
MySQL NDB Cluster 7.5 Release Notes

Abstract

This document contains release notes for the changes in each release of MySQL NDB Cluster that uses version 7.5 of the [NDB \(NDBCLUSTER\)](#) storage engine.

Each NDB Cluster 7.5 release is based on a mainline MySQL Server release and a particular version of the [NDB](#) storage engine, as shown in the version string returned by executing `SELECT VERSION()` in the `mysql` client, or by executing the `ndb_mgm` client `SHOW` or `STATUS` command; for more information, see [MySQL NDB Cluster 7.5](#).

For general information about features added in NDB Cluster 7.5, see [What is New in MySQL NDB Cluster 7.5](#). For a complete list of all bugfixes and feature changes in MySQL Cluster, please refer to the changelog section for each individual NDB Cluster release.

For additional MySQL 5.7 documentation, see the [MySQL 5.7 Reference Manual](#), which includes an overview of features added in MySQL 5.7 that are not specific to NDB Cluster ([What Is New in MySQL 5.7](#)), and discussion of upgrade issues that you may encounter for upgrades from MySQL 5.6 to MySQL 5.7 ([Changes Affecting Upgrades to MySQL 5.7](#)). For a complete list of all bugfixes and feature changes made in MySQL 5.7 that are not specific to [NDB](#), see [MySQL 5.7 Release Notes](#).

Updates to these notes occur as new product features are added, so that everybody can follow the development process. If a recent version is listed here that you cannot find on the download page (<http://dev.mysql.com/downloads/>), the version has not yet been released.

The documentation included in source and binary distributions may not be fully up to date with respect to release note entries because integration of the documentation occurs at release build time. For the most up-to-date release notes, please refer to the online documentation instead.

For legal information, see the [Legal Notices](#).

For help with using MySQL, please visit either the [MySQL Forums](#) or [MySQL Mailing Lists](#), where you can discuss your issues with other MySQL users.

For additional documentation on MySQL products, including translations of the documentation into other languages, and downloadable versions in variety of formats, including HTML and PDF formats, see the [MySQL Documentation Library](#).

Document generated on: 2017-01-04 (revision: 10541)

Table of Contents

Preface and Legal Notices	1
Changes in MySQL NDB Cluster 7.5.5 (5.7.16-ndb-7.5.5) (Not yet released, General Availability)	3
Changes in MySQL NDB Cluster 7.5.4 (5.7.16-ndb-7.5.4) (2016-10-18, General Availability)	3
Changes in MySQL NDB Cluster 7.5.3 (5.7.13-ndb-7.5.3) (2016-07-08, Release Candidate)	13
Changes in MySQL NDB Cluster 7.5.2 (5.7.12-ndb-7.5.2) (2016-06-01, Development Milestone)	14
Changes in MySQL NDB Cluster 7.5.1 (5.7.11-ndb-7.5.1) (2016-03-18, Development Milestone)	21
Changes in MySQL NDB Cluster 7.5.0 (5.7.10-ndb-7.5.0) (2016-02-05, Development Milestone)	24
Release Series Changelogs: MySQL Cluster NDB 7.5	31

Preface and Legal Notices

This document contains release notes for the changes in each release of MySQL NDB Cluster that uses version 7.5 of the [NDB](#) storage engine.

Legal Notices

Copyright © 1997, 2016, Oracle and/or its affiliates. All rights reserved.

This software and related documentation are provided under a license agreement containing restrictions on use and disclosure and are protected by intellectual property laws. Except as expressly permitted in your license agreement or allowed by law, you may not use, copy, reproduce, translate, broadcast, modify, license, transmit, distribute, exhibit, perform, publish, or display any part, in any form, or by any means. Reverse engineering, disassembly, or decompilation of this software, unless required by law for interoperability, is prohibited.

The information contained herein is subject to change without notice and is not warranted to be error-free. If you find any errors, please report them to us in writing.

If this is software or related documentation that is delivered to the U.S. Government or anyone licensing it on behalf of the U.S. Government, then the following notice is applicable:

U.S. GOVERNMENT END USERS: Oracle programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, delivered to U.S. Government end users are "commercial computer software" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, use, duplication, disclosure, modification, and adaptation of the programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, shall be subject to license terms and license restrictions applicable to the programs. No other rights are granted to the U.S. Government.

This software or hardware is developed for general use in a variety of information management applications. It is not developed or intended for use in any inherently dangerous applications, including applications that may create a risk of personal injury. If you use this software or hardware in dangerous applications, then you shall be responsible to take all appropriate fail-safe, backup, redundancy, and other measures to ensure its safe use. Oracle Corporation and its affiliates disclaim any liability for any damages caused by use of this software or hardware in dangerous applications.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Intel and Intel Xeon are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark of The Open Group.

This software or hardware and documentation may provide access to or information about content, products, and services from third parties. Oracle Corporation and its affiliates are not responsible for and expressly disclaim all warranties of any kind with respect to third-party content, products, and services unless otherwise set forth in an applicable agreement between you and Oracle. Oracle Corporation and its affiliates will not be responsible for any loss, costs, or damages incurred due to your access to or use of third-party content, products, or services, except as set forth in an applicable agreement between you and Oracle.

Documentation Accessibility

For information about Oracle's commitment to accessibility, visit the Oracle Accessibility Program website at <http://www.oracle.com/pls/topic/lookup?ctx=acc&id=docacc>.

Access to Oracle Support

Oracle customers that have purchased support have access to electronic support through My Oracle Support. For information, visit <http://www.oracle.com/pls/topic/lookup?ctx=acc&id=info> or visit <http://www.oracle.com/pls/topic/lookup?ctx=acc&id=trs> if you are hearing impaired.

This documentation is NOT distributed under a GPL license. Use of this documentation is subject to the following terms:

You may create a printed copy of this documentation solely for your own personal use. Conversion to other formats is allowed as long as the actual content is not altered or edited in any way. You shall not publish or distribute this documentation in any form or on any media, except if you distribute the documentation in a manner similar to how Oracle disseminates it (that is, electronically for download on a Web site with the software) or on a CD-ROM or similar medium, provided however that the documentation is disseminated together with the software on the same medium. Any other use, such as any dissemination of printed copies or use of this documentation, in whole or in part, in another publication, requires the prior written consent from an authorized representative of Oracle. Oracle and/or its affiliates reserve any and all rights to this documentation not expressly granted above.

Changes in MySQL NDB Cluster 7.5.5 (5.7.16-ndb-7.5.5) (Not yet released, General Availability)

MySQL Cluster NDB 7.5.5 is a new release of MySQL Cluster NDB 7.5, based on MySQL Server 5.7 and including features in version 7.5 of the [NDB](#) storage engine, as well as fixing recently discovered bugs in previous MySQL Cluster releases.

Obtaining MySQL Cluster NDB 7.5. MySQL Cluster NDB 7.5 source code and binaries can be obtained from <http://dev.mysql.com/downloads/cluster/>.

For an overview of changes made in MySQL Cluster NDB 7.5, see [What is New in MySQL NDB Cluster 7.5](#).

This release also incorporates all bugfixes and changes made in previous MySQL Cluster releases, as well as all bugfixes and feature changes which were added in mainline MySQL 5.7 through MySQL 5.7.16 (see [Changes in MySQL 5.7.16 \(2016-10-12, General Availability\)](#)).

Version 5.7.16-ndb-7.5.5 has no changelog entries, or they have not been published because the product version has not been released.

Changes in MySQL NDB Cluster 7.5.4 (5.7.16-ndb-7.5.4) (2016-10-18, General Availability)

MySQL Cluster NDB 7.5.4 is a new release of MySQL Cluster NDB 7.5, based on MySQL Server 5.7 and including features in version 7.5 of the [NDB](#) storage engine, as well as fixing recently discovered bugs in previous MySQL Cluster releases.

Obtaining MySQL Cluster NDB 7.5. MySQL Cluster NDB 7.5 source code and binaries can be obtained from <http://dev.mysql.com/downloads/cluster/>.

For an overview of changes made in MySQL Cluster NDB 7.5, see [What is New in MySQL NDB Cluster 7.5](#).

This release also incorporates all bugfixes and changes made in previous MySQL Cluster releases, as well as all bugfixes and feature changes which were added in mainline MySQL 5.7 through MySQL 5.7.16 (see [Changes in MySQL 5.7.16 \(2016-10-12, General Availability\)](#)).

- [Functionality Added or Changed](#)

- [Bugs Fixed](#)

Functionality Added or Changed

- **Important Change; Packaging:** Naming and organization of the RPMs provided for MySQL Cluster have been changed to align with those released for the MySQL server. All MySQL Cluster RPM package names are now prefixed with `mysql-cluster`. Data nodes are now installed using the `data-node` package; management nodes are now installed from the `management-server` package; and SQL nodes require the `server` and `common` packages. **Important:** SQL nodes must use the `mysql-cluster` version of these RPMs; the versions released for the standard MySQL server do not provide support for the NDB storage engine. All client programs, including both the `mysql` client and the `ndb_mgm` management client, are now included in the `client` RPM.

For more information, see [Installing NDB Cluster from RPM](#).

- Added the `--print-sql-log` option for the `ndb_restore` program included with the MySQL Cluster distribution. This option causes the program to log SQL statements to `stdout`.

Note that each table being restored in this fashion must have an explicitly defined primary key; the hidden primary key implemented by the NDB storage engine is *not* sufficient for this purpose. (Bug #13511949)

- For fully replicated tables, `ndb_desc` shows only nodes holding main fragment replicas for partitions; nodes with copy fragment replicas only are ignored. To make this information available in the `mysql` client, several new tables have been introduced in the `ndbinfo` information database. These tables are listed here, with brief descriptions:
 - `dict_obj_info` provides the names and types of database (DICT) objects in NDB, such as tables and indexes, as well as information about parent objects where applicable
 - `table_distribution_status` provides NDB table distribution status information
 - `table_fragments` provides information about the distribution of NDB table fragments
 - `table_info` provides information about logging, checkpointing, storage, and other options in force for each NDB table
 - `table_replicas` provides information about fragment replicas

For more information, see [The ndbinfo NDB Cluster Information Database](#). (Bug #81762, Bug #23547643)

- **MySQL NDB ClusterJ:** To help applications handle database errors better, a number of new features have been added to the `ClusterJDatastoreException` class:
 - A new method, `getCode()`, returns `code` from the `NdbError` object.
 - A new method, `getMySQLCode()`, returns `mysql_code` from the `NdbError` object.
 - A new subclass, `ClusterJDatastoreException.Classification`, gives users the ability to decode the result from `getClassification()`. The method `Classification.toString()` gives the name of the error classification as listed in [NDB Error Classifications](#).

(Bug #22353594)

Bugs Fixed

- **Important Change:** The default value of the `--ndb-default-column-format` server option has been changed from `DYNAMIC` to `FIXED`. This has been done for backwards compatibility. Only the default has been changed; setting this option to `DYNAMIC` continues to cause `DYNAMIC` to be used for `ROW_FORMAT` and `COLUMN_FORMAT` unless overridden. (Bug #24487363)
- **Important Change:** Event buffer status reporting has been improved by altering the semantics for calculating lag or slippage. Rather than defining this lag as the number of epochs behind, lag is now taken as the number of epochs completely buffered in the event buffer, but not yet consumed by the binlog injector thread. As part of this work, the default value for the `ndb_report_thresh_binlog_epoch_slip` system variable has been increased from 3 to 10. For more information, see the description of this variable in the documentation, as well as [Event Buffer Reporting in the Cluster Log](#). (Bug #22916457)

References: See also: Bug #22901309.

