

1.0.0

Cloudera Data Warehouse Runtime Release Notes

Date published: 2024-07-26

Date modified: 2024-08-15

CLOUDERA

<https://docs.cloudera.com/>

Legal Notice

© Cloudera Inc. 2025. All rights reserved.

The documentation is and contains Cloudera proprietary information protected by copyright and other intellectual property rights. No license under copyright or any other intellectual property right is granted herein.

Unless otherwise noted, scripts and sample code are licensed under the Apache License, Version 2.0.

Copyright information for Cloudera software may be found within the documentation accompanying each component in a particular release.

Cloudera software includes software from various open source or other third party projects, and may be released under the Apache Software License 2.0 (“ASLv2”), the Affero General Public License version 3 (AGPLv3), or other license terms. Other software included may be released under the terms of alternative open source licenses. Please review the license and notice files accompanying the software for additional licensing information.

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

Cloudera reserves the right to change any products at any time, and without notice. Cloudera assumes no responsibility nor liability arising from the use of products, except as expressly agreed to in writing by Cloudera.

Cloudera, Cloudera Altus, HUE, Impala, Cloudera Impala, and other Cloudera marks are registered or unregistered trademarks 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 CLOUDERA, CLOUDERA DOES NOT MAKE NOR GIVE ANY REPRESENTATION, WARRANTY, NOR COVENANT OF ANY KIND, WHETHER EXPRESS OR IMPLIED, IN CONNECTION WITH CLOUDERA TECHNOLOGY OR RELATED SUPPORT PROVIDED IN CONNECTION THEREWITH. CLOUDERA DOES NOT WARRANT THAT CLOUDERA PRODUCTS NOR SOFTWARE WILL OPERATE UNINTERRUPTED NOR THAT IT WILL BE FREE FROM DEFECTS NOR ERRORS, THAT IT WILL PROTECT YOUR DATA FROM LOSS, CORRUPTION NOR UNAVAILABILITY, NOR THAT IT WILL MEET ALL OF CUSTOMER’S BUSINESS REQUIREMENTS. WITHOUT LIMITING THE FOREGOING, AND TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, CLOUDERA 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.

Contents

- Cloudera Data Warehouse on premises Runtime Changelog..... 4**
 - Hive changelog for Cloudera Data Warehouse on premises.....4
 - Impala changelog for Cloudera Data Warehouse on premises.....8
 - Iceberg changelog for Cloudera Data Warehouse on premises..... 16
 - Hue changelog for Cloudera Data Warehouse on premises..... 18

Cloudera Data Warehouse on premises Runtime Changelog

Review the changes introduced in this release of Cloudera Data Warehouse on premises Runtime. This includes new features, improvements, and fixes for Hive, Impala, Iceberg, and Hue.

Hive changelog for Cloudera Data Warehouse on premises

Review the changes introduced in Hive for Cloudera Data Warehouse on premises.

2025.0.19.1000

- CDPD-57938:HIVE-29211: add LDAP callback handlers to enforce group filter for
- CDPD-83461: HIVE-29029: ClassCastException when compiling query with STACK UDTF in multiple UNION ALL branches
- DWX-21699:CVE-2023-2976: Update guava version to 32.1.2-jre
- DWX-21698:HIVE-28946:Upgrade commons-io to 2.16.1
- DWX-20754: HIVE-28938: Error in LATERAL VIEW with non native tables due to incorrect virtual columns
- CDPD-88550: HIVE-29118: Remove extra metadata calls if coloring is not needed
- CDPD-83530: HIVE-28962: Prevent committing outputs if an exception was thrown in the Tez processor
- DWX-21200:HIVE-28730: Upgrade aws-java-sdk to 1.12.720
- CDPD-85207:HIVE-28147:Upgrade commons-compress to 1.26.0

2025.0.19.1

- CDPD-77993: match parquet-hadoop version with downstream
- DWX-20989: Apply a workaround to bypass the nmap DNS resolution problem in metastore health check script
- CDPD-80655: HIVE-28878: Query with window function and redundant GROUP BY key fails during compilation
- DWX-20657: HIVE-28886: CalcitePlanner throws NullPointerException when column name case in ColumnStatisticsObj differs from the table schema
- DWX-20723: remove USER env from tez-entripoint.sh script
- CDPD-80424: Upgrade json-smart and resolve CVE-2015-5237 for Hive
- DWX-20242:HIVE-28856: Remove Jetty-Runner
- CDPD-80448: Upgrade com.fasterxml.jackson.core;jackson-core and resolve CVE-2020-36518 for Hive
- CDPD-80428: HIVE-28040: Upgrade netty to 4.1.116.Final due to CVEs
- CDPD-80450: HIVE-28859: Upgrade commons-logging to 1.3.5 to fix CVEs
- CDPD-80433: Upgrade avro and resolve CVE-2024-47561 for Hive
- CDPD-78222: HIVE-28775: HiveServer2: Implement HA Health Check endpoint on a separate port from WebUI
- DWX-20242:Exclude mina-core to fix CVE-2024-52046
- DWX-20179: Set default UTC timezone in chainguard based hive image
- DWX-18878: Chainguard base image update hive-arm
- DWX-20179: Set default UTC timezone in chainguard based hive image
- DWX-18878 (addendum): Fix date command for APP_SUBMIT_TIME_ENV
- DWX-18878: Chainguard base image update: hive
- CDPD-76830: HIVE-28778: Support Secure Zookeeper Access in Hive SecretManager
- DWX-17664 HIVE-28095: Hive Query History
- CDPD-79468: Excluded javax.el to remove CVE-2021-28170
- DWX-20248:Fix CVE-2024-23945 in hive image.

