Hortonworks Data Platform

Date of Publish: 2018-07-12

Contents

HDP 3.0.0 Release Notes

This document provides you with the latest information about the Hortonworks Data Platform (HDP) 3.0.0 release and its product documentation.

Component Versions

This section lists the official Apache versions of all the HDP 3.0.0 components. All components listed here are official Apache releases of the most recent stable versions available.

The Hortonworks approach is to provide patches only when necessary, to ensure the interoperability of components. Unless you are explicitly directed by Hortonworks Support to take a patch update, each of the HDP components should remain at the following package version levels, to ensure a certified and supported copy of HDP 3.0.0.

Official Apache versions for HDP 3.0.0:

- Apache Accumulo 1.7.0
- Apache Atlas 1.0.0
- Apache Calcite 1.16.0
- Apache DataFu 1.3.0
- Apache Druid 0.12.1 (incubating)
- Apache Hadoop 3.1.0
- Apache HBase 2.0.0
- Apache Hive 3.1.0
- Apache Kafka 1.0.1
- 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 1.1.0
- Apache Spark 2.3.1
- 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

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

Additional component versions:

- Apache Superset 0.23.3 (technical preview)

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

New Features

This section highlights new features in HDP 3.0.0.

Apache Components	Feature
Atlas	Core Capabilities • Show attributes for time-bound classification or business catalog mapping (ATLAS-2457) Support Business Terms and Categories • Migrate Atlas data in Titan graph DB to JanusGraph DB • Atlas Hbase hook to capture metadata and lineage • Tag propagation from object to child object or derivative asset (ATLAS-1821) • Address Storm Atlas hook compatibility with Storm 1.2 Security • Metadata Security: Authorization based on Classification or Business Catalog mapping (ATLAS-2459)
Druid	Integration • Kafka-Druid ingest. You can now map a Kafka topic to a Druid table. The events will be automatically ingested and available for querying in near real-time.
HDFS	Core Capabilities • Balance utilization of disks (varied capacities) inside a DataNode (HDFS-1312) • Reduce Storage Overhead of HDFS with Directory Level Reed Solomon Erasure Coding Encoding (HDFS-7285) • Support 2 Standby NameNodes for NameNode High Availability • NFS Gateway needs to work in front of ViewFS (file access to the unified namespace) • Expose Encrypted zones and Erasure coded zones via WebHDFS API (HDFS-11394, HDFS-13512) • Hive Support on Erasure Coded Directories (HDFS-7285) • Cloud Testing for HDFS: Cloud Failure Modes/Availability • Cloud: Connect/reattach to Elastic Block Volume to use centralized block storage for better TCO (vs. local disks)
HBase	Core Capabilities • Procedure V2. You can use Procedure V2 or procv2, which is an updated framework for executing multi-step HBase administrative operations when there is a failure, you can use this framework to implement all the master operations using procv2 to remove the need for tools like hbck in the future. Use procv2 to create, modify and delete tables. Other systems such as new AssignmentManager is implemented using procv2. • Fully off-heap read/write path. When you write data into HBase through Put operation, the cell objects do not enter JVM heap until the data is flushed to disk in an HFile. This helps to reduce total heap usage of a RegionServer and it copies less data making it more efficient. • Use of Netty for RPC layer and Async API. This replaces the old Java NIO RPC server with a Netty RPC server. Netty provides you the ability to easily provide an Asynchronous Java client API. • In-memory compactions. Periodic reorganization of the data in the Memstore can result in a reduction of overall I/O, that is data written and accessed from HDFS. The net performance increases when we keep more data in memory for a longer period of time. • Better dependency management. HBase now internally shades commonly-incompatible dependencies to prevent issues for downstream users. You can use shaded client jars that will reduce the burden on the existing applications. • Coprocessor and Observer API rewrite. Minor changes made to the API to remove ambiguous, misleading, and dangerous calls.

Apache Components	Feature
Hive	Core Capabilities • Workload management for LLAP. You can now run LLAP, in a multi-tenant environment without worrying about resource competition. • ACID v2 and ACID on by default. ACID v2 has performance improvements in both storage format and execution engine, there is either equal or better performance when compared to non-ACID tables. ACID on is enabled by default to allow full support for data updates. • Materialized view navigation. Hive's query engine now supports materialized view. The query engine will automatically use materialized view when they are available to speed up your queries. • Information schema. Hive now exposes the metadata of the database (tables, columns etc.) via Hive SQL interface directly. Integration • Hive Warehouse Connector for Spark. Hive Warehouse Connector allows you to connect Spark application with Hive data warehouses. The connector automatically handles ACID tables. • JDBC storage connector. You can now map any JDBC database's tables into Hive and query those tables in conjunction with other tables
Kafka	Core Capabilities • Kafka has been upgraded from 1.0.0 to 1.0.1 with some critical bug fixes. • A new feature was added to capture producer and topic partition level metrics. Security • Cache lastEntry in TimeIndex to avoid unnecessary disk access (KAFKA-6172) • AbstractIndex should cache index file to avoid unnecessary disk access during resize() (KAFKA-6175) • SSLTransportLayer should keep reading from socket until either the buffer is full or the socket has no more data (KAFKA-6258)
Knox	Usability • Admin UI along with service discovery and topology generation feature for simplifying and accelerating Knox configuration Security • Added SSO support for Zeppelin, YARN, MR2, HDFS, and Oozie • Added Knox Proxy support for YARN, Oozie, SHS (Spark History Server), HDFS, MR2, Livy, and SmartSense
Oozie	Core Capabilities • Upgrade Oozie baseline to 4.3.1 from 4.2 • Disable Oozie Hive Action Integration • Oozie support for Spark2

