set to true, running select query as hrt_qa user on external table fails due to permission denied to read warehouse/tablespace/managed directory
- [HIVE-20100](#): OpTraits : Select Optraits should stop when a mismatch is detected
- [HIVE-20085](#): review comments and fix drop table to honour external table purge property, enable ctas by default on external tables
- [HIVE-20039](#): Bucket pruning: Left Outer Join on bucketed table gives wrong result
- [HIVE-20085](#): Make all Druid Tables as EXTERNAL tables
- [HIVE-20067](#): fix InsertEvent on mm tables to not cause failing capability checks
- [HIVE-20025](#): Clean-up of event files created by HiveProtoLoggingHook
- [HIVE-17840](#): HiveMetaStore eats exception if transactionalListeners.notifyEvent fail
- [HIVE-20051](#): Skip authorization for temp tables
- [HIVE-19938](#): Upgrade scripts for information schema
- [HIVE-19326](#): stats auto gather: incorrect aggregation during UNION queries
- [HIVE-20059](#): Hive streaming should try shade prefix unconditionally on exception
- [HIVE-20028](#): Metastore client cache config is used incorrectly
- [HIVE-20038](#): Updates/deletes on non-bucketed acid tables throw NPE
- [HIVE-20002](#): Shipping jdbd-storage-handler dependency jars in LLAP
- [HIVE-19967](#): SMB Join : Need Optraits for PTFOperator ala GBY Op
- [HIVE-20004](#):HIVE-" Wrong scale used by ConvertDecimal64ToDecimal results in incorrect results
- [HIVE-19995](#): Aggregate row traffic for acid tables
- [HIVE-20009](#): Fix runtime stats for merge statement
- [HIVE-19989](#): Metastore uses wrong application name for HADOOP2 metrics
- [HIVE-20035](#): Serialize boolean as long for druid
- [HIVE-20013](#): Add an Implicit cast to date type for to_date function
- [HIVE-18140](#): Partitioned tables statistics can go wrong in basic stats mixed case

- [HIVE-18729](#): Druid Time column type
- [HIVE-19938](#): Upgrade scripts for information schema
- [HIVE-20010](#): Fix create view over literals
- [HIVE-19753](#): Strict managed tables mode in Hive
- [HIVE-19972](#): Followup to HIVE-19928 : Fix the check for managed table
- [HIVE-19481](#): Tablesample uses incorrect logic to pick files corresponding to buckets.
- [HIVE-19981](#): Managed tables converted to external tables by the HiveStrictManagedMigration utility should be set to delete data when the table is dropped
- [HIVE-19898](#): Disable TransactionalValidationListener when the table is not in the Hive catalog
- [HIVE-19904](#): Load data rewrite into Tez job fails for ACID
- [HIVE-19775](#): Schematool should use HS2 embedded mode in privileged auth mode
- [HIVE-19800](#): HIVE-19598 + adjustments. needs HIVE-19750 for post upgrade file rename to work properly
- [HIVE-19440](#): fix merge conflicts that got dropped
- [HIVE-19258](#): add originals support to MM tables
- [HIVE-19476](#): Fix failures in TestReplicationScenariosAcidTables, TestReplicationOnHDFSEncryptedZones and TestCopyUtils
- [HIVE-17657](#): export/import for MM tables is broken
- [HIVE-19312](#): MM tables don't work with BucketizedHIF
- [HIVE-19825](#): HiveServer2 leader selection shall use different zookeeper znode
- [HIVE-19728](#): beeline with USE_BEELINE_FOR_HIVE_CLI fails when trying to set hive.aux.jars.path
- [HIVE-19485](#): dump directory for non native tables should not be created
- [HIVE-19440](#): Make StorageBasedAuthorizer work with information schema
- [HIVE-19465](#): Upgrade ORC to 1.5.0
- [HIVE-19631](#): reduce epic locking in AbstractService
- [HIVE-19258](#): add originals support to MM tables
- [HIVE-19660](#): update branch-3 to be version 3.1 and fix storage-api mismatch
- [HIVE-14247](#): Disable parallel query execution within a session
- [HIVE-18193](#): Migrate existing ACID tables to use table Write ID instead of global txn ID
- [HIVE-19476](#): Fix failures in TestReplicationScenariosAcidTables, TestReplicationOnHDFSEncryptedZones and TestCopyUtils
- [HIVE-19435](#): Incremental replication cause data loss if a table is dropped followed by create and insert-into with different partition type
- [HIVE-19322](#): Broken test: TestNegativeMinimrCliDriver#testCliDriver[minimr_broken_pipe]
- [HIVE-17657](#): Export/import for MM tables is broken
- [HIVE-19304](#): Update templates.py based on config changes in YARN-7142 and YARN-8122
- [HIVE-19194](#): addendum for branch-3: fix imports
- [HIVE-19193](#): TestActivePassiveHA fails
- [HIVE-19014](#): Utilize YARN-8028 (queue ACL check) in Hive Tez session pool
- [HIVE-19072](#): incorrect token handling for LLAP plugin endpoint
- [HIVE-18971](#): add HS2 WM metrics for use in Grafana and such
- [HIVE-18952](#): Tez session disconnect and reconnect on HS2 HA failover INTERNAL COMMIT
- [HIVE-18717](#): Avoid transitive dependency on jetty 6.x
- [HIVE-18414](#): upgrade to tez-0.9.1
- [HIVE-18037](#): align kerberos keytab and principal with YARN Service
- [HIVE-14498](#): Freshness period for query rewriting using materialized views
- [HIVE-18096](#): add a user-friendly show plan command
- [HIVE-18037](#): plug-in llapstatus with YARN Service
- [HIVE-18037](#): fix issue with replacement of % %(queue.string)s with correct queue value
- [HIVE-15393](#): Update Guava version
- [HIVE-18150](#): Upgrade Spark Version to 2.2.0

HDP 3.0.0 and HDP 3.0.1 provided Hive 3.1.0 and no additional Apache patches.

Kafka

This release provides Kafka 2.0.0 and no additional Apache patches.

HDP 3.0.1 provided Kafka 1.1.1 and no additional Apache patches.

HDP 3.0.0 provided Kafka 1.0.1 and no additional Apache patches. A new feature was added to capture producer and topic partition level metrics. This feature is available in HDP 3.0.0. It is not available in Apache Kafka 1.0.1.

Knox

This release provides Knox 1.0.0 and no Apache patches.

HDP 3.0.0 and HDP 3.0.1 provided Knox 1.0.0 and no additional Apache patches.

Livy

This release provides Livy 0.5.0 and no additional patches.

HDP 3.0.0 and HDP 3.0.1 provided Livy 0.5.0 and no additional Apache patches.

Oozie

This release provides Oozie 4.3.1 and no additional Apache patches.

HDP 3.0.1 provided Oozie 4.3.1 and the following Apache patches:

- [OOZIE-3304](#): Parsing sharelib timestamps is not threadsafe (dionusos, matijhs via andras.piros)
- [OOZIE-3135](#): Configure log4j2 in SqoopMain
- [OOZIE-3072](#): oozie.service.HadoopAccessorService.action.configurations should overwrite default values set in Hadoop's configuration files (gezapeti)
- [OOZIE-2916](#): Set a job name for the MR Action's child job (asasvari)
- [OOZIE-2494](#): Cron syntax not handling DST properly (kmarion via andras.piros)
- [OOZIE-2726](#): Flaky test due to daylight saving changes (sasishsale, andras.piros via gezapeti)
- [OOZIE-3173](#): Coordinator job with frequency using cron syntax creates only one action in catchup mode (andras.piros)
- [OOZIE-2691](#): Show workflow action retry information in UI
- [OOZIE-3109](#): [log-streaming] Escape HTML-specific characters (dionusos via andras.piros)
- [OOZIE-2815](#): Oozie not always display job log
- [OOZIE-3075](#): Follow-up on OOZIE-3054: create the lib directory if it doesn't exist (pbacsko)
- [OOZIE-3054](#): Disable erasure coding for sharelib if Oozie runs on Hadoop 3 (pbacsko)
- [OOZIE-3048](#): Check El Functions for the coordinator action (satishsale)
- [OOZIE-3004](#): Forked action retry info is not working
- [OOZIE-2691](#): Show workflow action retry information in UI
- [OOZIE-3009](#): Number of Oozie tests executed dropped after OOZIE-2854 (andras.piros via asasvari)
- [OOZIE-2854](#): Oozie should handle transient database problems (andras.piros via gezapeti)
- [OOZIE-3026](#): fix openjpa enhancer stage during build for logging (dbdist13, andras.piros via pbacsko)
- [OOZIE-2816](#): Strip out the first command word from Sqoop action if its "sqoop" (harsh)
- [OOZIE-3113](#): Retry for ZK lock release (satishsale)
- [OOZIE-3208](#): "It should never happen" error messages should be more specific to root cause (kmarion via andras.piros)
- [OOZIE-2940](#): Possible NullPointerException in WorkflowActionBean (dionusos via gezapeti)
- [OOZIE-3297](#): Retry logic does not handle the exception from BulkJPAExecutor properly (andras.piros)
- [OOZIE-2733](#): change org.apache.hadoop.fs.permission.AccessControlException to org.apache.hadoop.security.AccessControlException (gezapeti)
- [OOZIE-3021](#): Error on job or SLA event listening: WARN logs instead of DEBUG (andras.piros via asasvari)

- [OOZIE-2387](#): Oozie is Unable to handle Spaces in file/archive tag. (asasvari via gezapeti)
- [OOZIE-2818](#): Can't overwrite oozie.action.max.output.data on a per-workflow basis (asasvari via pbacsko)
- [OOZIE-2041](#): Add an admin command to run the PurgeXCommand (abhishekbafna)
- [OOZIE-2875](#): Typo in ssh action twiki docs (Dongying Jiao via gezapeti)
- [OOZIE-1401](#): PurgeCommand should purge the workflow jobs w/o end_time (asasvari)
- [OOZIE-3156](#): Retry SSH action check when cannot connect to remote host (txsing via andras.piros)
- [OOZIE-3183](#): Better logging for SshActionExecutor and extended HA capability when calling to remote host (andras.piros)
- [OOZIE-2769](#): Extend FS action to allow setrep on a file (Artem Ervits via gezapeti)
- [OOZIE-3085](#): Improve logging in ActionExecutors: add log entry for start(), end() and kill() (kmarion via andras.piros)

HDP 3.0.0 provided Oozie 4.3.1 and the following Apache patches:

- [OOZIE-2802](#):: Spark action failure on Spark 2.1.0 due to duplicate sharelibs.
- [OOZIE-2658](#):: --driver-class-path can overwrite the classpath in SparkMain.
- [OOZIE-2787](#):: Oozie distributes application jar twice making the spark job fail.
- [OOZIE-2606](#):: Set spark.yarn.jars to fix Spark 2.0 with Oozie.
- [OOZIE-2436](#):: Fork/join workflow fails with "oozie.action.yarn.tag must not be null".
- [OOZIE-2578](#):: Oozie example distcp job fails to run within an encrypted zone with checksum match error.
- [OOZIE-2362](#):: SQL injection in BulkJPAExecutor.
- [OOZIE-2577](#):: Flaky tests TestCoordActionInputCheckXCommand.testTimeout and testTimeoutWithException.
- [OOZIE-2570](#):: remove -PtestPatchCompile from patch testing as there is no such profile.
- [OOZIE-2504](#):: Create a log4j.properties under HADOOP_CONF_DIR in Shell Action.
- [OOZIE-2567](#):: HCat connection is not closed while getting hcat cred.
- [OOZIE-2547](#):: Add mapreduce.job.cache.files to spark action.
- [OOZIE-2550](#):: Flaky tests in TestZKUUIDService.java.
- [OOZIE-2445](#):: Doc for - Specifying coordinator input datasets in more logical ways.
- [OOZIE-2541](#):: Possible resource leak in Hive2Credentials.
- [OOZIE-2563](#):: Pass spark-defaults.conf to spark action.
- [OOZIE-2556](#):: TestAbandonedCoordChecker.testCatchupJob is flaky.
- [OOZIE-2522](#):: There can be multiple coord submit from bundle in case of ZK glitch.
- [OOZIE-2553](#):: Cred tag is required for all actions in the workflow even if an action does not require it.
- [OOZIE-2503](#):: show ChildJobURLs to spark action.
- [OOZIE-2551](#):: Feature request: epoch timestamp generation.
- [OOZIE-2542](#):: Option to disable OpenJPA BrokerImpl finalization.
- [OOZIE-2447](#):: Illegal character 0x0 oozie client.
- [OOZIE-2548](#):: Flaky test TestZKLocksService.testLockReaper.
- [OOZIE-2546](#):: Improperly closed resources in OozieDBCLI.
- [OOZIE-2476](#):: When one of the action from fork fails with transient error, WF never joins .
- [OOZIE-2475](#):: Oozie does not cleanup action dir of killed actions.
- [OOZIE-2535](#):: User can't disable uber mode.
- [OOZIE-2482](#):: Pyspark job fails with Oozie.
- [OOZIE-2467](#):: Oozie can shutdown itself on long GC pause.
- [OOZIE-2537](#):: SqoopMain does not set up log4j properly.
- [OOZIE-2532](#):: patch apply does not handle binary files.
- [OOZIE-2330](#):: Spark action should take the global jobTracker and nameNode configs by default and allow file and archive elements.
- [OOZIE-2510](#):: TestLogStreamer.testStreamLog. fails in other timezones.
- [OOZIE-2531](#):: Prevent Spark trying for token which is already available.
- [OOZIE-2509](#):: SLA job status can stuck in running state.
- [OOZIE-2529](#):: Support adding secret keys to Credentials of Launcher.