- DWX-20087: Upgrade Spring to 5.3.39-cloudera-5.3.44
- CDPD-78626 Upgrade velocity-engine-core in order to fix commons-io cve
- CDPD-78632: HIVE-28610: Bump commons-io:commons-io from 2.12.0 to 2.14.0
- CDPD-78238: HIVE-28572: Support Distribute by and Cluster by clauses in CBO
- CDPD-71260: HIVE-27831: Set hive.cbo.fallback.strategy to NEVER by default
- CDPD-77934: HIVE-28532: Map Join Reuse cache allows to share hashtables for different join types
- CDPD-77867: HIVE-28589: Not null constraint does not enforced at invalid cast
- CDPD-75066: HIVE-28512 Retain properties in DDL_CTL_PARAMETERS_WHITELIST
- CDPD-74622: CDPD-73371: HIVE-28065: HIVE-27477: Upgrade Bouncy castle to bcprov-jdk18on
- CDPD-77834 (Part 2): Update the golden file of ds-q54.q and re-enable it
- CDPD-77905 HIVE-28700: MRCompactor may cause data loss when performing the major compaction
- CDPD-72237: HIVE-28259: Common table expression detection and rewrites using CBO
- CDPD-77705: HIVE-28408: Support ARRAY field access in CBO
- CDPD-75656: HIVE-28582: OOM when compiling query with many GROUP BY columns aliased multiple times
- HIVE-28018: Don't require HiveConf for JDBC Metadata calls
- HIVE-27874: Support datatype conversion on fetch threads
- HIVE-27872: Support multi-stream parallel fetch in JDBC driver
- HIVE-27873: Fix getOperationStatus and optimize fetch
- HIVE-27887: Provide reasonable defaults for ResultSetMetaData
- HIVE-27886: Configure Hadoop security for Kerberos when required
- HIVE-28315: Missing classes while using hive jdbc standalone jar
- CDPD-77713 HIVE-28669: Deadlock found when TxnStoreMutex trying to acquireLock
- DWX-19970: HIVE-28701: OTEL: Fix race condition due to unavailability of Query Id.
- DWX-19969: Split hive into hive-arm on CDWH to support multi-arch image generation
- DWX-19965: HIVE-28698: OTEL: Add Retry Policy in the Exporter.
- DWX-19627 HIVE-28620: Query result is cached in case of IOWD if the subquery is not trivial
- DWX-19928 HIVE-28623: Implement RuntimeContext to decouple runtime data from TezTask
- DWX-19773: Make commons-configuration dependency of atlas plugin available for hive
- CDPD-73977: Fix dependencies after Jersey upgrade
- CDPD-73371: HIVE-28455: Disable createDependencyReducedPom in maven-shade-plugin to not replace the original with dependency-reduced-pom.xml.
- CDPD-73227 : fixing hive build error
- CDPD-72360 : CDPD-69448: CDPD-67281 manage Jersey update with shaded Atlas - V2
- CDPD-72989: HIVE-27477: Upgrade maven-shade-plugin version from 3.2.4 to 3.5.0
- CDPD-77550: HIVE-28675: Maximize the removal of redundant columns from GROUP BY clauses
- CDPD-44551: HIVE-26952: Include property metastore.storage.schema.reader.impl by default
- DWX-19908: HIVE-28661: OTEL: Latency in retrieving query end time leads to thread interruption
- DWX-19907: HIVE-28660: OTEL: Modify span names to remove inconsistency in case of completed queries
- DWX-19786: HIVE-28648: OTEL: Add Execution Engine as part of the Query Data.
- DWX-19767: HIVE-28613: OTEL: Collect & expose JVM metrics for LLAP daemons.
- DWX-19766: HIVE-28542: OTEL: Implement OTEL Exporter to expose JVM details of HiveServer2
- DWX-19765: HIVE-28505: OTEL: Implement OTEL Exporter to expose query details from HiveServer2
- DWX-19911: HIVE-26564: Separate query live operation log and historical operation log
- DWX-19910: HIVE-24792: Potential thread leak in Operation
- DWX-19909: HIVE-24802: Show operation log at webui
- CDPD-69123: Upgrade Avro to latest 1.11.3
- CDPD-76936: HIVE-28245: Upgrade Spring to 5.3.39
- DWX-19749 HIVE-28638: Refactor stats handling in StatsRecordingThreadPool
- CDPD-76544: Upgrading library mysql-connector-j to version 8.2.0 for fixing CVE-2023-22102
- CDPD-77047: HIVE-28594 Handle the issue on HS2 WebUI's LDAP authentication

- CDPD-76855: HIVE-24657: Make Beeline Logging Dependencies Explicit
- CDPD-76853: HIVE-26638: Replace in-house CBO reduce expressions rules with Calcite's built-in classes
- CDPD-76854: HIVE-26626: Cut dependencies between HiveXxPullUpConstantsRule and HiveReduceExpressionsRule
- CDPD-76733: Upgrading avro library to fix cve CVE-2024-47561
- CDPD-75555: HIVE-28564: AssertionError for jdbc queries when JDBCExpandExpressionsRule produces non flat Filter conditions
- CDPD-75355: HIVE-28560: AssertionError in SqlImplementor when query over JDBC table contains LIMIT
- CDPD-75356: HIVE-24255: StorageHandler with select-limit query is returning 0 rows
- CDPD-76140: HIVE-27399: Add support for lateral views with CBO
- DWX-19668 HIVE-28616: OrcReader is not closed in OrcEncodedDataReader
- CDPD-71682: Backport HIVE-25719: Disable replication_metrics_ingest.q
- CDPD-71682: Backport HIVE-28354: Fix naming in negative LLAP CLI driver config
- CDPD-71682: Backport HIVE-25036: Disable script_broken_pipe2
- CDPD-71682: Backport HIVE-27992: Upgrade to tez 0.10.3
- CDPD-71682: Backport HIVE-24673: Migrate NegativeCliDriver and NegativeMinimrCliDriver to llap
- DWX-19635: HIVE-27357: Map-side SMB Join returns incorrect result when tables have different bucket size
- CDPD-75377: HIVE-28604 Allow HMS to configure the DataNucleus level 1 cache
- DWX-18611: Query-coordinator fails to start in WaitForLlap if custom jars are present
- CDPD-74482: HIVE-26968: Wrong results when shared work optimizer merges TS operator with different DPP edges
- CDPD-74481: HIVE-27375: SharedWorkOptimizer: Fix MapJoin cache reuse
- CDPD-74531: Metadata Summary creation fails for 7.3.1.0 based PVC runs with Oracle database
- CDPD-75809: HIVE-28571: Basic UNIONTYPE support in CBO
- DWX-18611: Query-coordinator fails to start in WaitForLlap if custom jars are present
- CDPD-74817 HIVE-28544: Ensure using UTF-8 encoding in some String/Char/Varchar related operations
- CDPD-75665: HIVE-28580 Create table using import command gives ddl pointing to incorrect location
- CDPD-75048 HIVE-13153: SessionID is appended to thread name twice
- CDPD-75152 HIVE-28569: ReExecuteLostAMQueryPlugin doesn't kick in if the message contains line break and diagnostics
- CDPD-75457: Backport HIVE-28249
- CDPD-58931: HIVE-28573: Unable to execute HPLSQL queries when hive authorization is enabled.
- CDPD-75556: Call createDmlFragment() when inserting into a partitioned table
- DWX-19504: HIVE-22948: QueryCache: Treat query cache locations as temporary storage
- CDPD-44060: HIVE-28416: Another process is rebuilding the materialized view is thrown for subsequent MV rebuilds when the current MV rebuild is aborted for the same MV.
- CDPD-75398: hive build failure in docker build
- CDPD-74292: HIVE-28540: Special characters in user DN escaped when querying LDAP
- DWX-19377: HIVE-28533: Fix compaction with custom pools
- CDPD-73714: HIVE-28483: CAST string to date should return null when format is invalid
- CDPD-74764: HIVE-28545 (Addendum) Remove global lock on HMSHandler.getMSForConf which is prone to deadlock
- CDPD-74764: HIVE-28545: Remove global lock on HMSHandler.getMSForConf which is prone to deadlock
- CDPD-57243: HIVE-28511: add column access information to HivePrivilegeObject
- CDPD-74640: HIVE-25907: IOW Directory queries fails to write data to final path when query result cache is enabled
- CDPD-74640: HIVE-21386: Extend the fetch task enhancement done in HIVE-21279 to make it work with query result cache
- CDPD-74539: Make MariaDB fall back to MySQL
- CDPD-74640: HIVE-28530: Set files in thread safe manner in HiveSequenceFileInputFormat
- CDPD-74479: HIVE-28435: Upgrade cron-utils to 9.2.1