- **NDB Cluster APIs:** The term “fragment count type” has been superseded by “partition balance”. This change affects `NDB_TABLE` options for `NDB` tables as well as in the `NDB` API. In `NDB_TABLE` table option syntax, the `FRAGMENT_COUNT_TYPE` keyword is replaced with `PARTITION_BALANCE`. In the `NDB` API, the `Table` methods `getFragmentCountType()` and `setFragmentCountType()` have been renamed to `getPartitionBalance()` and `setPartitionBalance()`, respectively; `getFragmentCountTypeString()` is renamed to `getPartitionBalanceString()`. In addition, `Object::FragmentCountType` has been renamed to `PartitionBalance`, and the names of its enumerated values have been updated to be consistent with the new nomenclature.

For more information on how these changes affect `NDB` API applications, see the indicated `Table` and `Object` member descriptions. For more information on the SQL-level changes made as part of this fix, see [Setting NDB_TABLE Options in Table Comments](#). (Bug #81761, Bug #23547525)

References: See also: Bug #83147, Bug #24733331.

- If more than 4096 seconds elapsed while calculating an internal `NdbDuration::microSec()` value, this could cause an assert warning that the calculation would overflow. We fix this to avoid any overflow or precision loss when converting from the internal “tick” format to microseconds and nanoseconds, by performing the calculation in two parts corresponding to seconds and fractions of a second. (Bug #24695026)
- The serial commit protocol—which commits each operation at each replica individually and serially, and is used by the `DBTC` kernel block (see [The DBTC Block](#)) for takeover and when a transaction is judged to have timed out during the `COMMIT` or `COMPLETE` phase—had no support for `LATE_COMMIT`, which is required for the `READ_BACKUP` and `FULLY_REPLICATED` protocols. (Bug #24681305)
- In some cases, `ALTER TABLE ... REORGANIZE PARTITION` could lead to an unplanned shutdown of the cluster. This was due to the fact that, for fully replicated tables, the log part ID was assumed to be the same as the partition ID. This worked when `FOR_RA_BY_LDM` was used, but not necessarily for the other partition balancing types. (Bug #24610551)
- Using `ALGORITHM=INPLACE` when changing any of a table's `NDB_TABLE` properties (see [Setting NDB_TABLE Options in Table Comments](#)) caused the server to fail. (Bug #24584741)
- Following successive `ALTER TABLE` statements updating `NDB_TABLE` properties (see [Setting NDB_TABLE Options in Table Comments](#)), the current values were not always shown by `SHOW CREATE TABLE` or `ndb_desc`. (Bug #24584690)
- Case-insensitivity of keywords such as `FULLY_REPLICATED` in `NDB_TABLE` comments was not honored. (Bug #24577931)

- An `ALTER TABLE` statement attempting to set both `FULLY_REPLICATED` and `PARTITION_BALANCE` (see [Setting NDB_TABLE Options in Table Comments](#)) failed with a garbled error message. (Bug #24577894)
- A number of dependencies between the binlog injector thread and the NDB utility thread—a recurring source of synchronization and other problems—were removed. The principal changes are listed here:
 - Moved the setup of binlog injector structures from the utility thread to the injector thread itself.
 - Removed sharing of some utility and injector thread structures between these threads.
 - Moved stopping of the utility thread from the injector thread into a common block in which other such threads are stopped.
 - Removed a number of hacks required by the previous design.
 - Removed some injector mutex locking and injector condition signaling which were made obsolete by the changes already listed.

(Bug #24496910)

References: See also: Bug #22204186.

- A late commit `ACK` signal used for `FULLY_REPLICATED` or `READ_BACKUP` tables caused the associated `ApiConnectionRecord` to have an invalid state. (Bug #24459817)

References: See also: Bug #24444861.

- Added missing error information for a failure occurring when tables on disk became full. (Bug #24425373)
- When `ndbmt` crashed, the resulting error log incorrectly specified the name of the trace for thread 0, appending the nonexistent suffix `_t0` to the file name. (Bug #24353408)
- Passing a nonexistent node ID to `CREATE NODEGROUP` led to random data node failures. (Bug #23748958)
- `DROP TABLE` followed by a node shutdown and subsequent master takeover—and with the containing local checkpoint not yet complete prior to the takeover—caused the LCP to be ignored, and in some cases, the data node to fail. (Bug #23735996)

References: See also: Bug #23288252.

- Removed an invalid assertion to the effect that all cascading child scans are closed at the time API connection records are released following an abort of the main transaction. The assertion was invalid because closing of scans in such cases is by design asynchronous with respect to the main transaction, which means that subscans may well take some time to close after the main transaction is closed. (Bug #23709284)
- Although arguments to the `DUMP` command are 32-bit integers, `ndb_mgmd` used a buffer of only 10 bytes when processing them. (Bug #23708039)
- The `READ_BACKUP` setting was not honored when performing scans on `BLOB` tables. (Bug #23703536)
- Setting `FULLY_REPLICATED=1` (see [Setting NDB_TABLE Options in Table Comments](#)) did not propagate to the internal `BLOB` part tables used for `BLOB` and `TEXT` columns. (Bug #23703343)
- The `READ_BACKUP` setting was not applied to unique indexes. (Bug #23702848)

- In `ReadCommitted` mode, `DBSPJ` read primary fragment replicas for tables with `READ_BACKUP` (see [Setting NDB_TABLE Options in Table Comments](#)), even when a local fragment was available. (Bug #23633848)
- `ALL_REPORT MemoryUsage` produced incorrect output when fully replicated tables were in use. (Bug #23539805)
- Ordered indexes did not inherit `READ_BACKUP` (see [Setting NDB_TABLE Options in Table Comments](#)) from an indexed table, which meant that ordered index scans continued to be routed to only to primary fragment replicas and never to backup fragment replicas.

Now `DBDICT` sets this property on ordered indexes from the table property when it distributes this information to instances of `DBTC` and `DBSPJ`. (Bug #23522027)

- Updates to a table containing a virtual column could cause the binary logging thread to fail. (Bug #23514050)
- A number of potential buffer overflow issues were found and fixed in the `NDB` codebase. (Bug #23152979)
- During an online upgrade from a MySQL Cluster NDB 7.3 release to an NDB 7.4 (or later) release, the failures of several data nodes running the lower version during local checkpoints (LCPs), and just prior to upgrading these nodes, led to additional node failures following the upgrade. This was due to lingering elements of the `EMPTY_LCP` protocol initiated by the older nodes as part of an LCP-plus-restart sequence, and which is no longer used in NDB 7.4 and later due to LCP optimizations implemented in those versions. (Bug #23129433)
- A `SIGNAL_DROPPED_REP` handler invoked in response to long message buffer exhaustion was defined in the `SPJ` kernel block, but not actually used. This meant that the default handler from `SimulatedBlock` was used instead in such cases, which shut down the data node. (Bug #23048816)

References: See also: Bug #23251145, Bug #23251423.

- When a data node has insufficient redo buffer during a system restart, it does not participate in the restart until after the other nodes have started. After this, it performs a takeover of its fragments from the nodes in its node group that have already started; during this time, the cluster is already running and user activity is possible, including DML and DDL operations.

During a system restart, table creation is handled differently in the `DIH` kernel block than normally, as this creation actually consists of reloading table definition data from disk on the master node. Thus, `DIH` assumed that any table creation that occurred before all nodes had restarted must be related to the restart and thus always on the master node. However, during the takeover, table creation can occur on non-master nodes due to user activity; when this happened, the cluster underwent a forced shutdown.

Now an extra check is made during system restarts to detect in such cases whether the executing node is the master node, and use that information to determine whether the table creation is part of the restart proper, or is taking place during a subsequent takeover. (Bug #23028418)

- `ndb_restore` set the `MAX_ROWS` attribute for a table for which it had not been set prior to taking the backup. (Bug #22904640)
- Whenever data nodes are added to or dropped from the cluster, the `NDB` kernel's Event API is notified of this using a `SUB_GCP_COMPLETE_REP` signal with either the `ADD` (add) flag or `SUB` (drop) flag set, as well as the number of nodes to add or drop; this allows `NDB` to maintain a correct count of `SUB_GCP_COMPLETE_REP` signals pending for every incomplete bucket. In addition to handling the bucket for the epoch associated with the addition or removal, it must also compensate for any later

incomplete buckets associated with later epochs. Although it was possible to complete such buckets out of order, there was no handling of these, leading a stall in to event reception.

This fix adds detection and handling of such out of order bucket completion. (Bug #20402364)

References: See also: Bug #82424, Bug #24399450.

- When performing online reorganization of tables, unique indexes were not included in the reorganization. (Bug #13714258)
- Under very high loads, neither the receive thread nor any user thread had sufficient capacity to handle poll ownership properly. This meant that, as the load and the number of active thread increased, it became more difficult to sustain throughput. This fix attempts to increase the priority of the receive thread and retains poll ownership if successful.

This fix requires sufficient permissions to be enabled. On Linux systems, this means ensuring that either the data node binary or the user it runs as has permission to change the nice level. (Bug #83217, Bug #24761073)

- Local reads of unique index and blob tables did not work correctly for fully replicated tables using more than one node group. (Bug #83016, Bug #24675602)
- The effects of an `ALTER TABLE` statement changing a table to use `READ_BACKUP` were not preserved after a restart of the cluster. (Bug #82812, Bug #24570439)
- Using `FOR_RP_BY_NODE` or `FOR_RP_BY_LDM` for `PARTITION_BALANCE` did not work with fully replicated tables. (Bug #82801, Bug #24565265)
- Changes to `READ_BACKUP` settings were not propagated to internal blob tables. (Bug #82788, Bug #24558232)
- The count displayed by the `c_exec` column in the `ndbinfo.threadstat` table was incomplete. (Bug #82635, Bug #24482218)
- The default `PARTITION_BALANCE` setting for `NDB` tables created with `READ_BACKUP=1` (see [Setting NDB_TABLE Options in Table Comments](#)) has been changed from `FOR_RA_BY_LDM` to `FOR_RP_BY_LDM`. (Bug #82634, Bug #24482114)
- The internal function `ndbcluster_binlog_wait()`, which provides a way to make sure that all events originating from a given thread arrive in the binary log, is used by `SHOW BINLOG EVENTS` as well as when resetting the binary log. This function waits on an injector condition while the latest global epoch handled by `NDB` is more recent than the epoch last committed in this session, which implies that this condition must be signalled whenever the binary log thread completes and updates a new latest global epoch. Inspection of the code revealed that this condition signalling was missing, and that, instead of being awakened whenever a new latest global epoch completes (~100ms), client threads waited for the maximum timeout (1 second).

This fix adds the missing injector condition signalling, while also changing it to a condition broadcast to make sure that all client threads are alerted. (Bug #82630, Bug #24481551)

- Fully replicated internal foreign key or unique index triggers could fire multiple times, which led to aborted transactions for an insert or a delete operation. This happened due to redundant deferred constraint triggers firing during pre-commit. Now in such cases, we ensure that only triggers specific to unique indexes are fired in this stage. (Bug #82570, Bug #24454378)
- Backups potentially could fail when using fully replicated tables due to their high usage (and subsequent exhaustion) of internal trigger resources. To compensate for this, the amount of memory reserved in the

NDB kernel for internal triggers has been increased, and is now based in part on the maximum number of tables. (Bug #82569, Bug #24454262)

References: See also: Bug #23539733.