- [OOZIE-1402](#):: Increase retry interval for non-progressing coordinator action with fix value .
- [OOZIE-2512](#):: ShareLibservice returns incorrect path for jar.
- [OOZIE-2471](#):: Show child job url tab for distcp.
- [OOZIE-2511](#):: SubWorkflow missing variable set from option if config-default is present in parent workflow.
- [OOZIE-2391](#):: spark-opts value in workflow.xml is not parsed properly.
- [OOZIE-2489](#):: XML parsing is vulnerable.
- [OOZIE-2485](#):: Oozie client keeps trying to use expired auth token.
- [OOZIE-2490](#):: Oozie can't set hadoop.security.token.service.use_ip.
- [OOZIE-2474](#):: <job-xml> is not being applied to the launcher job.
- [OOZIE-2486](#):: TestSLAEventsGetForFilterJPAExecutor is flakey.
- [OOZIE-2481](#):: Add YARN_CONF_DIR in the Shell action.
- [OOZIE-2492](#):: JSON security issue in js code.
- [OOZIE-2429](#):: TestEventGeneration test is flakey.
- [OOZIE-2466](#):: Repeated failure of TestMetricsInstrumentation.testSamplers.
- [OOZIE-2470](#):: Remove infinite socket timeouts in the Oozie email action.
- [OOZIE-2246](#):: CoordinatorInputCheckCommand does not behave properly when har file is one of data dependency and doesn't exist.
- [OOZIE-2446](#):: Job does not fail during submission if non existent credential is specified .
- [OOZIE-2283](#):: Documentation should not say that System.exit is not allowed in Java Action .
- [OOZIE-2400](#):: Workflow xml configuration parser cannot deal with namespace prefix.
- [OOZIE-2452](#):: Coordinator Functional Specification - EL Constants Typo.
- [OOZIE-2173](#):: DISCLAIMER.txt still refers to Apache Incubator.
- [OOZIE-2312](#):: oozie doesn't purge audit and error log.
- [OOZIE-2431](#):: Support hive.metastore.uris as alias for heat.metastore.uri.
- [OOZIE-1976](#):: Specifying coordinator input datasets in more logical ways.
- [OOZIE-2444](#):: Need conditional logic in bundles.
- [OOZIE-2394](#):: Oozie can execute command without holding lock.
- [OOZIE-1922](#):: MemoryLocksService fails if lock is acquired multiple times in same thread and released.
- [OOZIE-2432](#):: TestPurgeXCommand fails.
- [OOZIE-2434](#):: inconsistent coord action status and workflow job status.
- [OOZIE-2438](#):: Oozie client "jobs -filter" diagnostic message clarification.
- [OOZIE-2165](#):: Job log fetching can fail in Oozie HA mode when using doAs impersonation.
- [OOZIE-2430](#):: Add root logger for hive,sqoop action.
- [OOZIE-2441](#):: SubWorkflow action with propagate-configuration but no global section throws NPE on submit.
- [OOZIE-2370](#):: Docs for Coordinator Action Status Notification has wrong property name.
- [OOZIE-2419](#):: HBase credentials are not correctly proxied.
- [OOZIE-2439](#):: FS Action no longer uses name-node from global section or default NN.
- [OOZIE-2435](#):: TestCoordChangeXCommand is flakey.
- [OOZIE-2428](#):: TestSLAService, TestSLAEventGeneration flakey tests.
- [OOZIE-2437](#):: Exclude additional Hadoop dependencies from Spark Sharelib.
- [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-2380](#):: Oozie Hive action failed with wrong tmp path.
- [OOZIE-2222](#):: Oozie UI parent job should be clickable.
- [OOZIE-2407](#):: AbandonedService should not send mail if there is no abandoned coord.
- [OOZIE-2402](#):: oozie-setup.sh sharelib create takes a long time on large clusters.
- [OOZIE-2185](#):: Make oozie cli source conf/oozie-client-env.sh.
- [OOZIE-2413](#):: Kerberos credentials can expire if the KDC is slow to respond.
- [OOZIE-2411](#):: Add BCC to oozie email action.
- [OOZIE-2397](#):: LAST_ONLY and NONE don't properly handle READY actions.

- [OOZIE-2401](#):: Typo in twiki docs with FilesAchives instead of FilesArchives.
- [OOZIE-2168](#):: Oozie flow and action names have 50 char limit.
- [OOZIE-2346](#):: Add sub-workflow information like the super-parent id and workflow depth into the 'oozie.job.info' property.
- [OOZIE-2303](#):: Typo in documentation.
- [OOZIE-2328](#):: Coordinator endtime change should check if the last action is in database .
- [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-2251](#):: Expose instrumental matrices in Realtime Graphing tool.
- [OOZIE-2314](#):: Unable to kill old instance child job by workflow or coord rerun by Launcher .
- [OOZIE-2386](#):: org.apache.oozie.util.TestXConfiguration.testSubstituteVar is flakey.
- [OOZIE-2385](#):: org.apache.oozie.TestCoordinatorEngineStreamLog.testCoordLogStreaming is flakey.
- [OOZIE-2382](#):: org.apache.oozie.action.hadoop.TestPigMain.testPig_withNullExternalID is flakey.
- [OOZIE-2379](#):: org.apache.oozie.command.coord.TestCoordPushDependencyCheckXCommand.testRequeueOnException is flakey.
- [OOZIE-2378](#):: org.apache.oozie.client.TestOozieCLI.testAdminInstrumentation is flakey.
- [OOZIE-2377](#):: Hive2 Action should not propagate oozie.hive2.* properties to Beeline.
- [OOZIE-2376](#):: Default action configs not honored if no <configuration> section in workflow.
- [OOZIE-2365](#):: oozie fail to start when smtp password not set.
- [OOZIE-2360](#):: Spark Action fails due to missing mesos jar.
- [OOZIE-2345](#):: Parallel job submission for forked actions.
- [OOZIE-2358](#):: Coord rerun cleanup should reuse heat connections.
- [OOZIE-2356](#):: Add a way to enable/disable credentials in a workflow.
- [OOZIE-2355](#):: Hive2 Action doesn't pass along oozie configs to jobconf.
- [OOZIE-2318](#):: Provide better solution for specifying SSL truststore to Oozie Client.
- [OOZIE-2344](#):: Enabling 'oozie.action.jobinfo.enable' doesn't inject the job information into the map/reduce job's configuration..
- [OOZIE-2350](#):: Package changes for release.
- [OOZIE-2347](#):: Remove unnecessary new Configuration./new jobConf. calls from oozie.
- [OOZIE-2348](#):: Recovery service keeps on recovering coord action of suspended jobs.
- [OOZIE-2277](#):: Honor oozie.action.sharelib.for.spark in Spark jobs.
- [OOZIE-2322](#):: Oozie Web UI doesn't work with Kerberos in Internet Explorer 10 or 11 and curl .
- [OOZIE-2343](#):: Shell Action should take Oozie Action config and setup HADOOP_CONF_DIR.
- [OOZIE-2245](#):: Service to periodically check database schema.
- [OOZIE-2332](#):: Add ability to provide Hive and Hive 2 Action queries inline in workflows .
- [OOZIE-2329](#):: Make handling yarn restarts configurable.
- [OOZIE-2228](#):: Statustransit service doesn't pick bundle with suspend status.
- [OOZIE-2325](#):: Shell action fails if user overrides oozie.launcher.mapreduce.map.env.
- [OOZIE-2324](#):: A syntax error in the kill node causes the workflow to get stuck and other problems.
- [OOZIE-2309](#):: Enable the coord:dateOffset. function in /coordinator-app/datasets/dataset/@initial-instance.
- [OOZIE-2305](#):: Compile Oozie with Hive-1.2.0.
- [OOZIE-2320](#):: TestZKXLogStreamingService.testStreamingWithMultipleOozieServers_coordActionList is failing.
- [OOZIE-2293](#):: Oozie 4.0.1 build failed while building Catalog.
- [OOZIE-2308](#):: Add support for bundle:conf. function.
- [OOZIE-2315](#):: TestOozieCLI.testshareLibUpdate_withSecurity fails with Hadoop 2.
- [OOZIE-2291](#):: Hive2 workflow.xml.security should have "cred" in action tag instead of "hive2" tag.
- [OOZIE-2289](#):: hive-jdbc dependency in core/pom.xml should be compile.

- [OOZIE-2300](#):: TestAuthFilterAuthOozieClient.testClientAuthTokenCache fails with Hadoop 2.7.0 and later.
- [OOZIE-1837](#):: LauncherMainHadoopUtils sensitive to clock skew.
- [OOZIE-2187](#):: Add a way to specify a default JT/RM and NN.
- [OOZIE-2272](#):: Use Hadoop's CredentialProvider for passwords in oozie-site.
- [OOZIE-2287](#):: Add support for deleting heat partitions in fs action delete.
- [OOZIE-2285](#):: Change in concurrency should trigger coord action ready command.
- [OOZIE-2284](#):: HBaseCredentials should only add hbase-default.xml and hbase-site.xml to actionConf.
- [OOZIE-2286](#):: Update Log4j and Log4j-extras to latest 1.2.x release.
- [OOZIE-2250](#):: Show log for WAITING and TIMEDOUT coord actions.
- [OOZIE-2262](#):: Fix log streaming from other server with start/end filter.
- [OOZIE-2159](#):: 'oozie validate' command should be moved server-side.
- [OOZIE-2271](#):: Upgrade Tomcat to 6.0.44.
- [OOZIE-2266](#):: Fix 'total' actions returned in coordinator job info.
- [OOZIE-2264](#):: Fix coord:offset. to resolve correct data set.
- [OOZIE-2178](#):: fix javadoc to compile on JDK8.
- [OOZIE-2268](#):: Update ActiveMQ version for security and other fixes.
- [OOZIE-2215](#):: Support glob in FS EL function.
- [OOZIE-2270](#):: Audit log doesn't get loaded when "audit log" tab is clicked for coord job.
- [OOZIE-2257](#):: BundleStatusTransitXCommand doesn't update last modified time.
- [OOZIE-2260](#):: Only set "java.io.tmpdir" to "./tmp" for the AM.

Phoenix

This release provides Phoenix 5.0.0 and no additional Apache patches.

HDP 3.0.0 and HDP 3.0.1 provided Phoenix 5.0.0 and no additional Apache patches.

Pig

This release provides Pig 0.16.0 and the listed Apache patches.

HDP 3.0.0 and HDP 3.0.1 provided Pig 0.16.0 and the following Apache patches:

- [PIG-4947](#): LOAD with HBaseStorage using a mix of pure wildcards and prefixed wildcards results in empty maps for the pure wildcarded column families.
- [PIG-4948](#): Pig on Tez AM use too much memory on a small cluster.
- [PIG-4921](#): Kill running jobs on InterruptedException (partial).
- [PIG-4916](#): Pig on Tez fail to remove temporary HDFS files in some cases.
- [PIG-5287](#): Bump jython to 2.7.1.

