

HDInsight June 2018 refresh

Table of Contents

Summary.....	2
New features.....	2
Component versions.....	2
Technical Preview Features.....	3
Apache Patch Information	3
Hadoop.....	3
Atlas	Error! Bookmark not defined.
HBase	5
Hive	6
Kafka	9
Mahout.....	10
Oozie	11
Phoenix	11
Pig	12
Ranger	12
Slider	12
Spark	12
Sqoop	14
Storm.....	14
Tez.....	14
Zeppelin.....	14
ZooKeeper	14
Fixed Common Vulnerabilities and Exposures.....	15
CVE-2017-7676.....	15
CVE-2017-7677.....	15
CVE-2017-9799.....	16
CVE-2016-4970.....	16
CVE-2016-8746.....	17
CVE-2016-8751.....	17
Fixed Issues for Support	18

Behavioral changes.....	44
Known Issues.....	47
Deprecation	Error! Bookmark not defined.
Upgrading	50

Summary

Azure HDInsight is one of the most popular services amongst enterprise for open source Hadoop & Spark analytics on Azure. With the [plus 50 percent price cut on HDInsight](#), customers moving to the cloud are reaping more savings than ever.

New features

The new updates and capabilities fall in to the following categories:

1. **Update Hadoop and other open source projects** – In addition to, 1000+ bug fixes across 20+ open source projects, this update contains a new version of **Spark (2.3)** and **Kafka (1.0)**.
 - a. [New features in Spark 2.3](#)
 - b. [New features in Kafka 1.0](#)
2. **Update R Server 9.1 to Machine Learning Services 9.3** – With this release, we are providing data scientists and engineers with the best of open source enhanced with algorithmic innovations and ease of operationalization, all available in their preferred language with the speed of Apache Spark. This release expands upon the capabilities offered in R Server with added support for Python, leading to the cluster name change from R Server to ML Services.
3. **Support for Azure Data Lake Storage Gen2**– HDInsight will support the Preview release of Azure Data Lake Storage Gen2. In the available regions, customers will be able to choose an ADLS Gen2 account as the Primary or Secondary store for their HDInsight clusters.
4. **HDInsight Enterprise Security Package Updates (Preview)** – (Preview) [Virtual Network Service Endpoints](#) support for Azure Blob Storage, ADLS Gen1, Cosmos DB and Azure DB.

Component versions

The official Apache versions of all HDInsight 3.6 components are listed below. All components listed here are official Apache releases of the most recent stable versions available.

- Apache Hadoop 2.7.3
- Apache HBase 1.1.2
- Apache Hive 1.2.1

- Apache Hive 2.1.0
- Apache Kafka 1.0.0
- Apache Mahout 0.9.0+
- Apache Oozie 4.2.0
- Apache Phoenix 4.7.0
- Apache Pig 0.16.0
- Apache Ranger 0.7.0
- Apache Slider 0.92.0
- Apache Spark 2.2.0/2.3.0
- Apache Sqoop 1.4.6
- Apache Storm 1.1.0
- Apache TEZ 0.7.0
- Apache Zeppelin 0.7.3
- Apache ZooKeeper 3.4.6

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

Technical Preview Features

The following features are available within this release but are not ready for production deployment. We encourage you to explore these technical preview features in non-production environments and provide feedback on your experiences through the [MSDN forum](#) or [Stackoverflow \(#hdinsight\)](#).

Apache Patch Information

Hadoop

This release provides Hadoop Common 2.7.3 and the following Apache patches:

- [HADOOP-13190](#): Mention LoadBalancingKMSClientProvider in KMS HA documentation.
- [HADOOP-13227](#): AsyncCallHandler should use an event driven architecture to handle async calls.
- [HADOOP-14104](#): Client should always ask namenode for kms provider path.
- [HADOOP-14799](#): Update nimbus-jose-jwt to 4.41.1.
- [HADOOP-14814](#): Fix incompatible API change on FsServerDefaults to HADOOP-14104.
- [HADOOP-14903](#): Add json-smart explicitly to pom.xml.
- [HADOOP-15042](#): Azure PageBlobInputStream.skip() can return negative value when numberOfPagesRemaining is 0.
- [HADOOP-15255](#): Upper/Lower case conversion support for group names in LdapGroupsMapping.
- [HADOOP-15265](#): exclude json-smart explicitly from hadoop-auth pom.xml.

- [HDFS-7922](#): ShortCircuitCache#close is not releasing ScheduledThreadPoolExecutors.
- [HDFS-8496](#): Calling stopWriter() with FSDataSetImpl lock held may block other threads (cmccabe).
- [HDFS-10267](#): Extra "synchronized" on FsDataSetImpl#recoverAppend and FsDataSetImpl#recoverClose.
- [HDFS-10489](#): Deprecate dfs.encrypted.key.provider.uri for HDFS encryption zones.
- [HDFS-11384](#): Add option for balancer to disperse getBlocks calls to avoid NameNode's rpc.CallQueueLength spike.
- [HDFS-11689](#): New exception thrown by DFSClient%isHDFSEncryptionEnabled broke hacky hive code..
- [HDFS-11711](#): DN should not delete the block On "Too many open files" Exception.
- [HDFS-12347](#): TestBalancerRPCDelay#testBalancerRPCDelay fails very frequently.
- [HDFS-12781](#): After Datanode down, In Namenode UI Datanode tab is throwing warning message..
- [HDFS-13054](#): Handling PathsNotEmptyDirectoryException in DFSClient delete call..
- [HDFS-13120](#): Snapshot diff could be corrupted after concat.
- [YARN-3742](#): YARN RM will shut down if ZKClient creation times out.
- [YARN-6061](#): Add an UncaughtExceptionHandler for critical threads in RM.
- [YARN-7558](#): yarn logs command fails to get logs for running containers if UI authentication is enabled.
- [YARN-7697](#): Fetching logs for finished application fails even though log aggregation is complete.

HDP 2.6.4 provided Hadoop Common 2.7.3 and the following Apache patches:

- [HADOOP-13700](#): Remove unthrown IOException from TrashPolicy#initialize and #getInstance signatures.
- [HADOOP-13709](#): Ability to clean up subprocesses spawned by Shell when the process exits..
- [HADOOP-14059](#): typo in s3a rename(self, subdir) error message.
- [HADOOP-14542](#): Add IOUtils.cleanupWithLogger that accepts slf4j logger API.
- [HDFS-9887](#): WebHdfs socket timeouts should be configurable.
- [HDFS-9914](#): Fix configurable WebhDFS connect/read timeout.
- [MAPREDUCE-6698](#): Increase timeout on TestUnnecessaryBlockingOnHistoryFileInfo.testTwoThreadsQueryingDifferentJobOfSameUser.
- [YARN-4550](#): Some tests in TestContainerLaunch fails on non-english locale environment.
- [YARN-4717](#): TestResourceLocalizationService.testPublicResourceInitializesLocalDir fails Intermittently due to IllegalArgumentException from cleanup.
- [YARN-5042](#): Mount /sys/fs/cgroup into Docker containers as readonly mount.
- [YARN-5318](#): Fix intermittent test failure of TestRMAdminService#testRefreshNodesResourceWithFileSystemBasedConfigurationProvider.
- [YARN-5641](#): Localizer leaves behind tarballs after container is complete..
- [YARN-6004](#): Refactor TestResourceLocalizationService#testDownloadingResourcesOnContainer so that it is less than 150 lines.
- [YARN-6078](#): Containers stuck in Localizing state.

- [YARN-6805](#): NPE in LinuxContainerExecutor due to null PrivilegedOperationException exit code.

HBase

This release provides HBase 1.1.2 and the following Apache patches.

- [HBASE-13376](#): Improvements to Stochastic load balancer.
- [HBASE-13716](#): Stop using Hadoop's FSConstants.
- [HBASE-13848](#): Access InfoServer SSL passwords through Credential Provider API.
- [HBASE-13947](#): Use MasterServices instead of Server in AssignmentManager.
- [HBASE-14135](#): HBase Backup/Restore Phase 3: Merge backup images.
- [HBASE-14473](#): Compute region locality in parallel.
- [HBASE-14517](#): Show regionserver's version in master status page.
- [HBASE-14606](#): TestSecureLoadIncrementalHFiles tests timed out in trunk build on apache.
- [HBASE-15210](#): Undo aggressive load balancer logging at tens of lines per millisecond.
- [HBASE-15515](#): Improve LocalityBasedCandidateGenerator in Balancer.
- [HBASE-15615](#): Wrong sleep time when RegionServerCallable need retry.
- [HBASE-16135](#): PeerClusterZnode under rs of removed peer may never be deleted.
- [HBASE-16570](#): Compute region locality in parallel at startup.
- [HBASE-16810](#): HBase Balancer throws ArrayIndexOutOfBoundsException when regionserver is in /hbase/dRAINING znode and unloaded.
- [HBASE-16852](#): TestDefaultCompactSelection failed on branch-1.3.
- [HBASE-17387](#): Reduce the overhead of exception report in RegionActionResult for multi().
- [HBASE-17850](#): Backup system repair utility.
- [HBASE-17931](#): Assign system tables to servers with highest version.
- [HBASE-18083](#): Make large/small file clean thread number configurable in HFileCleaner.
- [HBASE-18084](#): Improve CleanerChore to clean from directory which consumes more disk space.
- [HBASE-18164](#): Much faster locality cost function and candidate generator.
- [HBASE-18212](#): In Standalone mode with local filesystem HBase logs Warning message:Failed to invoke 'unbuffer' method in class class org.apache.hadoop.fs.FSDataInputStream.
- [HBASE-18808](#): Ineffective config check in BackupLogCleaner#getDeletableFiles().
- [HBASE-19052](#): FixedFileTrailer should recognize CellComparatorImpl class in branch-1.x.
- [HBASE-19065](#): HRegion#bulkLoadHFiles() should wait for concurrent Region#flush() to finish.
- [HBASE-19285](#): Add per-table latency histograms.
- [HBASE-19393](#): HTTP 413 FULL head while accessing HBase UI using SSL. .
- [HBASE-19395](#): [branch-1] TestEndToEndSplitTransaction.testMasterOpsWhileSplitting fails with NPE.
- [HBASE-19421](#): branch-1 does not compile against Hadoop 3.0.0.
- [HBASE-19934](#): HBaseSnapshotException when read replicas is enabled and online snapshot is taken after region splitting.
- [HBASE-20008](#): [backport] NullPointerException when restoring a snapshot after splitting a region.