Apache Components	Feature
Phoenix	**Core Capabilities** • Query log. This is a new system table "SYSTEM.LOG" that captures information about queries that are being run against the cluster (client-driven). • Column encoding. This is new to HDP. You can use a custom encoding scheme of data in the HBase table to reduce the amount of space taken. This increases the performance due to less data to read and thereby reduces the storage. The performance gain is 30% and above for the sparse tables. • Supports GRANT and REVOKE commands. This provides automatic changes to indexes ACLs, if access changed for data table or view. • Support for sampling tables. • Supports atomic update (ON DUPLICATE KEY). • Supports snapshot scanners for MR-based queries. **Integration** • HBase 2.0 support. • Python driver for Phoenix Query Server. This Provides Python DB 2.0 API implementation. • Hive 3.0 support for Phoenix. This provides updated phoenix-hive StorageHandler for the new Hive version. • Spark 2.3 support for Phoenix. This provides updated phoenix-spark driver for new the Spark version. **Security** • Hardening of both the secondary indexes that includes Local and Global.
Ranger	**Security** • Time Bound Authorization Policies • Hive UDF execution authorization • Hive workload management authorization • RangerKafkaAuthorizer to support new operations and resources added in Kafka 1.0 • Read-only Ranger user roles for auditing purposes • Auditing for usersync operations • HDFS Federation support • Support metadata authorization changes in Atlas 1.0 • Add ability to specify passwords for admin accounts during ranger install • Add consolidated db schema script for all supported DB flavor **Ease of Use** • Show actual hive query in Ranger Audit UI • Group policies using Labels • Install and turn on Ranger and Atlas by default in HDP3
Spark	**Core Capabilities** • Spark 2.3.1 GA on HDP 3.0 • Structured Streaming support for ORC • Enable Security and ACLs in History Server • Support running Spark jobs in a Docker Container • Upgrade Spark/Zeppelin/Livy from HDP 2.6 to HDP 3.0 • Cloud: Spark testing with S3Guard/S3A Committers • Certification for the Staging Committer with Spark • Integrate with new Metastore Catalog feature • Beeline support for Spark Thrift Server **Integration** • Support per notebook interpreter configuration • Livy to support ACLs • Knox to proxy Spark History Server UI • Structured Streaming support for Hive Streaming library • Transparent write to Hive warehouse

Apache Components	Feature
Storm	Core Capabilities • Storm has been upgraded from 1.1.0 to 1.2.1. Storm 1.2.1 now supports all HDP 3.0 components including Hadoop/HDFS 3.0, HBase 2.0 and Hive 3.

YARN	Core Capabilities
	
	• Support intra-queue preemption to support balancing between apps from different users and priorities in the same queue
	• Support async-scheduling (vs. per node-heartbeat) for better response time in a large YARN cluster
	• Support generalized resource-placement in YARN: Affinity/Anti-affinity
	• Application Priority scheduling support in Capacity Scheduler
	• Expose a framework for more powerful apps to queues mappings (YARN-8016, YARN-3635)
	• Support Application timeout feature in YARN
	• Support GPU scheduling/isolation on YARN
	• Support Docker Containers running on YARN
	• Support assemblies on YARN (YARN-6613)
	• YARN Service framework - Slider functionality in YARN
	• Support a simplified services REST API on Slider / YARN
	• Simplified discovery of services via DNS
	• Support HDP on (YARN + Slider + Docker / YCloud) via CloudBreak integration
	• NodeManager should support automatic restart of service containers
	• Support auto-spawning of admin configured system-services
	• Migrate LLAP-on-Slider to LLAP-on-YARN-Service-Framework
	• Support dockerized Spark jobs on YARN
	
	Ease of Use
	
	• Timeline Service V2
	• Need resource manager web UI authorization control (users can only see their own jobs)
	• Support better classpath isolation for users, remove guava conflicts from user runtime
	• A more user-friendly and developer-friendly YARN web UI
	• YARN/MapReduce integration with SSO/Proxy (via Knox)
	
	Enterprise Readiness
	
	• Enable Capacity Scheduler preemption by default
	• Cgroup support for YARN in a non-secure cluster, and so LinuxContainerExecutor always on by default
	• Enable cgroups and CPU scheduling for YARN containers by default
	  Support for deleting queues without requiring a RM restart
	• Enhancements to STOP queue handling
	  Support Side by Side HDFS tarball based install of multiple Spark auxiliary services in YARN (YARN-1151)
	• Create a log aggregation tool (HAR files) to reduce NameNode load (YARN-4086)
	• YARN queue ACL support when doAs=false
	• Provide an API in YARN to get queue mapping result before application submitted

Zeppelin	Core Capabilities • Change Zeppelin UI across the board to not display stack traces • Zeppelin should have the option for user name case conversion (ZEPPELIN-3312) • Update to 0.8 release of Zeppelin Ease of Use • Zeppelin UI: Knox SSO for Quicklinked Web UIs

Deprecation Notices

This section points out any technology from previous releases that have been deprecated, moved, or removed from this release. Use this section 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

The following components are removed in this HDP release:

Component or Capability	Marked Deprecated as of	Status	Comments
Apache Falcon	HDP 2.6.0	Removed	Consider Data Lifecycle Manager as an alternative for Falcon use cases.
Apache Flume	HDP 2.6.0	Removed	Consider Hortonworks DataFlow as an alternative for Flume use cases.
Apache Mahout	HDP 2.6.0	Removed	Consider Apache Spark as an alternative depending on the workload.
Apache Slider	HDP 2.6.0	Removed	Apache Slider functionality is absorbed by Apache YARN.
Cascading	HDP 2.6.0	Removed	No replacement.
Hue	HDP 2.6.0	Removed	Consider Ambari Views as an alternative.

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

Unsupported Features

This section lists features that are not officially supported by HDP.

Although the features listed in the Technial Preview and Community Features sections exist within HDP, Hortonworks does not currently support them.