Ranger

This release provides Ranger 1.2.0 and the listed Apache patches.

- [RANGER-2292](#): Test case fix for RANGER-2276
- [RANGER-2244](#): Tomcat Security Vulnerability Alert. The version of the tomcat for ranger should upgrade to 7.0.91 or later.
- [RANGER-2288](#): Sqoop repository config missing 'Common Name for Certificate'
- [RANGER-2284](#): Unable to build image using docker
- [RANGER-2276](#): Email Address should be verified when Add New User in Ranger Admin
- [RANGER-2267](#): Add a icon to differentiate the status of the service
- [RANGER-2277](#): Kylin repository config missing 'Common Name for Certificate'
- [RANGER-2280](#): The emptyText of User Sync and Plugin Status should be reasonable
- [RANGER-2286](#): Ranger install may be prevented by leftover DB entry
- [RANGER-2283](#): User is getting total count of groups even if he is assigned to one group due to which pagination is breaking

- [RANGER-2279](#): Reduce the time spent changing passwords during Ranger Admin install
- [RANGER-2278](#): Unable to delete user if he has references in new ref tables
- [RANGER-2248](#): Sorting does not work in AbstractPredicateUtil.java
- [RANGER-2257](#): Add policyID to error message when click the Access log of Audit
- [RANGER-2258](#): Improve the policy list page to prompt users when the service is disabled
- [RANGER-2264](#): Kafka default policies for new resources are not showing up in UI when upgrade is done from older version
- [RANGER-2251](#): Need to provide options for making java heap size memory configurable in Ranger services
- [RANGER-2259](#): Need to provide appropriate permissions for unix-auth files.
- [RANGER-2252](#): Permission Kafka Admin should not be part of Topic resource in Ranger Kafka resource definition
- [RANGER-2209](#): Added Ranger service definition for ABFS
- [RANGER-2220](#): minify JavaScript files during packaging
- [RANGER-2238](#): String comparison should not use '==' in ServiceUtil.java
- [RANGER-2249](#): Ranger Audit not flushed immediately to hdfs
- [RANGER-2235](#): Modify the login session detail page as a modal
- [RANGER-2241](#): Removed mds and sha1 generation
- [RANGER-2222](#): Apache RangerKafkaPlugin support to handle Kafka Cluster as a new resource
- [RANGER-2210](#): Ranger support for Apache Kafka 2.0.0

HDP 3.0.0 and HDP 3.0.1 provided Ranger 1.1.0 and the following Apache patches.

- [RANGER-1715](#): Enhance Ranger Hive Plugin to support authorization on Hive replication Tasks
- [RANGER-1715](#): Enhance Ranger Hive Plugin to support authorization on Hive replication Tasks - repladmin permission
- [RANGER-1841](#): Audit log record for 'use dbName' hive command contains large number of tags
- [RANGER-1828](#): Good coding practice-add additional headers in ranger
- [RANGER-1833](#): Update Ranger to use 0.8.1 Atlas version
- [RANGER-1825](#): BUG-89494 : Ranger Tagsync start is failing
- [RANGER-91119](#): Exclude jackson jaxrs library from ranger-admin packaging
- [RANGER-1886](#): TagSync updates for changes in Atlas API
- [RANGER-1886](#): TagSync updates for changes in Atlas API - Part 2
- [RANGER-1883](#): Part 2 - All tag sources and sinks should use the same kerberos identity and reuse it
- [RANGER-1897](#): TagSync should replace use of V1 Atlas APIs with V2 APIs for efficient tag-download from Atlas
- [RANGER-1895](#): Simplify Storm dependencies
- [RANGER-1797](#): Tomcat Security Vulnerability Alert. The version of the tomcat for ranger should upgrade to 7.0.82.
- [RANGER-1906](#): Simplify Atlas plugin dependency management
- [RANGER-1898](#): Simplify Knox plugin dependency management
- [RANGER-1893](#): Update HBase dependency to 1.2.6
- [RANGER-1922](#): Simplify HBase dependencies
- [RANGER-1926](#): Upgrade Storm dependency to 1.1.1
- [RANGER-1905](#): Fix assembly creation - it's a black magic. As ranger-plugins-audit doesn't exclude guava from hadoop common, an older guava is leaked into embeddedwebserver classpath, which somehow filtered out by the assembly plugin Fix the atlas plugin assembly, so the guava is actually coming from the 'atlas-plugin's dependency, and not from the ranger-plugins-audit's dependency. And revert "RANGER-1858:ranger-admin start fail when policymgr_http_enabled equal to false."
- [RANGER-1984](#): Hbase audit log records may not show all tags associated with accessed column
- [RANGER-1999](#): Ranger policy engine updates to support list-of-values in access resource
- [RANGER-2000](#): Policy effective dates to support time-bound and temporary authorization
- [RANGER-1987](#): Upgrade to Storm 1.2.0 in Ranger

- [RANGER-2019](#): Handle upgrade scenario to rename the old ATLAS service def and use the new service def
- [RANGER-2019](#): Handle upgrade scenario to rename the old ATLAS service def and use the new service def
- [RANGER-2027](#): Evaluate grantor's group membership in the plugin for grant/revoke request
- [RANGER-2026](#): Update Hbase plugin to handle default namespace
- [RANGER-2018](#): Upgrade to Spring 4
- [RANGER-1994](#): Tomcat Security Vulnerability Alert. The version of the tomcat for ranger should upgrade to 7.0.85.
- [RANGER-2035](#): fix handling of null implClass in servicedefs (Oracle returns null for stored empty strings)
- [RANGER-2043](#): Ranger KMS KeyProvider and HSM KeyProvider should have more debug logs
- [RANGER-2115](#): Atlas service def creation failing in Ranger manual installation

Spark

This release provides Spark 2.3.2 and the listed Apache patches.

- [SPARK-21783](#): Turn on `native` ORC impl and PPD by default
- [SPARK-23456](#): Turn on `native` ORC impl and PPD by default
- [SPARK-23228](#): Add Python Created jsparkSession to JVM's defaultSession
- [SPARK-23510](#): Support Hive 2.2 and Hive 2.3 metastore
- [SPARK-23518](#): Avoid metastore access when the users only want to read and write data frames
- [SPARK-23635](#): Spark executor env variable is overwritten by same name AM env variable
- [SPARK-23787](#): Fix file download test in SparkSubmitSuite for Hadoop 2.9.
- [SPARK-23355](#): convertMetastore should not ignore table properties
- [SPARK-24110](#): Avoid UGI.loginUserFromKeytab in STS
- [SPARK-24149](#): Retrieve all federated namespaces tokens
- [SPARK-24209](#): Automatic retrieve proxyBase from Knox headers
- [SPARK-24312](#): Upgrade to 2.3.3 for Hive Metastore Client 2.3
- [SPARK-24377](#): make --py-files work in non pyspark application
- [SPARK-24479](#): Added config for registering streamingQueryListeners
- [SPARK-24518](#): Using Hadoop credential provider API to store password
- [SPARK-24660](#): Show correct error pages when downloading logs in SHS
- [SPARK-23654](#): remove jets3t as a dependency of spark
- [SPARK-25126](#): Avoid creating Reader for all orc files
- [SPARK-23679](#): Setting RM_HA_URLS for AmIpFilter to avoid redirect failure in YARN mode
- [SPARK-25306](#): Avoid skewed filter trees to speed up `createFilter` in ORC

HDP 3.0.0 and HDP 3.0.1 provided Spark 2.3.1 and the following Apache patches.

- [SPARK-24495](#): SortMergeJoin with duplicate keys wrong results.
- [SPARK-207](#): Remove hardcoded FS scheme from Spark archive.

(Backport from 2.3.2)

- [SPARK-24455](#): fix typo in TaskSchedulerImpl comment.
- [SPARK-24369](#): Correct handling for multiple distinct aggregations having the same argument set.
- [SPARK-24468](#): Handle negative scale when adjusting precision for decimal operations.
- [SPARK-23732](#): Fix source links in generated scaladoc.
- [SPARK-24502](#): flaky test: UnsafeRowSerializerSuite.
- [SPARK-24531](#): Remove version 2.2.0 from testing versions in HiveExternalCatalogVersionsSuite.
- [SPARK-24506](#): Add UI filters to tabs added after binding.
- [SPARK-23754](#): Move UDF stop iteration wrapping from driver to executor.
- Remove unnecessary sort in UnsafeInMemorySorterSuite.
- Fix typo in serializer exception.
- Revert [SPARK-21743](#) top-most limit should not cause memory leak.

- [SPARK-24531](#): Replace 2.3.0 version with 2.3.1.

(Backport from 2.4)

- [SPARK-21783](#): Turn on `native` ORC impl and PPD by default.
- [SPARK-23456](#): Turn on `native` ORC impl and PPD by default.
- [SPARK-23228](#): Add Python Created jsparkSession to JVM's defaultSession.
- [SPARK-23510](#): Support Hive 2.2 and Hive 2.3 metastore.
- [SPARK-23518](#): Avoid metastore access when the users only want to read and write data frames.
- [SPARK-23635](#): Spark executor env variable is overwritten by same name AM env variable.
- [SPARK-23787](#): Fix file download test in SparkSubmitSuite for Hadoop 2.9..
- [SPARK-23355](#) convertMetastore should not ignore table properties.
- [SPARK-24110](#) Avoid UGI.loginUserFromKeytab in STS.
- [SPARK-24149](#): Retrieve all federated namespaces tokens.
- [SPARK-24209](#): Automatic retrieve proxyBase from Knox headers.
- [SPARK-24312](#): Upgrade to 2.3.3 for Hive Metastore Client 2.3.
- [SPARK-24377](#): make --py-files work in non pyspark application.
- [SPARK-24479](#): Added config for registering streamingQueryListeners.
- [SPARK-24518](#): Using Hadoop credential provider API to store password.
- [SPARK-24660](#): Show correct error pages when downloading logs in SHS.

Sqoop

This release provides Sqoop 1.4.7 and no additional Apache patches.

HDP 3.0.0 and HDP 3.0.1 provided Sqoop 1.4.7 and no additional Apache patches.

Storm

This release provides Storm 1.2.1 and the patches listed for HDP 3.0.0 and HDP 3.0.1.

HDP 3.0.0 and HDP 3.0.1 provided Storm 1.2.1 and the following Apache patches:

- [STORM-1839](#): Kinesis Spout
- [STORM-1694](#): Kafka Spout Trident Implementation Using New Kafka Consumer API
- [STORM-1960](#): Add CORS support to STORM UI Rest api
- [STORM-1959](#): Add missing license header to KafkaPartitionOffsetLag
- [STORM-1950](#): Change response json of "Topology Lag" REST API to keyed by spoutId, topic, partition.
- [STORM-1833](#): Simple equi-join in storm-sql standalone mode
- [STORM-1866](#): Update Resource Aware Scheduler Documentation
- [STORM-1930](#): Kafka New Client API - Support for Topic Wildcards
- [STORM-1924](#): Adding conf options for Persistent Word Count Topology
- [STORM-1956](#): Disabling Backpressure by default
- [STORM-1934](#): Fix race condition between sync-supervisor and sync-processes
- [STORM-1919](#): Introduce FilterBolt on storm-redis
- [STORM-1742](#): More accurate 'complete latency'
- [STORM-1945](#): Fix NPE bugs on topology spout lag for storm-kafka-monitor
- [STORM-1888](#): add description for shell command
- [STORM-1902](#): add a simple & flexible FileNameFormat for storm-hdfs
- [STORM-1914](#): Storm Kafka Field Topic Selector
- [STORM-1907](#): PartitionedTridentSpoutExecutor has incompatible types that cause ClassCastException
- [STORM-1925](#): Remove Nimbus thrift call from Nimbus itself
- [STORM-1909](#): Update HDFS spout documentation
- [STORM-1136](#): Command line module to return kafka spout offsets lag and display in storm ui
- [STORM-1911](#): IClusterMetricsConsumer should use seconds to timestamp unit