- In the DBTC function `executeFullyReplicatedTrigger()` in the NDB kernel, an incorrect check of state led in some cases to failure handling when no failure had actually occurred. (Bug #82568, Bug #24454093)

References: See also: Bug #23539733.

- When returning from `LQHKEYREQ` with failure in `LQHKEYREF` in an internal trigger operation, no check was made as to whether the trigger was fully replicated, so that those triggers that were fully replicated were never handled. (Bug #82566, Bug #24453949)

References: See also: Bug #23539733.

- When `READ_BACKUP` had not previously been set, then was set to 1 as part of an `ALTER TABLE ... ALGORITHM=INPLACE` statement, the change was not propagated to internal unique index tables or `BLOB` tables. (Bug #82491, Bug #24424459)
- Distribution of MySQL privileges was incomplete due to the failure of the `mysql_cluster_move_privileges()` procedure to convert the `mysql.proxies_priv` table to NDB. The root cause of this was an `ALTER TABLE ... ENGINE NDB` statement which sometimes failed when this table contained illegal `TIMESTAMP` values. (Bug #82464, Bug #24430209)
- The internal variable `m_max_warning_level` was not initialized in `storage/ndb/src/kernel/blocks/thrman.cpp`. This sometimes led to node failures during a restart when the uninitialized value was treated as 0. (Bug #82053, Bug #23717703)
- Usually, when performing a system restart, all nodes are restored from redo logs and local checkpoints (LCPs), but in some cases some node might require a copy phase before it is finished with the system restart. When this happens, the node in question waits for all other nodes to start up completely before performing the copy phase. Notwithstanding the fact that it is thus possible to begin a local checkpoint before reaching start phase 4 in the `DBDIH` block, LCP status was initialized to `IDLE` in all cases, which could lead to a node failure. Now, when performing this variant of a system restart, the LCP status is no longer initialized. (Bug #82050, Bug #23717479)
- After adding a new node group online and executing `ALTER TABLE ... ALGORITHM=INPLACE REORGANIZE PARTITION`, partition IDs were not set correctly for new fragments.

In a related change done as part of fixing this issue, `ndb_desc -p` now displays rows relating to partitions in order of partition ID. (Bug #82037, Bug #23710999)

- When executing `STOP BACKUP` it is possible sometimes that a few bytes are written to the backup data file before the backup process actually terminates. When using `ODIRECT`, this resulted in the wrong error code being returned. Now in such cases, nothing is written to `O_DIRECT` files unless the alignment is correct. (Bug #82017, Bug #23701911)
- When transaction coordinator (TC) connection records were used up, it was possible to handle scans only for local checkpoints and backups, so that operations coming from the `DBUTIL` block—used for `ALTER TABLE ... REORGANIZE PARTITION` and other operations that reorganize metadata—were unnecessarily blocked. In addition, such operations were not always retried when TC records were exhausted. To fix this issue, a number of operation records are now earmarked for `DBUTIL` usage, as well as for LCP and backup usage so that these operations are also not negatively impacted by operations coming from `DBUTIL`.

For more information, see [The DBUTIL Block](#). (Bug #81992, Bug #23642198)

- Operations performing multiple updates of the same row within the same transaction could sometimes lead to corruption of lengths of page entries. (Bug #81938, Bug #23619031)
- During a node restart, a fragment can be restored using information obtained from local checkpoints (LCPs); up to 2 restorable LCPs are retained at any given time. When an LCP is reported to the `DIH` kernel block as completed, but the node fails before the last global checkpoint index written into this LCP has actually completed, the latest LCP is not restorable. Although it should be possible to use the older LCP, it was instead assumed that no LCP existed for the fragment, which slowed the restart process. Now in such cases, the older, restorable LCP is used, which should help decrease long node restart times. (Bug #81894, Bug #23602217)
- Optimized node selection (`ndb_optimized_node_selection` setting) was not respected by `ndb_data_node_neighbour` when this was enabled. (Bug #81778, Bug #23555834)
- `NDB` no longer retries a global schema lock if this has failed due to a timeout (default 3000ms) and there is the potential for this lock request to participate in a metadata lock-global schema lock deadlock. Now in such cases it selects itself as a “victim”, and returns the decision to the requestor of the metadata lock, which then handles the request as a failed lock request (preferable to remaining deadlocked indefinitely), or, where a deadlock handler exists, retries the metadata lock-global schema lock. (Bug #81775, Bug #23553267)
- Two issues were found in the implementation of hash maps—used by `NDB` for mapping a table row's hash value to a partition—for fully replicated tables:
 1. Hash maps were selected based on the number of fragments rather than the number of partitions. This was previously undetected due to the fact that, for other kinds of tables, these values are always the same.
 2. The hash map was employed as a partition-to-partition map, using the table row's hash value modulus the partition count as input.

This fix addresses both of the problems just described. (Bug #81757, Bug #23544220)

References: See also: Bug #81761, Bug #23547525, Bug #23553996.

- Using `mysqld` together with `--initialize` and `--ndbcluster` led to problems later when attempting to use `mysql_upgrade`. When running with `--initialize`, the server does not require `NDB` support, and having it enabled can lead to issues with `ndbinfo` tables. To prevent this from happening, using the `--initialize` option now causes `mysqld` to ignore the `--ndbcluster` option if the latter is also specified.

This issue affects upgrades from MySQL Cluster NDB 7.5.2 or 7.5.3 only. In cases where such upgrades fail for the reasons outlined previously, you can work around the issue by deleting all `.frm` files in the `data/ndbinfo` directory following a rolling restart of the entire cluster, then running `mysql_upgrade`. (Bug #81689, Bug #23518923)

References: See also: Bug #82724, Bug #24521927.

- While a `mysqld` was waiting to connect to the management server during initialization of the `NDB` handler, it was not possible to shut down the `mysqld`. If the `mysqld` was not able to make the connection, it could become stuck at this point. This was due to an internal wait condition in the utility and index statistics threads that could go unmet indefinitely. This condition has been augmented with a maximum timeout of 1 second, which makes it more likely that these threads terminate themselves properly in such cases.

In addition, the connection thread waiting for the management server connection performed 2 sleeps in the case just described, instead of 1 sleep, as intended. (Bug #81585, Bug #23343673)

- `ALTER TABLE ... ALGORITHM=INPLACE` on a fully replicated table did not copy the associated trigger ID, leading to a failure in the `DBDICT` kernel block. (Bug #81544, Bug #23330359)
- The list of deferred tree node lookup requests created when preparing to abort a `DBSPJ` request were not cleared when this was complete, which could lead to deferred operations being started even after the `DBSPJ` request aborted. (Bug #81355, Bug #23251423)

References: See also: Bug #23048816.

- Error and abort handling in `Dbspj::execTRANSID_AI()` was implemented such that its `abort()` method was called before processing of the incoming signal was complete. Since this method sends signals to the LDM, this partly overwrote the contents of the signal which was later required by `execTRANSID_AI()`. This could result in aborted `DBSPJ` requests cleaning up their allocated resources too early, or not at all. (Bug #81353, Bug #23251145)

References: See also: Bug #23048816.

- The read backup feature added in MySQL Cluster NDB 7.5.2 that makes it possible to read from backup replicas was not used for reads with lock, or for reads of `BLOB` tables or unique key tables where locks were upgraded to reads with lock. Now the `TCKEYREQ` and `SCAN_TABREQ` signals use a flag to convey information about such locks making it possible to read from a backup replica when a read lock was upgraded due to being the read of the base table for a `BLOB` table, or due to being the read for a unique key. (Bug #80861, Bug #23001841)
- Primary replicas of partitioned tables were not distributed evenly among node groups and local data managers.

As part of the fix for this issue, the maximum number of node groups supported for a single MySQL Cluster, which was previously not determined, is now set at 48 (`MAX_NDB_NODE_GROUPS`). (Bug #80845, Bug #22996305)

- Several object constructors and similar functions in the `NDB` codebase did not always perform sanity checks when creating new instances. These checks are now performed under such circumstances. (Bug #77408, Bug #21286722)
- An internal call to `malloc()` was not checked for `NULL`. The function call was replaced with a direct write. (Bug #77375, Bug #21271194)
- **NDB Cluster APIs:** Reuse of transaction IDs could occur when `Ndb` objects were created and deleted concurrently. As part of this fix, the NDB API methods `lock_ndb_objects()` and `unlock_ndb_objects` are now declared as `const`. (Bug #23709232)
- **NDB Cluster APIs:** When the management server was restarted while running an MGM API application that continuously monitored events, subsequent events were not reported to the application, with timeouts being returned indefinitely instead of an error.

This occurred because sockets for event listeners were not closed when restarting `mgmd`. This is fixed by ensuring that event listener sockets are closed when the management server shuts down, causing applications using functions such as `ndb_logevent_get_next()` to receive a read error following the restart. (Bug #19474782)

- **NDB Cluster APIs:** To process incoming signals, a thread which wants to act as a receiver must acquire polling rights from the transporter layer. This can be requested and assigned to a separate receiver thread, or each client thread can take the receiver role when it is waiting for a result.

When the thread acting as poll owner receives a sufficient amount of data, it releases locks on any other clients taken while delivering signals to them. This could make them runnable again, and the operating system scheduler could decide that it was time to wake them up, which happened at the expense of the poll owner threads, which were in turn excluded from the CPU while still holding polling rights on it. After this fix, polling rights are released by a thread before unlocking and signalling other threads. This makes polling rights available for other threads that are actively executing on this CPU.

This change increases concurrency when polling receiver data, which should also reduce latency for clients waiting to be woken up. (Bug #83129, Bug #24716756)

- **NDB Cluster APIs:** `libndbclient` and `libmysqlclient` exported conflicting symbols, resulting in a segmentation fault in debug builds on Linux. To fix this issue, the conflicting symbols in `libndbclient.so` are no longer publicly visible. Due to this change, the version number for `libndbclient.so` has been raised from 6.0.0 to 6.1.0. (Bug #83093, Bug #24707521)

References: See also: Bug #80352, Bug #22722555.

- **NDB Cluster APIs:** When NDB schema object ownership checks are enabled by a given `NdbTransaction`, objects used by this transaction are checked to make sure that they belong to the `NdbDictionary` owned by this connection. An attempt to create a `NdbOperation`, `NdbScanOperation`, or `NdbIndexScanOperation` on a table or index not belonging to the same connection fails.

This fix corrects a resource leak which occurred when the operation object to be created was allocated before checking schema object ownership and subsequently not released when the object creation failed. (Bug #81949, Bug #23623978)

References: See also: Bug #81945, Bug #23623251.

- **NDB Cluster APIs:** NDB API objects are allocated in the context of an `Ndb` object, or of an `NdbTransaction` object which is itself owned by an `Ndb` object. When a given `Ndb` object is destroyed, all remaining `NdbTransaction` objects are terminated, and all NDB API objects related to this `Ndb` object should be released at this time as well. It was found, when there remained unclosed `NdbTransaction` objects when their parent `Ndb` object was destroyed, leaks of objects allocated from the `NdbTransaction` objects could occur. (However, the `NdbTransaction` objects themselves did not leak.)

While it is advisable (and, indeed, recommended) to close an `NdbTransaction` explicitly as soon as its lifetime ends, the destruction of the parent `Ndb` object should be sufficient to release whatever objects are dependent on it. Now in cases such as described previously, the `Ndb` destructor checks to ensure that all objects derived from a given `Ndb` instance are truly released. (Bug #81945, Bug #23623251)

- **NDB Cluster APIs:** In some of the NDB API example programs included with the MySQL Cluster distribution, `ndb_end()` was called prior to calling the `Ndb_cluster_connection` destructor. This caused a segmentation fault in debug builds on all platforms. The example programs affected have also been extensively revised and refactored. See [NDB API Examples](#), for more information. (Bug #80352, Bug #22722555)

References: See also: Bug #83093, Bug #24707521.