Hive

This release provides Hive 1.2.1 and Hive 2.1.0 in addition to the following patches:

Hive 1.2.1 Apache patches:

- [HIVE-10697](#): ObjectInspectorConvertors#UnionConvertor does a faulty conversion.
- [HIVE-11266](#): count(*) wrong result based on table statistics for external tables.
- [HIVE-12245](#): Support column comments for an HBase backed table.
- [HIVE-12315](#): Fix Vectorized double divide by zero.
- [HIVE-12360](#): Bad seek in uncompressed ORC with predicate pushdown.
- [HIVE-12378](#): Exception on HBaseSerDe.serialize binary field.
- [HIVE-12785](#): View with union type and UDF to the struct is broken.
- [HIVE-14013](#): Describe table doesn't show unicode properly.
- [HIVE-14205](#): Hive doesn't support union type with AVRO file format.
- [HIVE-14421](#): FS.deleteOnExit holds references to _tmp_space.db files..
- [HIVE-15563](#): Ignore Illegal Operation state transition exception in SQLOperation.runQuery to expose real exception.
- [HIVE-15680](#): Incorrect results when hive.optimize.index.filter=true and same ORC table is referenced twice in query, in MR mode.
- [HIVE-15883](#): HBase mapped table in Hive insert fail for decimal.
- [HIVE-16232](#): Support stats computation for columns in QuotedIdentifier .
- [HIVE-16828](#): With CBO enabled, Query on partitioned views throws IndexOutOfBoundsException.
- [HIVE-17013](#): Delete request with a subquery based on select over a view.
- [HIVE-17063](#): insert overwrite partition onto a external table fail when drop partition first.
- [HIVE-17259](#): Hive JDBC does not recognize UNIONTYPE columns.
- [HIVE-17419](#): ANALYZE TABLE...COMPUTE STATISTICS FOR COLUMNS command shows computed stats for masked tables.
- [HIVE-17530](#): ClassCastException when converting uniontype.
- [HIVE-17621](#): Hive-site settings are ignored during HCatInputFormat split-calculation.
- [HIVE-17636](#): Add multiple_agg.q test for blobstores.
- [HIVE-17729](#): Add Database and Explain related blobstore tests.
- [HIVE-17731](#): add a backward compat option for external users to HIVE-11985.
- [HIVE-17803](#): With Pig multi-query, 2 HCatStorers writing to the same table will trample each other's outputs.
- [HIVE-17829](#): ArrayIndexOutOfBoundsException - HBASE-backed tables with Avro schema in Hive2.
- [HIVE-17845](#): insert fails if target table columns are not lowercase.
- [HIVE-17900](#): analyze stats on columns triggered by Compactor generates malformed SQL with > 1 partition column.
- [HIVE-18026](#): Hive webhcat principal configuration optimization.
- [HIVE-18031](#): Support replication for Alter Database operation.
- [HIVE-18090](#): acid heartbeat fails when metastore is connected via hadoop credential.

- [HIVE-18189](#): Hive query returning wrong results when set `hive.groupby.orderby.position.alias` to true.
- [HIVE-18258](#): Vectorization: Reduce-Side GROUP BY MERGEPARTIAL with duplicate columns is broken.
- [HIVE-18293](#): Hive is failing to compact tables contained within a folder that is not owned by identity running HiveMetaStore.
- [HIVE-18327](#): Remove the unnecessary HiveConf dependency for MiniHiveKdc.
- [HIVE-18341](#): Add repl load support for adding "raw" namespace for TDE with same encryption keys.
- [HIVE-18352](#): introduce a METADATAONLY option while doing REPL DUMP to allow integrations of other tools.
- [HIVE-18353](#): CompactorMR should call `jobclient.close()` to trigger cleanup (Prabhu Joseph via Thejas Nair).
- [HIVE-18390](#): IndexOutOfBoundsException when query a partitioned view in ColumnPruner.
- [HIVE-18429](#): Compaction should handle a case when it produces no output.
- [HIVE-18447](#): JDBC: Provide a way for JDBC users to pass cookie info via connection string.
- [HIVE-18460](#): Compactor doesn't pass Table properties to the Orc writer.
- [HIVE-18467](#): support whole warehouse dump / load + create/drop database events (Anishek Agarwal, reviewed by Sankar Hariappan) .
- [HIVE-18551](#): Vectorization: VectorMapOperator tries to write too many vector columns for Hybrid Grace.
- [HIVE-18587](#): insert DML event may attempt to calculate a checksum on directories.
- [HIVE-18613](#): Extend JsonSerDe to support BINARY type.
- [HIVE-18626](#): Repl load "with" clause does not pass config to tasks.
- [HIVE-18660](#): PCR doesn't distinguish between partition and virtual columns.
- [HIVE-18754](#): REPL STATUS should support 'with' clause.
- [HIVE-18754](#): REPL STATUS should support 'with' clause.
- [HIVE-18788](#): Clean up inputs in JDBC PreparedStatement.
- [HIVE-18794](#): Repl load "with" clause does not pass config to tasks for non-partition tables.
- [HIVE-18808](#): Make compaction more robust when stats update fails.
- [HIVE-18817](#): ArrayIndexOutOfBoundsException exception during read of ACID table..
- [HIVE-18833](#): Auto Merge fails when "insert into directory as orcfile".
- [HIVE-18879](#): Disallow embedded element in UDFXPathUtil needs to work if xercesImpl.jar in classpath.
- [HIVE-18907](#): Create utility to fix acid key index issue from HIVE-18817.

Hive 2.1.0 Apache Patches:

- [HIVE-14013](#): Describe table doesn't show unicode properly.
- [HIVE-14205](#): Hive doesn't support union type with AVRO file format.
- [HIVE-15563](#): Ignore Illegal Operation state transition exception in SQLOperation.runQuery to expose real exception.
- [HIVE-15680](#): Incorrect results when `hive.optimize.index.filter=true` and same ORC table is referenced twice in query, in MR mode.

- [HIVE-15883](#): HBase mapped table in Hive insert fail for decimal.
- [HIVE-16757](#): Remove calls to deprecated AbstractRelNode.getRows.
- [HIVE-16828](#): With CBO enabled, Query on partitioned views throws IndexOutOfBoundsException.
- [HIVE-17063](#): insert overwrite partition onto a external table fail when drop partition first.
- [HIVE-17259](#): Hive JDBC does not recognize UNIONTYPE columns.
- [HIVE-17530](#): ClassCastException when converting uniontype.
- [HIVE-17600](#): Make OrcFile's enforceBufferSize user-settable..
- [HIVE-17601](#): improve error handling in LlapServiceDriver.
- [HIVE-17613](#): remove object pools for short, same-thread allocations.
- [HIVE-17617](#): Rollup of an empty resultset should contain the grouping of the empty grouping set.
- [HIVE-17621](#): Hive-site settings are ignored during HCatInputFormat split-calculation.
- [HIVE-17629](#): CachedStore: Have a whitelist/blacklist config to allow selective caching of tables/partitions and allow read while prewarming.
- [HIVE-17636](#): Add multiple_agg.q test for blobstores.
- [HIVE-17702](#): incorrect isRepeating handling in decimal reader in ORC.
- [HIVE-17729](#): Add Database and Explain related blobstore tests.
- [HIVE-17731](#): add a backward compat option for external users to HIVE-11985.
- [HIVE-17803](#): With Pig multi-query, 2 HCatStorers writing to the same table will trample each other's outputs.
- [HIVE-17845](#): insert fails if target table columns are not lowercase.
- [HIVE-17900](#): analyze stats on columns triggered by Compactor generates malformed SQL with > 1 partition column.
- [HIVE-18006](#): Optimize memory footprint of HLLDenseRegister.
- [HIVE-18026](#): Hive webhcat principal configuration optimization.
- [HIVE-18031](#): Support replication for Alter Database operation.
- [HIVE-18090](#): acid heartbeat fails when metastore is connected via hadoop credential.
- [HIVE-18189](#): Order by position does not work when cbo is disabled.
- [HIVE-18258](#): Vectorization: Reduce-Side GROUP BY MERGEPARTIAL with duplicate columns is broken.
- [HIVE-18269](#): LLAP: Fast Llap io with slow processing pipeline can lead to OOM.
- [HIVE-18293](#): Hive is failing to compact tables contained within a folder that is not owned by identity running HiveMetaStore.
- [HIVE-18318](#): LLAP record reader should check interrupt even when not blocking.
- [HIVE-18326](#): LLAP Tez scheduler - only preempt tasks if there's a dependency between them.
- [HIVE-18327](#): Remove the unnecessary HiveConf dependency for MiniHiveKdc.
- [HIVE-18331](#): Add relogin when TGT expire and some logging/lambda.
- [HIVE-18341](#): Add repl load support for adding "raw" namespace for TDE with same encryption keys.
- [HIVE-18352](#): introduce a METADATAONLY option while doing REPL DUMP to allow integrations of other tools.
- [HIVE-18353](#): CompactorMR should call jobclient.close() to trigger cleanup.
- [HIVE-18384](#): ConcurrentModificationException in log4j2.x library.

- [HIVE-18390](#): IndexOutOfBoundsException when query a partitioned view in ColumnPruner.
- [HIVE-18447](#): JDBC: Provide a way for JDBC users to pass cookie info via connection string.
- [HIVE-18460](#): Compactor doesn't pass Table properties to the Orc writer.
- [HIVE-18462](#): (Explain formatted for queries with map join has columnExprMap with unformatted column name).
- [HIVE-18467](#): support whole warehouse dump / load + create/drop database events.
- [HIVE-18488](#): LLAP ORC readers are missing some null checks.
- [HIVE-18490](#): Query with EXISTS and NOT EXISTS with non-equi predicate can produce wrong result.
- [HIVE-18506](#): LlapBaseInputFormat - negative array index.
- [HIVE-18517](#): Vectorization: Fix VectorMapOperator to accept VRBs and check vectorized flag correctly to support LLAP Caching).
- [HIVE-18523](#): Fix summary row in case there are no inputs.
- [HIVE-18528](#): Aggregate stats in ObjectStore get wrong result.
- [HIVE-18530](#): Replication should skip MM table (for now).
- [HIVE-18548](#): Fix log4j import.
- [HIVE-18551](#): Vectorization: VectorMapOperator tries to write too many vector columns for Hybrid Grace.
- [HIVE-18577](#): SemanticAnalyzer.validate has some pointless metastore calls.
- [HIVE-18587](#): insert DML event may attempt to calculate a checksum on directories.
- [HIVE-18597](#): LLAP: Always package the log4j2 API jar for org.apache.log4j.
- [HIVE-18613](#): Extend JsonSerializer to support BINARY type.
- [HIVE-18626](#): Repl load "with" clause does not pass config to tasks.
- [HIVE-18643](#): don't check for archived partitions for ACID ops.
- [HIVE-18660](#): PCR doesn't distinguish between partition and virtual columns .
- [HIVE-18754](#): REPL STATUS should support 'with' clause.
- [HIVE-18788](#): Clean up inputs in JDBC PreparedStatement.
- [HIVE-18794](#): Repl load "with" clause does not pass config to tasks for non-partition tables.
- [HIVE-18808](#): Make compaction more robust when stats update fails.
- [HIVE-18815](#): Remove unused feature in HPL/SQL.
- [HIVE-18817](#): ArrayIndexOutOfBoundsException exception during read of ACID table..
- [HIVE-18833](#): Auto Merge fails when "insert into directory as orcfile".
- [HIVE-18879](#): Disallow embedded element in UDFXPathUtil needs to work if xercesImpl.jar in classpath.
- [HIVE-18944](#): Grouping sets position is set incorrectly during DPP.

Kafka

This release provides Kafka 1.0.0 and the following Apache patches.

- [KAFKA-4827](#): Kafka connect: error with special characters in connector name.
- [KAFKA-6118](#): Transient failure in kafka.api.SaslScramSslEndToEndAuthorizationTest.testTwoConsumersWithDifferentSaslCredentials.
- [KAFKA-6156](#): JmxReporter can't handle windows style directory paths.