- CDPD-74225: HIVE-28480: Disable SMB on partition hash generator mismatch across join branches in previous RS
- DWX-17905 HIVE-28165: HiveSplitGenerator: send splits through filesystem instead of RPC in case of big payload
- DWX-19345: HIVE-28502: Refactor method names that start with capital letters in PasswdAuthenticationProvider class
- CDPD-74278: HIVE-27080: Support project pushdown in JDBC storage handler even when filters are not pushed
- CDPD-74234: HIVE-28264: OOM/slow compilation when query contains SELECT clauses with nested expressions
- CDPD-71484 HIVE-27884: LLAP: Reuse FileSystem objects from cache across different tasks in the same LLAP daemon
- CDPD-74337: HIVE-26580: SHOW COMPACTIONS should support ordering and limiting functionality
- CDPD-72985: HIVE-27984: Addendum: Support backward compatibility of hms thrift struct about column stats
- DWX-19301: HIVE-28457: HS2 WEBUI: LDAP authentication
- CDPD-72985: HIVE-27984: Support backward compatibility of hms thrift struct about column stats
- CDPD-74205: HIVE-28484: SharedWorkOptimizer leaves residual unused operator tree that send DPP events to unknown operators
- HIVE-27405: Throw out the detail error Invalid partition name to the clients
- CDPD-73727: HIVE-28321: Support select alias in the having clause for CBO
- CDPD-23041: HIVE-24815: Remove IDXS Table from Metastore Schema
- CDPD-67834:HIVE-28375:Upgrade Nimbus-JOSE-JWT to 9.37.3
- CDPD-74120: HIVE-28333: Skip CBO for UNIQUEJOIN and TABLESAMPLE (#5307)
- CDPD-72007: HIVE-28310: Disable hive.optimize.join.disjunctive.transitive.predicates.pushdown by default
- CDPD-73669: HIVE-28456: ObjectStore updatePartitionColumnStatisticsInBatch can cause connection starvation
- DWX-18831: Enable Multi-arch docker builds for Hive on cdwh stack
- CDPD-73304: HIVE-27730: Bump org.xerial.snappy:snappy-java from 1.1.10.1 to 1.1.10.4
- CDPD-73303: HIVE-28313: Support ambiguous column references when CBO fallback is disabled
- CDPD-71596: HIVE-28409: Column lineage when creating view is missing if atlas HiveHook is set
- CDPD-72968: Add ifExists field to dropCatalogRequest
- CDPD-73269: HIVE-28431: Fix RexLiteral to ExprNode conversion if the literal is an empty string
- CDPD-73046 HIVE-28028: Remove duplicated proto reader/writer classes introduced in HIVE-19288
- CDPD-70951:Upgrade Aircompressor to 0.27 due to CVE-2024-36114
- CDPD-66081:derby version to 10.14.3.0-cloudera1
- CDPD-62654: HIVE-27746: Hive Metastore should send single AlterPartitionEvent with list of partitions
- CDPD-72840: HIVE-28143: After HIVE-27492 fix, some HPLSQL built-in functions like trim, lower are not working when used in insert statement.
- CDPD-72492: HIVE-28270: Fix missing partition paths bug on drop_database
- DWX-18297: Moving hms-module build to be part of full-build profile
- CDPD-72605: HIVE-28371: Optimize add partitions authorization in HiveMetaStore
- CDPD-72491: HIVE-28338: HiveMetaStoreClient connection count is not correct
- CDPD-72470: HIVE-28314: Support non-boolean WHERE conditions in CBO
- CDPD-72495: HIVE-28302: Let SUM UDF return NULL when all rows have non-numeric texts
- CDPD-13332: Backport HIVE-25532: Fixing authorization for Kill Query command.
- CDPD-13332: Backport HIVE-24751: Kill trigger in workload manager fails with "No privilege" exception when authorization is disabled
- CDPD-13332: HIVE-23084: Implement kill query in multiple HS2 environment
- CDPD-71122: HIVE-28286: Add filtering support for get_table metas
- CDPD-71687 HIVE-27972: Set 'tez' as default value in hive.execution.engine
- CDPD-71687 HIVE-24707: Apply Sane Default for Tez Containers as Last Resort
- CDPD-67597:Upgrade postgresql version in packaging

- CDPD-71394 HIVE-28324: HIVE_CLUSTER_ID in env: unified way to mark a cluster - addendum empty string default
- CDPD-71394 HIVE-28324: HIVE_CLUSTER_ID in env: unified way to mark a cluster
- CDPD-71387: HIVE-28325: Slow compilation without owner information with Ranger
- CDPD-70956: HIVE-28285: Exception when querying JDBC tables with Hive/DB column types mismatch
- CDPD-67649: HIVE-28098: Fails to copy empty column statistics of materialized CTE
- CDPD-71104: HIVE-28082: HiveAggregateReduceFunctionsRule could generate an inconsistent result
- CDPD-67648: HIVE-28080: Propagate statistics from a source table to the materialized CTE
- CDPD-71103: HIVE-28254: CBO (Calcite Return Path): Multiple DISTINCT leads to wrong results
- CDPD-71184 HIVE-28312: Save query error message on Driver level
- CDPD-71188 HIVE-27023: Add setting to prevent tez session from being opened during startup
- DWX-18549: ASYNC_PROFILE_HOME env var points to invalid location
- DWX-18550 HIVE-28305: ProfileServlet: add html to output formats and prepare for profiler 3.0
- CDPD-71107: Update some defaults in download_dep_tars.sh
- CDPD-69970: HIVE-27498: Support custom delimiter in SkippingTextInputFormat
- CDPD-64754: HIVE-28202: Incorrect projected column size after ORC upgrade to v1.6.7
- DWX-18331: HIVE-28121: Use direct SQL for transactional altering table parameter
- CDPD-69326: HIVE-28287: Attempt make the scratch directory writable before failing.
- CDPD-69080: HIVE-28190: Fix MaterializationRebuild lock heartbeat
- CDPD-66418: Backport HIVE-27481: Addendum: Fix post-refactor issues
- CDPD-64851: Backport HIVE-27481: TxnHandler cleanup
- CDPD-64654: Backport HIVE-27406: CompactionTxnHandler cleanup
- CDPD-43503: HIVE-26566: Upgrade H2 database version to 2.1.214
- CDPD-69885: HIVE-28253: Unable to set the value for hplsql.onerror in hplsql mode.
- CDPD-69724: HIVE-28247: Execute immediate `select count(\*) from tbl` throwing ClassCastException in hplsql mode.
- CDPD-69648: HIVE-28215: Signalling CONDITION HANDLER is not working in HPLSQL.
- CDPD-69647: HIVE-28214: HPLSQL not using the hive variables passed through beeline using --hivevar option
- CDPD-62373: HIVE-27755: Quote identifiers in SQL emitted by SchemaTool for MySQL
- CDPD-24384 Backport HIVE-28173: Fixed staging dir issues on materialized views on HDFS encrypted tables
- DWX-17033: Fixed NPE for UA non BinaryPredicate join condition.
- CDPD-69227 HIVE-28212: MiniHS2: use a base folder which is more likely writable on the local FS

Impala changelog for Cloudera Data Warehouse on premises

Review the changes introduced in Impala for Cloudera Data Warehouse on premises.