- [STORM-1849](#): HDFSFileTopology should use the 3rd argument as topologyName
- [STORM-1906](#): Window count/length of zero should be disallowed
- [STORM-1893](#): Support OpenTSDB for storing timeseries data.
- [STORM-1723](#): Introduce ClusterMetricsConsumer
- [STORM-1700](#): Introduce 'whitelist' / 'blacklist' option to MetricsConsumer
- [STORM-1698](#): Asynchronous MetricsConsumerBolt
- [STORM-1705](#): Cap number of retries for a failed message
- [STORM-1884](#): Prioritize pendingPrepare over pendingCommit
- [STORM-1575](#): fix TwitterSampleSpout NPE on close
- [STORM-1874](#): Update logger private permissions
- [STORM-1865](#): update command line client document
- [STORM-1771](#): HiveState should flushAndClose before closing old or idle Hive connections
- [STORM-1882](#): Expose TextFileReader public
- [STORM-1873](#): Implement alternative behaviour for late tuples
- [STORM-1719](#): Introduce REST API: Topology metric stats for stream
- [STORM-1887](#): Fixed help message for set_log_level command
- [STORM-1878](#): Flux can now handle IStatefulBolts
- [STORM-1864](#): StormSubmitter should throw respective exceptions and log respective errors for registered submitter hook invocation
- [STORM-1868](#): Modify TridentKafkaWordCount to run in distributed mode
- [STORM-1859](#): Ack late tuples in windowed mode
- [STORM-1851](#): Fix default nimbus impersonation authorizer config
- [STORM-1848](#): Make KafkaMessageId and Partition serializable to support
- [STORM-1862](#): Flux ShellSpout and ShellBolt can't emit to named streams
- [STORM-1728](#): TransactionalTridentKafkaSpout error
- [STORM-1850](#): State Checkpointing Documentation update
- [STORM-1674](#): Idle KafkaSpout consumes more bandwidth than needed
- [STORM-1842](#): Forward references in storm.thrift cause tooling issues
- [STORM-1730](#): LocalCluster#shutdown() does not terminate all storm threads/thread pools.
- [STORM-1841](#): Address a few minor issues in windowing and doc
- [STORM-1709](#): Added group by support in storm sql standalone mode
- [STORM-1720](#): Support GEO in storm-redis

Tez

This release provides Tez 0.9.1 and the listed Apache patches.

- [TEZ-3976](#): Batch ShuffleManager error report events
- [TEZ-3984](#): Shuffle: Out of Band DME event sending causes errors
- [TEZ-3980](#): ShuffleRunner: the wake loop needs to check for shutdown
- [TEZ-3958](#): Add internal vertex priority information into the tez dag.dot debug information
- [TEZ-3978](#): DAGClientServer Socket exception when localhost name lookup failures
- [TEZ-3974](#): Correctness regression of TEZ-955 in TEZ-2937
- [TEZ-3934](#): LegacySpeculator sometime issues wrong number of speculative attempts
- [TEZ-3942](#): RPC getTask writable optimization invalid in hadoop 2.8+
- [TEZ-3965](#): TestMROOutput: Fix the hard-coded /tmp/output paths
- [TEZ-3916](#): Add hadoop-azure-datalake jar to azure profile
- [TEZ-3970](#): NullPointerException in Tez ShuffleHandler Ranged Fetch
- [TEZ-3964](#): Inflater not closed in some places
- [TEZ-3912](#): Fetchers should be more robust to corrupted inputs
- [TEZ-3963](#): Possible InflaterInputStream leaked in TezCommonUtils and related classes
- [TEZ-3954](#): Reduce Tez Shuffle Handler Memory needs for holding TezIndexRecords

- [TEZ-3960](#): Better error handling in proto history logger and add doAs support.
- [TEZ-3962](#): Configuration decode leaks an Inflater object
- [TEZ-3953](#): Restore API-compat for DAGClient for TEZ-3951
- [TEZ-3951](#): TezClient wait too long for the DAGClient for prewarm; tries to shut down the wrong DAG
- [TEZ-3944](#): TestTaskScheduler times-out on Hadoop3
- [TEZ-3938](#): Task attempts failing due to not making progress
- [TEZ-3949](#): TestATSHistoryV15 is failing with hadoop3+
- [TEZ-3946](#): NoClassDefFoundError, org.apache.hadoop.mapred.ShuffleHandler.
- [TEZ-3929](#): Upgrade Jersey to 1.19
- [TEZ-3947](#): TestATSHistoryWithACLs fails with Hadoop3 and Jersey 1.19
- [TEZ-3943](#): TezClient leaks DAGClient for prewarm
- [TEZ-3939](#): Remove performance hit of precondition check in AM for register running task attempt
- [TEZ-3940](#): Reduce time to convert TaskFinishedEvent to string
- [TEZ-3902](#): Upgrade to netty-3.10.5.Final.jar
- [TEZ-3935](#): DAG aware scheduler should release unassigned new containers rather than hold them
- [TEZ-3937](#): Empty partition BitSet to byte[] conversion creates one extra byte in rounding error
- [TEZ-3824](#): MRCombiner creates new JobConf copy per spill
- [TEZ-3933](#): Remove sleep from test TestExceptionPropagation
- [TEZ-3911](#): Optional min/max/avg aggr. task counters reported to HistoryLoggingService at final counter aggr
- [TEZ-3932](#): TaskSchedulerManager can throw NullPointerException during DAGAppMaster container cleanup race
- [TEZ-3931](#): TestExternalTezServices fails on Hadoop3
- [TEZ-3930](#): TestDagAwareYarnTaskScheduler fails on Hadoop 3
- [TEZ-3927](#): TestReduceProcessor fails on Hadoop 3.x
- [TEZ-3924](#): TestDefaultSorter fails intermittently due random keys and interaction with RLE and partition collisions
- [TEZ-3914](#): Recovering a large DAG fails to size limit exceeded
- [TEZ-3887](#): Tez Shuffle Handler should support Index Cache configuration
- [TEZ-3914](#): Recovering a large DAG fails to size limit exceeded
- [TEZ-3817](#): DAGs can hang after more than one uncaught Exception during doTransition.
- [TEZ-3915](#): Create protobuf based history event logger.
- [TEZ-3909](#): DAG can hang if vertex with no tasks is killed
- [TEZ-3907](#): Improve log message to include the location the writers decide to spill output
- [TEZ-3874](#): NPE in TezClientUtils when "yarn.resourcemanager.zk-address" is present in Configuration.
- [TTEZ-3892](#): getClient API for TezClient
- [TEZ-3897](#): Tez Local Mode hang for vertices with broadcast input.
- [TEZ-3888](#): Update Jetty to org.eclipse.jetty 9.x
- [TEZ-3898](#): TestTezCommonUtils fails when compiled against hadoop version >= 2.8
- [TEZ-3896](#): TestATSV15HistoryLoggingService#testNonSessionDomains is failing
- [TEZ-3893](#): Tez Local Mode can hang for cases.
- [TEZ-3894](#): Tez intermediate outputs implicitly rely on permissive umask for shuffle
- [TEZ-3895](#): Missing name for local mode task scheduler service async request handler thread
- [TEZ-3770](#): DAG-aware YARN task scheduler
- [TEZ-3880](#): Do not count rejected tasks as killed in vertex progress
- [TEZ-3877](#): Delete unordered spill files once merge is done

HDP 3.0.0 and HDP 3.0.1 provided Tez 0.9.1 and no additional Apache patches.

Zeppelin

This release provides Zeppelin 0.8.0 and no additional Apache patches.

HDP 3.0.0 and HDP 3.0.1 provided Zeppelin 0.8.0 and no additional Apache patches.

ZooKeeper

This release provides ZooKeeper 3.4.6 and the patches listed for HDP 3.0.0 and HDP 3.0.1

HDP 3.0.1 provided ZooKeeper 3.4.6 with the following Apache patches:

- [ZOOKEEPER-2662](#): Export a metric for txn log sync times
- [ZOOKEEPER-3019](#): Add a metric to track number of slow fsyncs

HDP 3.0.0 provided ZooKeeper 3.4.6 and the following Apache patches:

- [ZOOKEEPER-2141](#): ACL cache in DataTree never removes entries
- [ZOOKEEPER-1897](#): ZK Shell/Cli not processing commands (stack via michim)
- [ZOOKEEPER-1926](#): Unit tests should only use build/test/data for data
- [ZOOKEEPER-2064](#): Prevent resource leak in various classes
- [ZOOKEEPER-2053](#): Make scripts use ZOOKEEPER_HOME if set.
- [ZOOKEEPER-1876](#): Add support for installing windows services in .cmd scripts
- [ZOOKEEPER-1848](#): [WINDOWS] Java NIO socket channels does not work with Windows ipv6 on JDK6
- [ZOOKEEPER-1574](#): Set svn:eol-style property to native.
- [ZOOKEEPER-1575](#): adding .gitattributes to prevent CRLF and LF mismatches for source and text files (Raja Aluri via michim)
- [ZOOKEEPER-1888](#): ZkCli.cmd commands fail with "'java' is not recognized as an internal or external command" (Ivan Mitic via michim)
- [ZOOKEEPER-1791](#): ZooKeeper package includes unnecessary jars that are part of the package. (mahadev)
- [ZOOKEEPER-1506](#): Re-try DNS hostname -> IP resolution if node connection fails
- [ZOOKEEPER-1952](#): Default log directory and file name can be changed
- [ZOOKEEPER-706](#): large numbers of watches can cause session re-establishment to fail (Chris Thunes)
- [ZOOKEEPER-1576](#): Zookeeper cluster - failed to connect to cluster if one of the provided IPs causes java.net.UnknownHostException (Edward Ribeiro via camille)

Fixed Common Vulnerabilities and Exposures

Common Vulnerabilities and Exposures (CVE) that are addressed in this release.

CVE-2018-1331

Component: Storm

Summary: An attacker with access to a secure storm cluster in some cases could execute arbitrary code as a different user.

Severity: Moderate

Vendor: Hortonworks

Versions Affected: HDP 3.0.0, HDP 3.0.1, HDP 2.6.x and HDF 3.2 or earlier

Users Affected: Users with Storm deployed in a secure cluster.

Impact: See [STORM-3026](#). An attacker with access to a secure storm cluster in some cases could execute arbitrary code as a different user.

Recommended Action: Upgrade to HDP 3.1 or HDF 3.3. After upgrading, the following configs needs to be set to enforce these ACL checks.`storm.nimbus.zookeeper.acls.check: true``storm.nimbus.zookeeper.acls.fixup: true`

CVE-2018-1332

Component: Storm

Summary: In a secure Storm cluster an attacker could impersonate another user when communicating with some Storm Daemons.

Severity: Moderate

Vendor: Hortonworks

Versions Affected: HDP 3.0.0, HDP 2.6.x, HDF 3.2 and earlier

Users Affected: Users with Storm deployed in a secure cluster.

Impact: See [STORM-3027](#). The affected Storm versions expose a vulnerability that could allow a user to impersonate another user when communicating with some Storm Daemons.

Recommended Action: Upgrade to HDP 3.1.0 or HDF 3.3.

CVE-2018-11777

Component: Hive/Hive2

Summary: Local resources on HiveServer2 machines are not properly protected against malicious user if Ranger or SQL Standard Authorizer is not in use.

Severity: Important

Vendor: Hortonworks

Versions Affected: HDP 1.0.0 to HDP 2.6.5, HDP 3.0.0, and HDP 3.0.1

Users Affected: This affects only configurations of HDP where Ranger or SQL Standard Authorization is not enabled.

Impact: Local resource on HiveServer2 machine will be read/written by arbitrary Hive user if Ranger or SQL Standard Authorization is not in use.

Recommended Action: It is recommended to upgrade a HDP version with the fix if HiveServer2 is used, and Ranger or SQL Standard Authorizer is not in use. Admin needs to specify the following entries in the hiveserver2-site.xml file:

```
<property>
  <name>hive.security.authorization.enabled</name>
  <value>true</value>
</property>
<property>
  <name>hive.security.authorization.manager</name>
  <value>org.apache.hadoop.hive.ql.security.authorization.plugin.fallback.
  FallbackHiveAuthorizerFactory</value>
</property>
```

CVE-2018-1314

Component: Hive/Hive2

Summary: Hive "EXPLAIN" operation does not check for necessary authorization of involved entities in a query. An unauthorized user can do "EXPLAIN" on arbitrary table or view and expose table metadata and statistics.