Technical Preview Features

The following features are available within HDP 3.0.0 but are not ready for production deployment. Hortonworks encourages you to explore these technical preview features in non-production environments and provide feedback on your experiences through the Hortonworks Community Forums.

Table 1: Technical Previews

Component	Feature
Spark	Native Atlas connector or hook for Spark to capture metadata and lineage
YARN	Introduced in a previous release: • Cloud Service Metrics for YARN (YARN-4488) • Cloud Service Health for YARN • In-place upgrade of long-running applications in YARN
Superset	Apache Superset is a modern, enterprise-ready business intelligence web application

Upgrading to HDP 3.0.0

Instructions to upgrade from a previous version of HDP.

If you already have HDP 2.6.x installed, upgrading your cluster to HDP 3.0.0 means:

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

Before you begin

Consider this before you start your upgrade.

Ensure that you know which HDP components you need to upgrade at your installation.

Upgrade

Upgrade information and instructions.

You can upgrade your cluster using Ambari, see the *Ambari Upgrade Guide* for instructions .

Related Information
Ambari Upgrade Guide
Obtain the HDP Repositories
Register and Install HDP Version

Behavioral Changes

Behavioral changes denote a marked change in behavior from the previously released version to this version of software. In HDP 3.0.0, behavioral changes affect the following Hadoop components.

Table 2: Behavioral Changes

Hortonworks Bug ID	Apache Component	Apache JIRA	Summary	Details
RMP-10355	HDFS	HDFS-13081	Now a secure DataNode can be started without privileged ports by setting SASL for RPC and SSL for HTTP.	Scenario:New blacklist based RPC encryption negotiation is introduced to allow flexible deployment of RPC encryption. Previous behavior: In a secure cluster, Datanode required privileged RPC/HTTP ports even if SASL QoP is configured for RPC and SSL is configured for HTTP, respectively. SASL QoP encryption negotiation supports WhitelistBasedTrustedChannelResolver: only client/server in the whitelist file are trusted (unencrypted). Everything else is encrypted by default. New behavior: DFS-13081: In secure cluster, DataNode can be configured to use non-privileged RPC port, secured by SASL QoP and non-privileged HTTP port, secured by SSL. HDFS-13060: SASL QoP encryption negotiation now supports BlacklistBasedTrustedChannelResolver: only client/server NOT in the blacklist file are trusted(unencrypted). Workaround/Expected Customer Action: Configure the secure Dataode by setting SASL (QOP) for RPC and SSL for http without privilege ports. Remove HDFS_DATANODE_SECURE_USER from hadoop-env to remove the privileged port. Configure RPC wire encryption using blacklist or whitelist approach as required.
RMP-9793	Oozie	N/A	Oozie will not run Hive actions anymore. Please use Hive2 actions instead.	Scenario: Hive CLI is removed in HDP 3.0.0 because it is a security risk. You can only use HiveServer2 actions now . Oozie should make this explicit by disallowing Hive action. Previous behavior: You do use a Hive action. New behavior: Hive CLI is removed, you can only use HiveServer2 actions. Workaround/Expected Customer Action: Modify old workflows to use Hive2 actions.
BUG-104280	Oozie	Oozie-1624	Oozie can be configured to do not use dependencies from share lib which matches for the defined pattern.	Scenario: It is possible to have different versions of a specific dependency on share lib in different directories. When it causes incompatibility failures user can exclude the unwanted ones. Previous behavior: You can use a Hive action. New behavior: In your job.properties file, you can now provide a pattern for different actions. For example, oozie.action.sharelib.for.hive2.exclude=./unwanted-1.0.0.jar* means that any file that matches will not present on distributed cache so will not be available in the started MapReduce job when user runs their Hive2 action. Workaround/Expected Customer Action: Exclude pattern can be defined for more action.

Hortonworks Bug ID	Apache Component	Apache JIRA	Summary	Details
BUG-95538	HBase	N/A	Delete markers no longer incorrectly masks newer data.	Scenario: copyTable command with allCells option shows deleted row from original table Previous behavior: A delete marker would eclipse all writes to that same row until a compaction occurs, at which time the compaction runs over that row. New behavior: A delete marker will only eclipse writes to that same row which are at a timestamp older than the delete marker's timestamp. Workaround/Expected Customer Action: Restructure applications to account for this change.
RMP-11005	Ranger, Atlas	Ranger-2019	Ranger's Atlas service definition has been updated.	Scenario: In Ranger, the Atlas service definition has been updated. When looking at the home page of Ranger (Resource Based Policies / Service Manager), there will appear to be two Atlas services. Previous behavior: There was one Atlas service labeled "Atlas". New Behavior: The old service definition is labeled "Atlas-V1", and the new service definition is labeled "Atlas".

Apache Patch Information

The following sections list patches in each HDP 3.0.0 component beyond what was fixed in the base version of the Apache component.

Accumulo

This release provides Accumulo 1.7.0 with no additional Apache patches.

Atlas

This release provides Atlas 1.0.0 and the listed 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 with no additional Apache patches.

DataFu

This release provides DataFu 1.3.0 with no additional Apache patches.

Hadoop

This release provides Hadoop Common 3.1.0 and the listed 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 Takanobu 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 datandoes 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 Yiqun 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 2.0.0 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 cod

Hive

This release provides Hive 3.1.0 with no additional Apache patches.

Kafka

This release provides Kafka 1.0.1 with 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 with no Apache patches.

Oozie

This release provides Oozie 4.3.1 and the listed 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 hcat.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 hcat 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 hcat 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 with no additional Apache patches.

Pig

This release provides Pig 0.16.0 and the listed 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 1.1.0 and the listed 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 reource
- 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 2.3.1 and the listed Apache patches.

- SPARK-24495: SortMergeJoin with duplicate keys wrong results.
- SPARK-207: Remove hardcode 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.