Changes in MySQL NDB Cluster 7.5.3 (5.7.13-ndb-7.5.3) (2016-07-08, Release Candidate)

MySQL Cluster NDB 7.5.3 is a new release of MySQL Cluster NDB 7.5, based on MySQL Server 5.7 and including features in version 7.5 of the [NDB](#) storage engine, as well as fixing recently discovered bugs in previous MySQL Cluster releases.

Obtaining MySQL Cluster NDB 7.5. MySQL Cluster NDB 7.5 source code and binaries can be obtained from <http://dev.mysql.com/downloads/cluster/>.

For an overview of changes made in MySQL Cluster NDB 7.5, see [What is New in MySQL NDB Cluster 7.5](#).

This release also incorporates all bugfixes and changes made in previous MySQL Cluster releases, as well as all bugfixes and feature changes which were added in mainline MySQL 5.7 through MySQL 5.7.13 (see [Changes in MySQL 5.7.13 \(2016-06-02, General Availability\)](#)).

- [Functionality Added or Changed](#)
- [Bugs Fixed](#)

Functionality Added or Changed

- **Important Change:** It is now possible to set [READ_BACKUP](#) for an existing table online using an SQL statement such as `ALTER TABLE ... ALGORITHM=INPLACE, COMMENT="NDB_TABLE=READ_BACKUP=1"`. See [Setting NDB_TABLE Options in Table Comments](#), for further information about the [READ_BACKUP](#) option. (Bug #80858, Bug #23001617)

References: See also: Bug #18435416.

- Added three new tables to the [ndbinfo](#) information database to provide running information about locks and lock attempts in an active MySQL Cluster. These tables, with brief descriptions, are listed here:
 - [cluster_locks](#): Current lock requests which are waiting for or holding locks; this information can be useful when investigating stalls and deadlocks. Analogous to [cluster_operations](#).
 - [locks_per_fragment](#): Counts of lock claim requests, and their outcomes per fragment, as well as total time spent waiting for locks successfully and unsuccessfully. Analogous to [operations_per_fragment](#) and [memory_per_fragment](#).
 - [server_locks](#): Subset of cluster transactions—those running on the local `mysqld`, showing a connection ID per transaction. Analogous to [server_operations](#).

For more information, see [The ndbinfo NDB Cluster Information Database](#).

- The [NDB](#) storage engine now supports generated columns (see [CREATE TABLE and Generated Columns](#)) as well as indexes on stored generated columns.

Indexes on virtual generated columns of [NDB](#) tables are not supported.

`ALTER TABLE` statements that add stored generated columns to [NDB](#) tables cannot be performed online.

Bugs Fixed

- The [ndbinfo](#) `cpustat_1sec` and `cpustat_20sec` tables did not provide any history information. (Bug #23520271)

- During shutdown, the `mysqld` process could sometimes hang after logging `NDB Util: Stop ... NDB Util: Wakeup`. (Bug #23343739)
References: See also: Bug #21098142.
- During expansion or reduction of a hash table, allocating a new overflow page in the `DBACC` kernel block caused the data node to fail when it was out of index memory. This could sometimes occur after a large table had increased or decreased very rapidly in size. (Bug #23304519)
References: This issue is a regression of: Bug #13436216.
- When tracking time elapsed from sending a `SCAN_FRAGREQ` signal to receiving a corresponding `SCAN_FRAGCONF`, the assumption was made in the `DBTC` kernel block that a `SCAN_FRAGCONF` can occur only after sending a `SCAN_FRAGREQ` or `SCAN_NEXTREQ` signal, which is not always the case: It is actually possible that a local query handler can, immediately after sending a `SCAN_FRAGCONF`, send an additional `SCAN_FRAGCONF` signal upon reporting that the scan is closed. This problem is fixed by ensuring that the timer value is initialized each time before use. (Bug #81907, Bug #23605817, Bug #23306695)
- Table indexes were listed in the output of `ndb_desc` in a nondeterministic order that could vary between platforms. Now these indexes are ordered by ID in the output. (Bug #81763, Bug #23547742)
- Following a restart of the cluster, the first attempt to read from any of the `ndbinfo` `cpustat`, `cpustat_50ms`, `cpustat_1sec`, or `cpustat_20sec` tables generated a warning to the effect that columns were missing from the table. Subsequently, the `thread_sleeping` and `spin_time` columns were found to be missing from each of these tables. (Bug #81681, Bug #23514557)
References: See also: Bug #23305078.
- Using a `ThreadConfig` parameter value with a trailing comma led to an assertion. (Bug #81588, Bug #23344374)
- **NDB Cluster APIs:** The scan lock takeover issued by `NdbScanOperation::lockCurrentTuple()` did not set the operation type for the takeover operation. (Bug #23314028)

Changes in MySQL NDB Cluster 7.5.2 (5.7.12-ndb-7.5.2) (2016-06-01, Development Milestone)

MySQL Cluster NDB 7.5.2 is a new release of MySQL Cluster NDB 7.5, based on MySQL Server 5.7 and including features in version 7.5 of the `NDB` storage engine, as well as fixing recently discovered bugs in previous MySQL Cluster releases.

Obtaining MySQL Cluster NDB 7.5. MySQL Cluster NDB 7.5 source code and binaries can be obtained from <http://dev.mysql.com/downloads/cluster/>.

For an overview of changes made in MySQL Cluster NDB 7.5, see [What is New in MySQL NDB Cluster 7.5](#).

This release also incorporates all bugfixes and changes made in previous MySQL Cluster releases, as well as all bugfixes and feature changes which were added in mainline MySQL 5.7 through MySQL 5.7.12 (see [Changes in MySQL 5.7.12 \(2016-04-11, General Availability\)](#)).

- [Functionality Added or Changed](#)
- [Bugs Fixed](#)

Functionality Added or Changed

- **Performance:** A deficiency in event buffer memory allocation was identified as inefficient and possibly leading to undesirable results. This could happen when additional memory was allocated from the operating system to buffer received event data even when memory had already been allocated but remained unused. This is fixed by allocating the event buffer memory directly from the page allocation memory manager (`mmap()`), where such functionality is offered by the operating system, allowing for direct control over the memory such that it is in fact returned to the system when released.

This reimplemention avoids the tendencies of the existing one to approach worst-case memory usage, maintenance of data structures for a worst-case event buffer event count, and useless caching of free memory in unusable positions. This work should also help minimize the runtime costs of buffering events, minimize heap fragmentation, and avoid OS-specific problems due to excessive numbers of distinct memory mappings.

In addition, the relationship between epochs and internal `EventData` objects is now preserved throughout the event lifecycle, reception to consumption, thus removing the need for iterating, and keeping in synch, two different lists representing the epochs and their `EventData` objects.

As part of this work, better reporting on the relevant event buffer metrics is now provided in the cluster logs.

References: See also: Bug #21651536, Bug #21660947, Bug #21661297, Bug #21673318, Bug #21689380, Bug #21809959.

- **Important Change; NDB Replication:** The `ndb_binlog_index` table now uses the `InnoDB` storage engine. When upgrading previous versions to the current release or to a later one, use `mysql_upgrade` with `--force --upgrade-system-tables` to have it perform `ALTER TABLE ... ENGINE=INNODB` on this table. Use of the `MyISAM` storage engine for this table continues to be supported for backward compatibility.

One benefit of this change is that it is now possible to depend on transactional behavior and lock-free reads for this table, which should help alleviate concurrency issues during purge operations and log rotation, and improve the availability of this table.

Due to differences in storage requirements for `MyISAM` and `InnoDB` tables, users should be aware that `ndb_binlog_index` may now take up more disk space than was previously required.

For more information, see [NDB Cluster Replication Schema and Tables](#).

- `ndb_restore` now performs output logging for specific stages of its operation. (Bug #21097957)
- An improvement in the hash index implementation used by MySQL Cluster data nodes means that partitions may now contain more than 16 GB of data for fixed columns, and the maximum partition size for fixed column data is now increased to 128 TB. The previous limitation originated with the `DBACC` block in the `NDB` kernel using only 32-bit references to the fixed-size part of a row handled in the `DBTUP` block, even though 45-bit references were already in use elsewhere in the kernel outside the `DBACC` block; all such references in `DBACC` now use 45-bit pointers instead.

As part of this work, error messages returned by the `DBACC` kernel block that were overly generic have now been improved by making them more specific. (Bug #17465232)

- The `NDB` storage engine supports the MySQL `JSON` data type and MySQL `JSON` functions implemented in MySQL 5.7.8 and later. This support is subject to the limitation that a single `NDB` table can have at most 3 `JSON` columns.
- A number of changes and improvements were made to the handling of send threads by `NDB`. The changes are summarized with brief descriptions in the following list:

- Decreased resource requirements for send threads, making sure that a given configuration using send threads outperforms the same configuration without send threads.
- Made use of otherwise idle threads (other than receiver threads) as send threads without incurring extra CPU resources in low-load situations.
- Improved response time for write transactions.
- Provided for handling of bad configuration data in a more graceful manner.
- Made it possible to measure CPU usage with improved real-time reporting from data nodes on their resource usage.

As part of implementing the last item of those just listed, a number of new tables providing information about CPU and thread activity by node, thread ID, and thread type have been added to the `ndbinfo` information database. These tables are listed here:

- `cpustat`: Provides per-second, per-thread CPU statistics
- `cpustat_50ms`: Shows raw per-thread CPU statistics data, gathered every 50ms
- `cpustat_1sec`: Provides raw per-thread CPU statistics data, gathered each second
- `cpustat_20sec`: Displays raw per-thread CPU statistics data, gathered every 20 seconds
- `threads`: Shows names and descriptions of thread types

For more information, see [The ndbinfo NDB Cluster Information Database](#).

- Rewrote the implementation of the `NDB` storage engine's global schema lock functionality to make use of the metadata lock hooks implemented in the MySQL Server in MySQL 5.7.11.
- A number of improvements provide additional read scalability for `NDB` by making it possible to read tables locally. It is now possible to enable reads from any replica, rather than from the primary replica only. This is disabled by default to remain compatible with previous behavior, but can be enabled for a given SQL node using the `ndb_read_backup` system variable added in this release.

It also becomes possible to be more flexible about the assignment of partitions by setting a fragment count type. Possible count types are one per node, one per node group, one per Local Data Manager (LDM) per node (this is the previous assignment scheme and still the default), and one per LDM per node group. This setting can be controlled for individual tables by means of a `FRAGMENT_COUNT_TYPE` option embedded in an `NDB_TABLE` comment in `CREATE TABLE` or `ALTER TABLE`.

Tables using one of the two per-node-group settings for the fragment count type can also be fully replicated. This requires that the table's fragment count type is `ONE_PER_NODE_GROUP` or `ONE_PER_LDM_PER_NODE_GROUP`, and can be enabled using the option `FULLY_REPLICATED=1` within in an `NDB_TABLE` comment. The option can be enabled by default for all new `NDB` tables using the `ndb_fully_replicated` system variable added in this release.

Settings for table-level `READ_BACKUP` are also supported using the `COMMENT="NDB_TABLE=..."` syntax. It is also possible (and often preferable) to set multiple options in one comment within a single `CREATE TABLE` or `ALTER TABLE` statement. For more information and examples, see [Setting NDB_TABLE Options in Table Comments](#).