Severity: Important

Vendor: Hortonworks

Versions Affected: HDP 1.0.0 to HDP 2.6.5, HDP 3.0.0, and HDP 3.0.1

Impact: Hive metadata and statistics is not secure against unauthorized Hive user.

Recommended Action: Upgrading to an HDP version with the fix will address the problem. HDP versions with fix:

- 3.0.1.3 (If current version is HDP 3.0.x)
- 2.6.5.54 (If current version is HDP HDP-2.6.5.0)

- 2.6.5.1003 (If current version is HDP 2.6.5.100* versions released for Data Lifecycle Manager support)

CVE-2018-8008

Component: Storm

Summary: Apache Storm arbitrary file write vulnerability.

Severity: Important

Vendor: The Apache Software Foundation

Versions Affected: HDP 3.0.0, HDP 2.6.5 and earlier

Impact: Apache Storm version 1.0.6 and earlier, 1.2.1 and earlier, and version 1.1.2 and earlier expose an arbitrary file write vulnerability, that can be achieved using a specially crafted zip archive (affects other archives as well, bzip2, tar, xz, war, cpio, 7z), that holds path traversal filenames. So when the filename gets concatenated to the target extraction directory, the final path ends up outside of the target folder.

Recommended Action: Upgrade to HDP 3.0.1 or HDP 3.1.0.

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-108745	HIVE-20315	Vectorization: Fix more NULL / Wrong Results issues and avoid unnecessary casts/conversions
BUG-109279	HIVE-20311	add txn stats checks to some more paths
BUG-111143	CALCITE-2563	Materialized view rewriting may swap columns in equivalent classes incorrectly
BUG-111233	PHOENIX-4884	INSTR function should work seamlessly with literal and non-literal arguments
BUG-111235	PHOENIX-3746	NTH_VALUE returns 'null' when aggregating results of aggregate sub-query
BUG-111360	HIVE-20625	Regex patterns not working in SHOW MATERIALIZED VIEWS '<pattern>'
BUG-111366	HIVE-20617	Fix type of constants in IN expressions to have correct type
BUG-111946	HIVE-20677	JDBC storage handler ordering problem - single split flag
BUG-112133	HIVE-20618	During join selection BucketMapJoin might be chosen for non bucketed tables
BUG-112351	HIVE-20712	HivePointLookupOptimizer should extract deep cases
BUG-112358	ATLAS-2892 , ATLAS-2894	Fix for Atlas Tag Deletion via Knox issue
BUG-113446	HIVE-20792	Inserting timestamp with zones truncates the data
BUG-115301	HIVE-20988	Wrong results for group by queries with primary key on multiple columns

Other

Hortonworks Bug ID	Apache JIRA	Summary
BUG-109785	HIVE-20430 , HIVE-20446	CachedStore: bug / UT fixes when CachedStore is enabled
BUG-110118	N/A	Disable DDL API in webhcat
BUG-110503	YARN-8964	[UI2] YARN ui2 should use clusters/{cluster name} for all ATSV2 REST APIs

Hortonworks Bug ID	Apache JIRA	Summary
BUG-110627	HIVE-20409	Hive ACID: Update/delete/merge does not clean hdfs staging directory
BUG-110714	SPARK-23243 , SPARK-24415 , SPARK-24704 , SPARK-24889 , SPARK-24909 , SPARK-25124 , SPARK-25164 , SPARK-25231 , SPARK-25234 , SPARK-25273 , SPARK-25313 , SPARK-25330 , SPARK-25357 , SPARK-25368 , SPARK-25371 , SPARK-25402	Update Spark2 with critical fixes (9/12)
BUG-110908	HIVE-20542	Incremental REPL DUMP progress information log message is incorrect.
BUG-111491	HBASE-21219	Hbase incremental backup fails with null pointer exception
BUG-112401	HIVE-20718	Add perf cli driver with constraints
BUG-112561	HIVE-20684 , HIVE-20686 , HIVE-20698 , HIVE-20700	Backport multiple hive-druid fixes
BUG-112721	HIVE-20696	msck_*.q tests are broken
BUG-112943	HIVE-20807	Refactor LlapStatusServiceDriver
BUG-112948	N/A	Fix type change assertion error from calcite
BUG-113018	HIVE-20838	Timestamps with timezone are set to null when using the streaming API
BUG-113027	HIVE-20788	Extended SJ reduction may backtrack columns incorrectly when creating filters
BUG-113338	HIVE-20966	Support bootstrap replication to a target with hive.strict.managed.tables enabled.
BUG-113340	HIVE-20966	Support incremental replication to a target with hive.strict.managed.tables enabled.
BUG-113463	N/A	Set default tez.runtime.shuffle.batch.wait to 1 second
BUG-113690	KNOX-1098	Livy proxyUser should be added when not present
BUG-113799	HIVE-20744 , HIVE-20920	Join reorder issue in q17 and q29 after enabling constraints
BUG-113824	HIVE-20820	MV partition on clause position
BUG-113915	HIVE-20869	Fix test results file
BUG-114169	HBASE-21457	BackupUtils#getWALFilesOlderThan refers to wrong FileSystem
BUG-114180	HBASE-21466	WALProcedureStore uses wrong FileSystem if wal.dir is not under rootdir
BUG-114181	N/A	Update GCS connector to released version 1.9.10
BUG-114351	CALCITE-2669	RelMdTableReferences should check whether references inferred from input are null for Union/Join operators
BUG-114398	HIVE-19701	getDelegationTokenFromMetaStore doesn't need to be synchronized
BUG-114441	HIVE-20918	Flag to enable/disable pushdown of computation from Calcite into JDBC connection
BUG-114443	CALCITE-2673	Default implementation of SqlDialect should not support pushing all functions
BUG-114453	HIVE-20775	Factor cost of each SJ reduction when costing a follow-up reduction
BUG-114503	KNOX-1602	JsonFilterReader should handle strings, numbers, booleans, and null at root
BUG-114518	KNOX-1424	knox renaming "app" to "/gateway/ui_ns1/zeppelin/app"
BUG-114689	ATLAS-2975	Hive hook generates duplicate column_lineage entities

Hortonworks Bug ID	Apache JIRA	Summary
BUG-114751	HIVE-20949	Improve PKFK cardinality estimation in physical planning
BUG-114755	KNOX-1516 , KNOX-1612 , KNOX-1616	Upgrade Jetty to 9.4.12
BUG-114779	N/A	Ambari-View through Knox throws "IllegalArgumentException" while downloading file with Space in name
BUG-114957	HIVE-20775	Make HIVE-20775 deterministic
BUG-115019	HBASE-21523 , HBASE-21524	Chatty DEBUG logging in Master log
BUG-115023	YARN-9067	YARN Resource Manager is running OOM because of leak of Configuration Object
BUG-115099	HIVE-20951	LLAP: Set Xms to 50% always
BUG-115218	HADOOP-15959	Changes in kerberos user to local mapping permit unauthed access
BUG-115240	N/A	Memory leak in HiveIsolatedClassLoader
BUG-115270	CALCITE-2720	Presence of an empty aggregation should not lead to Exceptions
BUG-115392	KNOX-1649	Provide new service definitions for Ambari to support trusted proxy
BUG-115394	KNOX-1645	Improve JSESSIONID handling
BUG-115494	ATLAS-2982 , ATLAS-2984	Import fails to create classification-def
BUG-113792	RANGER-2278	Unable to delete user if he has references in new ref tables.
BUG-112785	RANGER-2252	Permission "Kafka Admin" should not be part of Topic resource in Ranger Kafka resource definition
BUG-112594	RANGER-2249	Ranger Audit not flushed immediately to HDFS
BUG-113510	RANGER-2273	Allow service admin and delegated admin user to view list of users and groups though they have 'USER' role

Performance

Hortonworks Bug ID	Apache JIRA	Summary
BUG-105250	HIVE-19985 , HIVE-20418 , ORC-389	Benchmark - LLAP performance issue for fully compacted ACID tables
BUG-108445	HIVE-20260	NDV of a column shouldn't be scaled when row count is changed by filter on another column
BUG-108446	HIVE-20439	Use the inflated memory limit during join selection for llap
BUG-108973	HIVE-20339	Vectorization: Lift unneeded restriction causing some PTF with RANK not to be vectorized
BUG-109318	HIVE-19097	related equals and in operators may cause inaccurate stats estimations
BUG-109408	HIVE-20433	Implicit String to Timestamp conversion is slow
BUG-109491	HIVE-20432	Rewrite BETWEEN to IN for integer types for stats estimation
BUG-109497	HIVE-20418	LLAP IO may not handle ORC files that have row index disabled correctly for queries with no columns selected
BUG-109794	N/A	Tez: lost data event problems with tez.runtime.enable.final-merge.in.output=false
BUG-109973	HIVE-20491 , HIVE-20503	Websales output is broadcasted in tpeds q49
BUG-110095	HIVE-20537 , HIVE-20612	Join order change in query 24 causes 3x performance degradation
BUG-110253	HIVE-20491	Fix mapjoin size estimations for Fast implementation
BUG-110836	N/A	spark llap query execution is stuck for more than 7 hrs

Hortonworks Bug ID	Apache JIRA	Summary
BUG-110938	HIVE-20296	Improve HivePointLookupOptimizerRule to be able to extract from more sophisticated contexts
BUG-110959	HIVE-20558	Change default of hive.hashtable.key.count.adjustment to 0.99
BUG-110960	N/A	Set hive.llap.memory.oversubscription.max.executors.per.query to be 1/3rd of num of llap daemons
BUG-111004	HIVE-20660	Group by statistics estimation could be improved by bounding the total number of rows to source table
BUG-111028	HIVE-20103	WM: Only Aggregate DAG counters if at least one is used
BUG-111075	CALCITE-1413 , CALCITE-2327 , CALCITE-2438 , CALCITE-2444 , CALCITE-2462 , CALCITE-2469 , CALCITE-2494 , CALCITE-2527 , CALCITE-2556 , CALCITE-2567 , CALCITE-2581 , CALCITE-2586 , CALCITE-2611 , CALCITE-2615 , HIVE-14431 , HIVE-20692 , HIVE-20710	Fold where coalesce(ws_qty,0)>0 and coalesce(cs_qty, 0)>0
BUG-111076	CALCITE-2555	RexSimplify: >=(true, null) could be simplified to null
BUG-111115	N/A	Change default value of hive.tez.llap.min.reducer.per.executor
BUG-111182	N/A	TPCDS query29 regresses due to join order change
BUG-111516	HIVE-20622	Backwards Compat: Support non-zero padded date and timestamp entries
BUG-111518	HIVE-20321	Vectorization: Cut down memory size of 1 col VectorHashKeyWrapper to <1 CacheLine
BUG-111582	HIVE-20636	Improve number of null values estimation after outer join
BUG-111589	HIVE-20623	Shared work: Extend sharing of map-join cache entries in LLAP
BUG-111743	HIVE-20702	Account for overhead from datastructure aware estimations during mapjoin selection
BUG-111936	HIVE-20704	Extend HivePreFilteringRule to support other functions
BUG-112007	HIVE-20326	Create constraints with RELY as default instead of NO RELY
BUG-112010	HIVE-20880	Update default value for hive.stats.filter.in.min.ratio
BUG-112069	N/A	Configure default HDP value for hive.llap.task.scheduler.locality.delay
BUG-112134	RANGER-2203 , RANGER-2218	Review and update database schema for ranger policies to minimize database queries/updates
BUG-112350	HIVE-20704	Extend HivePreFilteringRule to support other functions
BUG-112527	HIVE-20683	Add ability to pass in Bloom filter from Hive Queries
BUG-112619	KNOX-1521	Implement UrlRewriteRequestStream methods for performance
BUG-112764	HIVE-20765 , HIVE-20766	TPCDS Query compiles 10x slower in the latest benchmark run
BUG-112827	HIVE-20762	NOTIFICATION_LOG cleanup interval is hardcoded as 60s and is too small.
BUG-112839	HIVE-20767	Multiple project between join operators may affect join reordering using constraints
BUG-112951	HIVE-20804	Further improvements to group by optimization with constraints
BUG-113294	CALCITE-2631	General improvements in case simplification
BUG-113295	CALCITE-2537	Make sure assertions are enabled before calling VolcanoPlanner#validate