Livy

This release provides Livy 0.5.0 and no additional patches.

Sqoop

This release provides Sqoop 1.4.7 with no additional Apache patches.

Storm

This release provides Storm 1.2.1 and the listed 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 forregistered 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 with no additional Apache patches.

Zeppelin

This release provides Zeppelin 0.8.0 with no additional Apache patches.

ZooKeeper

This release provides ZooKeeper 3.4.6 and the listed Apache.

- 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

This section covers all Common Vulnerabilities and Exposures (CVE) that are addressed in this release.

No CVE fixes apply to HDP 3.0.0.

Fixed Issues

Represents selected issues that were previously logged through 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 the Hortonworks Quality Engineering team.

Incorrect Results

Hortonworks Bug ID	Apache JIRA	Summary
BUG-91853	N/A	Select count(*) returns empty (null) instead of 0 If the filter does not match any row
BUG-93880	ATLAS-2462	Sqoop import for all tables throws NPE for no table provided in command.
BUG-94623	HIVE-12505	Spark job completes successfully but there is an HDFS disk quota full error reported in driver log as below
BUG-96151	RANGER-2091	Ranger returns 302 instead of 401 if URL doesn't match policymgr_external_url

Other

Hortonworks Bug ID	Apache JIRA	Summary
BUG-85267	N/A	hadoop-env.sh is sourced multiple times during startup with HDFS services
BUG-90621	N/A	Druid need to declare logrotate pkg as a dependency
BUG-95261	HADOOP-15185	Update ADLS connector to use the current version of ADLS SDK

Performance

Hortonworks Bug ID	Apache JIRA	Summary
BUG-104547	HIVE-10815	Add HIVE-10815 changes in spark2
BUG-73569	N/A	TPCDS Query 28 - query scans the same big-table ~6x times
BUG-79774	HIVE-17116	Vectorization: Add infrastructure for vectorization of ROW__ID struct
BUG-81266	N/A	Perf: Ensure ROW__ID struct vectorizes
BUG-89714	N/A	Ranger table 'x_auth_sess' grows rapidly and never purged
BUG-95670	HIVE-18577, HIVE-18643	When run update/delete query on ACID partitioned table, HS2 read all each partitions.

Potential Data Loss

Hortonworks Bug ID	Apache JIRA	Summary
BUG-80027	HIVE-16177, HIVE-16688	Non-Acid to acid conversion doesn't handle _copy_N files

Query Failure

Hortonworks Bug ID	Apache JIRA	Summary
BUG-94506	PHOENIX-4525	Integer overflow in GroupBy execution

Security

Hortonworks Bug ID	Apache JIRA	Summary
BUG-101452	OOZIE-2427	[Kerberos] Authentication failure for the javascript resources under /ext-2.2
BUG-103083	HBASE-20406	HBase Thrift HTTP - Shouldn't handle TRACE/OPTIONS methods
BUG-72076	N/A	Best coding practices for encryption algorithm used
BUG-89037	RANGER-1868	SCA : Session Fixation
BUG-89041	RANGER-1814	SCA : Unreleased Resource: Streams
BUG-89145	RANGER-1828	Add XSS related headers in Ranger
BUG-94772	RANGER-1956	Remove command line option 'get' from CredentialBuilder API
BUG-95909	RANGER-1960	HBase auth does not take table namespace into consideration for deleting snapshot
BUG-98840	N/A	Use java io filesystem methods to access local files

Stability

Hortonworks Bug ID	Apache JIRA	Summary
BUG-103721	HIVE-19686	Creation of spark catalog in hive db fails
BUG-103722	N/A	SPARK: move spark catalog creation from thrift server to history server
BUG-103758	ZEPPELIN-3499	Deadlock between Interpreter restart and JobProgressPoller
BUG-103815	N/A	Matplotlib with Python3.4 fails
BUG-103907	N/A	IsolatedClientLoader should receive Apache Hadoop version
BUG-104547	HIVE-10815	Add HIVE-10815 changes in spark2
BUG-59028	YARN-5195	[YARN-5195] RM crashed with NPE while handling APP_ATTEMPT_REMOVED event
BUG-61924	TEZ-1248	Tez: disable slow-start & auto-reducer parallelism when reducer-count is small
BUG-63473	N/A	JDBC dependencies are not present out-of-the box in Zeppelin interpreter configurations
BUG-66580	HIVE-14947, HIVE-16542, HIVE-16832	ACID: Multi-insert and ROW__ID uniqueness within a partition
BUG-70994	LIVY-272	Add progressInfo to Statement
BUG-72871	LIVY-304	matplotlib doesn't work in livy pyspark interpreter
BUG-72874	LIVY-305	Pot doesn't work in SparkRInterpreter of livy
BUG-72969	LIVY-307	Paragraph becomes unresponsive when no resources for executors.
BUG-74988	AMBARI-20106, TEZ-3634	[LLAP] : intermittent task container time out - long GC pauses
BUG-75560	SPARK-15848	Ref EAR-3782 / Spark unable to read partitioned table in avro format and column name in upper case
BUG-76330	N/A	TPC-DS Query70 in Spark-2.1 throws grouping_id error
BUG-76859	SPARK-19439	[SPARK-19439][PYSPARK][SQL] PySpark's registerJavaFunction Should Support UDAFs
BUG-79762	N/A	Failure to handle hive tables with multi-directory locations
BUG-80659	KNOX-1330, ZEPPELIN-3491	Zeppelin favicon is not displayed (going thru Knox)
BUG-82358	N/A	Obsolete
BUG-83808	ZEPPELIN-2742	Unify Zeppelin's Storage Layer