This release also introduces the `ndb_data_node_neighbour` system variable, which is intended to be employed with transaction hinting (and fully-replicated tables), and to provide the current SQL node with

the ID of a “nearby” data node to use. See the description of this variable in the documentation for more information.

References: See also: Bug #18435416.

- Made the following enhancements and additions to the `ThreadConfig` multithreaded data node (`ndbmtd`) configuration parameter:
 - Added support for non-exclusive CPU locking on FreeBSD and Windows using `cpubind` and `cpuset`.
 - Added support for exclusive CPU locking on Solaris, using `cpubind_exclusive` and `cpuset_exclusive`, which are added in this release.
 - Added thread prioritization using `thread_prio`, which is added in this release. `thread_prio` is supported on Linux, FreeBSD, Windows, and Solaris, but the exact effects of this setting are platform-specific; see the documentation for details.
 - Added support for setting `realtime` on Windows platforms.

For more information, see the description of the `ThreadConfig` parameter in the online documentation.

- **NDB Cluster APIs:** Added the `Ndb::setEventBufferQueueEmptyEpoch()` method, which makes it possible to enable queuing of empty events (event type `TE_EMPTY`). (Bug #22157845)
- **NDB Cluster APIs:** `Ndb_free_list_t` is a template used in the implementation of the NDB API to create a list keeping released API objects such as `NdbTransaction`, `NdbOperation`, and `NdbRecAttr`. One drawback to this template is that released API objects are kept in the list for the lifetime of the owning `Ndb` object, such that a transient peak in the demand for any object causes an effective leak in memory that persists until this `Ndb` object itself has been released.

This work adds statistics to each `Ndb_free_list` instance which samples the usage of objects maintained by the list; now, when objects are released, they can be released into the free list, or deallocated, based on the collected usage statistics.

- **MySQL NDB ClusterJ:** To make it easier for ClusterJ to handle fatal errors that require the `SessionFactory` to be closed, a new public method in the `SessionFactory` interface, `getConnectionPoolSessionCounts()`, has been created. When it returns zeros for all pooled connections, it means all sessions have been closed, at which point the `SessionFactory` can be closed and reopened. See [Reconnecting to an NDB Cluster](#) for more detail. (Bug #22353594)

Bugs Fixed

- **Performance:** A performance problem was found in an internal polling method `do_poll()` where the polling client did not check whether it had itself been woken up before completing the poll. Subsequent analysis showed that it is sufficient that only some clients in the polling queue receive data. `do_poll()` can then signal these clients and give up its polling rights, even if the maximum specified wait time (10 ms) has not expired.

This change allows `do_poll()` to continue polling until either the maximum specified wait time has expired, or the polling client itself has been woken up (by receiving what it was waiting for). This avoids unnecessary thread switches between client threads and thus reduces the associated overhead by as much as 10% in the API client, resulting in a significant performance improvement when client threads perform their own polling. (Bug #81229, Bug #23202735)

- **Incompatible Change:** When the data nodes are only partially connected to the API nodes, a node used for a pushdown join may get its request from a transaction coordinator on a different node, without

(yet) being connected to the API node itself. In such cases, the `NodeInfo` object for the requesting API node contained no valid info about the software version of the API node, which caused the `DBSPJ` block to assume (incorrectly) when aborting to assume that the API node used `NDB` version 7.2.4 or earlier, requiring the use of a backward compatibility mode to be used during query abort which sent a node failure error instead of the real error causing the abort.

Now, whenever this situation occurs, it is assumed that, if the `NDB` software version is not yet available, the API node version is greater than 7.2.4. (Bug #23049170)

- **Important Change:** When started with the `--initialize` option, `mysqld` no longer enables the `NDBCLUSTER` storage engine plugin. This change was needed to prevent attempted initialization of system databases as distributed (rather than as specific to individual SQL nodes), which could result in a metadata lock deadlock. This fix also brings the behavior of `--initialize` in this regard into line with that of the discontinued `--bootstrap` option, which started a minimal `mysqld` instance without enabling `NDB`. (Bug #22758238)
- **Microsoft Windows:** `ndb_mgmd` failed to start on 32-bit Windows platforms, due to an issue with calling dynamically loaded functions; such issues were also likely to occur with other `NDB` programs using `ndb_init()`. It was found that all of the functions used are already supported in targeted versions of Windows, so this problem is fixed by removing the dynamic loading of these functions and using the versions provided by the Windows header files instead. (Bug #80876, Bug #23014820)
- **Microsoft Windows:** When building MySQL Cluster on Windows using more than one parallel build job it was sometimes possible for the build to fail because `host_info.exe` could not be installed. To fix this problem, the `install_mcc` target is now always built prior to the `host_info` target. (Bug #80051, Bug #22566368)
- **Microsoft Windows:** Performing `ANALYZE TABLE` on a table having one or more indexes caused `ndbmtd` to fail with an `InvalidAttrInfo` error due to signal corruption. This issue occurred consistently on Windows, but could also be encountered on other platforms. (Bug #77716, Bug #21441297)
- Reserved send buffer for the loopback transporter, introduced in MySQL Cluster NDB 7.4.8 and used by API and management nodes for administrative signals, was calculated incorrectly. (Bug #23093656, Bug #22016081)

References: This issue is a regression of: Bug #21664515.

- During a node restart, re-creation of internal triggers used for verifying the referential integrity of foreign keys was not reliable, because it was possible that not all distributed TC and LDM instances agreed on all trigger identities. To fix this problem, an extra step is added to the node restart sequence, during which the trigger identities are determined by querying the current master node. (Bug #23068914)

References: See also: Bug #23221573.

- Following the forced shutdown of one of the 2 data nodes in a cluster where `NoOfReplicas=2`, the other data node shut down as well, due to arbitration failure. (Bug #23006431)
- Aborting a `CREATE LOGFILE GROUP` statement which had failed due to lack of shared global memory was not performed correctly, causing node failure. In addition, the transaction in which this occurred was not rolled back correctly, also causing any subsequent `CREATE LOGFILE GROUP` to fail. (Bug #22982618)
- The `ndbinfo.tc_time_track_stats` table uses histogram buckets to give a sense of the distribution of latencies. The sizes of these buckets were also reported as `HISTOGRAM BOUNDARY INFO` messages during data node startup; this printout was redundant and so has been removed. (Bug #22819868)

- Online upgrades from previous versions of MySQL Cluster to MySQL Cluster NDB 7.5 were not possible due to missing entries in the matrix used to test upgrade compatibility between versions. (Bug #22024947)
- A failure occurred in `DBTUP` in debug builds when variable-sized pages for a fragment totalled more than 4 GB. (Bug #21313546)
- Restoration of metadata with `ndb_restore -m` occasionally failed with the error message `Failed to create index...` when creating a unique index. While diagnosing this problem, it was found that the internal error `PREPARE_SEIZE_ERROR` (a temporary error) was reported as an unknown error. Now in such cases, `ndb_restore` retries the creation of the unique index, and `PREPARE_SEIZE_ERROR` is reported as NDB Error 748 `Busy during read of event table`. (Bug #21178339)

References: See also: Bug #22989944.

- `mysqld` did not shut down cleanly when executing `ndb_index_stat`. (Bug #21098142)

References: See also: Bug #23343739.

- The following improvements were made to the data node error logging mechanism:
 - Increased the message slot size 499 bytes to 999 bytes to prevent log messages from overwriting one another or from being truncated.
 - Added a `Trace file name` field to the output. This field contains the trace file name (without any path) and trace file number for the thread causing the trace.
 - `ndbmt` trace files are also now shown in the error log.

(Bug #21082710)

- `DBDICT` and `GETTABINFOREQ` queue debugging were enhanced as follows:
 - Monitoring by a data node of the progress of `GETTABINFOREQ` signals can be enabled by setting `DictTrace >= 2`.
 - Added the `ApiVerbose` configuration parameter, which enables NDB API debug logging for an API node where it is set greater than or equal to 2.
 - Added `DUMP` code 1229 which shows the current state of the `GETTABINFOREQ` queue. (See `DUMP 1229`.)

See also [The DBDICT Block](#). (Bug #20368450)

References: See also: Bug #20368354.

- When a write to the `ndb_binlog_index` table failed during a MySQL Server shutdown, `mysqld` killed the NDB binary logging thread. (Bug #81166, Bug #23142945)
- Memory associated with table descriptions was not freed by the internal table information method `NdbDictInterface::parseTableInfo()`. (Bug #81141, Bug #23130084)
- Improved memory usage by the internal `TransporterFacade` constructor when performing mutex array initialization. (Bug #81134, Bug #23127824)
- Fixed a memory leak that occurred when an error was raised in `ha_ndbcluster::get_metadata()` or one of the functions which this method calls. (Bug #81045, Bug #23089566)

- An internal function used to validate connections failed to update the connection count when creating a new `Ndb` object. This had the potential to create a new `Ndb` object for every operation validating the connection, which could have an impact on performance, particularly when performing schema operations. (Bug #80750, Bug #22932982)
- A table scan on an `NDB` table using neither an ordered index nor any Disk Data columns normally uses an `ACC` scan. If this happened while scanning an unique but unordered index which shrank (due to rows being deleted) after the scan started and then grew again (rows inserted), a single row that had been neither deleted nor inserted could be scanned twice. (Bug #80733, Bug #22926938)
- On OS X, `ndb_config` failed when an empty string was used for the `--host` option. (Bug #80689, Bug #22908696)
- Starting a backup in the `ndb_mgm` client after creating a large number of tables caused a forced shutdown of the cluster. (Bug #80640, Bug #228849958)
- When an SQL node was started, and joined the schema distribution protocol, another SQL node, already waiting for a schema change to be distributed, timed out during that wait. This was because the code incorrectly assumed that the new SQL node would also acknowledge the schema distribution even though the new node joined too late to be a participant in it.

As part of this fix, printouts of schema distribution progress now always print the more significant part of a bitmask before the less significant; formatting of bitmasks in such printouts has also been improved. (Bug #80554, Bug #22842538)

- The `ndb_print_file` utility failed consistently on Solaris 9 for SPARC. (Bug #80096, Bug #22579581)
- The MySQL Cluster Auto-Installer failed to work in various ways on different platforms. (Bug #79853, Bug #22502247)
- The internal function `ndbrequire()`, which, like `assert()`, evaluates a given expression and terminates the process if the expression does not evaluate as true, now includes the failed expression in its output to the error logs. (Bug #77021, Bug #21128318)
- Trying to drop a table during an ongoing backup failed with the error message `Unknown table`; now, it fails with `Unable to alter table as backup is in progress`. (Bug #47301, Bug #11755512)

References: See also: Bug #44695, Bug #11753280.

- **NDB Disk Data:** The following improvements were made to logging during restarts by data nodes using MySQL Cluster Disk Data:
 - The total amount of undo log to be applied by the data node is now provided as the total number of pages present in the log. This is a worst case estimate.
 - Progress information is now provided at regular intervals (once for each 30000 records) as the undo log is applied. This information is supplied as the number of records and number of undo log pages applied so far during the current restart.

(Bug #22513381)

- **NDB Cluster APIs:** Deletion of `Ndb` objects used a disproportionately high amount of CPU. (Bug #22986823)
- **NDB Cluster APIs:** Executing a transaction with an `NdbIndexOperation` based on an obsolete unique index caused the data node process to fail. Now the index is checked in such cases, and if it cannot be used the transaction fails with an appropriate error. (Bug #79494, Bug #22299443)

- `mysql_upgrade` failed to upgrade the `sys` schema if a `sys` database directory existed but was empty. (Bug #81352, Bug #23249846, Bug #22875519)

Changes in MySQL NDB Cluster 7.5.1 (5.7.11-ndb-7.5.1) (2016-03-18, Development Milestone)

MySQL Cluster NDB 7.5.1 is a new release of MySQL Cluster NDB 7.5, based on MySQL Server 5.7 and including features in version 7.5 of the `NDB` storage engine, as well as fixing recently discovered bugs in previous MySQL Cluster releases.