Hortonworks Bug ID	Apache JIRA	Summary
BUG-113300	CALCITE-2384	Performance issue in getPulledUpPredicates
BUG-113390	RANGER-2279	Reduce the time spent changing passwords during Ranger Admin install
BUG-113592	CALCITE-2668 , CALCITE-2669 , HIVE-20744	Query 72 regresses with constraint join reordering patch HIVE-20744
BUG-113691	HIVE-20954	Vector RS operator is not using uniform hash function for TPC-DS query 95
BUG-113805	HIVE-20873	Use Murmur hash for VectorHashKeyWrapperTwoLong to reduce hash collision
BUG-113825	HIVE-20821	Rewrite SUM0 into SUM + COALESCE combination
BUG-113854	N/A	Query4 and 11 regressing in latest run due to missing CASE expression rewrite
BUG-113936	HIVE-20926	Semi join reduction hint fails when bloom filter entries are high or when there are no stats
BUG-114131	HADOOP-15846	ABFS: fix mask related bugs in setAcl, modifyAclEntries and removeAclEntries.
BUG-114166	PHOENIX-5010	Don't build client guidepost cache when phoenix.stats.collection.enabled is disabled
BUG-114696	HIVE-20944	Not validate stats during query compilation
BUG-75812	HIVE-20646	Partition filter condition is not pushed down to metastore query if it has IS NOT NULL.
BUG-82439	HIVE-17040	Join elimination in the presence of FK relationship
BUG-82628	HIVE-17043	Remove non unique columns from group by keys if not referenced later
BUG-83684	HIVE-17044	Transform LEFT/RIGHT OUTER into INNER join in the presence of FK relationship

Potential Data Loss

Hortonworks Bug ID	Apache JIRA	Summary
BUG-111887	HIVE-20680	Bootstrap is missing partitions in replicated DB when retry after kill query.
BUG-112474	HIVE-20719	SELECT statement fails after UPDATE with hive.optimize.sort.dynamic.partition optimization and vectorization on
BUG-113384	RANGER-2268	Optimize policy and tags migration to new schema

Query Failure

Hortonworks Bug ID	Apache JIRA	Summary
BUG-105553	HIVE-20853	Expose ShuffleHandler.registerDag in the llap daemon API
BUG-107930	HIVE-19674	Group by Decimal Constants push down to Druid tables.
BUG-109306	HIVE-20399	CTAS w/a custom table location that is not fully qualified fails for MM tables
BUG-109475	HIVE-20554	Unable to drop an external table after renaming it.
BUG-109814	HIVE-19748	Add appropriate null checks to DecimalColumnStatsAggregator
BUG-110331	HIVE-20496	Vectorization: Vectorized PTF IllegalStateException
BUG-110692	HIVE-20522	HiveFilterSetOpTransposeRule may throw assertion error due to nullability of fields
BUG-110809	HIVE-20541	REPL DUMP on external table with add partition event throws NoSuchElementException.
BUG-110881	N/A	spark hive warehouse connector failed to insert into table if column in table are more than 25.
BUG-110909	ATLAS-2827	Unable to persist vertex when indexed string attribute is really long
BUG-111126	HIVE-20524	Schema Evolution checking is broken in going from Hive version 2 to version 3 for ALTER TABLE VARCHAR to DECIMAL

Hortonworks Bug ID	Apache JIRA	Summary
BUG-111150	HIVE-20563	Vectorization: CASE WHEN expression fails when THEN/ELSE type and result type are different
BUG-111224	HIVE-20632	Query with get_splits UDF fails if materialized view is created on queried table.
BUG-111522	PHOENIX-4874	psql doesn't support date/time with values smaller than milliseconds
BUG-111995	HIVE-20652	JdbcStorageHandler push join of two different datasource to jdbc driver
BUG-111996	HIVE-20669	JdbcStorageHandler push union of two different datasource to jdbc driver
BUG-112003	HIVE-20705	Vectorization: Native Vector MapJoin doesn't support Complex Big Table values
BUG-112111	PHOENIX-4859	Using local index in where statement for join (only rhs table) query fails
BUG-112626	CALCITE-2444	Handle IN expressions when converting SqlNode to SQL
BUG-113293	CALCITE-2639	FilterReduceExpressionsRule causes ArithmeticException at execution time
BUG-113305	HIVE-20817	Reading Timestamp datatype via HiveServer2 gives errors
BUG-113376	HIVE-20835	Interaction between constraints and MV rewriting may create loop in Calcite planner
BUG-114525	HIVE-20937	Postgres jdbc query fail with "LIMIT must not be negative"
BUG-115132	CALCITE-1866 , CALCITE-2188 , CALCITE-2206 , CALCITE-2309 , CALCITE-2715	Backport multiple fixes for CALCITE JDBC pushdown
BUG-96841	HIVE-20510 , HIVE-20540	Bucketing: _bucket_number is marked as a constant, not a virtual column

Security

Hortonworks Bug ID	Apache JIRA	Summary
BUG-108625	RANGER-2213	Tomcat Security Vulnerability Alert. The version of the tomcat for ranger should upgrade to 7.0.90.
BUG-110103	RANGER-2204	Ranger Admin's admin log event for changing Audit Logging of a policy doesn't show the actual changes
BUG-110208	HIVE-20644	Avoid exposing sensitive information through a Hive Runtime exception
BUG-110275	N/A	Upgrade commons-collection4 to 4.1
BUG-110890	N/A	Fix a security issue createRDDFromFile - large object ser/deserialization between R / JVM
BUG-111435	RANGER-2208	Code improvement to fetch User/Group information and Service Config details
BUG-111654	HIVE-20651	JdbcStorageHandler password should be encrypted
BUG-112218	RANGER-2245	Exclude Jetty libraries
BUG-112239	HIVE-20731	keystore file in JdbcStorageHandler should be authorized
BUG-112568	N/A	[Druid] Fix Apache Commons Vulnerability
BUG-112620	KNOX-1429	HadoopAuthFilter avoid logging sensitive values
BUG-112890	RANGER-2263	Remove unnecessary explicit dependency for apache commons compress Jar in Ranger
BUG-113125	HIVE-20420	Provide a fallback authorizer when no other authorizer is in use
BUG-113214	RANGER-2269	Implement best coding practices for validating user input
BUG-113433	RANGER-2270	Restrict tag module access to unprivileged users
BUG-113786	ACCUMULO-3460	Monitor should not allow HTTP TRACE
BUG-113849	RANGER-2244	Tomcat Security Vulnerability Alert. The version of the tomcat for ranger should upgrade to 7.0.91 or later.

Hortonworks Bug ID	Apache JIRA	Summary
BUG-115501	N/A	Remove Livy Session API from Knox Service Definition

Stability

Hortonworks Bug ID	Apache JIRA	Summary
BUG-102391	YARN-8742	[UI2] Container logs on Application / Service pages on UI2 are not available many case, improve error messages in such cases.
BUG-102482	YARN-8628	[UI2] Few duplicated or inconsistent information displayed in UI2
BUG-107921	YARN-8520	Document best practice for user management
BUG-108561	HIVE-20349	Implement Retry Logic in HiveDruidSplit for Scan Queries
BUG-109074	TEZ-3980	ShuffleRunner: the wake loop needs to check for shutdown
BUG-109879	N/A	HiveServer2 Interactive fails to start
BUG-110149	N/A	HDFS: diskbalance test fails (disk balancer not executed due to threshold value)
BUG-110186	N/A	Hive upgrade with data deployed seems to be failing at ConvertTable
BUG-110400	HIVE-20511	REPL DUMP is leaking metastore connections
BUG-110668	HIVE-20513	Vectorization: Improve Fast Vector MapJoin Bytes Hash Tables
BUG-110962	N/A	Create SSD cache dir if it doesnt exist already
BUG-111119	RANGER-2186	Increment service-specific policy and tag versions after update transaction is committed
BUG-111239	HADOOP-15744	AbstractContractAppendTest fails against HDFS on HADOOP-15407 branch
BUG-111427	HADOOP-15748	S3 listing inconsistency can raise NPE in globber
BUG-111586	YARN-8758	Support getting PreemptionMessage when using AMRMClientAsync
BUG-111738	HADOOP-15694 , HADOOP-15704 , HADOOP-15714 , HADOOP-15715 , HADOOP-15719 , HADOOP-15740 , HADOOP-15744 , HADOOP-15745 , HADOOP-15769 , HADOOP-15773 , HADOOP-15778	ABFS related fixes
BUG-111834	HIVE-20095	Fix feature to push computation to jdbc external tables
BUG-111953	HIVE-20720	Add partition column option to JDBC handler
BUG-111973	HADOOP-15739	ABFS: remove unused maven dependencies and add used undeclared dependencies
BUG-111974	HADOOP-15723	ABFS: Ranger Support
BUG-111975	HADOOP-15801	ABFS: Fixing skipUserGroupMetadata in AzureBlobFileSystemStore
BUG-112055	HIVE-20682	Async query execution can potentially fail if shared sessionHive is closed by master thread.
BUG-112065	HADOOP-15795	Make HTTPS the default protocol for ABFS
BUG-112066	HADOOP-15792	typo in AzureBlobFileSystem.getIsNamespaceEnabled
BUG-112067	HADOOP-15793	ABFS: Skip unsupported test cases when non namespace enabled in ITestAzureBlobFileSystemAuthorization
BUG-112260	YARN-8853	[UI2] Application Attempts tab is not shown correctly when there are no attempts
BUG-112347	N/A	hbase:meta location in ZooKeeper set to OPENING which precludes Master from assigning it

Hortonworks Bug ID	Apache JIRA	Summary
BUG-112427	YARN-8868	Set HTTPOnly attribute to Cookie
BUG-112578	RANGER-2242	JiSQL utility is failing Oracle UDF
BUG-112686	RANGER-2250	Service configs fields are not showing for atlas service form page
BUG-112689	HADOOP-15809	ABFS: better exception handling when making getAccessToken call
BUG-112690	HADOOP-15825	ABFS: Enable some tests for namespace not enabled account using OAuth
BUG-112698	HADOOP-15851	Disable wildfly logs to the console
BUG-112839	HIVE-20767	Multiple project between join operators may affect join reordering using constraints
BUG-112909	YARN-8535	Fix DistributedShell unit tests
BUG-112945	HADOOP-15850	CopyCommitter#concatFileChunks should check that the blocks per chunk is not 0
BUG-113276	HADOOP-15823	ABFS: Stop requiring client ID and tenant ID for MSI
BUG-113277	HADOOP-15839	Review + update cloud store sensitive keys in hadoop.security.sensitive-config-keys
BUG-113464	HIVE-20834	Hive QueryResultCache entries keeping reference to SemanticAnalyzer from cached query
BUG-113617	HDFS-13941 , HDFS-14026	make storageId in BlockPoolTokenSecretManager.checkAccess optional
BUG-114131	HADOOP-15846	ABFS: fix mask related bugs in setAcl, modifyAclEntries and removeAclEntries.
BUG-114139	N/A	Backport Druid bloom filter caching Fix [Druid PR #6488]
BUG-114140	HIVE-20101	fix druid-bloom-filter thread-safety

Supportability

Hortonworks Bug ID	Apache JIRA	Summary
BUG-104177	N/A	Add Knox service definition for SMM
BUG-106354	N/A	Option to remove the server-name from HTTP-header response (HDP 3.1.x)
BUG-107096	N/A	HS2 and hivemetastore log do not log to configured log location
BUG-108450	HIVE-20130	Better logging for information schema synchronizer
BUG-109899	HIVE-20455	Log spew from security.authorization.PrivilegeSynchronizer.run
BUG-110039	HIVE-20476	CopyUtils used by REPL LOAD and EXPORT/IMPORT operations ignore distcp error.
BUG-110041	HIVE-20607	TxnHandler should use PreparedStatement to execute direct SQL queries.
BUG-111565	HIVE-20612	Create new join multi-key correlation flag for CBO
BUG-112136	RANGER-2251	Need to provide options for making java heap size memory configurable in Ranger services
BUG-112588	CALCITE-2622	RexFieldCollation toString method is not deterministic
BUG-112621	KNOX-1522	Add HA support for HadoopAuthProvider
BUG-112959	HIVE-20718	Add perf cli driver with constraints
BUG-114132	HADOOP-15812	ABFS: Improve AbfsRestOperationException format to ensure full msg can be displayed on console
BUG-113513	RANGER-2272	Ensure that case of resource-definition names and access-type names in Ranger policy is the same as in service-definition after successful validation