2025.0.19.1000

- CDPD-89883: IMPALA-14289: Suppress data race in ThreadTokenAvailableCb
- CDPD-91121: IMPALA-14466: Remote client should not cache admissionD's IP when retrying AdmitQuery RPC
- IMPALA-13680: Avoid flush() when closing TSSLSocket
- IMPALA-14433: Fix OpenTelemetry Tracing Deadlock
- CDPD-89834: IMPALA-14392: Fix a crash in AdmissionD in GetQueryStatus PrintId()
- CDPD-89351: Fix test_admission_state_map_mem_leak fail in CDWH-2025.0.19.1000
- IMPALA-14312: Fix Issues with ssl_minimum_version Flag Validator
- CDPD-88743: IMPALA-14276: Fix memory leak by removing AdmissionState on rejection
- IMPALA-14280: (Addendum) Waits for updating active catalogd address
- IMPALA-14283: Invalidate the cache when served by a new catalogd
- IMPALA-14227: (Addendum) Add more tests for catalogd HA warm failover

- IMPALA-14220 (part 2): Delay AcceptRequest until catalog is stable
- IMPALA-14227: In HA failover, passive catalogd should apply pending HMS events before being active
- IMPALA-14221: Avoid hard-coding table list in TestWarmupCatalog
- IMPALA-14214: Compile OpenTelemetry-cpp Against STDLIB
- IMPALA-14217: Fixes Incompatibility with OpenSSL 1.0
- IMPALA-13235: [Patch 3 of 5] - Consume OpenTelemetry C++ SDK
- IMPALA-13237: [Patch 2 of 5] - Add OpenSSL Utility Function to Validate PEM Bundles
- IMPALA-13235: Add Reusable GFlag Validators
- IMPALA-13235: Consume Latest Toolchain Builds
- IMPALA-13952: Update curl version to 8.14.1
- IMPALA-14134: Switch to newer versions of zlib / cloudflare zlib
- IMPALA-14067: Bump glog version to 0.6.0 in Impala
- IMPALA-13479: Patch gperftools to remove 1GB limit on thread caches
- IMPALA-14138: Manually disable block location loading via Hadoop config
- CDPD-81013 IMPALA-14194: Fix empty file handling with recent Ozone versions
- CDPD-54682: IMPALA-12057: Track removed coordinators to reject queued queries early
- IMPALA-14220: CatalogServer::IsActive must not hold catalog_lock_
- IMPALA-14223: Cleanup subdirectories in INSERT OVERWRITE
- IMPALA-14224: Cleanup subdirectories in TRUNCATE
- IMPALA-13778: Update ignored_dir_prefix_list
- CDPD-83059: IMPALA-14074: Warmup metadata cache in catalogd for critical tables
- CDPD-77579: IMPALA-13813: OAuth/JWT Avoid key verification on every rpc call
- IMPALA-14130: Remove wait_num_tables arg in start-impala-cluster.py
- IMPALA-13802: Ignore error during postgres init
- IMPALA-14125: Avoid downloading maven from archive.apache.org
- IMPALA-13802: move postgresql init to the end of bin/bootstrap_system.sh

2025.0.19.1

- IMPALA-14070: Use checkedMultiply in SortNode.java
- CDPD-82599: IMPALA-14006: Bound max_instances in CreateInputCollocatedInstances
- CDPD-82303: IMPALA-14001: Start EXEC_TIME_LIMIT_S timer after backend execution begins
- CDPD-82600: IMPALA-13991: Skip CROSS_JOIN rewrite if subquery is in disjunctive
- DWX-20913: IMPALA-14029: Add Kerberos utilities to Docker image build
- IMPALA-13873: Missing equivalence conjunct in aggregation node with inline views
- IMPALA-13987: Fix stress_catalog_init_delay_ms check in RELEASE
- IMPALA-13850 (part 1): Wait until CatalogD active before resetting
- CDPD-79495: IMPALA-13759: Fix Hive ACID INSERT OVERWRITE base detection
- CDPD-81690: Ensure that OpenSSL is not using a FIPS profile for Chainguard
- IMPALA-13937: Use simpler chmod syntax to set +t on /var/tmp in Docker build
- DWX-18882: Rebase Impala Docker images on Chainguard for release builds
- IMPALA-13825: Extend Docker container build to custom base images
- IMPALA-13826: Migrate from imp to importlib in the config generator
- IMPALA-13881: Fix Workload Management Statement Expression Limit Exceeded Errors
- IMPALA-13812: Fail query for certain errors related to AI functions
- IMPALA-13724: Add hostnames for Docker host and gateway to Impala containers
- CDPD-79675: Use the toolchain mirror in us-west-2 for stack builds
- CDPD-79500: Change internal PyPI mirror to Nexus for Releng builds
- IMPALA-13671: Add Additional Debug Output
- IMPALA-13817: Impala fails to start if 'ai_endpoint' and 'ai_additional_platforms' are not set in the right order
- IMPALA-13792: Cross compile AI functions

- IMPALA-13798: Cleanup host-level remote scratch dir on shutdown
- IMPALA-13789: Defer creating Path objects in loading file metadata
- IMPALA-13783: Fix huge temp list created in catalog-update thread
- CDPD-79908: IMPALA-13804: Use redacted statement in live table
- IMPALA-13803: Fix hiveserver2_protocol_version Values in Workload Management
- IMPALA-13786: Skip rewriting expr of Hive auto-generated label
- IMPALA-13772: Fix Workload Management DMLs Timeouts
- IMPALA-13771: Fix heap-use-after-free in Cluster Membership Manager
- CDPD-78626: IMPALA-13740: (Addendum) Set velocity-engine-core.version downstream
- CDPD-78626: IMPALA-13740: Update velocity-engine-core to 2.4.1
- IMPALA-12785: Add commands to control event-processor status (#387)
- CDPD-79424: Fix missing execute_query_using_vector
- IMPALA-13201: System Table Queries Execute When Admission Queues are Full
- IMPALA-13736: Fix Use-After-Free in ExecutorGroup.RemoveExecutor
- CDPD-78505: IMPALA-13703: Cancel running queries before shutdown deadline
- CDPD-78024: IMPALA-13677: Support remote scratch directory cleanup at Impala daemon startup
- IMPALA-13594: Read Puffin stats also from older snapshots
- IMPALA-12648: Add KILL QUERY statement
- CDPD-75589: Update the golden files for ExternalFeTest
- CDPD-78207: Disable ESTIMATE_DUPLICATE_IN_PREAGG in downstream
- IMPALA-13678: Validate remote_submit_time against coordinator time
- IMPALA-13500: (Addendum) Do not assert CatalogD's partition id
- IMPALA-13666: Provide a non-null fileMetadataStats for HdfsPartition
- IMPALA-2945: Account for duplicate keys on multiple nodes preAgg
- IMPALA-13644: Generalize and move getPerInstanceNdvForCpuCosting
- IMPALA-13565: Add general AI platform support to ai_generate_text
- IMPALA-13253: Add option to enable keepalive for client connections
- IMPALA-13592: Set IV length before setting IV in OpenSsl
- IMPALA-13656: MERGE redundantly accumulates memory in HDFS WRITER
- IMPALA-13288: OAuth AuthN Support for Impala
- IMPALA-13662: Bump the ARM toolchain to support ARM builds for RHEL 9
- CDPD-77932: Use Ozone from CDP rather than the Ozone parcel
- CDPD-77710 (Part 3): Provide a non-null fileMetadataStats for HdfsPartition
- IMPALA-13154: Update metrics when loading an HDFS table
- IMPALA-13618: (Addendum) Remove IMPALA_COMMONS_IO_VERSION downstream
- CDPD-77821: IMPALA-13664: Lower datanucleus.connectionPool.maxPoolSize to 20
- IMPALA-13655: UPDATE redundantly accumulates memory in HDFS WRITER
- IMPALA-13403: Refactor the checks of skip reloading file metadata for ALTER_TABLE events
- IMPALA-13126: Obtain table read lock in EP to process partitioned event
- IMPALA-13458: Fix installing curl on Red Hat variants for dockerised tests
- CDPD-77756: Fix incorrect conflict resolution of IMPALA-13620 (#320)
- IMPALA-13361: Add INSERT * and UPDATE SET * syntax for MERGE statement
- IMPALA-13620: Refresh compute_table_stats.py script
- IMPALA-13086: Lower AggregationNode estimate using stats predicate
- IMPALA-13518: Show target name of COMMIT_TXN events in logs
- IMPALA-13305: Better thrift compatibility checks based on pyparsing
- IMPALA-13465: Trace TupleId further to reduce Agg cardinality
- IMPALA-13638: Translate apostrophe to underscore in Prometheus metric names.
- IMPALA-13641: Lazily init Parquet column read counters
- CDPD-77637: Add back a public static method removed in IMPALA-13622