Obtaining MySQL Cluster NDB 7.5. MySQL Cluster NDB 7.5 source code and binaries can be obtained from <http://dev.mysql.com/downloads/cluster/>.

For an overview of changes made in MySQL Cluster NDB 7.5, see [What is New in MySQL NDB Cluster 7.5](#).

This release also incorporates all bugfixes and changes made in previous MySQL Cluster releases, as well as all bugfixes and feature changes which were added in mainline MySQL 5.7 through MySQL 5.7.11 (see [Changes in MySQL 5.7.11 \(2016-02-05, General Availability\)](#)).

- [Functionality Added or Changed](#)
- [Bugs Fixed](#)

Functionality Added or Changed

- **Important Change:** When creating or adding columns to `NDB` tables, the default value used for both the `COLUMN_FORMAT` and `ROW_FORMAT` options is now `DYNAMIC` rather than `FIXED`.

This change does not affect the row format or column format used by existing tables. New columns added to such tables use the new defaults, and existing columns are changed to use these as well, provided that the `ALTER TABLE` statement in question uses `ALGORITHM=COPY`. Note that this cannot be done implicitly if `mysqld` is run with `--ndb-allow-copying-alter-table=FALSE` (the default is `TRUE`).

A new MySQL server option `--ndb-default-column-format` is added for backwards compatibility; set this to `FIXED` to force the old defaults to be used for `COLUMN_FORMAT` and `ROW_FORMAT`.



Note

This behavior change is superseded in MySQL Cluster NDB 7.5.4.

- Scans have been improved by replacing the `DIH_SCAN_GET_NODES_REQ` signal, formerly used for communication between the `DBTC` and `DBDIH` kernel blocks in `NDB`, with the `DIGETNODESREQ` signal, which supports direct execution and allows for the elimination of `DIH_SCAN_GET_NODES_REF` and `DIH_SCAN_GET_NODES_CONF`, as well as for `DIH_SCAN_TAB_REQ` and `DIH_SCAN_TAB_COMPLETE_REP` signals to `EXECUTE_DIRECT`. This enables higher scalability of data nodes when used for scan operations by decreasing the use of CPU resources for scan operations by an estimated five percent in some cases. This should also improve response times, which could help prevent issues with overload of the main threads.

As part of these changes, scans made in the `BACKUP` kernel block have also been improved and made more efficient.

References: See also: Bug #80640, Bug #22884995.

- The following improvements have been made to event buffer reporting in the cluster log:
 - Each report now identifies the API node that sent it.
 - The fields shown in the report have been improved and expanded. Percentages are now better specified and used only when appropriate. For improved clarity, the `apply_epoch` and `latest_epoch` fields have been renamed to `latest_consumed_epoch` and `latest_buffered_epoch`, respectively. The ID of the `Ndb` object serving as the source of the report is now shown, as is the reason for making the report (as the `report_reason` field).
 - The frequency of unnecessary reporting has been reduced by limiting reports to those showing only significant changes in event buffer usage.
 - The MGM API adds a new `NDB_LE_EventBufferStatus2` event type to handle the additional information provided by the new event buffer reporting. The `NDB_LE_EventBufferStatus` event type used in older versions of MySQL Cluster is now deprecated, and will eventually be removed.

For more information, see [Event Buffer Reporting in the Cluster Log](#), as well as [The `ndb_logevent` Structure](#).

Bugs Fixed

- **Important Change:** The minimum value for the `BackupDataBufferSize` data node configuration parameter has been lowered from 2 MB to 512 KB. The default and maximum values for this parameter remain unchanged. (Bug #22749509)
- During node failure handling, the request structure used to drive the cleanup operation was not maintained correctly when the request was executed. This led to inconsistencies that were harmless during normal operation, but these could lead to assertion failures during node failure handling, with subsequent failure of additional nodes. (Bug #22643129)
- The previous fix for a lack of mutex protection for the internal `TransporterFacade::deliver_signal()` function was found to be incomplete in some cases. (Bug #22615274)

References: This issue is a regression of: Bug #77225, Bug #21185585.

- Compilation of MySQL with Visual Studio 2015 failed in `ConfigInfo.cpp`, due to a change in Visual Studio's handling of spaces and concatenation. (Bug #22558836, Bug #80024)
- When setup of the binary log as an atomic operation on one SQL node failed, this could trigger a state in other SQL nodes in which they appeared to detect the SQL node participating in schema change distribution, whereas it had not yet completed binary log setup. This could in turn cause a deadlock on the global metadata lock when the SQL node still retrying binary log setup needed this lock, while another `mysqld` had taken the lock for itself as part of a schema change operation. In such cases, the second SQL node waited for the first one to act on its schema distribution changes, which it was not yet able to do. (Bug #22494024)
- Duplicate key errors could occur when `ndb_restore` was run on a backup containing a unique index. This was due to the fact that, during restoration of data, the database can pass through one or more inconsistent states prior to completion, such an inconsistent state possibly having duplicate values for a column which has a unique index. (If the restoration of data is preceded by a run with `--disable-indexes` and followed by one with `--rebuild-indexes`, these errors are avoided.)

Added a check for unique indexes in the backup which is performed only when restoring data, and which does not process tables that have explicitly been excluded. For each unique index found, a warning is now printed. (Bug #22329365)

- When setting up event logging for `ndb_mgmd` on Windows, MySQL Cluster tries to add a registry key to `HKEY_LOCAL_MACHINE`, which fails if the user does not have access to the registry. In such cases `ndb_mgmd` logged the error `Could neither create or open key`, which is not accurate and which can cause confusion for users who may not realize that file logging is available and being used. Now in such cases, `ndb_mgmd` logs a warning `Could not create or access the registry key needed for the application to log to the Windows EventLog. Run the application with sufficient privileges once to create the key, or add the key manually, or turn off logging for that application.` An error (as opposed to a warning) is now reported in such cases only if there is no available output at all for `ndb_mgmd` event logging. (Bug #20960839)
- `NdbDictionary` metadata operations had a hard-coded 7-day timeout, which proved to be excessive for short-lived operations such as retrieval of table definitions. This could lead to unnecessary hangs in user applications which were difficult to detect and handle correctly. To help address this issue, timeout behaviour is modified so that read-only or short-duration dictionary interactions have a 2-minute timeout, while schema transactions of potentially long duration retain the existing 7-day timeout.

Such timeouts are intended as a safety net: In the event of problems, these return control to users, who can then take corrective action. Any reproducible issue with `NdbDictionary` timeouts should be reported as a bug. (Bug #20368354)

- Optimization of signal sending by buffering and sending them periodically, or when the buffer became full, could cause `SUB_GCP_COMPLETE_ACK` signals to be excessively delayed. Such signals are sent for each node and epoch, with a minimum interval of `TimeBetweenEpochs`; if they are not received in time, the `SUMA` buffers can overflow as a result. The overflow caused API nodes to be disconnected, leading to current transactions being aborted due to node failure. This condition made it difficult for long transactions (such as altering a very large table), to be completed. Now in such cases, the `ACK` signal is sent without being delayed. (Bug #18753341)
- MySQL Cluster did not compile correctly with Microsoft Visual Studio 2015, due to a change from previous versions in the VS implementation of the `_vsnprintf()` function. (Bug #80276, Bug #22670525)
- When setting CPU spin time, the value was needlessly cast to a boolean internally, so that setting it to any nonzero value yielded an effective value of 1. This issue, as well as the fix for it, apply both to setting the `SchedulerSpinTimer` parameter and to setting `spintime` as part of a `ThreadConfig` parameter value. (Bug #80237, Bug #22647476)
- Processing of local checkpoints was not handled correctly on Mac OS X, due to an uninitialized variable. (Bug #80236, Bug #22647462)
- A logic error in an `if` statement in `storage/ndb/src/kernel/blocks/dbacc/DbaccMain.cpp` rendered useless a check for determining whether `ZREAD_ERROR` should be returned when comparing operations. This was detected when compiling with `gcc` using `-Werror=logical-op`. (Bug #80155, Bug #22601798)

References: This issue is a regression of: Bug #21285604.

- Suppressed a `CMake` warning that was caused by use of an incorrectly quoted variable name. (Bug #80066, Bug #22572632)
- When using `CREATE INDEX` to add an index on either of two `NDB` tables sharing circular foreign keys, the query succeeded but a temporary table was left on disk, breaking the foreign key constraints. This issue was also observed when attempting to create an index on a table in the middle of a chain of foreign keys—that is, a table having both parent and child keys, but on different tables. The problem did not occur when using `ALTER TABLE` to perform the same index creation operation; and subsequent

analysis revealed unintended differences in the way such operations were performed by `CREATE INDEX`.

To fix this problem, we now make sure that operations performed by a `CREATE INDEX` statement are always handled internally in the same way and at the same time that the same operations are handled when performed by `ALTER TABLE` or `DROP INDEX`. (Bug #79156, Bug #22173891)

- The `PortNumber` SCI, SHM, and TCP configuration parameters, which were deprecated in MySQL 5.1.3, have now been removed and are no longer accepted in configuration files.

This change does not affect the `PortNumber` management node configuration parameter, whose behavior remains unaltered. (Bug #77405, Bug #21280456)

- For busy servers, client connection or communication failure could occur if an I/O-related system call was interrupted. The `mysql_options()` C API function now has a `MYSQL_OPT_RETRY_COUNT` option to control the number of retries for interrupted system calls. (Bug #22336527)

References: See also: Bug #22389653.

Changes in MySQL NDB Cluster 7.5.0 (5.7.10-ndb-7.5.0) (2016-02-05, Development Milestone)

MySQL Cluster NDB 7.5.0 is a new release of MySQL Cluster NDB 7.5, based on MySQL Server 5.7 and including features in version 7.5 of the `NDB` storage engine, as well as fixing recently discovered bugs in previous MySQL Cluster releases.

Obtaining MySQL Cluster NDB 7.5. MySQL Cluster NDB 7.5 source code and binaries can be obtained from <http://dev.mysql.com/downloads/cluster/>.

For an overview of changes made in MySQL Cluster NDB 7.5, see [What is New in MySQL NDB Cluster 7.5](#).

This release also incorporates all bugfixes and changes made in previous MySQL Cluster releases, as well as all bugfixes and feature changes which were added in mainline MySQL 5.7 through MySQL 5.7.10 (see [Changes in MySQL 5.7.10 \(2015-12-07, General Availability\)](#)).