- [KAFKA-6164](#): ClientQuotaManager threads prevent shutdown when encountering an error loading logs.
- [KAFKA-6167](#): Timestamp on streams directory contains a colon, which is an illegal character.
- [KAFKA-6179](#): RecordQueue.clear() does not clear MinTimestampTracker's maintained list.
- [KAFKA-6185](#): Selector memory leak with high likelihood of OOM in case of down conversion.
- [KAFKA-6190](#): GlobalKTable never finishes restoring when consuming transactional messages.
- [KAFKA-6210](#): IllegalArgumentException if 1.0.0 is used for inter.broker.protocol.version or log.message.format.version.
- [KAFKA-6214](#): Using standby replicas with an in memory state store causes Streams to crash.
- [KAFKA-6215](#): KafkaStreamsTest fails in trunk.
- [KAFKA-6238](#): Issues with protocol version when applying a rolling upgrade to 1.0.0.
- [KAFKA-6260](#): AbstractCoordinator not clearly handles NULL Exception.
- [KAFKA-6261](#): Request logging throws exception if acks=0.
- [KAFKA-6274](#): Improve KTable Source state store auto-generated names.

Mahout

In HDP-2.3.x and 2.4.x, instead of shipping a specific Apache release of Mahout, we synchronized to a particular revision point on Apache Mahout trunk. This revision point is after the 0.9.0 release, but before the 0.10.0 release. This provides a large number of bug fixes and functional enhancements over the 0.9.0 release, but provides a stable release of the Mahout functionality before the complete conversion to new Spark-based Mahout in 0.10.0.

The revision point chosen for Mahout in HDP 2.3.x and 2.4.x is from the "mahout-0.10.x" branch of Apache Mahout, as of 19 December 2014, revision 0f037cb03e77c096 in GitHub.

In HDP-2.5.x and 2.6.x, we removed the "commons-httpclient" library from Mahout because we view it as an obsolete library with possible security issues., and upgraded the Hadoop-Client in Mahout to version 2.7.3, the same version used in HDP-2.5. As a result:

- Previously compiled Mahout jobs will need to be recompiled in the HDP-2.5 or 2.6 environment.
- There is a small possibility that some Mahout jobs may encounter "ClassNotFoundException" or "could not load class" errors related to "org.apache.commons.httpclient", "net.java.dev.jets3t", or related class name prefixes. If these errors happen, you may consider whether to manually install the needed jars in your classpath for the job, if the risk of security issues in the obsolete library is acceptable in your environment.
- There is an even smaller possibility that some Mahout jobs may encounter crashes in Mahout's hbase-client code calls to the hadoop-common libraries, due to binary compatibility problems. Regrettably, there is no way to resolve this issue except revert to the HDP-2.4.2 version of Mahout, which may have security issues. Again, this should be very unusual, and is unlikely to occur in any given Mahout job suite.

Oozie

This release provides Oozie 4.2.0 with the following Apache patches.

- [OOZIE-2571](#): Add spark.scala.binary.version Maven property so that Scala 2.11 can be used.
- [OOZIE-2606](#): Set spark.yarn.jars to fix Spark 2.0 with Oozie.
- [OOZIE-2658](#): --driver-class-path can overwrite the classpath in SparkMain.
- [OOZIE-2787](#): Oozie distributes application jar twice making the spark job fail.
- [OOZIE-2792](#): Hive2 action is not parsing Spark application ID from log file properly when Hive is on Spark.
- [OOZIE-2799](#): Setting log location for spark sql on hive.
- [OOZIE-2802](#): Spark action failure on Spark 2.1.0 due to duplicate sharelibs.
- [OOZIE-2923](#): Improve Spark options parsing.
- [OOZIE-3109](#): SCA : Cross-Site Scripting: Reflected.
- [OOZIE-3139](#): Oozie validates workflow incorrectly.
- [OOZIE-3167](#): Upgrade tomcat version on Oozie 4.3 branch.

Phoenix

This release provides Phoenix 4.7.0 and the following Apache patches:

- [PHOENIX-1751](#): Perform aggregations, sorting, etc, in the preScannerNext instead of postScannerOpen.
- [PHOENIX-2714](#): Correct byte estimate in BaseResultIterators and expose as interface.
- [PHOENIX-2724](#): Query with large number of guideposts is slower compared to no stats.
- [PHOENIX-2855](#): Workaround Increment TimeRange not being serialized for HBase 1.2.
- [PHOENIX-3023](#): Slow performance when limit queries are executed in parallel by default.
- [PHOENIX-3040](#): Don't use guideposts for executing queries serially.
- [PHOENIX-3112](#): Partial row scan not handled correctly.
- [PHOENIX-3240](#): ClassCastException from Pig loader.
- [PHOENIX-3452](#): NULLS FIRST/NULL LAST should not impact whether GROUP BY is order preserving.
- [PHOENIX-3469](#): Incorrect sort order for DESC primary key for NULLS LAST/NULLS FIRST.
- [PHOENIX-3789](#): Execute cross region index maintenance calls in postBatchMutateIndispensably.
- [PHOENIX-3865](#): IS NULL does not return correct results when first column family not filtered against.
- [PHOENIX-4290](#): Full table scan performed for DELETE with table having immutable indexes.
- [PHOENIX-4373](#): Local index variable length key can have trailing nulls while upserting.
- [PHOENIX-4466](#): java.lang.RuntimeException: response code 500 - Executing a spark job to connect to phoenix query server and load data.
- [PHOENIX-4489](#): HBase Connection leak in Phoenix MR Jobs.
- [PHOENIX-4525](#): Integer overflow in GroupBy execution.
- [PHOENIX-4560](#): ORDER BY with GROUP BY doesn't work if there is WHERE on pk column.

- [PHOENIX-4586](#): UPSERT SELECT doesn't take in account comparison operators for subqueries..
- [PHOENIX-4588](#): Clone expression also if it's children have Determinism.PER_INVOCATION.

Pig

This release provides Pig 0.16.0 with the following Apache patches.

- [PIG-5159](#): Fix Pig not saving grunt history.
- [PIG-5175](#): Upgrade jruby to 1.7.26.

Ranger

This release provides Ranger 0.7.0 and the following Apache patches:

- [RANGER-1805](#): Code improvement to follow best practices in js.
- [RANGER-1960](#): Take snapshot's table name into consideration for deletion.
- [RANGER-1982](#): Error Improvement for Analytics Metric of Ranger Admin and Ranger KMS.
- [RANGER-1984](#): Hbase audit log records may not show all tags associated with accessed column.
- [RANGER-1988](#): Fix insecure randomness.
- [RANGER-1990](#): Add One-way SSL MySQL support in Ranger Admin.
- [RANGER-2006](#): Fix problems detected by static code analysis in ranger usersync for ldap sync source.
- [RANGER-2008](#): Policy evaluation is failing for multiline policy conditions.

Slider

This release provides Slider 0.92.0 with no additional Apache patches.

Spark

This release provides Spark 2.3.0 and the following Apache patches:

- [SPARK-13587](#): Support virtualenv in pyspark.
- [SPARK-19964](#): Avoid reading from remote repos in SparkSubmitSuite.
- [SPARK-22882](#): ML test for structured streaming: ml.classification.
- [SPARK-22915](#): Streaming tests for spark.ml.feature, from N to Z.
- [SPARK-23020](#): Fix another race in the in-process launcher test.
- [SPARK-23040](#): Returns interruptible iterator for shuffle reader.
- [SPARK-23173](#): Avoid creating corrupt parquet files when loading data from JSON.
- [SPARK-23264](#): Fix scala.MatchError in literals.sql.out.
- [SPARK-23288](#): Fix output metrics with parquet sink.
- [SPARK-23329](#): Fix documentation of trigonometric functions.
- [SPARK-23406](#): Enable stream-stream self-joins for branch-2.3.
- [SPARK-23434](#): Spark should not warn `metadata directory` for a HDFS file path.
- [SPARK-23436](#): Infer partition as Date only if it can be casted to Date.

- [SPARK-23457](#): Register task completion listeners first in ParquetFileFormat.
- [SPARK-23462](#): improve missing field error message in `StructType`.
- [SPARK-23490](#): Check storage.locationUri with existing table in CreateTable.
- [SPARK-23524](#): Big local shuffle blocks should not be checked for corruption.
- [SPARK-23525](#): Support ALTER TABLE CHANGE COLUMN COMMENT for external hive table.
- [SPARK-23553](#): Tests should not assume the default value of `spark.sql.sources.default`.
- [SPARK-23569](#): Allow pandas_udf to work with python3 style type-annotated functions.
- [SPARK-23570](#): Add Spark 2.3.0 in HiveExternalCatalogVersionsSuite.
- [SPARK-23598](#): Make methods in BufferedRowIterator public to avoid runtime error for a large query.
- [SPARK-23599](#): Add a UUID generator from Pseudo-Random Numbers.
- [SPARK-23599](#): Use RandomUUIDGenerator in Uuid expression.
- [SPARK-23601](#): Remove .md5 files from release.
- [SPARK-23608](#): Add synchronization in SHS between attachSparkUI and detachSparkUI functions to avoid concurrent modification issue to Jetty Handlers.
- [SPARK-23614](#): Fix incorrect reuse exchange when caching is used.
- [SPARK-23623](#): Avoid concurrent use of cached consumers in CachedKafkaConsumer (branch-2.3).
- [SPARK-23624](#): Revise doc of method pushFilters in Datasource V2.
- [SPARK-23628](#): calculateParamLength should not return 1 + num of expressions.
- [SPARK-23630](#): Allow user's hadoop conf customizations to take effect.
- [SPARK-23635](#): Spark executor env variable is overwritten by same name AM env variable.
- [SPARK-23637](#): Yarn might allocate more resource if a same executor is killed multiple times.
- [SPARK-23639](#): Obtain token before init metastore client in SparkSQL CLI.
- [SPARK-23642](#): AccumulatorV2 subclass isZero scaladoc fix.
- [SPARK-23644](#): Use absolute path for REST call in SHS.
- [SPARK-23645](#): Add docs RE `pandas_udf` with keyword args.
- [SPARK-23649](#): Skipping chars disallowed in UTF-8.
- [SPARK-23658](#): InProcessAppHandle uses the wrong class in getLogger.
- [SPARK-23660](#): Fix exception in yarn cluster mode when application ended fast.
- [SPARK-23670](#): Fix memory leak on SparkPlanGraphWrapper.
- [SPARK-23671](#): Fix condition to enable the SHS thread pool.
- [SPARK-23691](#): Use sql_conf util in PySpark tests where possible.
- [SPARK-23695](#): Fix the error message for Kinesis streaming tests.
- [SPARK-23706](#): spark.conf.get(value, default=None) should produce None in PySpark.
- [SPARK-23728](#): Fix ML tests with expected exceptions running streaming tests.
- [SPARK-23729](#): Respect URI fragment when resolving globs.
- [SPARK-23759](#): Unable to bind Spark UI to specific host name / IP.
- [SPARK-23760](#): CodegenContext.withSubExprEliminationExprs should save/restore CSE state correctly.
- [SPARK-23769](#): Remove comments that unnecessarily disable Scalastyle check.
- [SPARK-23788](#): Fix race in StreamingQuerySuite.
- [SPARK-23802](#): PropagateEmptyRelation can leave query plan in unresolved state.