- IMPALA-13637: Add `ENABLE_TUPLE_ANALYSIS_IN_AGGREGATE` option
- IMPALA-12141: EP shouldn't fail while releasing write lock if the lock is not held previously
- IMPALA-13362: Implement `WHEN NOT MATCHED BY SOURCE` syntax for `MERGE` statement
- IMPALA-13596: Add warnings and exceptions to reading of fair-scheduler file.
- IMPALA-13622: Fix negative cardinality bug in `AggregationNode.java`
- IMPALA-13628: Use `impala::Thread` in `PeriodicCounterUpdater`
- IMPALA-13619: (Addendum) Set `commons-lang3` downstream
- IMPALA-13526: Fix Agg node creation order in `DistributedPlanner`
- IMPALA-13551: Produce the shell tarball by pip installing `impala-shell`
- IMPALA-13368: Fixup Redhat detection for Python ≥ 3.8
- IMPALA-13064: Install service for linux packaging
- IMPALA-13588: Update Puffin reading doc after IMPALA-13370
- IMPALA-13590: Use `CacheLineAligned` instead of `CACHELINE_ALIGNED` for `PerFilterState`
- IMPALA-13598: `OPTIMIZE` redundantly accumulates memory in `HDFS WRITER`
- IMPALA-13544: Addendum: fixed assert message.
- IMPALA-13536: Fix Workload Management Init with Catalog HA
- IMPALA-13597: Upgrade `critique-gerrit-review.py` to Python3
- IMPALA-13535: Add script to restore stats on `PlannerTest`
- IMPALA-889: Add `trim()` function matching ANSI SQL definition
- IMPALA-13589: `SELECT INPUT__FILE__NAME` can crash Impala
- IMPALA-13585: Make `pip_download.py` interruptible
- IMPALA-13567: Update `RowsRead` counter more frequently
- IMPALA-13448: Log cause when failing to flush lineage events, audit events or profiles
- IMPALA-13370: Read Puffin stats from `metadata.json` property if available
- IMPALA-13511: Addendum, fixed wrong case statement
- IMPALA-13556: Log `GetRuntimeProfile` and `GetExecSummary` at `VLOG_QUERY`
- IMPALA-13540: Calcite planner: fix wrong results for set operators
- IMPALA-13558: Workaround Python 2 tarfile issue by patching `tarfile.py`
- IMPALA-13511: Calcite planner: support sub-second datetime parts
- IMPALA-13513: Support decode function
- IMPALA-13543: `single_node_perf_run.py` must accept `tpcds_partitioned`
- IMPALA-13179: Make non-deterministic functions ineligible for tuple caching
- IMPALA-5792: Eliminate duplicate beeswax python code
- IMPALA-13541: Calcite planner, declare `explicit_cast` in operator table.
- IMPALA-13516: Fix handling of cast functions
- IMPALA-13509: Copy rows directly to `OutboundRowBatch` during hash partitioning
- IMPALA-12957: Support reading Inf and NaN from JSON
- IMPALA-13482: Bug fixes for lag/coalesce in analytic function.
- IMPALA-13502: Clean up around constructors
- IMPALA-13148: Show the number of in-progress Catalog operations
- IMPALA-13507: Allow disabling glog buffering via `with_args` fixture
- IMPALA-13327: Fix wrong unpack dir name for Apache HBase
- IMPALA-13441: Support explain statements in Impala planner
- CDPD-75833: Make `impala-config-branch.sh` work with ZSH
- IMPALA-12390 (part 4): Enable unnecessary-value-param
- IMPALA-12737: Query columns in workload management tables.
- IMPALA-13340: Fix missing partitions in `COPY TESTCASE` of LocalCatalog mode
- IMPALA-13497: Add `TupleCacheBytesWritten/Read` to the profile
- IMPALA-12758: Fix catalogd not setting `prev_id` for reloaded partitions
- IMPALA-13477: Set `request_pool` in `QueryStateRecord` for CTAS query

- CDPD-75628: Replace the datasketches.version property in java/pom.xml
- IMPALA-13395: Adds USE_APACHE_COMPONENTS=true in all-build-options job
- IMPALA-13461: Added rules to make tpch queries work.
- IMPALA-13247: Support Reading Puffin files for the current snapshot
- IMPALA-13445: Ignore num partition for unpartitioned writes
- IMPALA-13462: Added support for functions used in tpch
- IMPALA-13459: Handle duplicate table in same query
- IMPALA-13405: Do tuple analysis to lower AggregationNode cardinality
- IMPALA-13455: Put expressions in CNF form for performance.
- IMPALA-13412: Use the BYTES type for tuple cache entries-in-use-bytes
- IMPALA-13430: Too many RelNodes created for "IN" literals
- IMPALA-13446: Bump CDP GBN to 58457853 to get Ranger improvements
- IMPALA-11298: Allow proxy users to share hs2 session from different hosts or realms
- IMPALA-13393: Remove old javax.el config
- IMPALA-12908: (Addendum) use RUNTIME_FILTER_WAIT_TIME_MS for tuple cache TPC testing
- IMPALA-13429: Calcite planner crashing with outer join
- IMPALA-11943: Mark utf8 string functions with IR_ALWAYS_INLINE
- IMPALA-12146: Fix incorrect host memory reserved when the executor quits abnormally
- IMPALA-12216: Print timestamp for impala-shell errors
- IMPALA-13426: Log Java debug sleeps at debug
- IMPALA-13197: Implement Analytic Exprs for Calcite
- IMPALA-13181: Disable tuple caching for locations with limits
- IMPALA-13408: use a specific flag for the topic prefix cluster identifier.
- IMPALA-13186: Tag query option scope for tuple cache
- IMPALA-13427: Make connect timeout of statestore HA RPC tunable
- IMPALA-11729: Speed up start-impala-cluster.py
- IMPALA-13407: Codegen fails with struct in TOP-N
- IMPALA-13406: Switch to curl 8.10.1 to resolve CVEs
- IMPALA-13185: Include runtime filter source in key
- IMPALA-13396: Unify tmp dir management in CustomClusterTestSuite
- IMPALA-12908: Add correctness check for tuple cache
- IMPALA-12686: Switch to toolchain with basic debug info
- IMPALA-13312: Use client address from X-Forwarded-For Header in Ranger Audit Logs
- IMPALA-13384: Only install gcovr deps for coverage builds
- IMPALA-12939: Bound IMPALA_BUILD_THREADS for cgroups and memory
- IMPALA-11265: Part1: Clear GroupContentFiles once used
- CDPD-74806: Force updating Impala package version
- IMPALA-13022: Added infrastructure for implicit casting of functions
- IMPALA-13302: Restore registering all conjuncts
- CDPD-74442: Provide default implementation for IMPALA-12876
- IMPALA-13233: Improve display of instance-level skew in query timeline
- IMPALA-12876: Add catalogVersion and loaded timestamp in query profiles
- IMPALA-12594: Add flag to tune KrpcDataStreamSender mem estimate
- IMPALA-13322: Fix alter on SystemTables
- IMPALA-13378: Verify tuple ids in descriptor table received in executor side
- IMPALA-13182: Support uploading additional jars
- IMPALA-13371: Avoid throwing exception in FindFileInPath()
- IMPALA-13156: Investigation: Set explicit credential provider for S3 builds
- IMPALA-7086: Cache timezone in *_utc_timestamp()
- IMPALA-11431: Avoid getting stats for synthetic column row__id from HMS (#207)