Hortonworks Bug ID	Apache JIRA	Summary
BUG-83957	N/A	Uber bug to track difference in result for spark2 vs hive
BUG-84804	N/A	[Knox Proxy][Spark2] "Show incomplete applications" link redirection does not work
BUG-84810	SPARK-23644	[Knox Proxy][Spark2] complete application is not shown on Spark HS page
BUG-87427	SPARK-21912	Creating ORC datasource table should check invalid column names
BUG-88053	N/A	Denial of Service
BUG-89295	ZEPPELIN-2970	Insecure Randomness
BUG-90314	ATLAS-2214	Search using trait attributes - "!=" and "is not null" together and in order in string,boolean tag attribute filter comparison
BUG-93099	N/A	Druid-Hive(Secure) When Broker and Router running on different nodes-- queries with expression fails with org.apache.hive.druid.com.fasterxml.jackson.core.JsonParseException: Unexpected character ('<' (code 60)):
BUG-93116	RANGER-1957	Ranger Usersync is not syncing users or groups periodically when incremental sync is enabled.
BUG-93215	MAPREDUCE-6415	[MAPREDUCE-6415] Create a tool to combine aggregated logs into HAR files
BUG-93791	N/A	Livy Interpreter Restart for a user should only restart that users session.
BUG-93817	ZEPPELIN-3114	Ensure Apache Zeppelin has all the scenarios covered/fixed identified in BUG-93529
BUG-94384	CALCITE-2086	HTTP 413 error while connecting to PQS via ODBC driver
BUG-94731	N/A	Using timezone with positive offset from UTC throws error Hive/Druid
BUG-94799	HIVE-6091	STS doesn't clean .pipeout files
BUG-96290	HDFS-11701	NPE from Unresolved Host causes permanent DFSInputStream failures
BUG-97052	HIVE-17403	ORC concatenate can cause data loss
BUG-97090	N/A	Remove dependency on commons-beanutils:commons-beanutils before version 1.9.2 for Ambari Metrics Hadoop Sink
BUG-97091	N/A	Remove dependency on commons-beanutils:commons-beanutils before version 1.9.2 for Ambari Metrics Collector
BUG-97383	ZEPPELIN-3405	Zeppelin fails to display the User home page if user belongs to roles with space in its name.
BUG-97701	ZOOKEEPER-2227, ZOOKEEPER-2693, ZOOKEEPER-2726	DOS attack on wchp/wchc four letter words (4lw) : ZOOKEEPER-2693
BUG-98417	SPARK-24518	Backport: Passwords are stored in plaintext and world readable in spark config files
BUG-99735	YARN-8106	LogAggregationIndexedFileController - Can not get log meta from the log

Supportability

Hortonworks Bug ID	Apache JIRA	Summary
BUG-100410	OOZIE-3192	Oozie runtime directories are not getting removed
BUG-104114	ZEPPELIN-3513	Zeppelin Note Creation Failure is not shown on UI
BUG-105076	HIVE-19862	Hive Server check failed because of refused connection to Hive Server 2
BUG-76309	N/A	backport Fix #3517 f'No Such File' error when execute script out of druid installatio
BUG-87343	HIVE-18031	Support replication for Alter Database operation.
BUG-90454	HIVE-17840	HiveMetaStore consumes the exception if transactionalListeners.notifyEvent fail
BUG-92699	HIVE-18832	Support change management for trashing data files from ACID tables.
BUG-93116	RANGER-1957	Ranger Usersync is not syncing users or groups periodically when incremental sync is enabled.

Hortonworks Bug ID	Apache JIRA	Summary
BUG-98512	RANGER-2021, RANGER-2064	Ranger "Login Sessions" Audits impossible to browse due to 'rangerusersync'
BUG-98550	N/A	Add beacon service definition to Knox for HDP 3.0.0

Upgrade

Hortonworks Bug ID	Apache JIRA	Summary
BUG-104913	N/A	Hive schematool upgrade fails on EU from HDP-2.6 to HDP-3.0
BUG-105665	N/A	Hive Upgrade should run kinit before table conversion if needed

Usability

Hortonworks Bug ID	Apache JIRA	Summary
BUG-100570	ATLAS-2557	Fix to allow to lookup hadoop ldap groups when are groups from UGI are wrongly set or are not empty
BUG-103493	RANGER-2116	Fix policy condition widget's usability on resource and tag based policies authoring pages
BUG-103667	HIVE-19688	Need to update workhouse location in Hive CTLGS table after NN HA enabled
BUG-104374	HIVE-19477	Hiveserver2 does not emit open_connections metric in http mode, HBase Stochastic LB and HFile cleaner fixes, HDFS Datanode must not delete blocks on too many open files exception
BUG-104547	HIVE-10815	Add HIVE-10815 changes in spark2
BUG-105716	HBASE-20723	'WALSplitter uses the rootDir, which is walDir, as the recovered edits root path' to HDP-2.6.5
BUG-55893	HIVE-15858	beeline ^C doesn't close the session
BUG-78509	N/A	Created By is not displayed in the admin audit log pop-up
BUG-87848	RANGER-1773	Update Ranger KMS REST APIs to incorporate enunciate documentation
BUG-90314	ATLAS-2214	Search using trait attributes - "!=" and "is not null" together and in order in string,boolean tag attribute filter comparison
BUG-90513	N/A	Property "ranger.add-yarn-authorization=false" needs to be manually added to Custom ranger-yarn-security via Ambari
BUG-91364	AMBARI-22506	Incorrect pie chart distribution
BUG-92272	RANGER-1912	Ranger setup fails with mariadb/mysql when binary logging is enabled
BUG-92685	HIVE-18192	Introduce Write ID per table to replace global transaction ID
BUG-96378	SPARK-23355	convertMetastore should not ignore table properties
BUG-99254	HIVE-19396	Incorrect operation name in post-hook context: 'QUERY' for 'ANALYZE TABLE' command

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	Description of the problem or behavior SSL is deprecated and its use in production is not recommended. Use TLS. Workaround In Ambari: Use ssl.enabled.protocols=TLSv1|TLSv1.1|TLSv1.2 and security.server.disabled.protocols=SSL|SSLv2|SSLv3. For help configuring TLS for other components, contact customer support. Documentation will be provided in a future release.