Upgrade

Hortonworks Bug ID	Apache JIRA	Summary
BUG-104299	N/A	GCS GA for HDP 3.1.0
BUG-110593	N/A	Add Surrogate Keys function to Hive
BUG-111794	AMBARI-24718	STS fails after start, after stack upgrade from 3.0.1 to 3.1.0
BUG-111947	N/A	HWC: Build HWC against Spark 2.3.2
BUG-112756	KNOX-1526	Upgrade Knox to Jetty 9.2.26
BUG-112937	RANGER-2260	Atlas servicedef version change patch should update atlas access type def for tag def also.
BUG-113133	RANGER-2264	Kafka default policies for new resources are not showing up in UI when upgrade is done from older version
BUG-113283	N/A	Upgrade Jetty version to jetty-9.4.12.v20180830 or 9.3.25.v20180904
BUG-113388	HIVE-20638	Upgrade version of Jetty to 9.3.25.v20180904

Usability

Hortonworks Bug ID	Apache JIRA	Summary
BUG-104558	N/A	Support More formats for Hive Druid Kafka ingestion
BUG-106410	HIVE-20193	cbInfo is not present in the explain plan json
BUG-106869	N/A	Enable workload management by default
BUG-108987	HIVE-20352	Vectorization: Support grouping function
BUG-109045	OOZIE-3326	Sqoop Action should support tez delegation tokens for hive-import
BUG-109077	HIVE-20367	Vectorization: Support streaming for PTF AVG, MAX, MIN, SUM
BUG-109279	HIVE-20311	add txn stats checks to some more paths
BUG-109362	PHOENIX-4822	The configuration "phoenix.query.dateFormatTimeZone" doesn't work on the client
BUG-109494	N/A	HWC: Support auto-creation of Hive table from Spark DataFrame schema
BUG-110103	RANGER-2204	Ranger Admin's admin log event for changing Audit Logging of a policy doesn't show the actual changes
BUG-110261	N/A	Name for the 'global' resource for temporary UDF should be more intuitive and informative from End user's viewpoint.
BUG-110593	N/A	Add Surrogate Keys function to Hive
BUG-110599	RANGER-2229	Perform graceful terminate with retries before doing forceful kill for usersync and tagsync
BUG-110666	LIVY-498	SparkR interpreter fails to execute statement with Windows CRLF on *nix machine
BUG-110709	HIVE-20734	Beeline: When beeline-site.xml is and hive CLI redirects to beeline, it should use the system username/dummy password instead of prompting for one
BUG-111022	RANGER-2224	'drop temporary function <udf>' command should be handled by 'global' resource and 'Temporary UDF Admin' permission.
BUG-111298	HIVE-20593	Load Data for partitioned ACID tables fails with bucketId out of range: -1
BUG-111377	N/A	Allow any udfs with 0 arguments or with constant arguments as part of default clause
BUG-111443	KNOX-1440	WEBHDFS: retry should not be attempted for datanode write failures
BUG-111591	HIVE-20590	Allow merge statement to have column schema
BUG-111747	RANGER-2207	Allow resources to appear in column mask policies without being visible in access policies
BUG-111763	HIVE-20703	Put dynamic sort partition optimization under cost based decision

Hortonworks Bug ID	Apache JIRA	Summary
BUG-111853	MAPREDUCE-7125	JobResourceUploader creates LocalFileSystem when it's not necessary
BUG-112136	RANGER-2251	Need to provide options for making java heap size memory configurable in Ranger services
BUG-112381	KNOX-1518	Large HDFS file downloads are incomplete when content is gzipped
BUG-112396	HIVE-20716	Set default value for hive.cbo.stats.correlated.multi.key.joins to true
BUG-112482	HIVE-20763	Add google cloud storage (gs) to the exim uri schema whitelist
BUG-112636	HIVE-20377 , HIVE-20481 , HIVE-20561 , HIVE-20639 , HIVE-20735	Kafka Storage Handler
BUG-112686	RANGER-2250	Service configs fields are not showing for atlas service form page
BUG-112788	N/A	As Idempotent write, describe broker configs, alter broker configs operations don't work at topic level but at cluster level in kafka 2.0, this needs to be documented
BUG-112824	HIVE-20477	OptimizedSql is not shown if the expression contains INs
BUG-112840	RANGER-2247	Ranger Plugin for HDFS throws StringIndexOutOfBoundsException exception when policy resource is "\"
BUG-112990	N/A	Unable to update kafka policies in ranger UI
BUG-113125	HIVE-20420	Provide a fallback authorizer when no other authorizer is in use
BUG-113391	RANGER-2275	Make db_setup retry delay configurable
BUG-113537	N/A	HWC: When writing a DataFrame to existing Hive table, match columns by name (if necessary)
BUG-113600	N/A	Fix Mapbox viz for superset
BUG-113833	HIVE-20734	Beeline: When beeline-site.xml is and hive CLI redirects to beeline, it should use the system username/dummy password instead of prompting for one
BUG-114061	N/A	HWC: NPE when closing tasks that retrieved no data
BUG-114336	CALCITE-2668	Support for left/right outer join in RelMdExpressionLineage
BUG-114966	ATLAS-2965	Duplicate entities are created when when same unique attributes are given but different unassigned guids
BUG-115076	HIVE-20978	"hive.jdbc.*" should add to sqlStdAuthSafeVarNameRegexes
BUG-81627	HIVE-18453	ACID: Add "CREATE TRANSACTIONAL TABLE" syntax to unify ACID ORC & Parquet support
BUG-92695	N/A	Replicate Truncate operation on ACID table
BUG-112993	RANGER-2262	Improvement of export to excel from report listing page for Oracle database.

Known Issues

Summary of known issues for this release.


Hortonworks Bug ID	Apache JIRA	Apache component	Summary
BUG-79238	N/A	Documentation, HBase, HDFS, Hive, MapReduce, Zookeeper	<p>Description of the problem or behavior</p> <p>SSL is deprecated and its use in production is not recommended. Use TLS.</p> <p>Workaround</p> <p>In Ambari: Use <code>ssl.enabled.protocols=TLSv1 TLSv1.1 TLSv1.2</code> and <code>security.server.disabled.protocols=SSL SSLv2 SSLv3</code>. For help configuring TLS for other components, contact customer support.</p> <p>Documentation will be provided in a future release.</p>
BUG-106494	N/A	Documentation, Hive	<p>Description of Problem</p> <p>When you partition a Hive column of type double, if the column value is 0.0, the actual partition directory is created as "0". An AIOB exception occurs.</p> <p>Associated error message</p> <pre data-bbox="867 684 1620 1598"> 2018-06-28T22:43:55,498 ERROR 441773a0-851c-4b25-9e47-729183946a26 main exec.StatsTask: Failed to run stats task org.apache.hadoop.hive.ql.metadata.HiveException: java.lang.IndexOutOfBoundsException: Index: 8, Size: 8 at org.apache.hadoop.hive.ql.metadata.Hive.setPartit ~hive-exec-4.0.0-SNAPSHOT.jar:4.0.0- SNAPSHOT at org.apache.hadoop.hive.ql.stats.ColStatsProcesso ~hive-exec-4.0.0-SNAPSHOT.jar:4.0.0- SNAPSHOT at org.apache.hadoop.hive.ql.stats.ColStatsProcesso ~hive-exec-4.0.0-SNAPSHOT.jar:4.0.0- SNAPSHOT at org.apache.hadoop.hive.ql.exec.StatsTask.execut hive-exec-4.0.0-SNAPSHOT.jar:4.0.0- SNAPSHOT at org.apache.hadoop.hive.ql.exec.Task.executeTask hive-exec-4.0.0-SNAPSHOT.jar:4.0.0- SNAPSHOT at org.apache.hadoop.hive.ql.exec.TaskRunner.runSec hive-exec-4.0.0-SNAPSHOT.jar:4.0.0- SNAPSHOT at org.apache.hadoop.hive.ql.Driver.launchTask(Driv hive-exec-4.0.0-SNAPSHOT.jar:4.0.0- SNAPSHOT at org.apache.hadoop.hive.ql.Driver.execute(Driver hive-exec-4.0.0-SNAPSHOT.jar:4.0.0- SNAPSHOT at org.apache.hadoop.hive.ql.Driver.run </pre> <p>Workaround</p> <p>Do not partition columns of type double.</p>

Hortonworks Bug ID	Apache JIRA	Apache component	Summary
BUG-106379	N/A	Documentation, Hive	<p>Description of the Problem</p> <p>The upgrade process fails to perform necessary compaction of ACID tables and can cause permanent data loss.</p> <p>Workaround</p> <p>If you have ACID tables in your Hive metastore, enable ACID operations in Ambari or set Hive configuration properties to enable ACID. If ACID operations are disabled, the upgrade process does not convert ACID tables. This causes permanent loss of data; you cannot recover data in your ACID tables later.</p>
BUG-106286	N/A	Documentation, Hive	<p>Description of the Problem</p> <p>The upgrade process might fail to make a backup of the Hive metastore, which is critically important.</p> <p>Workaround</p> <p>Manually make a manual backup of your Hive metastore database before upgrading. Making a backup is especially important if you did not use Ambari to install Hive and create the metastore database, but highly recommended in all cases. Ambari might not have the necessary permissions to perform the backup automatically. The upgrade can succeed even if the backup fails, so having a backup is critically important.</p>
BUG-101082	N/A	Documentation, Hive	<p>Description of the problem or behavior</p> <p>When running Beeline in batch mode, queries killed by the Workload Management process can on rare occasions mistakenly return success on the command line.</p> <p>Workaround</p> <p>There is currently no workaround.</p>
BUG-103495	HBASE-20634 , HBASE-20680 , HBASE-20700	HBase	<p>Description of the problem or behavior</p> <p>Because the region assignment is refactored in HBase, there are unclear issues that may affect the stability of this feature. If you rely on RegionServer Groups feature, you are recommended to wait until a future HDP 3.x release, which will return the stability of this features as it was available in HBase 1.x/HDP 2.x releases.</p> <p>Workaround</p> <p>There is currently no workaround.</p>
BUG-98727	N/A	HBase	<p>Description of the problem or behavior</p> <p>Because the region assignment is refactored in HBase, there are unclear issues that may affect the stability of this feature. If you rely on Region replication feature, you are recommended to wait until a future HDP 3.x release, which will return the stability of this features as it was available in HBase 1.x/HDP 2.x releases.</p> <p>Workaround</p> <p>There is currently no workaround.</p>
BUG-105983	N/A	HBase	<p>Description of the problem or behavior</p> <p>An HBase service (Master or RegionServer) stops participating with the rest of the HBase cluster.</p> <p>Associated error message</p> <p>The service's log contains stack traces that contain "Kerberos principal name does NOT have the expected hostname part..."</p> <p>Workaround</p> <p>Retrying the connection solves the problem.</p>