- [SPARK-23806](#): Broadcast.unpersist can cause fatal exception when used with dynamic allocation.
- [SPARK-23808](#): Set default Spark session in test-only spark sessions.
- [SPARK-23809](#): Active SparkSession should be set by getOrCreate.
- [SPARK-23816](#): Killed tasks should ignore FetchFailures.
- [SPARK-23822](#): Improve error message for Parquet schema mismatches.
- [SPARK-23823](#): Keep origin in transformExpression.
- [SPARK-23827](#): StreamingJoinExec should ensure that input data is partitioned into specific number of partitions.
- [SPARK-23838](#): Running SQL query is displayed as "completed" in SQL tab.
- [SPARK-23881](#): Fix flaky test JobCancellationSuite."interruptible iterator of shuffle reader".

Sqoop

This release provides Sqoop 1.4.6 with no additional Apache patches.

Storm

This release provides Storm 1.1.1 and the following Apache patches:

- [STORM-2652](#): Exception thrown in JmsSpout open method.
- [STORM-2841](#): testNoAcksIfFlushFails UT fails with NullPointerException.
- [STORM-2854](#): Expose IEventLogger to make event logging pluggable.
- [STORM-2870](#): FileBasedEventLogger leaks non-daemon ExecutorService which prevents process to be finished.
- [STORM-2960](#): Better to stress importance of setting up proper OS account for Storm processes.

Tez

This release provides Tez 0.7.0 and the following Apache patches:

- [TEZ-1526](#): LoadingCache for TezTaskID slow for large jobs.

Zeppelin

This release provides Zeppelin 0.7.3 with no additional Apache patches.

- [ZEPPELIN-3072](#): Zeppelin UI becomes slow/unresponsive if there are too many notebooks.
- [ZEPPELIN-3129](#): Zeppelin UI Doesnt logs out in IE.
- [ZEPPELIN-903](#): Replace CXF with Jersey2.

ZooKeeper

This release provides ZooKeeper 3.4.6 and the following Apache patches:

- [ZOOKEEPER-1256](#): ClientPortBindTest is failing on Mac OS X.

- [ZOOKEEPER-1901](#): [JDK8] Sort children for comparison in AsyncOps tests.
- [ZOOKEEPER-2423](#): Upgrade Netty version due to security vulnerability (CVE-2014-3488).
- [ZOOKEEPER-2693](#): DOS attack on wchp/wchc four letter words (4lw).
- [ZOOKEEPER-2726](#): Patch for introduces potential race condition.

Fixed Common Vulnerabilities and Exposures

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

CVE-2017-7676

Summary: Apache Ranger policy evaluation ignores characters after '*' wildcard character

Severity: Critical

Vendor: Hortonworks

Versions Affected: HDInsight 3.6 versions including Apache Ranger versions 0.5.x/0.6.x/0.7.0

Users affected: Environments that use Ranger policies with characters after '*' wildcard character – like my*test, test*.txt

Impact: Policy resource matcher ignores characters after '*' wildcard character, which can result in unintended behavior.

Fix detail: Ranger policy resource matcher was updated to correctly handle wildcard matches.

Recommended Action: Upgrade to HDI 3.6 (with Apache Ranger 0.7.1+).

CVE-2017-7677

Summary: Apache Ranger Hive Authorizer should check for RWX permission when external location is specified

Severity: Critical

Vendor: Hortonworks

Versions Affected: HDInsight 3.6 versions including Apache Ranger versions 0.5.x/0.6.x/0.7.0

Users affected: Environments that use external location for hive tables

Impact: In environments that use external location for hive tables, Apache Ranger Hive Authorizer should check for RWX permission for the external location specified for create table.

Fix detail: Ranger Hive Authorizer was updated to correctly handle permission check with external location.

Recommended Action: Users should upgrade to HDI 3.6 (with Apache Ranger 0.7.1+).

CVE-2017-9799

Summary: Potential execution of code as the wrong user in Apache Storm

Severity: Important

Vendor: Hortonworks

Versions Affected: HDP 2.4.0, HDP-2.5.0, HDP-2.6.0

Users affected: Users who use Storm in secure mode and are using blobstore to distribute topology based artifacts or using the blobstore to distribute any topology resources.

Impact: Under some situations and configurations of storm it is theoretically possible for the owner of a topology to trick the supervisor to launch a worker as a different, non-root, user. In the worst case, this could lead to secure credentials of the other user being compromised. This vulnerability only applies to Apache Storm installations with security enabled.

Mitigation: Upgrade to HDP-2.6.2.1 as there are currently no workarounds.

CVE-2016-4970

Summary: handler/ssl/OpenSslEngine.java in Netty 4.0.x before 4.0.37.Final and 4.1.x before 4.1.1.Final allows remote attackers to cause a denial of service (infinite loop)

Severity: Moderate

Vendor: Hortonworks

Versions Affected: HDP 2.x.x since 2.3.x

Users Affected: All users that use HDFS.

Impact: Impact is low as Hortonworks does not use OpenSslEngine.java directly in Hadoop codebase.

Recommended Action: Upgrade to HDP 2.6.3.

CVE-2016-8746

Summary: Apache Ranger path matching issue in policy evaluation

Severity: Normal

Vendor: Hortonworks

Versions Affected: All HDP 2.5 versions including Apache Ranger versions 0.6.0/0.6.1/0.6.2

Users affected: All users of the ranger policy admin tool.

Impact: Ranger policy engine incorrectly matches paths in certain conditions when a policy contains wildcards and recursive flags.

Fix detail: Fixed policy evaluation logic

Recommended Action: Users should upgrade to HDP 2.5.4+ (with Apache Ranger 0.6.3+) or HDP 2.6+ (with Apache Ranger 0.7.0+)

CVE-2016-8751

Summary: Apache Ranger stored cross site scripting issue

Severity: Normal

Vendor: Hortonworks

Versions Affected: All HDP 2.3/2.4/2.5 versions including Apache Ranger versions 0.5.x/0.6.0/0.6.1/0.6.2

Users affected: All users of the ranger policy admin tool.

Impact: Apache Ranger is vulnerable to a Stored Cross-Site Scripting when entering custom policy conditions. Admin users can store some arbitrary javascript code execute when normal users login and access policies.

Fix detail: Added logic to sanitize the user input.

Recommended Action: Users should upgrade to HDP 2.5.4+ (with Apache Ranger 0.6.3+) or HDP 2.6+ (with Apache Ranger 0.7.0+)

Fixed Issues for Support

Fixed issues represent selected issues that were previously logged via Hortonworks Support, but are now addressed in the current release. These issues may have been reported in previous versions within the Known Issues section; meaning they were reported by customers or identified by Hortonworks Quality Engineering team.

Incorrect Results

Hortonworks Bug ID	Apache JIRA	Summary
BUG-100019	YARN-8145	yarn rmdadmin -getGroups doesn't return updated groups for user
BUG-100058	PHOENIX-2645	Wildcard characters do not match newline characters
BUG-100266	PHOENIX-3521 , PHOENIX-4190	Results wrong with local indexes
BUG-88774	HIVE-17617 , HIVE-18413 , HIVE-18523	query36 failing, row count mismatch
BUG-89765	HIVE-17702	incorrect isRepeating handling in decimal reader in ORC
BUG-92293	HADOOP-15042	Azure PageBlobInputStream.skip() can return negative value when numberOfPagesRemaining is 0
BUG-92345	ATLAS-2285	UI : Renamed saved search with date attribute.
BUG-92563	HIVE-17495 , HIVE-18528	Aggregate stats in ObjectStore get wrong result
BUG-92957	HIVE-11266	count(*) wrong result based on table statistics for external tables
BUG-93097	RANGER-1944	Action filter for Admin Audit is not working
BUG-93335	HIVE-12315	vectorization_short_regress.q has a wrong result issue for a double calculation
BUG-93415	HIVE-18258 , HIVE-18310	Vectorization: Reduce-Side GROUP BY MERGEPARTIAL with duplicate columns is broken
BUG-93939	ATLAS-2294	Extra parameter "description" added when creating a type

Hortonworks Bug ID	Apache JIRA	Summary
BUG-94007	PHOENIX-1751 , PHOENIX-3112	Phoenix Queries returns Null values due to HBase Partial rows
BUG-94266	HIVE-12505	Insert overwrite in same encrypted zone silently fails to remove some existing files
BUG-94414	HIVE-15680	Incorrect results when hive.optimize.index.filter=true and same ORC table is referenced twice in query
BUG-95048	HIVE-18490	Query with EXISTS and NOT EXISTS with non-equi predicate can produce wrong result
BUG-95053	PHOENIX-3865	IS NULL does not return correct results when first column family not filtered against
BUG-95476	RANGER-1966	Policy engine initialization does not create context enrichers in some cases
BUG-95566	SPARK-23281	Query produces results in incorrect order when a composite order by clause refers to both original columns and aliases
BUG-95907	PHOENIX-3451 , PHOENIX-3452 , PHOENIX-3469 , PHOENIX-4560	Fixing issues with ORDER BY ASC when query has aggregation
BUG-96389	PHOENIX-4586	UPSERT SELECT doesn't take in account comparison operators for subqueries.
BUG-96602	HIVE-18660	PCR doesn't distinguish between partition and virtual columns
BUG-97686	ATLAS-2468	[Basic Search] Issue with OR cases when NEQ is used with numeric types
BUG-97708	HIVE-18817	ArrayIndexOutOfBoundsException exception during read of ACID table.
BUG-97864	HIVE-18833	Auto Merge fails when "insert into directory as orcfile"
BUG-97889	RANGER-2008	Policy evaluation is failing for multiline policy conditions.