Hortonworks Bug ID	Apache JIRA	Apache component	Summary

BUG-106494	N/A	Documentation, Hive	**Description of Problem**
			
			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.
			
			Associated error message
			
			`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.setPartition ~hive-exec-4.0.0-SNAPSHOT.jar:4.0.0-SNAPSHOT at`
			
			`org.apache.hadoop.hive.ql.stats.ColStatsProcessor.pe ~hive-exec-4.0.0-SNAPSHOT.jar:4.0.0-SNAPSHOT at`
			
			`org.apache.hadoop.hive.ql.stats.ColStatsProcessor.pr ~hive-exec-4.0.0-SNAPSHOT.jar:4.0.0-SNAPSHOT at`
			
			`org.apache.hadoop.hive.ql.exec.StatsTask.execute(Sta hive-exec-4.0.0-SNAPSHOT.jar:4.0.0-SNAPSHOT at`
			
			`org.apache.hadoop.hive.ql.exec.Task.executeTask(Task hive-exec-4.0.0-SNAPSHOT.jar:4.0.0-SNAPSHOT at`
			
			`org.apache.hadoop.hive.ql.exec.TaskRunner.runSequent hive-exec-4.0.0-SNAPSHOT.jar:4.0.0-SNAPSHOT at`
			
			`org.apache.hadoop.hive.ql.Driver.launchTask(Driver.j hive-exec-4.0.0-SNAPSHOT.jar:4.0.0-SNAPSHOT at`
			
			`org.apache.hadoop.hive.ql.Driver.execute(Driver.java hive-exec-4.0.0-SNAPSHOT.jar:4.0.0-SNAPSHOT at`
			
			`org.apache.hadoop.hive.ql.Driver.run`
			
			Workaround
			
			Do not partition columns of type double.

Hortonworks Bug ID	Apache JIRA	Apache component	Summary
BUG-101082	N/A	Documentation, Hive	Description of the problem or behavior When running Beeline in batch mode, queries killed by the Workload Management process can on rare occasions mistakenly return success on the command line. Workaround There is currently no workaround.
BUG-106379	N/A	Documentation Hive	Description of the Problem The upgrade process fails to perform necessary compaction of ACID tables and can cause permanent data loss. Workaround 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.
BUG-106286	N/A	Documentation Hive	Description of the Problem The upgrade process might fail to make a backup of the Hive metastore, which is critically important. Workaround 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.
BUG-98628	HBASE-20530	HBase	Description of the problem or behavior When running the restore of an incremental backup, the restore task may fail with the error "java.io.IOException: No input paths specified in job". This only happens intermittently. Workaround Because exact causes of the error are unknown, there is no known workaround. Re-running the restore task may succeed; it may not.
BUG-103495	N/A	HBase	Description of the problem or behavior 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. Workaround There is currently no workaround.
BUG-98727	N/A	HBase	Description of the problem or behavior 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. Workaround There is currently no workaround.

Hortonworks Bug ID	Apache JIRA	Apache component	Summary
BUG-105983	N/A	HBase	Description of the problem or behavior An HBase service (Master or RegionServer) stops participating with the rest of the HBase cluster. Associated error message The service's log contains stack traces that contain "Kerberos principal name does NOT have the expected hostname part..." Workaround Retrying the connection solves the problem.
BUG-94954	HBASE-20552	HBase	Description of the problem or behavior After a rolling restart of HBase, the HBase master may not correctly assign out all Regions to the cluster. Associated error message TThere are regions in transition, including hbase:meta, which result in "Region is not online on RegionServer" messages on Master or RegionServer or messages around errors in ServerCrashProcudure in the Master. Workaround Restart the HBase Master.
BUG-96402	HIVE-18687	Hive	Description of the problem or behavior 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. Workaround 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.
BUG-88614	N/A	Hive	Description of the problem or behavior RDMBS schema for Hive metastore contains an index HL_TXNID_INDEX defined as CREATE INDEX HL_TXNID_INDEX ON HIVE_LOCKS USING hash (HL_TXNID); Hash indexes are not recommended by PostgreSQL. For more information, see https://www.postgresql.org/docs/9.4/static/indexes-types.html Workaround It's recommended that this index is changed to type BTREE.
BUG-101836	HIVE-19416, HIVE-19820	Hive	Description of the problem or behavior Statistics-based optimizations for metadata-only queries, such as count, count(distinct <partcol>), do not currently work for managed tables.
BUG-101836	N/A	Hive	Description of the problem or behavior Statistics-based optimizations for metadata-only queries, such as count, count(distinct <partcol>), do not currently work for managed tables. Workaround There is currently no workaround.

Hortonworks Bug ID	Apache JIRA	Apache component	Summary

BUG-107434	N/A	Hive	**Description of the problem or behavior**
			
			Tables that have buckets must be recreated when you upgrade a cluster to HDP 3.0. The hash function for bucketing has changed in HDP 3.0, causing a problem in certain operations, such as INSERT and JOIN. In these cases, Hive does not handle queries correctly if you mix old and new tables in the same query. To avoid this problem, recreate bucketed tables using the following workaround after upgrading, but before running any queries on the cluster.
			
			Workaround:
			
			1. Identify tables that have obsolete buckets. You can identify these tables by the TBLPROPERTY bucketing_version. If bucketing_version is 1, you need to recreate the table.
			