Hortonworks Bug ID	Apache JIRA	Apache component	Summary
BUG-96402	HIVE-18687	Hive	<p>Description of the problem or behavior</p> <p>When HiveServer2 is running in HA (high-availability) mode in HDP 3.0.0, resource plans are loaded in-memory by all HiveServer2 instances. If a client makes changes to a resource plan, the changes are reflected (pushed) only in the HiveServer2 to which the client is connected.</p> <p>Workaround</p> <p>In order for the resource plan changes to be reflected on all HiveServer2 instances, all HiveServer2 instances has to be restarted so that they can reload the resource plan from metastore.</p>
BUG-88614	N/A	Hive	<p>Description of the problem or behavior</p> <p>RDMBS schema for Hive metastore contains an index HL_TXNID_INDEX defined as</p> <pre>CREATE INDEX HL_TXNID_INDEX ON HIVE_LOCKS USING hash (HL_TXNID);</pre> <p>Hash indexes are not recommended by PostgreSQL. For more information, see https://www.postgresql.org/docs/9.4/static/indexes-types.html</p> <p>Workaround</p> <p>It's recommended that this index is changed to type BTREE.</p>
BUG-120655	N/A	Hive	<p>Description of the problem or behavior</p> <p>The Hive Warehouse Connector does not support non-ORC file formats for writes.</p> <p>Workaround</p> <p>There is currently no workaround.</p>
BUG-120579	N/A	Hive, HDFS	<p>Description of the problem or behavior</p> <p>If you need to use HDFS Transparent Data Encryption (TDE) with HDP 3.1.0, contact Hortonworks Support.</p>
BUG-60904	KNOX-823	Knox	<p>Description of the problem or behavior</p> <p>When Ambari is being proxied by Apache Knox, the QuickLinks are not rewritten to go back through the gateway. 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. Future release will make these reflect the gateway appropriately.</p> <p>Workaround</p> <p>There is currently no workaround.</p>
BUG-107399	N/A	Knox	<p>Description of the problem or behavior</p> <p>After upgrade from previous HDP versions, certain topology deployments may return a 503 error. This includes, but may not be limited to, knoxsso.xml for the KnoxSSO enabled services.</p> <p>Workaround</p> <p>When this is encountered, a minor change through Ambari (whitespace even) to the knoxsso topology (or any other with this issue) and restart of the Knox gateway server should eliminate the issue.</p>

Hortonworks Bug ID	Apache JIRA	Apache component	Summary
BUG-110463	KNOX-1434	Knox	<p>Description of the problem or behavior</p> <p>Visiting Knox Admin UI in any browser (Firefox / Chrome) sets the HTTP Strict Transport Security (HSTS) header for the host where Knox is running. Any subsequent request to other service on the same host (e.g. Graphana, Ranger etc.) over HTTP would get redirected to HTTPS due to this header.</p> <p>Please note that, this HSTS header is disabled in all Knox topologies by default.</p> <p>For more information, see https://knox.apache.org/books/knox-1-1-0/user-guide.html#HTTP+Strict+Transport+Security</p> <p>Impact</p> <p>All the non-SSL requests to other services get redirected automatically to HTTPS and would result in SSL errors like: SSL_ERROR_RX_RECORD_TOO_LONG or some other error.</p> <p>Workaround</p> <p>Use the manager.xml topology and remove the setting from the WebAppSec provider. You can do this using the Knox Admin UI. After you have removed the setting, close your browser or clear the cookies.</p>
BUG-106266	OOZIE-2769 , OOZIE-3085 , OOZIE-3156 , OOZIE-3183	Oozie	<p>Description of the problem or behavior</p> <p>When check() method of SshActionExecutor gets invoked, Oozie will execute the command "ssh <host-ip> ps -p <pid>" to determine whether the SSH action completes or not. However if the connection to the host fails during the action status check, the command will return with an error code, but the action status will be determined as OK, which may not be correct.</p> <p>Associated error message</p> <p>SSH command exits with the exit status of the remote command or with 255 if an error occurred.</p> <p>Workaround</p> <p>Retrying the connection solves the problem.</p>
BUG-95909	RANGER-1960	Ranger	<p>Description of problem or behavior</p> <p>Delete snapshot operation fails even if the user has Administrator privilege because the namespace is not considered in the Authorization flow for HBase Ranger plugin.</p> <p>Associated error message</p> <p>ERROR: org.apache.hadoop.hbase.security.AccessDeniedException: Insufficient permissions for user '&lt;username&gt;'; (action=admin)</p> <p>Workaround</p> <p>For the delete snapshot operation to succeed, you need to be system-wide Administrator privileges.</p>
BUG-89714	N/A	Ranger	<p>Description of the problem or behavior</p> <p>Sudden increase in Login Session audit events from Ranger Usersync and Ranger Tagsync.</p> <p>Associated error message</p> <p>If policy storage DB size increases suddenly, then periodically backup and purge 'x_auth_sess' table periodically.</p> <p>Workaround</p> <p>Take a backup of the policy DB store and purge 'x_auth_sess' table from Ranger DB schema.</p>

Hortonworks Bug ID	Apache JIRA	Apache component	Summary
BUG-101227	N/A	Spark	<p>Description of the problem or behavior</p> <p>When Spark Thriftserver has to run several queries concurrently, some of them can fail with a timeout exception when performing broadcast join.</p> <p>Associated error message</p> <pre>Caused by: java.util.concurrent.TimeoutException: Futures timed out after [300 seconds] at scala.concurrent.impl.Promise \$DefaultPromise.ready(Promise.scala:219) at scala.concurrent.impl.Promise \$DefaultPromise.result(Promise.scala:223) at scala.concurrent.Await\$\$anonfun \$result\$1.apply(package.scala:107) at scala.concurrent.BlockContext \$DefaultBlockContext \$.blockOn(BlockContext.scala:53) at scala.concurrent.Await \$.result(package.scala:107) at org.apache.spark.sql.execution.joins.BroadcastH</pre> <p>Workaround</p> <p>You can resolve this issue by increasing the spark.sql.broadcastTimeout value.</p>
BUG-109979	N/A	Spark	<p>Description of the problem or behavior</p> <p>YARN NodeManagers fail to start after a Spark patch upgrade due to YarnShuffleService CNF.</p> <p>Workaround</p> <p>To resolve this problem you must:</p> <p>Replace "{{spark2_version}}" with "\${hdp.version}" in "yarn.nodemanager.aux-services.spark2_shuffle.classpath" property value. For example, old value "{{stack_root}}/{{spark2_version}}/spark2/aux/*" -> new value "{{stack_root}}/\${hdp.version}/spark2/aux/*"</p>

Hortonworks Bug ID	Apache JIRA	Apache component	Summary
BUG-65977	SPARK-14922	Spark	<p>Description of the problem or behavior</p> <p>Since Spark 2.0.0, `DROP PARTITION BY RANGE` is not supported grammatically. In other words, only `=` is supported while `<`, `>`, `<=`, `>=` aren't.</p> <p>Associated error message</p> <pre>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) -----^^^</pre> <p>Workaround</p> <p>To drop partition, use the exact match with `=`.</p> <pre>scala> sql("alter table t drop partition (b=0) ").show</pre>
BUG-114383	N/A	Storm	<p>Description of the problem or behavior</p> <p>Submitting a topology to Storm fails. You see an error when you submit a topology to Storm.</p> <p>Associated error message</p> <p>The following error message is displayed when submitting a topology with the stack trace org.apache.storm.hack:</p> <pre>Exception in thread "main" java.lang.IllegalArgumentException at org.apache.storm.hack.shade.org.objectweb.asm. Source) at org.apache.storm.hack.shade.org.objectweb.asm. Source) at org.apache.storm.hack.shade.org.objectweb.asm. Source)</pre> <p>Workaround</p> <p>Find if `client.jartransformer.class` is present in the Storm configuration using the Ambari user interface. If the config is present, please set to ` ` and restart Storm service to take effect.</p> <p> Note: The value that you have to provide is one whitespace, Ambari restricts the configuration to have a value, but does not allow multiple whitespaces.</p>

Hortonworks Bug ID	Apache JIRA	Apache component	Summary
BUG-106917	N/A	Sqoop	<p>Description of the problem or behavior</p> <p>In HDP 3, managed Hive tables must be transactional (hive.strict.managed.tables=true). Transactional tables with Parquet format are not supported by Hive. Hive imports with --as-parquetfile must use external tables by specifying --external-table-dir.</p> <p>Associated error message</p> <pre>Table db.table failed strict managed table checks due to the following reason: Table is marked as a managed table but is not transactional.</pre> <p>Workaround</p> <p>When using --hive-import with --as-parquetfile, users must also provide --external-table-dir with a fully qualified location of the table:</p> <pre>sqoop import ... --hive-import --as-parquetfile --external-table-dir hdfs:///path/to/table</pre>
BUG-102672	N/A	Sqoop	<p>Description of the problem or behavior</p> <p>In HDP 3, managed Hive tables must be transactional (hive.strict.managed.tables=true). Writing transactional table with HCatalog is not supported by Hive. This leads to errors during HCatalog Sqoop imports if the specified Hive table does not exist or is not external.</p> <p>Associated error message</p> <p>Store into a transactional table db.table from Pig/Mapreduce is not supported</p> <p>Workaround</p> <p>Before running the HCatalog import with Sqoop, the user must create the external table in Hive. The --create-hcatalog-table does not support creating external tables.</p>
BUG-109607	N/A	YARN	<p>Description of the problem or behavior</p> <p>With wire encryption enabled with containerized Spark on YARN with Docker, Spark submit fails in "cluster" deployment mode. Spark submit in "client" deployment mode works successfully.</p> <p>Associated error message</p> <p>Store into a transactional table db.table from Pig/Mapreduce is not supported.</p> <p>Workaround</p> <p>There is currently no workaround.</p>

Hortonworks Bug ID	Apache JIRA	Apache component	Summary
BUG-110192	N/A	YARN	<p>Description of the problem or behavior</p> <p>When YARN is installed and configured with KNOX SSO alone, Application Timeline Server web endpoint blocks remote REST calls from YARN UI and displays a 401 Unauthorized error.</p> <p>Associated error message</p> <p>401 Unauthorized error.</p> <p>Workaround</p> <p>Administrator needs to configure Knox authentication handler for Timeline Server and existing hadoop level configuration.</p> <p>Administrator needs to tune the following cluster specific configurations. Values for the last two property is in the hadoop.authentication.* properties file.</p> <pre> <property> <name>yarn.timeline-service.http- authentication.type</name> <value>org.apache.hadoop.security.authentication value> </property> <property> <name>yarn.timeline-service.http- authentication.authentication.provider.url</ name> <value>https://ctr- e138-1518143905142-455650-01-000002.hwx.site:44 gateway/knoxssso/api/v1/webssso/< value> </property> <property> <name>yarn.timeline-service.http- authentication.public.key.pem</name> <value>public.key.pem</value> </property> </pre>
RMP-11408	ZEPPELIN-2170	Zeppelin	<p>Description of the problem or behavior</p> <p>Zeppelin does not show all WARN messages thrown by spark-shell at the Zeppelin's notebook level.</p> <p>Workaround</p> <p>There is currently no workaround for this.</p>

Legal Information

Hortonworks Data Platform 3.1.0

Copyright information for Hortonworks Data Platform components may be found within the documentation accompanying each component in a particular release.

Hortonworks Data Platform incorporates software from various open source projects released primarily under the Apache Software License 2.0 (“ASLv2”). Other software included may be released under the terms of alternative ASLv2 compatible open source licenses. Please review the license and notice files accompanying the software for additional licensing information.

Please visit the Hortonworks Data Platform page for more information on Hortonworks Data Platform technology. For more information on Hortonworks support services, please visit either the Support or Sales page. Feel free to contact us directly to discuss your specific needs.

Hortonworks reserves the right to change any products described herein at any time, and without notice. Hortonworks assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by Hortonworks.

Trademark: Hortonworks Data Platform is/are trademark/s of Hortonworks, Inc. in the United States and other countries. All other trademarks are the property of their respective owners.

Disclaimer: EXCEPT AS EXPRESSLY PROVIDED IN A WRITTEN AGREEMENT WITH HORTONWORKS, HORTONWORKS DOES NOT MAKE OR GIVE ANY REPRESENTATION, WARRANTY, OR COVENANT OF ANY KIND, WHETHER EXPRESS OR IMPLIED, IN CONNECTION WITH HORTONWORKS TECHNOLOGY OR RELATED SUPPORT PROVIDED IN CONNECTION THEREWITH. HORTONWORKS DOES NOT WARRANT THAT Hortonworks Data Platform WILL OPERATE UNINTERRUPTED OR THAT IT WILL BE FREE FROM DEFECTS OR ERRORS, THAT IT WILL PROTECT YOUR DATA FROM LOSS, CORRUPTION OR UNAVAILABILITY, OR THAT IT WILL MEET ALL OF CUSTOMER'S BUSINESS REQUIREMENTS. WITHOUT LIMITING THE FOREGOING, AND TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, HORTONWORKS EXPRESSLY DISCLAIMS ANY AND ALL IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, QUALITY, NON-INFRINGEMENT, TITLE, AND FITNESS FOR A PARTICULAR PURPOSE AND ANY REPRESENTATION, WARRANTY, OR COVENANT BASED ON COURSE OF DEALING OR USAGE IN TRADE.