- [Functionality Added or Changed](#)
- [Bugs Fixed](#)

Functionality Added or Changed

- **Important Change:** Previously, the `ndbinfo` information database included lookup tables that used the `MyISAM` storage engine. This dependency on `MyISAM` has now been removed. (Bug #20075747)
- **Important Change:** Previously, the `NDB` scheduler always optimized for speed against throughput in a predetermined manner (this was hard coded); this balance can now be set using the `SchedulerResponsiveness` data node configuration parameter. This parameter accepts an integer in the range of 0-10 inclusive, with 5 as the default. Higher values provide better response times relative to throughput. Lower values provide increased throughput, but impose longer response times. (Bug #78531, Bug #21889312)
- **Important Change:** A number of MySQL Cluster data node configuration parameters were deprecated in earlier versions of MySQL Cluster, and have been removed with this release. These parameters include `Id`, `NoOfDiskPagesToDiskDuringRestartTUP`, `NoOfDiskPagesToDiskDuringRestartACC`, `NoOfDiskPagesToDiskAfterRestartACC`,

`NoOfDiskPagesToDiskAfterRestartTUP`, `ReservedSendBufferMemory`, `MaxNoOfIndexes`, and `Discless` (use `Diskless` instead), as well as `DiskCheckpointSpeed` and `DiskCheckpointSpeedInRestart`. The archaic and unused `ByteOrder` computer configuration parameter has also been removed, as well as the unused `MaxNoOfSavedEvents` management node configuration parameter. These parameters are no longer supported; most of them already did not have (or no longer had) any effect. Trying to use any of these parameters in a MySQL Cluster configuration file now results in an error.

For more information, see [What is New in MySQL NDB Cluster 7.5](#). (Bug #77404, Bug #21280428)

- **Important Change:** The `ExecuteOnComputer` configuration parameter for management, data, and API nodes is now deprecated, and is subject to removal in a future MySQL Cluster release. For all types of MySQL Cluster nodes, you should now use the `HostName` parameter exclusively for identifying hosts in the cluster configuration file.

This information is also now displayed in the output of `ndb_config --configinfo --xml`. (Bug #53052, Bug #11760628)

- **Important Change:** The `ndbinfo` database can now provide default and current information about MySQL Cluster node configuration parameters as a result of the following changes:
 1. The `config_params` table has been enhanced with additional columns providing information about each configuration parameter, including its type, default, and maximum and minimum values (where applicable).
 2. A new `config_values` table has been added. A row in this table shows the current value of a parameter on a given node.

You can obtain values of MySQL Cluster configuration parameters by name using a join on these two tables such as the one shown here:

```
SELECT  p.param_name AS Name,
        v.node_id AS Node,
        p.param_type AS Type,
        p.param_default AS 'Default',
        v.config_value AS Current
FROM    config_params p
JOIN    config_values v
ON      p.param_number = v.config_param
WHERE   p.param_name IN ('NodeId', 'HostName', 'DataMemory', 'IndexMemory');
```

- Deprecated MySQL Cluster node configuration parameters are now indicated as such by `ndb_config --configinfo --xml`. For each parameter currently deprecated, the corresponding `<param/>` tag in the XML output now includes the attribute `deprecated="true"`. (Bug #21127135)
- Added the `--ndb-cluster-connection-pool-nodeids` option for `mysqld`, which can be used to specify a list of nodes by node ID for connection pooling. The number of node IDs in the list must equal the value set for `--ndb-cluster-connection-pool`. (Bug #19521789)
- Added the `PROMPT` command in the `ndb_mgm` client. This command has the syntax `PROMPT string`, which sets the client's prompt to `string`. Issuing the command without an argument causes the prompt to be reset to the default (`ndb_mgm>`). See [Commands in the NDB Cluster Management Client](#), for more information. (Bug #18421338)
- When the `--database` option has not been specified for `ndb_show_tables`, and no tables are found in the `TEST_DB` database, an appropriate warning message is now issued. (Bug #50633, Bug #11758430)

- The **NDB** storage engine now uses the improved records-per-key interface for index statistics introduced for the optimizer in MySQL 5.7. Some improvements due to this change are listed here:
 - The optimizer can now choose better execution plans for queries on **NDB** tables in many cases where a less optimal join index or table join order would previously have been chosen.
 - **EXPLAIN** now provides more accurate row estimates than previously.
 - Improved cardinality estimates can be obtained from **SHOW INDEX**.
- **NDB Replication:** Normally, **RESET SLAVE** causes all entries to be deleted from the `mysql.ndb_apply_status` table. This release adds the `ndb_clear_apply_status` system variable, which makes it possible to override this behavior. This variable is **ON** by default; setting it to **OFF** keeps **RESET SLAVE** from purging the `ndb_apply_status` table. (Bug #12630403)

Bugs Fixed

- **Incompatible Change; NDB Cluster APIs:** The `pollEvents2()` method now returns -1, indicating an error, whenever a negative value is used for the time argument. (Bug #20762291)
- **Important Change:** The options controlling behavior of **NDB** programs with regard to the number and timing of successive attempts to connect to a management server have changed as listed here:
 - The minimum value for the `--connect-retry-delay` option common to all **NDB** programs has been changed from 0 to 1; this means that all **NDB** programs now wait at least 1 second between successive connection attempts, and it is no longer possible to set a waiting time equal to 0.
 - The semantics for the `--connect-retries` option have changed slightly, such that the value of this option now sets the number of times an **NDB** program tries to connect to a management server. Setting this option to 0 now causes the program to attempt the connection indefinitely, until it either succeeds or is terminated by other means (such as **kill**).
 - In addition, the default for the `--connect-retries` option for the `ndb_mgm` client has been changed from 3 to 12, so that the minimum, maximum, and default values for this option when used with `ndb_mgm` are now exactly the same as for all other **NDB** programs.

The `ndb_mgm --try-reconnect` option, although deprecated in MySQL Cluster NDB 7.4, continues to be supported as a synonym for `ndb_mgm --connect-retries` to provide backwards compatibility. The default value for `--try-reconnect` has also been changed from 3 to 12, respectively, so that this option continues to behave in the exactly in the same way as `--connect-retries`.

(Bug #22116937)

- **Important Change:** In previous versions of MySQL Cluster, other DDL operations could not be part of `ALTER ONLINE TABLE ... RENAME ...` (This was disallowed by the fix for BUG#16021021.) MySQL Cluster NDB 7.5 makes the following changes:
 - Support for the **ONLINE** and **OFFLINE** keywords, which was deprecated in MySQL Cluster NDB 7.3, is now removed, and use of these now causes a syntax error; the **NDB** storage engine now accepts *only* `ALGORITHM = DEFAULT`, `ALGORITHM = COPY`, and `ALGORITHM = INPLACE` to specify whether the **ALTER** operation is copying or in-place, just as in the standard MySQL Server.
 - **NDB** now allows `ALTER TABLE ... ALGORITHM=COPYING RENAME`.
- (Bug #20804269, Bug #76543, Bug #20479917, Bug #75797)

References: See also: Bug #16021021.

- **Important Change; NDB Cluster APIs:** `Ndb::pollEvents()` is now compatible with the `TE_EMPTY`, `TE_INCONSISTENT`, and `TE_OUT_OF_MEMORY` event types introduced in MySQL Cluster NDB 7.4.3. For detailed information about this change, see the description of this method in the *MySQL Cluster API Developer Guide*. (Bug #20646496)
- **Important Change; NDB Cluster APIs:** Added the method `Ndb::isExpectingHigherQueuedEpochs()` to the NDB API to detect when additional, newer event epochs were detected by `pollEvents2()`.

The behavior of `Ndb::pollEvents()` has also been modified such that it now returns `NDB_FAILURE_GCI` (equal to `~(Uint64) 0`) when a cluster failure has been detected. (Bug #18753887)

- **Important Change; NDB Cluster APIs:** To release the memory used for dropped event operations, the event API formerly depended on `pollEvents()` and `nextEvent()` to consume all events possibly referring to the dropped events. This dependency between `dropEventOperation()` and the first two methods required the entire event buffer to be read before attempting to release event operation memory (that is, until successive calls to `pollEvents()` and `nextEvent()` returned no more events).

A related cleanup issue arose following the reset of the event buffer (when all event operations had previously been dropped), and the event buffer was truncated by the first `createEventOperation()` call subsequent to the reset.

To fix these problems, the event buffer is now cleared when the last event operation is dropped, rather than waiting for a subsequent create operation which might or might not occur. Memory taken up by dropped event operations is also now released when the event queue has been cleared, which removes the hidden requirement for consuming all events to free up memory. In addition, event operation memory is now released as soon as all events referring to the operation have been consumed, rather than waiting for the entire event buffer to be consumed. (Bug #78145, Bug #21661297)

- **Important Change; NDB Cluster APIs:** The MGM API error-handling functions `ndb_mgm_get_latest_error()`, `ndb_mgm_get_latest_error_msg()`, and `ndb_mgm_get_latest_error_desc()` each failed when used with a `NULL` handle. You should note that, although these functions are now null-safe, values returned in this case are arbitrary and not meaningful. (Bug #78130, Bug #21651706)
- **Important Change; NDB Cluster APIs:** The following NDB API methods were not actually implemented and have been removed from the sources:
 - `Datafile` methods: `getNode()`, `setNode()`, and `getFileNo()`
 - `Undofile` methods: `getNode()`, `setNode()`, and `getFileNo()`
 - `Table` methods: `getObjectType()` and `setObjectType()`
- A serious regression was inadvertently introduced in MySQL Cluster NDB 7.4.8 whereby local checkpoints and thus restarts often took much longer than expected. This occurred due to the fact that the setting for `MaxDiskWriteSpeedOwnRestart` was ignored during restarts and the value of `MaxDiskWriteSpeedOtherNodeRestart`, which is much lower by default than the default for `MaxDiskWriteSpeedOwnRestart`, was used instead. This issue affected restart times and performance only and did not have any impact on normal operations. (Bug #22582233)
- The epoch for the latest restorable checkpoint provided in the cluster log as part of its reporting for `EventBufferStatus` events (see [NDB Cluster : Messages in the Cluster Log](#)) was not well defined and thus unreliable; depending on various factors, the reported epoch could be the one currently being consumed, the one most recently consumed, or the next one queued for consumption.

This fix ensures that the latest restorable global checkpoint is always regarded as the one that was most recently completely consumed by the user, and thus that it was the latest restorable global checkpoint that existed at the time the report was generated. (Bug #22378288)

- Added the `--ndb-allow-copying-alter-table` option for `mysqld`. Setting this option (or the equivalent system variable `ndb_allow_copying_alter_table`) to `OFF` keeps `ALTER TABLE` statements from performing copying operations. The default value is `ON`. (Bug #22187649)

References: See also: Bug #17400320.

- Attempting to create an `NDB` table having greater than the maximum supported combined width for all `BIT` columns (4096) caused data node failure when these columns were defined with `COLUMN_FORMAT DYNAMIC`. (Bug #21889267)
- Creating a table with the maximum supported number of columns (512) all using `COLUMN_FORMAT DYNAMIC` led to data node failures. (Bug #21863798)
- In a MySQL Cluster with multiple LDM instances, all instances wrote to the node log, even inactive instances on other nodes. During restarts, this caused the log to be filled with messages from other nodes, such as the messages shown here:

```
2015-06-24 00:20:16 [ndbd] INFO      -- We are adjusting Max Disk Write Speed,
a restart is ongoing now
...
2015-06-24 01:08:02 [ndbd] INFO      -- We are adjusting Max Disk Write Speed,
no restarts ongoing anymore
```

Now this logging is performed only by the active LDM instance. (Bug #21362380)

- Backup block states were reported incorrectly during backups. (Bug #21360188)

References: See also: Bug #20204854, Bug #21372136.

- For a timeout in `GET_TABINFOREQ` while executing a `CREATE INDEX` statement, `mysqld` returned Error 4243 (`Index not found`) instead of the expected Error 4008 (`Receive from NDB failed`).