Hortonworks Bug ID	Apache JIRA	Summary
BUG-98655	RANGER-2066	Hbase column family access is authorized by a tagged column in the column family
BUG-99883	HIVE-19073 , HIVE-19145	StatsOptimizer may mangle constant columns

Other

Hortonworks Bug ID	Apache JIRA	Summary
BUG-100267	HBASE-17170	HBase is also retrying DoNotRetryIOException because of class loader differences.
BUG-92367	YARN-7558	"yarn logs" command fails to get logs for running containers if UI authentication is enabled.
BUG-93159	OOZIE-3139	Oozie validates workflow incorrectly
BUG-93936	ATLAS-2289	Embedded kafka/zookeeper server start/stop code to be moved out of KafkaNotification implementation
BUG-93942	ATLAS-2312	Use ThreadLocal DateFormat objects to avoid simultaneous use from multiple threads
BUG-93946	ATLAS-2319	UI : Deleting a tag which at 25+ position in the tag list in both Flat and Tree structure needs a refresh to remove the tag from the list.
BUG-94618	YARN-5037 , YARN-7274	Ability to disable elasticity at leaf queue level
BUG-94901	HBASE-19285	Add per-table latency histograms
BUG-95259	HADOOP-15185 , HADOOP-15186	Update adls connector to use the current version of ADLS SDK
BUG-95619	HIVE-18551	Vectorization: VectorMapOperator tries to write too many vector columns for Hybrid Grace
BUG-97223	SPARK-23434	Spark should not warn `metadata directory` for a HDFS file path

Performance

Hortonworks Bug ID	Apache JIRA	Summary
BUG-83282	HBASE-13376 , HBASE-14473 , HBASE-15210 , HBASE-15515 , HBASE-16570 , HBASE-16810 , HBASE-18164	Fast locality computation in balancer
BUG-91300	HBASE-17387	Reduce the overhead of exception report in RegionActionResult for multi()
BUG-91804	TEZ-1526	LoadingCache for TezTaskID slow for large jobs
BUG-92760	ACCUMULO-4578	Cancel compaction FATE operation does not release namespace lock
BUG-93577	RANGER-1938	Solr for Audit setup doesn't use DocValues effectively
BUG-93910	HIVE-18293	Hive is failing to compact tables contained within a folder that is not owned by identity running HiveMetaStore
BUG-94345	HIVE-18429	Compaction should handle a case when it produces no output
BUG-94381	HADOOP-13227 , HDFS-13054	Handling RequestHedgingProxyProvider RetryAction order: FAIL < RETRY < FAILOVER_AND_RETRY.
BUG-94432	HIVE-18353	CompactorMR should call jobclient.close() to trigger cleanup
BUG-94869	PHOENIX-4290 , PHOENIX-4373	Requested row out of range for Get on HRegion for local indexed salted phoenix table.
BUG-94928	HDFS-11078	Fix NPE in LazyPersistFileScrubber
BUG-94964	HIVE-18269 , HIVE-18318 , HIVE-18326	Multiple LLAP fixes
BUG-95669	HIVE-18577 , HIVE-18643	When run update/delete query on ACID partitioned table, HS2 read all each partitions.
BUG-96390	HDFS-10453	ReplicationMonitor thread could stuck for long time due to the race between replication and delete of same file in a large cluster.

Hortonworks Bug ID	Apache JIRA	Summary
BUG-96625	HIVE-16110	Revert of "Vectorization: Support 2 Value CASE WHEN instead of fall back to VectorUDFAdaptor"
BUG-97109	HIVE-16757	Use of deprecated getRows() instead of new estimateRowCount(RelMetadataQuery..) has serious performance impact
BUG-97110	PHOENIX-3789	Execute cross region index maintenance calls in postBatchMutateIndispensably
BUG-98833	YARN-6797	TimelineWriter does not fully consume the POST response
BUG-98931	ATLAS-2491	Update Hive hook to use Atlas v2 notifications

Potential Data Loss

Hortonworks Bug ID	Apache JIRA	Summary
BUG-95613	HBASE-18808	Ineffective config check in BackupLogCleaner#getDeletableFiles()
BUG-97051	HIVE-17403	Fail concatenation for unmanaged and transactional tables
BUG-97787	HIVE-18460	Compactor doesn't pass Table properties to the Orc writer
BUG-97788	HIVE-18613	Extend JsonSerDe to support BINARY type

Query Failure

Hortonworks Bug ID	Apache JIRA	Summary
BUG-100180	CALCITE-2232	Assertion error on AggregatePullUpConstantsRule while adjusting Aggregate indices
BUG-100422	HIVE-19085	FastHiveDecimal abs(0) sets sign to +ve
BUG-100834	PHOENIX-4658	IllegalStateException: requestSeek cannot be called on ReversedKeyValueHeap
BUG-102078	HIVE-17978	TPCDS queries 58 and 83 generate exceptions in vectorization.

Hortonworks Bug ID	Apache JIRA	Summary
BUG-92483	HIVE-17900	analyze stats on columns triggered by Compactor generates malformed SQL with > 1 partition column
BUG-93135	HIVE-15874 , HIVE-18189	Hive query returning wrong results when set hive.groupby.orderby.position.alias to true
BUG-93136	HIVE-18189	Order by position does not work when cbo is disabled
BUG-93595	HIVE-12378 , HIVE-15883	HBase mapped table in Hive insert fail for decimal and binary columns
BUG-94007	PHOENIX-1751 , PHOENIX-3112	Phoenix Queries returns Null values due to HBase Partial rows
BUG-94144	HIVE-17063	insert overwrite partition onto a external table fail when drop partition first
BUG-94280	HIVE-12785	View with union type and UDF to `cast` the struct is broken
BUG-94505	PHOENIX-4525	Integer overflow in GroupBy execution
BUG-95618	HIVE-18506	LlapBaseInputFormat - negative array index
BUG-95644	HIVE-9152	CombineHiveInputFormat: Hive query is failing in Tez with java.lang.IllegalArgumentException exception
BUG-96762	PHOENIX-4588	Clone expression also if it's children have Determinism.PER_INVOCATION
BUG-97145	HIVE-12245 , HIVE-17829	Support column comments for an HBase backed table
BUG-97741	HIVE-18944	Grouping sets position is set incorrectly during DPP
BUG-98082	HIVE-18597	LLAP: Always package the log4j2 API jar for org.apache.log4j
BUG-99849	N/A	Create a new table from a file wizard tries to use default database

Security

Hortonworks Bug ID	Apache JIRA	Summary
BUG-100436	RANGER-2060	Knox proxy with knox-sso is not working for ranger
BUG-101038	SPARK-24062	Zeppelin %Spark interpreter "Connection refused" error, "A secret key must be specified..." error in HiveThriftServer
BUG-101359	ACCUMULO-4056	Update version of commons-collection to 3.2.2 when released
BUG-54240	HIVE-18879	Disallow embedded element in UDFXPathUtil needs to work if xercesImpl.jar in classpath
BUG-79059	OOZIE-3109	Escape log-streaming's HTML-specific characters
BUG-90041	OOZIE-2723	JSON.org license is now CatX
BUG-93754	RANGER-1943	Ranger Solr authorization is skipped when collection is empty or null
BUG-93804	HIVE-17419	ANALYZE TABLE...COMPUTE STATISTICS FOR COLUMNS command shows computed stats for masked tables
BUG-94276	ZEPPELIN-3129	Zeppelin UI does not log out in IE
BUG-95349	ZOOKEEPER-1256 , ZOOKEEPER-1901	Upgrade netty
BUG-95483	N/A	Fix for CVE-2017-15713
BUG-95646	OOZIE-3167	Upgrade tomcat version on Oozie 4.3 branch
BUG-95823	N/A	Knox: Upgrade Beanutils
BUG-95908	RANGER-1960	HBase auth does not take table namespace into consideration for deleting snapshot
BUG-96191	FALCON-2322 , FALCON-2323	Upgrade Jackson and Spring versions to avoid security vulnerabilities
BUG-96502	RANGER-1990	Add One-way SSL MySQL support in Ranger Admin
BUG-96712	FLUME-3194	upgrade derby to the latest (1.14.1.0) version

Hortonworks Bug ID	Apache JIRA	Summary
BUG-96713	FLUME-2678	Upgrade xalan to 2.7.2 to take care of CVE-2014-0107 vulnerability
BUG-96714	FLUME-2050	Upgrade to log4j2 (when GA)
BUG-96737	N/A	Use java io filesystem methods to access local files
BUG-96925	N/A	Upgrade Tomcat from 6.0.48 to 6.0.53 in Hadoop
BUG-96977	FLUME-3132	Upgrade tomcat jasper library dependencies
BUG-97022	HADOOP-14799 , HADOOP-14903 , HADOOP-15265	Upgrading Nimbus-JOSE-JWT library with version above 4.39
BUG-97101	RANGER-1988	Fix insecure randomness
BUG-97178	ATLAS-2467	Dependency upgrade for Spring and nimbus-jose-jwt
BUG-97180	N/A	Upgrade Nimbus-jose-jwt
BUG-98038	HIVE-18788	Clean up inputs in JDBC PreparedStatement
BUG-98353	HADOOP-13707	Revert of "If kerberos is enabled while HTTP SPNEGO is not configured, some links cannot be accessed"
BUG-98372	HBASE-13848	Access InfoServer SSL passwords through Credential Provider API
BUG-98385	ATLAS-2500	Add additional headers to Atlas response.
BUG-98564	HADOOP-14651	Update okhttp version to 2.7.5
BUG-99440	RANGER-2045	Hive table columns with no explicit allow policy are listed with 'desc table' command
BUG-99803	N/A	Oozie should disable HBase dynamic class loading

Stability

Hortonworks Bug ID	Apache JIRA	Summary
BUG-100040	ATLAS-2536	NPE in Atlas Hive Hook
BUG-100057	HIVE-19251	ObjectStore.getNextNotification with LIMIT should use less memory
BUG-100072	HIVE-19130	NPE is thrown when REPL LOAD applied drop partition event.
BUG-100073	N/A	too many close_wait connection from hiveserver to data node
BUG-100319	HIVE-19248	REPL LOAD doesn't throw error if file copy fails.
BUG-100352	N/A	CLONE - RM purging logic scans /registry znode too frequently
BUG-100427	HIVE-19249	Replication: WITH clause is not passing the configuration to Task correctly in all cases
BUG-100430	HIVE-14483	java.lang.ArrayIndexOutOfBoundsException org.apache.orc.impl.TreeReaderFactory\$BytesColumnVectorUtil.commonReadByteArrays
BUG-100432	HIVE-19219	Incremental REPL DUMP should throw error if requested events are cleaned-up.
BUG-100448	SPARK-23637 , SPARK-23802 , SPARK-23809 , SPARK-23816 , SPARK-23822 , SPARK-23823 , SPARK-23838 , SPARK-23881	Update Spark2 to 2.3.0+ (4/11)

Hortonworks Bug ID	Apache JIRA	Summary
BUG-100740	HIVE-16107	JDBC: HttpClient should retry one more time on NoHttpResponseException
BUG-100810	HIVE-19054	Hive Functions replication fails
BUG-100937	MAPREDUCE-6889	Add Job#close API to shutdown MR client services.
BUG-101065	ATLAS-2587	Set read ACL for /apache_atlas/active_server_info znode in HA for Knox proxy to read.
BUG-101093	STORM-2993	Storm HDFS bolt throws ClosedChannelException when Time rotation policy is used
BUG-101181	N/A	PhoenixStorageHandler doesn't handle AND in predicate correctly
BUG-101266	PHOENIX-4635	HBase Connection leak in org.apache.phoenix.hive.mapreduce.PhoenixInputFormat
BUG-101458	HIVE-11464	lineage info missing if there are multiple outputs
BUG-101485	N/A	hive metastore thrift api is slow and causing client timeout
BUG-101628	HIVE-19331	Hive incremental replication to cloud failed.
BUG-102048	HIVE-19381	Hive Function Replication to cloud fails with FunctionTask
BUG-102064	N/A	Hive Replication [onprem to onprem] tests failed in ReplCopyTask
BUG-102137	HIVE-19423	Hive Replication [Onprem to Cloud] tests failed in ReplCopyTask
BUG-102305	HIVE-19430	HS2 and hive metastore OOM dumps
BUG-102361	N/A	multiple insert results in single insert replicated to target hive cluster (onprem - s3)