- IMPALA-13347: Fixes TSAN Thread Leak of Workload Management Thread
- IMPALA-13344: Analyze new rewrite exprs
- CDPD-74199: Produce Java dependency tree
- IMPALA-12165: Add option for split debug information (-gsplit-dwarf)
- IMPALA-12906: Incorporate scan range information into the tuple cache key
- IMPALA-12363: Upgrade RE2 to 2023-03-01
- IMPALA-13328: Fix missing krb5-config in building impala_quickstart_client docker image
- IMPALA-12737: Refactor the Workload Management Initialization Process.
- IMPALA-12867: Filter files to OPTIMIZE based on file size
- IMPALA-13311: Hive3 INSERT failed by ClassNotFoundException: org.apache.tez.runtime.api.Event
- IMPALA-13310 Add the value of the http 'X-Forwarded-For' header to the runtime profile
- IMPALA-13082: Fix downstream build scripts to use IMPALA_JACKSON_VERSION
- IMPALA-13082: Use separate versions for jackson vs jackson-databind
- IMPALA-13317: Enhance tpc_sort_key for wider name support
- IMPALA-13303: FileSystemUtil.listFiles() should handle non-recursive case
- IMPALA-13291: Filter dmesg messages by date
- IMPALA-13274: Filter out illegal output for certain join nodes
- IMPALA-12856: Event processor should ignore processing partition with empty partition values
- IMPALA-13262: Do not always migrate inferred predicates into inline view
- IMPALA-13313: Fix ExpireQueries deadlock
- IMPALA-12954: Implement Sorting capability for Calcite planner
- IMPALA-13196 (part2): fix some urls in template files
- IMPALA-10408: Support build using Apache components
- IMPALA-13296: Check column compatibility earlier for table migration
- IMPALA-13301: Replace the aircompressor.version Property in java/pom.xml
- IMPALA-13301: Upgrade aircompressor to 0.27
- IMPALA-13240: Add gerrit comments for Thrift/FlatBuffers changes
- IMPALA-12947: Implement Calcite Planner Union and Value RelNodes
- IMPALA-13246: Smallify strings during broadcast exchange
- IMPALA-13293: Fix too long wait for initial catalog update
- IMPALA-13207: Update error message for operations on blacklist dbs
- CDPD-73367: Apply hadoop-source.patch
- IMPALA-13294: Add support for long polling to avoid client side wait
- IMPALA-13276: Revise the documentation of 'RUNTIME_FILTER_WAIT_TIME_MS'
- IMPALA-12737: Add HS2 support to the InternalServer class.
- IMPALA-13115: Add query id to error messages
- IMPALA-13071: Update the doc of Impala components
- IMPALA-13252: (Addendum) PrintId cancel query
- IMPALA-13279: Upgrade gcovr to 7.2
- IMPALA-13272: Analytic function of collections can lead to crash
- IMPALA-13252: Always use PrintId for TUniqueId
- IMPALA-13271: Correct the documentation with respect to granting privileges on URI
- IMPALA-13270: Fix IllegalStateException on runtime filter
- IMPALA-13230: Dump stacktrace for impala-shell when it receives SIGUSR1
- IMPALA-13264: Use the toolchain's gcov for bin/coverage_helper.sh
- CDPD-72923: Build and publish debug docker images
- IMPALA-13243: Patch pom.xml with Dropwizard Metrics version for CDP install.
- IMPALA-13189: Reset the database on clearing imported query profiles
- IMPALA-13256: Support more than 2G rows for COUNT(*) on jdbc table
- IMPALA-12771: Impala catalogd events-skipped may mark the wrong number

- IMPALA-12277: Fix NullPointerException for partitioned inserts when partition list is stale
- IMPALA-13077: Fix selectivity estimation for SEMI JOIN
- IMPALA-13252: Consistently use PrintId to print TUniqueId
- IMPALA-13214: Skip wait_until_connected when shell exits
- IMPALA-13243: Update Dropwizard Metrics to 4.2.x
- IMPALA-12857: Add flag to enable merge-on-read even if tables are configured with copy-on-write
- IMPALA-13226: Rename TupleCacheInfo.finalize() to finalizeHash()
- IMPALA-13208: Add cluster id to the membership and request-queue topic names
- IMPALA-13231: Gitignore auto-generated files for ranger
- IMPALA-13194: Fast-serialize position delete records
- IMPALA-13161: Fix column index overflow in DelimitedTextParser
- IMPALA-13209: Optimize ConvertRowBatchTime in ExchangeNode
- IMPALA-13088, IMPALA-13109: Use RoaringBitmap instead of sorted vector of int64s
- CDPD-72675: Impala - Upgrade Apache Derby to internal version
- IMPALA-13001: Support graceful and force shutdown for impala.sh
- IMPALA-13014: Upgrade Maven to 3.9.6
- IMPALA-13193: RuntimeFilter on parquet dictionary should evaluate NULL values
- IMPALA-12786: Optimize count(*) for JSON scans
- IMPALA-13203: Rewrite 'id = 0 OR false' as expected
- IMPALA-13196: Fully qualify urls in www/query_timeline
- IMPALA-9441, IMPALA-13170: Ops listing dbs/tables should handle db not exists
- IMPALA-12093: impala-shell to preserve all cookies
- IMPALA-13028: Strip dynamic link libraries in Linux DEB/RPM packages
- IMPALA-6311: Lower max_filter_error_rate to 10%
- IMPALA-13180: Upgrade postgresql to 42.5.6
- IMPALA-12800: Implement hashCode everywhere
- IMPALA-13106: Support larger imported query profile sizes through compression
- IMPALA-13175: Upgrade Spring Framework to 5.3.37
- IMPALA-13168: Add README file for setting up Trino
- IMPALA-13076 Add pstack and jstack to Impala Redhat docker images
- IMPALA-13136: Refactor AnalyzedFunctionCallExpr (for Calcite)
- IMPALA-13169: Specify cluster id before starting HiveServer2
- IMPALA-12940: Added filtering capability for Calcite planner
- IMPALA-12370: Allow converting timestamps to UTC when writing to Kudu
- IMPALA-13137: Add additional client fetch metrics columns to the queries page
- IMPALA-13159: Fix query cancellation caused by statestore failover
- IMPALA-13150: Possible buffer overflow in StringVal::CopyFrom()
- IMPALA-13152: Avoid NaN, infinite, and negative ProcessingCost
- IMPALA-12921, IMPALA-12985: Support running Impala with locally built Ranger (#76)
- IMPALA-13075: Cap memory usage for ExprValuesCache at 256KB
- IMPALA-13138: Never smallify existing StringValue objects, only new ones during DeepCopy
- IMPALA-12712: Invalidate metadata on table should set better createEventId
- IMPALA-13131: Azure OpenAI API expects 'api-key' instead of 'Authorization' in the request header
- IMPALA-12562: Cast double and float to string with exact precision
- IMPALA-12800: Add cache for isTrueWithNullSlots() evaluation
- IMPALA-11871: Skip permissions loading and check on HDFS if Ranger is enabled
- IMPALA-12800: Use HashMap for ExprSubstitutionMap lookups
- IMPALA-13146: Download NodeJS from native toolchain
- IMPALA-12935: First pass on Calcite planner functions
- IMPALA-13130: Prioritize EndDataStream messages