			   Example: SHOW TBLPROPERTIES x;
			
			``
			+---------------------------
			+-------------+ |
			                             prpt_name
			
				prpt_value |
				
				 +---------------------------
				+-------------+
					bucketing_version
					
						1
						
								numFiles	1
									            |
									                                …
									``
									
									   Proceed to the next step.
									2. Using a CTAS statement, create a new table. For example, create table x_new based on the old table x:
									
									   Example: CREATE TABLE x_new AS SELECT * FROM x;
									3. Verify that table x_new has all the data.
									4. Drop old table x.
									5. Create a schema named the same as the old table x.
									
									   Example: CREATE TABLE x …
									6. Insert the data from table x_new into the table x schema.
									
									   Example: INSERT INTO x SELECT * FROM x_new;
									7. Verify that the data is in table x.
									8. Drop the table x_new.

BUG-120655	N/A	Hive	**Description of the problem or behavior** The Hive Warehouse Connector does not support non-ORC file formats for writes. Workaround There is currently no workaround.
BUG-60904	KNOX-823	Knox	**Description of the problem or behavior** 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. Workaround There is currently no workaround.

Hortonworks Bug ID	Apache JIRA	Apache component	Summary
BUG-107399	N/A	Knox	Description of the problem or behavior 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. Workaround 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.
BUG-91996	LIVY-299	Livy Zeppelin	Description of the problem or behavior Livy Spark interpreter will only print the output of the last code line on the output. For example, if we submitted: print(10) print(11) Only "11" will be printed out, the output of first line "10" will be ignored. Workaround If you want to see the output of a particular line, then it must be the last line in the code block.
BUG-106266	OOZIE-3156	Oozie	Description of the problem or behavior 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. 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. Associated error message Ssh command exits with the exit status of the remote command or with 255 if an error occurred Workaround Retrying the connection solves the problem.
BUG-107236	N/A	Ranger	Description of the problem or behavior Atlas Rest sync source is not supported for tagsync. Workaround Using Kafka is recommended.

Hortonworks Bug ID	Apache JIRA	Apache component	Summary
BUG-101227	N/A	Spark	Description of the problem or behavior When Spark Thriftserver has to run several queries concurrently, some of them can fail with a timeout exception when performing broadcast join. Associated error message ``` 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.BroadcastHashJoi ``` Workaround You can resolve this issue by increasing the spark.sql.broadcastTimeout value.
BUG-100187	SPARK-23942	Spark	Description of the problem or behavior In Spark, users can register QueryExecutionListener to add callbacks for query executions, for example, for an action such as collect, foreach and show in DataFrame. You may use spark.session().listenerManager().register(...) or spark.sql.queryExecutionListeners configuration to set the query execution listener. This usually works in other API languages as well; however, this was a bug in collect Python API - the callback was not being called. Now, it is being called correctly within Spark side. Workaround Workaround is to manually call the callbacks right after collect in Python API with a try-catch.
BUG-109607	N/A	Spark	Description of the problem or behavior 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. Workaround There is currently no workaround.

Hortonworks Bug ID	Apache JIRA	Apache component	Summary

BUG-65977	SPARK-14922	Spark	Description of the problem or behavior
			
			Since Spark 2.0.0, `DROP PARTITION BY RANGE` is not supported grammatically. In other words, only '=' is supported while `<', '>', '<=', '>=' aren't.
			
			Associated error message
			
			``
			scala> sql("alter table t drop partition (b<1)
			                              ").show
			org.apache.spark.sql.catalyst.parser.ParseException: mismatched
			                      input '<'
			expecting {')', ','}(line 1, pos 31) == SQL == alter table t
			                        drop partition (b<1) ----------------------------^^^
			``
			
			Workaround
			
			To drop partition, use the exact match with '='.
			
			``
			scala>
			                        sql("alter table
			t drop partition (b=0) ").show
			``

BUG-110970	N/A	Spark	Description of the problem or behavior For long-running SparkSQL jobs on a Kerberized cluster, some JDBC clients may randomly fail with a "no token found in cache" error after the delegation token expiry period. Workaround 1. On the Ambari dashboard, select Spark2 > Configs > Custom spark2-thrift-sparkconf and add the following properties: ``` spark.yarn.keytab spark.yarn.principal ``` 2. Set the value of these properties to match the values of the spark.history.kerberos.keytab and spark.history.kerberos.principal properties in Spark2 > Configs > Advanced spark2-defaults. 3. Click Save, then restart Spark and any other services that require a restart.

Hortonworks Bug ID	Apache JIRA	Apache component	Summary
| BUG-106917 | N/A | Sqoop | Description of the problem or behavior<br><br>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.<br><br>Associated error message<br><br>```
Table db.table failed strict managed
  table checks due to the
                     following reason:
  Table is marked as a managed table but
  is not
                     transactional.
```<br><br>Workaround<br><br>When using --hive-import with --as-parquetfile, users must also provide --external-table-dir with a fully qualified location of the table:<br><br>```
sqoop import ... --hive-import
                     --as-
parquetfile --external-table-dir
                     hdfs:///path/to/
table
``` |
BUG-102672	N/A	Sqoop	Description of the problem or behavior 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. Associated error message Store into a transactional table db.table from Pig/Mapreduce is not supported Workaround 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.
RMP-11408	ZEPPELIN-2170	Zeppelin	Description of the problem or behavior Zeppelin does not show all WARN messages thrown by spark-shell at the Zeppelin's notebook level. Workaround There is currently no workaround for this.
BUG-91364	AMBARI-22506	Zeppelin	Description of the problem or behavior Pie charts in Zeppelin does not display the correct distribution as per the provided data. This occurs when there is a "," in data i.e. there is number formatting applied to the data. Workaround Add a manual configuration setting in Zeppelin's JDBC interpreter setting. Add "phoenix.phoenix.query.numberFormat" with value "#.#".
N/A	N/A	N/A	Description of the problem or behavior Open JDK 8u242 is not supported as it causes Kerberos failure. Workaround Use a different version of Open JDK.