Hortonworks Bug ID	Apache JIRA	Summary
BUG-87624	N/A	Enabling storm event logging causes workers to continuously die
BUG-88929	HBASE-15615	Wrong sleep time when RegionServerCallable need retry
BUG-89628	HIVE-17613	remove object pools for short, same-thread allocations
BUG-89813	N/A	SCA : Code Correctness: Non-Synchronized Method Overrides Synchronized Method
BUG-90437	ZEPPELIN-3072	Zeppelin UI becomes slow/unresponsive if there are too many notebooks
BUG-90640	HBASE-19065	HRegion#bulkLoadHFiles() should wait for concurrent Region#flush() to finish
BUG-91202	HIVE-17013	Delete request with a subquery based on select over a view
BUG-91350	KNOX-1108	NiFiHaDispatch not failing over
BUG-92054	HIVE-13120	propagate doAs when generating ORC splits
BUG-92373	FALCON-2314	Bump TestNG version to 6.13.1 to avoid BeanShell dependency
BUG-92381	N/A	testContainerLogsWithNewAPI and testContainerLogsWithOldAPI UT fails
BUG-92389	STORM-2841	testNoAcksIfFlushFails UT fails with NullPointerException
BUG-92586	SPARK-17920 , SPARK-20694 , SPARK-21642 , SPARK-22162 , SPARK-22162	Update Spark2 up-to-date to 2.2.1 (Jan. 16)

Hortonworks Bug ID	Apache JIRA	Summary
	22289 , SPAR K- 22373 , SPAR K- 22495 , SPAR K- 22574 , SPAR K- 22591 , SPAR K- 22595 , SPAR K- 22601 , SPAR K- 22603 , SPAR K- 22607 , SPAR K- 22635 , SPAR K- 22637 , SPAR K- 22653 , SPAR K- 22654 , SPAR K- 22686 , SPAR K- 22688 , SPAR K- 22817 , SPAR K- 22862 , SPAR K- 22889 , SPAR K- 22972 , SPAR K- 22975 , SPAR K- 22982 , SPAR	

Hortonworks Bug ID	Apache JIRA	Summary
	K-22983 , SPAR K-22984 , SPAR K-23001 , SPAR K-23038 , SPAR K-23095	
BUG-92680	ATLAS-2288	NoClassDefFoundError Exception while running import-hive script when hbase table is created via Hive
BUG-92760	ACCUMULO-4578	Cancel compaction FATE operation does not release namespace lock
BUG-92797	HDFS-10267 , HDFS-8496	Reducing the datanode lock contentions on certain use cases
BUG-92813	FLUME-2973	Deadlock in hdfs sink
BUG-92957	HIVE-11266	count(*) wrong result based on table statistics for external tables
BUG-93018	ATLAS-2310	In HA, the passive node redirects the request with wrong URL encoding
BUG-93116	RANGER-1957	Ranger Usersync is not syncing users or groups periodically when incremental sync is enabled.
BUG-93361	HIVE-12360	Bad seek in uncompressed ORC with predicate pushdown
BUG-93426	CALCITE-2086	HTTP/413 in certain circumstances due to large Authorization headers
BUG-93429	PHOENIX-3240	ClassCastException from Pig loader
BUG-93485	N/A	Cannot get table mytestorg.apache.hadoop.hive.ql.metadata.InvalidTableException: Table not found when running analyze table on columns in LLAP

Hortonworks Bug ID	Apache JIRA	Summary
BUG-93512	PHOENIX-4466	java.lang.RuntimeException: response code 500 - Executing a spark job to connect to phoenix query server and load data
BUG-93550	N/A	Zeppelin %spark.r does not work with spark1 due to scala version mismatch
BUG-93910	HIVE-18293	Hive is failing to compact tables contained within a folder that is not owned by identity running HiveMetaStore
BUG-93926	ZEPPELIN-3114	Notebooks and interpreters are not getting saved in zeppelin after >1d stress testing
BUG-93932	ATLAS-2320	classification "*" with query throws 500 Internal server exception.
BUG-93948	YARN-7697	NM goes down with OOM due to leak in log-aggregation (part#1)
BUG-93965	ATLAS-2229	DSL search : orderby non-string attribute throws exception
BUG-93986	YARN-7697	NM goes down with OOM due to leak in log-aggregation (part#2)
BUG-94030	ATLAS-2332	Creation of type with attributes having nested collection datatype fails
BUG-94080	YARN-3742 , YARN-6061	Both RM are in standby in secure cluster
BUG-94081	HIVE-18384	ConcurrentModificationException in log4j2.x library
BUG-94168	N/A	Yarn RM goes down with Service Registry is in wrong state ERROR
BUG-94330	HADOOP-13190 , HADOOP-14104 , HADOOP-14814 , HDFS-	HDFS should support for multiple KMS Uris

Hortonworks Bug ID	Apache JIRA	Summary
	10489 , HDFS-11689	
BUG-94345	HIVE-18429	Compaction should handle a case when it produces no output
BUG-94372	ATLAS-2229	DSL query : hive_table name = ["t1","t2"] throws invalid DSL query exception
BUG-94381	HADOOP-13227 , HDFS-13054	Handling RequestHedgingProxyProvider RetryAction order: FAIL < RETRY < FAILOVER_AND_RETRY.
BUG-94432	HIVE-18353	CompactorMR should call jobclient.close() to trigger cleanup
BUG-94575	SPARK-22587	Spark job fails if fs.defaultFS and application jar are different url
BUG-94791	SPARK-22793	Memory leak in Spark Thrift Server
BUG-94928	HDFS-11078	Fix NPE in LazyPersistFileScrubber
BUG-95013	HIVE-18488	LLAP ORC readers are missing some null checks
BUG-95077	HIVE-14205	Hive doesn't support union type with AVRO file format
BUG-95200	HDFS-13061	SaslDataTransferClient#checkTrustAndSend should not trust a partially trusted channel
BUG-95201	HDFS-13060	Adding a BlacklistBasedTrustedChannelResolver for TrustedChannelResolver
BUG-95284	HBASE-19395	[branch-1] TestEndToEndSplitTransaction.testMasterOpsWhileSplitting fails with NPE
BUG-95301	HIVE-18517	Vectorization: Fix VectorMapOperator to accept VRBs and check vectorized flag correctly to support LLAP Caching

Hortonworks Bug ID	Apache JIRA	Summary
BUG-95542	HBASE-16135	PeerClusterZnode under rs of removed peer may never be deleted
BUG-95595	HIVE-15563	Ignore Illegal Operation state transition exception in SQLOperation.runQuery to expose real exception.
BUG-95596	YARN-4126 , YARN-5750	TestClientRMService fails
BUG-96019	HIVE-18548	Fix log4j import
BUG-96196	HDFS-13120	Snapshot diff could be corrupted after concat
BUG-96289	HDFS-11701	NPE from Unresolved Host causes permanent DFSInputStream failures
BUG-96291	STORM-2652	Exception thrown in JmsSpout open method
BUG-96363	HIVE-18959	Avoid creating extra pool of threads within LLAP
BUG-96390	HDFS-10453	ReplicationMonitor thread could stuck for long time due to the race between replication and delete of same file in a large cluster.
BUG-96454	YARN-4593	Deadlock in AbstractService.getConfig()
BUG-96704	FALCON-2322	ClassCastException while submitAndSchedule feed
BUG-96720	SLIDER-1262	Slider functests are failing in Kerberized environment
BUG-96931	SPARK-23053 , SPARK-23186 , SPARK-23230 , SPARK-23358	Update Spark2 up-to-date (Feb. 19)

Hortonworks Bug ID	Apache JIRA	Summary
	K-23376 , SPARK-23391	
BUG-97067	HIVE-10697	ObjectInspectorConvertors#UnionConvertor does a faulty conversion
BUG-97244	KNOX-1083	HttpClient default timeout should be a sensible value
BUG-97459	ZEPPELIN-3271	Option for disabling scheduler
BUG-97511	KNOX-1197	AnonymousAuthFilter is not added when authentication=Anonymous in service
BUG-97601	HIVE-17479	Staging directories do not get cleaned up for update/delete queries
BUG-97605	HIVE-18858	System properties in job configuration not resolved when submitting MR job
BUG-97674	OOZIE-3186	Oozie is unable to use configuration linked using jceks://file/...
BUG-97743	N/A	java.lang.NoClassDefFoundError exception while deploying storm topology
BUG-97756	PHOENIX-4576	Fix LocalIndexSplitMergeIT tests failing in master branch
BUG-97771	HDFS-11711	DN should not delete the block On "Too many open files" Exception
BUG-97869	KNOX-1190	Knox SSO support for Google OIDC is broken.
BUG-97879	PHOENIX-4489	HBase Connection leak in Phoenix MR Jobs
BUG-98392	RANGER-2007	ranger-tagsync's Kerberos ticket fails to renew
BUG-98484	N/A	Hive Incremental Replication to Cloud not working

Hortonworks Bug ID	Apache JIRA	Summary
BUG-98533	HBASE-19934 , HBASE-20008	Hbase snapshot restore is failing due to Null pointer exception
BUG-98555	PHOENIX-4662	NullPointerException in TableResultIterator.java on cache resend
BUG-98579	HBASE-13716	Stop using Hadoop's FSConstants
BUG-98705	KNOX-1230	Many Concurrent Requests to Knox causes URL Mangling
BUG-98983	KNOX-1108	NiFiHaDispatch not failing over
BUG-99107	HIVE-19054	Function replication shall use "hive.repl.replica.functions.root.dir" as root
BUG-99145	RANGER-2035	Errors accessing servicedefs with empty implClass with Oracle backend
BUG-99160	SLIDER-1259	Slider does not work in multi homed environments
BUG-99239	ATLAS-2462	Sqoop import for all tables throws NPE for no table provided in command
BUG-99301	ATLAS-2530	Newline at the beginning of the name attribute of a hive_process and hive_column_lineage
BUG-99453	HIVE-19065	Metastore client compatibility check should include syncMetaStoreClient
BUG-99521	N/A	ServerCache for HashJoin is not re-created when iterators are re-instantiated
BUG-99590	PHOENIX-3518	Memory Leak in RenewLeaseTask
BUG-99618	SPARK-23599 , SPARK-23806	Update Spark2 to 2.3.0+ (3/28)

Hortonworks Bug ID	Apache JIRA	Summary
BUG-99672	ATLAS-2524	Hive hook with V2 notifications - incorrect handling of 'alter view as' operation
BUG-99809	HBASE-20375	Remove use of getCurrentUserCredentials in hbase-spark module

Supportability

Hortonworks Bug ID	Apache JIRA	Summary
BUG-87343	HIVE-18031	Support replication for Alter Database operation.
BUG-91293	RANGER-2060	Knox proxy with knox-sso is not working for ranger
BUG-93116	RANGER-1957	Ranger Usersync is not syncing users or groups periodically when incremental sync is enabled.
BUG-93577	RANGER-1938	Solr for Audit setup doesn't use DocValues effectively
BUG-96082	RANGER-1982	Error Improvement for Analytics Metric of Ranger Admin and Ranger Kms
BUG-96479	HDFS-12781	After Datanode down, In Namenode UI Datanode tab is throwing warning message.
BUG-97864	HIVE-18833	Auto Merge fails when "insert into directory as orcfile"
BUG-98814	HDFS-13314	NameNode should optionally exit if it detects FsImage corruption