- IMPALA-13119: Fix cost_ initialization at CostingSegment.java
- IMPALA-13134: DDL hang with SYNC_DDL enabled when Catalogd is changed to standby status
- IMPALA-12705: Add /catalog_ha_info page on Statestore to show catalog HA information
- CDPD-71872: IMPALA-13120: Load failed table without need for manual invalidate (#99)
- CDPD-59574: Update exception path for Avro 1.11 (#52)
- IMPALA-12680: Fix NullPointerException during AlterTableAddPartitions
- IMPALA-13129: Move runtime filter skipping at registerRuntimeFilter
- IMPALA-13111: Fix the calculation of fragment ids for impala-gdb.py
- IMPALA-13057: Incorporate tuple/slot information into tuple cache key
- IMPALA-13108: Update version to 4.5.0-SNAPSHOT
- IMPALA-13107: Don't start query on executor if instance number equals 0
- IMPALA-8042: Assign BETWEEN selectivity for discrete-unique column
- IMPALA-13105: Fix multiple imported query profiles fail to import/clear at once
- IMPALA-13034: Add logs and counters for HTTP profile requests blocking client fetches
- IMPALA-13102: Normalize invalid column stats from HMS
- IMPALA-13083: Clarify REASON_MEM_LIMIT_TOO_LOW_FOR_RESERVATION
- IMPALA-11735: Handle CREATE_TABLE event when the db is invisible to the impala server user
- IMPALA-12800: Skip $O(n^2)$ ExprSubstitutionMap::verify() for release builds
- IMPALA-13020 (part 2): Split out external vs internal Thrift max message size
- IMPALA-13020 (part 1): Change thrift_rpc_max_message_size to int64_t
- IMPALA-12559 (part 2): Fix build issue for different versions of openssl
- IMPALA-12559: Support x5c Parameter for RSA JSON Web Keys
- IMPALA-13038: Support profile tab for imported query profiles
- IMPALA-12607: Bump the GBN and fetch events specific to the db/table from the metastore
- IMPALA-13040: Add waiting mechanism in UpdateFilterFromRemote
- IMPALA-10451: Fix avro table loading failures caused by HIVE-24157
- IMPALA-11499: Refactor UrlEncode function to handle special characters
- IMPALA-12934: Added Calcite parsing files to Impala
- IMPALA-13018: Block push down of conjuncts with implicit casting on base columns for jdbc tables
- IMPALA-13058: Init first_arrival_time_ and completion_time_ with -1
- IMPALA-13061: Create query live as external table
- IMPALA-13054: Avoid revisiting children in QueryStateExpanded
- IMPALA-13009: Fix catalogd not sending deletion updates for some dropped partitions
- IMPALA-12977: add search and pagination to /hadoop-varz
- IMPALA-13044: Upgrade bouncycastle to 1.78
- IMPALA-13031: Enhancing logging for spilling configuration with local buffer directory details
- IMPALA-13049: Add dependency management for log4j2 to use 2.18.0
- IMPALA-12910: Support running TPCH/TPCDS queries for JDBC tables
- IMPALA-13045: Wait for impala_query_live to exist
- IMPALA-12684: Enable IMPALA_COMPRESSED_DEBUG_INFO by default
- IMPALA-13012: Lower default query_log_max_queued
- IMPALA-13005: Create Query Live table in HMS
- IMPALA-13024: Ignore slots if using default pool and empty group
- IMPALA-12872: Use Calcite for optimization - part 1: simple queries
- IMPALA-12950: Improve error message in case of out-of-range numeric conversions
- IMPALA-12543: Detect self-events before finishing DDL
- IMPALA-12933: Avoid fetching unnecessary events of unwanted types
- IMPALA-12657: Improve ProcessingCost of ScanNode and NonGroupingAggregator
- IMPALA-12988: Calculate an unbounded version of CpuAsk
- IMPALA-12938: add-opens for platform.cgroupv1

- IMPALA-13016: Fix ambiguous row_regex that check for no-existence
- IMPALA-12980: Translate CpuAsk into admission control slots
- IMPALA-12874: Identify active and standby catalog and statestore in the web debug endpoint
- IMPALA-12998: Add SHOW_METADATA_TABLES to ignored DDL
- IMPALA-12990: Fix impala-shell handling of unset rows_deleted
- IMPALA-11495: Add glibc version and effective locale to the Web UI
- IMPALA-12963: Return parent PID when children spawned
- IMPALA-12999: Add log4j.properties to the DEB/RPM packages
- IMPALA-12986: Base64Encode fails if the 'out_len' output parameter is passed with certain values
- IMPALA-12564: Prevent Hive loading libfusesupport.so in the minicluster during TSAN runs
- IMPALA-5323: Support BINARY columns in Kudu tables
- IMPALA-12965: Add debug query option RUNTIME_FILTER_IDS_TO_SKIP
- IMPALA-12978: Fix impala-shell's live progress with older Impalas
- IMPALA-12920: Support ai_generate_text built-in function for OpenAI's chat completion AP

Iceberg changelog for Cloudera Data Warehouse on premises

Review the changes introduced in Iceberg for Cloudera Data Warehouse on premises.

2025.0.19.1000

Hive - Iceberg changes

No new features or fixes.

Impala - Iceberg changes

- CDPD-87405: IMPALA-14185: Error unnesting nested array from Iceberg with DELETE files
- CDPD-85228/IMPALA-14154: IllegalStateException with Iceberg table with DELETE
- IMPALA-14014: Fix COMPUTE STATS with TABLESAMPLE clause

2025.0.19.1-74

Hive - Iceberg changes

- CDPD-72164: HIVE-28276: Iceberg: Make Iceberg split threads configurable when table scanning
- CDPD-72045: HIVE-28368: Iceberg: Unable to read PARTITIONS Metadata table.
- CDPD-71812: HIVE-28353: Iceberg: Reading *Files Metadata table files if the column is of TIMESTAMP type.
- CDPD-70374: HIVE-28275: Iceberg: Add support for 'If Not Exists' and 'or Replace' for Create Tag.
- CDPD-72046: HIVE-28299: Iceberg: Optimize show partitions through column projection
- DWX-18658: HIVE-28256: Iceberg: Major QB Compaction on partition level with evolution
- CDPD-71472: HIVE-28323: Iceberg: Allow reading tables irrespective whether they were created with hive engine enabled or not.
- DWX-18477: HIVE-28282: Merging into iceberg table fails with copy on write when values clause has a function call
- CDPD-70373: HIVE-28274: Iceberg: Add support for 'If Not Exists' and 'or Replace' for Create Branch.
- CDPD-70435: HIVE-27880: Iceberg: Support creating a branch on an empty table
- CDPD-70188: HIVE-28278: Iceberg: Stats: IllegalStateException Invalid file: file length 0
- CDPD-69309: HIVE-28132: Iceberg: Add support for Replace Tag.
- CDPD-69704: HIVE-28266: Iceberg: select count(*) from data_files metadata tables gives wrong result
- CDPD-69311: HIVE-28225: Iceberg: Delete on entire table fails on COW mode.