Table 3: Technical Service Bulletin

Technical Service Bulletin	Apache JIRA	Apache component	Summary
TSB-327	HDFS-5698	HDFS	CVE-2018-11768: HDFS FSImage Corruption (potential DoS, file/dir takeover)
			In very large clusters, the in-memory format to store the user, group, acl, and extended attributes may exceed the size of the on disk format, causing corruption of fsImage.
			For the latest update on this issue, see the corresponding Knowledge article: CVE-2018-11768: HDFS FSImage Corruption (potential DoS, file/dir takeover)
TSB-405	N/A	N/A	Impact of LDAP Channel Binding and LDAP signing changes in Microsoft Active Directory
			Microsoft has introduced changes in LDAP Signing and LDAP Channel Binding to increase the security for communications between LDAP clients and Active Directory domain controllers. These optional changes will have an impact on how 3rd party products integrate with Active Directory using the LDAP protocol.
			Workaround
			Disable LDAP Signing and LDAP Channel Binding features in Microsoft Active Directory if they are enabled
			For the latest update on this issue see the corresponding Knowledge article:TSB-2021 405: Impact of LDAP Channel Binding and LDAP signing changes in Microsoft Active Directory
TSB-406	N/A	HDFS	CVE-2020-9492 Hadoop filesystem bindings (ie: webhdfs) allows credential stealing
			WebHDFS clients might send SPNEGO authorization header to remote URL without proper verification. A maliciously crafted request can trigger services to send server credentials to a webhdfs path (ie: webhdfs://…) for capturing the service principal
			For the latest update on this issue see the corresponding Knowledge article: TSB-2021 406: CVE-2020-9492 Hadoop filesystem bindings (ie: webhdfs) allows credential stealing

Technical Service Bulletin	Apache JIRA	Apache component	Summary
TSB-450	HBASE-21000	HBase	Default limits for PressureAwareCompactionThroughputController are too low HDP and CDH releases suffer from low compaction throughput limits, which cause storefiles to back up faster than compactions can re-write them. For the latest update on this issue see the corresponding Knowledge article: Cloudera Customer Advisory: Default limits for PressureAwareCompactionThroughputController are too low
TSB-463	N/A	HBase	HBase Performance Issue The HDFS short-circuit setting dfs.client.read.shortcircuit is overwritten to disabled by hbase-default.xml. HDFS short-circuit reads bypass access to data in HDFS by using a domain socket (file) instead of a network socket. This alleviates the overhead of TCP to read data from HDFS which can have a meaningful improvement on HBase performance (as high as 30-40%). For the latest update on this issue see the corresponding Knowledge article: TSB 2021-463: HBase Performance Issue
TSB-480/2	HIVE-24224	Hive	Hive ignores the property to skip a header or footer in a compressed file Incorrect results can occur running SELECT queries if count value is greater than 0. For the latest update on this issue see the corresponding Knowledge article: TSB 2021-480.2: Hive ignores the property to skip a header or footer in a compressed file

Technical Service Bulletin	Apache JIRA	Apache component	Summary
TSB-494	• HBASE-20781 • HBASE-25459 • PHOENIX-5250	HBase	Accumulated WAL Files Cannot be Cleaned up When Using Phoenix Secondary Global Indexes The Write-ahead-log (WAL) files for Phoenix tables that have secondary global indexes defined on them, cannot be automatically cleaned up by HBase, leading to excess storage usage and possible error due to filling up the storage. Workaround Perform rolling restart of HBase if the number of znodes under hbase-secure/splitWAL in ZooKeeper is greater than 8000. For the latest update on this issue see the corresponding Knowledge article: TSB 2021-494: Accumulated WAL Files Cannot be Cleaned up When Using Phoenix Secondary Global Indexes
TSB-497	N/A	Solr	CVE-2021-27905: Apache Solr SSRF vulnerability with the Replication handler The Apache Solr ReplicationHandler (normally registered at "/replication" under a Solr core) has a "masterUrl" (also "leaderUrl" alias) parameter. The "masterUrl" parameter is used to designate another ReplicationHandler on another Solr core to replicate index data into the local core. To help prevent the CVE-2021-27905 SSRF vulnerability, Solr should check these parameters against a similar configuration used for the "shards" parameter. For the latest update on this issue, see the corresponding Knowledge article: TSB 2021-497: CVE-2021-27905: Apache Solr SSRF vulnerability with the Replication handler
TSB-512	N/A	HBase	HBase MOB data loss HBase tables with the MOB feature enabled may encounter problems which result in data loss. For the latest update on this issue, see the corresponding Knowledge article: TSB 2021-512: HBase MOB data loss

Documentation Errata

This section contains late additions or corrections to the product documentation.

Using Apache HBase to store and access data

Component Affected: HBase

The following feature is shipped with HDP 3.0.0 GA, and it is not documented in What's New in Apache HBase.

Backup/restore

You can use the built-in tooling in HBase to create full and incremental backups of the HBase data.

Enable Cgroups

Component Affected: YARN

Description of the problem

Incorrect recommendation for setting the value of yarn.nodemanager.linux-container-executor.cgroups.strict-resourceusage

Set Flexible or Strict CPU limits

Strict limits are not enabled (set to false) by default.

Correction

Do not set the value of yarn.nodemanager.linux-container-executor.cgroups.strict-resourceusage to true. Property: yarn.nodemanager.linux-container-executor.cgroups.strict-resourceusageValue: false. For example:

```
<property>
<name>yarn.nodemanager.linux-container-executor.cgroups.strict-resource-
usage</name>
<value>true</value>
</property>
```

Legal Information

Hortonworks Data Platform 3.0.0

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 # 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.