Upgrade

Hortonworks Bug ID	Apache JIRA	Summary
BUG-100134	SPARK-22919	Revert of "Bump Apache httpclient versions"
BUG-95823	N/A	Knox: Upgrade Beanutils
BUG-96751	KNOX-1076	Update nimbus-jose-jwt to 4.41.2
BUG-97864	HIVE-18833	Auto Merge fails when "insert into directory as orcfile"

Hortonworks Bug ID	Apache JIRA	Summary
BUG-99056	HADOOP-13556	Change Configuration.getPropsWithPrefix to use getProps instead of iterator
BUG-99378	ATLAS-2461 , ATLAS-2554	Migration utility to export Atlas data in Titan graph DB

Usability

Hortonworks Bug ID	Apache JIRA	Summary
BUG-100045	HIVE-19056	IllegalArgumentException in FixAcidKeyIndex when ORC file has 0 rows
BUG-100139	KNOX-1243	Normalize the required DNs that are Configured in KnoxToken Service
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-100646	ATLAS-2102	Atlas UI Improvements: Search results page
BUG-100737	HIVE-19049	Add support for Alter table add columns for Druid
BUG-100750	KNOX-1246	Update service config in Knox to support latest configurations for Ranger.
BUG-100965	ATLAS-2581	Regression with V2 Hive hook notifications : Moving table to a different database
BUG-84413	ATLAS-1964	UI : Support to order columns in Search table
BUG-90570	HDFS-11384 , HDFS-12347	Add option for balancer to disperse getBlocks calls to avoid NameNode's rpc.CallQueueLength spike
BUG-90584	HBASE-19052	FixedFileTrailer should recognize CellComparatorImpl class in branch-1.x
BUG-90979	KNOX-1224	Knox Proxy HADispatcher to support Atlas in HA.
BUG-91293	RANGER-2060	Knox proxy with knox-sso is not working for ranger
BUG-92236	ATLAS-2281	Saving Tag/Type attribute filter queries with null/not null filters.

Hortonworks Bug ID	Apache JIRA	Summary
BUG-92238	ATLAS-2282	Saved favorite search appears only on refresh after creation when there are 25+ favorite searches .
BUG-92333	ATLAS-2286	Pre-built type 'kafka_topic' should not declare 'topic' attribute as unique
BUG-92678	ATLAS-2276	Path value for hdfs_path type entity is set to lower case from hive-bridge.
BUG-93097	RANGER-1944	Action filter for Admin Audit is not working
BUG-93135	HIVE-15874 , HIVE-18189	Hive query returning wrong results when set hive.groupby.orderby.position.alias to true
BUG-93136	HIVE-18189	Order by position does not work when cbo is disabled
BUG-93387	HIVE-17600	Make OrcFile's "enforceBufferSize" user-settable.
BUG-93495	RANGER-1937	Ranger tagsync should process ENTITY_CREATE notification, to support Atlas import feature
BUG-93512	PHOENIX-4466	java.lang.RuntimeException: response code 500 - Executing a spark job to connect to phoenix query server and load data
BUG-93801	HBASE-19393	HTTP 413 FULL head while accessing HBase UI using SSL.
BUG-93804	HIVE-17419	ANALYZE TABLE...COMPUTE STATISTICS FOR COLUMNS command shows computed stats for masked tables
BUG-93932	ATLAS-2320	classification "*" with query throws 500 Internal server exception.
BUG-93933	ATLAS-2286	Pre-built type 'kafka_topic' should not declare 'topic' attribute as unique
BUG-93938	ATLAS-2283 , ATLAS-2295	UI updates for classifications
BUG-93941	ATLAS-2296 , ATLAS-2307	Basic search enhancement to optionally exclude sub-type entities and sub-classification-types
BUG-93944	ATLAS-2318	UI : Clicking on child tag twice , parent tag is selected

Hortonworks Bug ID	Apache JIRA	Summary
BUG-93946	ATLAS-2319	UI : Deleting a tag which at 25+ position in the tag list in both Flat and Tree structure needs a refresh to remove the tag from the list.
BUG-93977	HIVE-16232	Support stats computation for column in QuotedIdentifier
BUG-94030	ATLAS-2332	Creation of type with attributes having nested collection datatype fails
BUG-94099	ATLAS-2352	Atlas server should provide configuration to specify validity for Kerberos DelegationToken
BUG-94280	HIVE-12785	View with union type and UDF to `cast` the struct is broken
BUG-94332	SQOOP-2930	Sqoop job exec not overriding the saved job generic properties
BUG-94428	N/A	Dataplane Profiler Agent REST API Knox support
BUG-94514	ATLAS-2339	UI : Modifications in "columns" in Basic search result view affects DSL also.
BUG-94515	ATLAS-2169	Delete request fails when hard delete is configured
BUG-94518	ATLAS-2329	Atlas UI Multiple Hovers appears if user click on another tag which is incorrect
BUG-94519	ATLAS-2272	Save the state of dragged columns using save search API.
BUG-94627	HIVE-17731	add a backward compat option for external users to HIVE-11985
BUG-94786	HIVE-6091	Empty pipeout files are created for connection create/close
BUG-94793	HIVE-14013	Describe table doesn't show unicode properly
BUG-94900	OOZIE-2606 , OOZIE-2658 , OOZIE-2787 , OOZIE-2802	Set spark.yarn.jars to fix Spark 2.0 with Oozie
BUG-94901	HBASE-19285	Add per-table latency histograms

Hortonworks Bug ID	Apache JIRA	Summary
BUG-94908	ATLAS-1921	UI : Search using entity and trait attributes : UI doesn't perform range check and allows providing out of bounds values for integral and float data types.
BUG-95086	RANGER-1953	improvement on user-group page listing
BUG-95193	SLIDER-1252	Slider agent fails with SSL validation errors with python 2.7.5-58
BUG-95314	YARN-7699	queueUsagePercentage is coming as INF for getApp REST api call
BUG-95315	HBASE-13947 , HBASE-14517 , HBASE-17931	Assign system tables to servers with highest version
BUG-95392	ATLAS-2421	Notification updates to support V2 data structures
BUG-95476	RANGER-1966	Policy engine initialization does not create context enrichers in some cases
BUG-95512	HIVE-18467	support whole warehouse dump / load + create/drop database events
BUG-95593	N/A	Extend Oozie DB utils to support Spark2 sharelib creation
BUG-95595	HIVE-15563	Ignore Illegal Operation state transition exception in SQLOperation.runQuery to expose real exception.
BUG-95685	ATLAS-2422	Export: Support type-based Export
BUG-95798	PHOENIX-2714 , PHOENIX-2724 , PHOENIX-3023 , PHOENIX-3040	Don't use guideposts for executing queries serially
BUG-95969	HIVE-16828 , HIVE-17063 , HIVE-18390	Partitioned view fails with FAILED: IndexOutOfBoundsException Index: 1, Size: 1
BUG-96019	HIVE-18548	Fix log4j import
BUG-96288	HBASE-14123 , HBASE-14135 , HBASE-17850	Backport Hbase Backup/Restore 2.0
BUG-96313	KNOX-1119	Pac4J OAuth/OpenID Principal Needs to be Configurable

Hortonworks Bug ID	Apache JIRA	Summary
BUG-96365	ATLAS-2442	User with read-only permission on entity resource not able perform basic search
BUG-96479	HDFS-12781	After Datanode down, In Namenode UI Datanode tab is throwing warning message.
BUG-96502	RANGER-1990	Add One-way SSL MySQL support in Ranger Admin
BUG-96718	ATLAS-2439	Update Sqoop hook to use V2 notifications
BUG-96748	HIVE-18587	insert DML event may attempt to calculate a checksum on directories
BUG-96821	HBASE-18212	In Standalone mode with local filesystem HBase logs Warning message:Failed to invoke 'unbuffer' method in class class org.apache.hadoop.fs.FSDataInputStream
BUG-96847	HIVE-18754	REPL STATUS should support 'with' clause
BUG-96873	ATLAS-2443	Capture required entity attributes in outgoing DELETE messages
BUG-96880	SPARK-23230	When hive.default.fileformat is other kinds of file types, create textfile table cause a serde error
BUG-96911	OOZIE-2571 , OOZIE-2792 , OOZIE-2799 , OOZIE-2923	Improve Spark options parsing
BUG-97100	RANGER-1984	Hbase audit log records may not show all tags associated with accessed column
BUG-97110	PHOENIX-3789	Execute cross region index maintenance calls in postBatchMutateIndispensably
BUG-97145	HIVE-12245 , HIVE-17829	Support column comments for an HBase backed table
BUG-97409	HADOOP-15255	Upper/Lower case conversion support for group names in LdapGroupsMapping
BUG-97535	HIVE-18710	extend inheritPerms to ACID in Hive 2.X
BUG-97742	OOZIE-1624	Exclusion pattern for sharelib JARs

Hortonworks Bug ID	Apache JIRA	Summary
BUG-97744	PHOENIX-3994	Index RPC priority still depends on the controller factory property in hbase-site.xml
BUG-97787	HIVE-18460	Compactor doesn't pass Table properties to the Orc writer
BUG-97788	HIVE-18613	Extend JsonSerDe to support BINARY type
BUG-97899	HIVE-18808	Make compaction more robust when stats update fails
BUG-98038	HIVE-18788	Clean up inputs in JDBC PreparedStatement
BUG-98383	HIVE-18907	Create utility to fix acid key index issue from HIVE-18817
BUG-98388	RANGER-1828	Good coding practice-add additional headers in ranger
BUG-98392	RANGER-2007	ranger-tagsync's Kerberos ticket fails to renew
BUG-98533	HBASE-19934 , HBASE-20008	Hbase snapshot restore is failing due to Null pointer exception
BUG-98552	HBASE-18083 , HBASE-18084	Make large/small file clean thread number configurable in HFileCleaner
BUG-98705	KNOX-1230	Many Concurrent Requests to Knox causes URL Mangling
BUG-98711	N/A	NiFi dispatch can't use two-way SSL without service.xml modifications
BUG-98880	OOZIE-3199	Let system property restriction configurable
BUG-98931	ATLAS-2491	Update Hive hook to use Atlas v2 notifications
BUG-98983	KNOX-1108	NiFiHaDispatch not failing over
BUG-99088	ATLAS-2511	Provide options to selectively import database / tables from Hive into Atlas
BUG-99154	OOZIE-2844 , OOZIE-2845 , OOZIE-2858 , OOZIE-2885	Spark query failed with "java.io.FileNotFoundException: hive-site.xml (Permission denied)" exception
BUG-99239	ATLAS-2462	Sqoop import for all tables throws NPE for no table provided in command
BUG-99636	KNOX-1238	Fix Custom Truststore Settings for Gateway