- DWX-17603: HIVE-28077: Iceberg: Major QB Compaction on partition level
- CDPD-68139: HIVE-28131: Iceberg: Add support for Replace Branch.

Impala - Iceberg changes

- IMPALA-13932: Add file path and position-based duplicate check for IcebergMergeNode
- IMPALA-13825: Extend Docker container build to custom base images
- IMPALA-13854: IcebergPositionDeleteChannel uses incorrect capacity
- IMPALA-13853: Don't adjust Iceberg field IDs for data files that don't have complex types
- IMPALA-13737: Directly load file metadata via IcebergFileMetadataLoader
- IMPALA-13789: Defer creating Path objects in loading file metadata
- IMPALA-13772: Fix Workload Management DMLs Timeouts
- IMPALA-13768: Redundant Iceberg delete records are shuffled around which cause error "Invalid file path arrived at builder"
- IMPALA-13594: Read Puffin stats also from older snapshots
- CDPD-78207: Disable ESTIMATE_DUPLICATE_IN_PREAGG in downstream
- IMPALA-13205: Do not include Iceberg position fields for MERGE statements with INSERT merge clauses
- IMPALA-13656: MERGE redundantly accumulates memory in HDFS WRITER
- IMPALA-13324: Enable statement rewrite for merge queries for IcebergMergeImpl
- IMPALA-13655: UPDATE redundantly accumulates memory in HDFS WRITER
- IMPALA-13501: Clean up uncommitted Iceberg files after validation check failure
- IMPALA-13086: Lower AggregationNode estimate using stats predicate
- IMPALA-13305: Better thrift compatibility checks based on pyparsing
- IMPALA-13589: SELECT INPUT__FILE__NAME can crash Impala
- IMPALA-11265: Part2: Store Iceberg file descriptors in encoded format
- IMPALA-13370: Read Puffin stats from metadata.json property if available
- IMPALA-13495: Make exceptions from the Calcite planner easier to classify
- IMPALA-13484: Don't call alter_table() on HMS when loading Iceberg table
- IMPALA-13325: Use RowBatch::CopyRows in IcebergDeleteNode
- IMPALA-13467: Fix partition list size calculation for empty Iceberg scan nodes
- IMPALA-13247: Support Reading Puffin files for the current snapshot
- IMPALA-12861: Fix mixed file format listing for Iceberg tables
- IMPALA-13463: Impala should ignore case of Iceberg schema elements
- IMPALA-13425: Iceberg tables crash server with Calcite planner
- IMPALA-13364: Schema resolution doesn't work for migrated partitioned Iceberg tables that have complex types
- IMPALA-13220: Docs for Iceberg DROP PARTITION
- IMPALA-11265: Part1: Clear GroupContentFiles once used
- IMPALA-12732: Add support for MERGE statements for Iceberg tables
- IMPALA-13254: Optimize REFRESH for Iceberg tables
- IMPALA-12867: Filter files to OPTIMIZE based on file size
- IMPALA-13274: Filter out illegal output for certain join nodes
- IMPALA-13296: Check column compatibility earlier for table migration
- IMPALA-12850: Add better error message for REFRESH iceberg_tbl PARTITION(...)
- IMPALA-12857: Add flag to enable merge-on-read even if tables are configured with copy-on-write
- IMPALA-13088, IMPALA-13109: Use RoaringBitmap instead of sorted vector of int64s
- IMPALA-13085: Add warning and NULL out DECIMAL values in Iceberg metadata tables
- IMPALA-13079: Add support for FLOAT/DOUBLE in Iceberg metadata tables
- IMPALA-11499: Refactor UrlEncode function to handle special characters

- IMPALA-13035: Querying metadata tables from non-Iceberg tables throws IllegalArgumentException
- IMPALA-12973,IMPALA-11491,IMPALA-12651: Support BINARY nested in complex types in select list
- IMPALA-13002: Iceberg V2 tables with Avro delete files aren't read properly
- IMPALA-12543: Detect self-events before finishing DDL
- IMPALA-12990: Fix impala-shell handling of unset rows_deleted
- IMPALA-13003: Handle Iceberg AlreadyExistsException
- IMPALA-13006: Restrict Iceberg tables to Parquet
- IMPALA-12996: Add support for DATE in Iceberg metadata tables
- IMPALA-12810: Simplify IcebergDeleteNode and IcebergDeleteBuilder
- IMPALA-12991: Eliminate unnecessary SORT for Iceberg DELETES
- IMPALA-12970: Fix ConcurrentModificationException for Iceberg table scans

Hue changelog for Cloudera Data Warehouse on premises

Review the changes introduced in Hue for Cloudera Data Warehouse on premises.

2025.0.19.1000

- CDPD-90930: fix key generation bug that caused frontend cache to fail
- DWX-21802: fix the undefined check resolving as true for scheduleTask tab
- CDPD-82819: Improve kt_renewer logging for Kerberos Ticket Renewal Failures
- CDPD-81151: remove redundant Error keyword from alert
- CDPD-85071| CDPD-88219: Improve file extension restrictions check for old upload operations
- DWX-21190: Assign filebrowser permissions to the default group

2025.0.19.1-74

- CDPD-72754: vLLM integration with LLama
- DWX-15075: Queries not getting listed in Impala job browser
- CDPD-76594: [Importer] Update Iceberg table syntax for Impala
- CDPD-76409: [ozone] Fix edge cases when volume name start with service_id in path
- CDPD-76282: [Filebrowser] Fix support for reading Parquet files
- CDPD-72496: [fb] Stop file uploads with restrictive file types for old upload handlers
- CDPD-71994: [core] Add the inline documents search back in the documents page
- CDPD-69626: Make semantic search top-k configurable
- CDPD-68357: Multi database support - UI change
- CDPD-61994: [sqlai] Skip loading of sample data for views
- CDPD-58397: Integrate with CML model deployment
- CDPD-57994: [Importer] Fix importer for Parquet, ORC, and Avro formats
- CDPD-73741: Upgrade openai python package to 1.44.1
- CDPD-71098: [Hive] Support complex data type sampling
- CDPD-64387: Add foreign keys to table metadata passed to LLM
- CDPD-69627: Table metadata caching issue
- CDPD-69732: include commons-lang3 library
- CDPD-77746: Fix Metrics for Oracle DPI errors
- CDPD-80268: [ui-admin] hide log out button for Knox
- CDPD-79289: [conf] Update default value for RESTRICT_FILE_EXTENSIONS to None