Hortonworks Bug ID	Apache JIRA	Summary
BUG-99650	KNOX-1223	Zeppelin's Knox proxy doesn't redirect /api/ticket as expected
BUG-99804	OOZIE-2858	HiveMain, ShellMain and SparkMain should not overwrite properties and config files locally
BUG-99805	OOZIE-2885	Running Spark actions should not need Hive on the classpath
BUG-99806	OOZIE-2845	Replace reflection-based code which sets variable in HiveConf
BUG-99807	OOZIE-2844	Increase stability of Oozie actions when log4j.properties is missing or not readable
RMP-9995	AMBARI-22222	Switch druid to use /var/druid directory instead of /apps/druid on local disk

Behavioral changes

Apache Component	Apache JIRA	Summary	Details
Spark 2.3	N/A	Changes as documented in the Apache Spark release notes	<ul style="list-style-type: none"> There is a `Deprecation` document and a `Change of behavior` guide https://spark.apache.org/releases/spark-release-2-3-0.html#deprecations For SQL part, there is another detailed `Migration` guide (from 2.2 to 2.3), http://spark.apache.org/docs/latest/sql-programming-guide.html#upgrading-from-spark-sql-22-to-23
Spark	HIVE-12505	Spark job completes successfully but there is an HDFS disk quota full error	<p>Scenario: Running insert overwrite when a quota is set on the Trash folder of the user who runs the command.</p> <p>Previous Behavior: The job succeeds even though it fails to move the data to the Trash. The result can wrongly contain some of the data previously present in the table.</p> <p>New Behavior: When the move to the Trash folder fails, the files are permanently deleted.</p>
Kafka 1.0	N/A	Changes as documented in the Apache Spark release notes	http://kafka.apache.org/10/documentation.html#upgrade_100_notable

Hive/ Ranger		Additional ranger hive policies required for INSERT OVERWRITE	<p>Scenario:Additional ranger hive policies required for INSERT OVERWRITE</p> <p>Previous behavior: Hive INSERT OVERWRITE queries succeed as usual.</p> <p>New behavior: Hive INSERT OVERWRITE queries are unexpectedly failing after upgrading to HDP-2.6.x with the error:</p> <p>Error while compiling statement: FAILED: HiveAccessControlException Permission denied: user jdoe does not have WRITE privilege on /tmp/*(state=42000,code=40000)</p> <p>As of HDP-2.6.0, Hive INSERT OVERWRITE queries require a Ranger URI policy to allow write operations, even if the user has write privilege granted through HDFS policy.</p> <p>Workaround/Expected Customer Action:</p> <ol style="list-style-type: none"> 1. Create a new policy under the Hive repository. 2. In the dropdown where you see Database, select URI. 3. Update the path (Example: /tmp/*) 4. Add the users and group and save. 5. Retry the insert query.
HDFS	N/A	HDFS should support for multiple KMS Uris	<p>Previous Behavior:dfs.encrypted.key.provider.uri property was used to configure the KMS provider path.</p> <p>New Behavior:dfs.encrypted.key.provider.uri is now deprecated in favor of hadoop.security.key.provider.path to configure the KMS provider path.</p>
Zeppelin	ZEPPELIN-3271	Option for disabling scheduler	<p>Component Affected: Zeppelin-Server</p> <p>Previous Behavior: In previous releases of Zeppelin, there was no option for disabling scheduler.</p> <p>New Behavior: By default, users will no longer see scheduler, as it is disabled by default.</p> <p>Workaround/Expected Customer Action: If you want to enable scheduler, you will need to add azeppelin.notebook.cron.enable with value oftrue under custom zeppelin site in Zeppelin settings from Ambari.</p>

Known Issues

- **Spark 2.3**
 - [\[SPARK-23523\]](#)[SQL] Incorrect result caused by the rule OptimizeMetadataOnlyQuery
 - [\[SPARK-23406\]](#) Bugs in stream-stream self-joins
 - Spark sample notebooks are not available when Azure Data Lake Storage (Gen2) is default storage of the cluster.
- **Enterprise Security Package**
 - Spark Thrift Server does not accept connections from ODBC clients.
 - Workaround steps
 - Wait for about 15 minutes after cluster creation.
 - Check ranger UI for existence of hivesampletable_policy.
 - Restart Spark service.
 - STS connection should work now.
- **Workaround for Ranger service check failure**

([RANGER-1607](#): Workaround for Ranger service check failure while upgrading to HDP 2.6.2 from previous HDP versions.

Note

Only when Ranger is SSL enabled. Only when Ranger is SSL enabled.

This issue arises when attempting to upgrade to HDP-2.6.1 from previous HDP versions through Ambari. Ambari uses a curl call to do a service check to Ranger service in Ambari. If the JDK version used by Ambari is JDK-1.7, the curl call will fail with the below error:

curl: (35) error:14077410:SSL routines:SSL23_GET_SERVER_HELLO:sslv3 alert handshake failure

The reason for this error is the tomcat version used in Ranger is Tomcat-7.0.7*. Using JDK-1.7 conflicts with default ciphers provided in Tomcat-7.0.7*.

You can resolve this issue in two ways:

- Update the JDK used in Ambari from JDK-1.7 to JDK-1.8 (see the section [Change the JDK Version](#) in the Ambari Reference Guide).
- If you want to continue supporting a JDK-1.7 environment:
 1. Add the property `ranger.tomcat.ciphers` in the `ranger-admin-site` section in your Ambari Ranger configuration with the below value:
`SSL_RSA_WITH_RC4_128_MD5, SSL_RSA_WITH_RC4_128_SHA, TLS_RSA_WITH_AES_128_CBC_SHA, SSL_RSA_WITH_3DES_EDE_CBC_SHA`
 2. If your environment is configured for Ranger-KMS, add the property `ranger.tomcat.ciphers` in the `ranger-kms-site` section in your Ambari Ranger configuration with the below value:

SSL_RSA_WITH_RC4_128_MD5, SSL_RSA_WITH_RC4_128_SHA,
TLS_RSA_WITH_AES_128_CBC_SHA, SSL_RSA_WITH_3DES_EDE_CBC_SHA

Note

The noted values are working examples and may not be indicative of your environment. Ensure that the way you set these properties matches how your environment is configured.

The noted values are working examples and may not be indicative of your environment. Ensure that the way you set these properties matches how your environment is configured.

- **RangerUI: Escape of policy condition text entered in the policy form**

Component Affected: Ranger

Description of Problem

If a user wants to create policy with custom policy conditions and the expression or text contains special characters, then policy enforcement will not work. Special characters are converted into ASCII before saving the policy into the database.

Special Characters: & < > " ` '

For example, the condition `tags.attributes['type']='abc'` would get converted to the following once the policy is saved: User can see the policy condition with these chars by opening policy in edit mode.

```
tags.attds[&#x27;dsds&#x27;]=&#x27;cssdfs&#x27;
```

You can see the policy condition with these characters by opening the policy in edit mode.

Workaround

Option #1: Create/Update policy via Ranger Rest API

REST URL: `http://<host>:6080/service/plugins/policies`

Creating policy with policy condition:

The below example will create policy with tags as `tags-test` and assign it to `public` group with policy condition `astags.attr['type']=='abc'` by selecting all hive component permission like `select,update,create,drop,alter,index,lock,all`.

Example:

```
curl -H "Content-Type: application/json" -X POST
http://localhost:6080/service/plugins/policies -u admin:admin -d
'{"policyType": "0", "name": "P100", "isEnabled": true, "isAuditEnabled": true, "description":
"", "resources": {"tag": {"values": ["tags-
test"], "isRecursive": "", "isExcludes": false}}, "policyItems": {"groups": ["public"], "condition
s": [{"type": "accessed-after-expiry", "values": []}, {"type": "tag-
expression", "values": ["tags.attr['type']=='abc']}], "accesses": [{"type": "hive:select", "isAll
owed": true}, {"type": "hive:update", "isAllowed": true}, {"type": "hive:create", "isAllowed": t
rue}, {"type": "hive:drop", "isAllowed": true}, {"type": "hive:alter", "isAllowed": true}, {"type":
"hive:index", "isAllowed": true}, {"type": "hive:lock", "isAllowed": true}, {"type": "hive:all", "is
Allowed": true}]}, "denyPolicyItems": [], "allowExceptions": [], "denyExceptions": [], "service
": "tagdev"}
```

Update existing policy with policy condition:

The below example will update policy with tags as `tags-test` and assign it to `public` group with policy condition `astags.attr['type']=='abc'` by selecting all hive component permission like `select,update,create,drop,alter,index,lock,all`.

REST URL: `http://<host-name>:6080/service/plugins/policies/<policy-id>`

Example:

```
curl -H "Content-Type: application/json" -X PUT
http://localhost:6080/service/plugins/policies/18 -u admin:admin -d
'{"id": "18", "guid": "ea78a5ed-07a5-447a-978d-
e636b0490a54", "isEnabled": true, "createdBy": "Admin", "updatedBy": "Admin", "createTi
me": "1490802077000", "updateTime": "1490802077000", "version": "1", "service": "tagdev", "na
me": "P0101", "policyType": "0", "description": "", "resourceSignature": "e5fdb911a25aa7f77a
f5a9546938d9ed", "isAuditEnabled": true, "resources": {"tag": {"values": ["tags"], "isExclude
s": false, "isRecursive": false}}, "policyItems": {"accesses": [{"type": "hive:select", "isAllowed
": true}, {"type": "hive:update", "isAllowed": true}, {"type": "hive:create", "isAllowed": true}, {
"type": "hive:drop", "isAllowed": true}, {"type": "hive:alter", "isAllowed": true}, {"type": "hive
:index", "isAllowed": true}, {"type": "hive:lock", "isAllowed": true}, {"type": "hive:all", "isAllow
ed": true}], "users": [], "groups": ["public"], "conditions": [{"type": "ip-
```

```
range","values":["tags.attributes['type']=abc"]}], "delegateAdmin":false}], "denyPolicyItems": [], "allowExceptions": [], "denyExceptions": [], "dataMaskPolicyItems": [], "rowFilterPolicyItems": []}'
```

Option #2: Apply Javascript changes

Steps to update JS file :

1. Find out PermissionList.js file under /usr/hdp/current/ranger-admin
2. Find out definition of renderPolicyCondition function (line no:404).
3. Remove below line from that function i.e under display function(line no:434)

```
val = _.escape(val);//Line No:460
```

After removing the above line, the Ranger UI will allow you to create policies with policy condition that can contain special characters and policy evaluation will be successful for the same policy.

Deprecation

- **OMS Portal:** We have removed the link from HDInsight resource page that was pointing to OMS portal. Log Analytics initially used its own portal called the OMS portal to manage its configuration and analyze collected data. All functionality from this portal has been moved to the Azure portal where it will continue to be developed. HDInsight has deprecated the support for OMS portal. Customers will use HDInsight Log Analytics integration in Azure portal.
- **Spark 2.3**
 - <https://spark.apache.org/releases/spark-release-2-3-0.html#deprecations>

Upgrading

All of these features are available in HDInsight 3.6. To get the latest version of Spark, Kafka and R Server (Machine Learning Services), please choose the Spark, Kafka, ML Services version when you [create a HDInsight 3.6 cluster](#). To get support for ADLS, you can choose the ADLS storage type as an option. Existing clusters will not be upgraded to these versions automatically.

All new clusters created after June 2018 will automatically get the 1000+ bug fixes across all the open source projects. Please follow [this](#) guide for best practices around upgrading to a newer HDInsight version.