The fix for this bug also fixes similar timeout issues for a number of other signals that are sent the `DBDICT` kernel block as part of DDL operations, including `ALTER_TAB_REQ`, `CREATE_INDX_REQ`, `DROP_FK_REQ`, `DROP_INDX_REQ`, `INDEX_STAT_REQ`, `DROP_FILE_REQ`, `CREATE_FILEGROUP_REQ`, `DROP_FILEGROUP_REQ`, `CREATE_EVENT`, `WAIT_GCP_REQ`, `DROP_TAB_REQ`, and `LIST_TABLES_REQ`, as well as several internal functions used in handling `NDB` schema operations. (Bug #21277472)

References: See also: Bug #20617891, Bug #20368354, Bug #19821115.

- Previously, multiple send threads could be invoked for handling sends to the same node; these threads then competed for the same send lock. While the send lock blocked the additional send threads, work threads could be passed to other nodes.

This issue is fixed by ensuring that new send threads are not activated while there is already an active send thread assigned to the same node. In addition, a node already having an active send thread assigned to it is no longer visible to other, already active, send threads; that is, such a node is longer added to the node list when a send thread is currently assigned to it. (Bug #20954804, Bug #76821)

- Queueing of pending operations when the redo log was overloaded (`DefaultOperationRedoProblemAction` API node configuration parameter) could lead to timeouts when data nodes ran out of redo log space (`P_TAIL_PROBLEM` errors). Now when the redo log is full, the node aborts requests instead of queuing them. (Bug #20782580)

References: See also: Bug #20481140.

- An `NDB` event buffer can be used with an `Ndb` object to subscribe to table-level row change event streams. Users subscribe to an existing event; this causes the data nodes to start sending event data signals (`SUB_TABLE_DATA`) and epoch completion signals (`SUB_GCP_COMPLETE`) to the `Ndb` object. `SUB_GCP_COMPLETE_REP` signals can arrive for execution in concurrent receiver thread before completion of the internal method call used to start a subscription.

Execution of `SUB_GCP_COMPLETE_REP` signals depends on the total number of `SUMA` buckets (sub data streams), but this may not yet have been set, leading to the present issue, when the counter used for tracking the `SUB_GCP_COMPLETE_REP` signals (`TOTAL_BUCKETS_INIT`) was found to be set to erroneous values. Now `TOTAL_BUCKETS_INIT` is tested to be sure it has been set correctly before it is used. (Bug #20575424, Bug #76255)

References: See also: Bug #20561446, Bug #21616263.

- `NDB` statistics queries could be delayed by the error delay set for `ndb_index_stat_option` (default 60 seconds) when the index that was queried had been marked with internal error. The same underlying issue could also cause `ANALYZE TABLE` to hang when executed against an `NDB` table having multiple indexes where an internal error occurred on one or more but not all indexes.

Now in such cases, any existing statistics are returned immediately, without waiting for any additional statistics to be discovered. (Bug #20553313, Bug #20707694, Bug #76325)

- Memory allocated when obtaining a list of tables or databases was not freed afterward. (Bug #20234681, Bug #74510)

References: See also: Bug #18592390, Bug #72322.

- Added the `BackupDiskWriteSpeedPct` data node parameter. Setting this parameter causes the data node to reserve a percentage of its maximum write speed (as determined by the value of `MaxDiskWriteSpeed`) for use in local checkpoints while performing a backup. `BackupDiskWriteSpeedPct` is interpreted as a percentage which can be set between 0 and 90 inclusive, with a default value of 50. (Bug #20204854)

References: See also: Bug #21372136.

- After restoring the database schema from backup using `ndb_restore`, auto-discovery of restored tables in transactions having multiple statements did not work correctly, resulting in `Deadlock found when trying to get lock; try restarting transaction` errors.

This issue was encountered both in the `mysql` client, as well as when such transactions were executed by application programs using Connector/J and possibly other MySQL APIs.

Prior to upgrading, this issue can be worked around by executing `SELECT TABLE_NAME, TABLE_SCHEMA FROM INFORMATION_SCHEMA.TABLES WHERE ENGINE = 'NDBCLUSTER'` on all SQL nodes following the restore operation, before executing any other statements. (Bug #18075170)

- Using `ndb_mgm STOP -f` to force a node shutdown even when it triggered a complete shutdown of the cluster, it was possible to lose data when a sufficient number of nodes were shut down, triggering a cluster shutdown, and the timing was such that `SUMA` handovers had been made to nodes already in the process of shutting down. (Bug #17772138)
- When using a sufficiently large value for `TransactionDeadlockDetectionTimeout` and the default value for `sort_buffer_size`, executing `SELECT * FROM ndbinfo.cluster_operations ORDER BY transid` with multiple concurrent conflicting or deadlocked transactions, each transaction

having several pending operations, caused the SQL node where the query was run to fail. (Bug #16731538, Bug #67596)

- The `ndbinfo.config_params` table is now read-only. (Bug #11762750, Bug #55383)
- `NDB` failed during a node restart due to the status of the current local checkpoint being set but not as active, even though it could have other states under such conditions. (Bug #78780, Bug #21973758)
- `ndbmttd` checked for signals being sent only after a full cycle in `run_job_buffers`, which is performed for all job buffer inputs. Now this is done as part of `run_job_buffers` itself, which avoids executing for extended periods of time without sending to other nodes or flushing signals to other threads. (Bug #78530, Bug #21889088)
- When attempting to enable index statistics, creation of the required system tables, events and event subscriptions often fails when multiple `mysqld` processes using index statistics are started concurrently in conjunction with starting, restarting, or stopping the cluster, or with node failure handling. This is normally recoverable, since the affected `mysqld` process or processes can (and do) retry these operations shortly thereafter. For this reason, such failures are no longer logged as warnings, but merely as informational events. (Bug #77760, Bug #21462846)
- It was possible to end up with a lock on the send buffer mutex when send buffers became a limiting resource, due either to insufficient send buffer resource configuration, problems with slow or failing communications such that all send buffers became exhausted, or slow receivers failing to consume what was sent. In this situation worker threads failed to allocate send buffer memory for signals, and attempted to force a send in order to free up space, while at the same time the send thread was busy trying to send to the same node or nodes. All of these threads competed for taking the send buffer mutex, which resulted in the lock already described, reported by the watchdog as `Stuck in Send`. This fix is made in two parts, listed here:
 1. The send thread no longer holds the global send thread mutex while getting the send buffer mutex; it now releases the global mutex prior to locking the send buffer mutex. This keeps worker threads from getting stuck in send in such cases.
 2. Locking of the send buffer mutex done by the send threads now uses a try-lock. If the try-lock fails, the node to make the send to is reinserted at the end of the list of send nodes in order to be retried later. This removes the `Stuck in Send` condition for the send threads.

(Bug #77081, Bug #21109605)

- **NDB Disk Data:** A unique index on a column of an `NDB` table is implemented with an associated internal ordered index, used for scanning. While dropping an index, this ordered index was dropped first, followed by the drop of the unique index itself. This meant that, when the drop was rejected due to (for example) a constraint violation, the statement was rejected but the associated ordered index remained deleted, so that any subsequent operation using a scan on this table failed. We fix this problem by causing the unique index to be removed first, before removing the ordered index; removal of the related ordered index is no longer performed when removal of a unique index fails. (Bug #78306, Bug #21777589)
- **NDB Replication:** While the binary log injector thread was handling failure events, it was possible for all `NDB` tables to be left indefinitely in read-only mode. This was due to a race condition between the binary log injector thread and the utility thread handling events on the `ndb_schema` table, and to the fact that, when handling failure events, the binary log injector thread places all `NDB` tables in read-only mode until all such events are handled and the thread restarts itself.

When the binary log inject thread receives a group of one or more failure events, it drops all other existing event operations and expects no more events from the utility thread until it has handled all of the failure events and then restarted itself. However, it was possible for the utility thread to continue

attempting binary log setup while the injector thread was handling failures and thus attempting to create the schema distribution tables as well as event subscriptions on these tables. If the creation of these tables and event subscriptions occurred during this time, the binary log injector thread's expectation that there were no further event operations was never met; thus, the injector thread never restarted, and NDB tables remained in read-only as described previously.

To fix this problem, the `Ndb` object that handles schema events is now definitely dropped once the `ndb_schema` table drop event is handled, so that the utility thread cannot create any new events until after the injector thread has restarted, at which time, a new `Ndb` object for handling schema events is created. (Bug #17674771, Bug #19537961, Bug #22204186, Bug #22361695)

- **NDB Cluster APIs:** The binary log injector did not work correctly with `TE_INCONSISTENT` event type handling by `Ndb::nextEvent()`. (Bug #22135541)

References: See also: Bug #20646496.

- **NDB Cluster APIs:** While executing `dropEvent()`, if the coordinator `DBDICT` failed after the subscription manager (`SUMA` block) had removed all subscriptions but before the coordinator had deleted the event from the system table, the dropped event remained in the table, causing any subsequent drop or create event with the same name to fail with NDB error 1419 `Subscription already dropped` or error 746 `Event name already exists`. This occurred even when calling `dropEvent()` with a nonzero force argument.

Now in such cases, error 1419 is ignored, and `DBDICT` deletes the event from the table. (Bug #21554676)

- **NDB Cluster APIs:** Creation and destruction of `Ndb_cluster_connection` objects by multiple threads could make use of the same application lock, which in some cases led to failures in the global dictionary cache. To alleviate this problem, the creation and destruction of several internal NDB API objects have been serialized. (Bug #20636124)
- **NDB Cluster APIs:** When an `Ndb` object created prior to a failure of the cluster was reused, the event queue of this object could still contain data node events originating from before the failure. These events could reference “old” epochs (from before the failure occurred), which in turn could violate the assumption made by the `nextEvent()` method that epoch numbers always increase. This issue is addressed by explicitly clearing the event queue in such cases. (Bug #18411034)

References: See also: Bug #20888668.

- **NDB Cluster APIs:** `Ndb::pollEvents()` and `pollEvents2()` were slow to receive events, being dependent on other client threads or blocks to perform polling of transporters on their behalf. This fix allows a client thread to perform its own transporter polling when it has to wait in either of these methods.

Introduction of transporter polling also revealed a problem with missing mutex protection in the `ndbcluster_binlog` handler, which has been added as part of this fix. (Bug #79311, Bug #20957068, Bug #22224571)

- **NDB Cluster APIs:** After the initial restart of a node following a cluster failure, the cluster failure event added as part of the restart process was deleted when an event that existed prior to the restart was later deleted. This meant that, in such cases, an Event API client had no way of knowing that failure handling was needed. In addition, the GCI used for the final cleanup of deleted event operations, performed by `pollEvents()` and `nextEvent()` when these methods have consumed all available events, was lost. (Bug #78143, Bug #21660947)

Release Series Changelogs: MySQL Cluster NDB 7.5

This section contains unified changelog information for the MySQL Cluster NDB 7.5 release series.

For changelogs covering individual MySQL Cluster NDB 7.5 releases, see [NDB Cluster Release Notes](#).

For general information about features added in MySQL Cluster NDB 7.5, see [What is New in MySQL NDB Cluster 7.5](#).

For an overview of features added in MySQL 5.7 that are not specific to MySQL Cluster, see [What Is New in MySQL 5.7](#). For a complete list of all bugfixes and feature changes made in MySQL 5.7 that are not specific to MySQL Cluster, see the MySQL 5.7 [Release Notes](#).

- [Changes in MySQL NDB Cluster 7.5.4 \(5.7.16-ndb-7.5.4\)](#)
- [Changes in MySQL NDB Cluster 7.5.3 \(5.7.13-ndb-7.5.3\)](#)
- [Changes in MySQL NDB Cluster 7.5.2 \(5.7.12-ndb-7.5.2\)](#)
- [Changes in MySQL NDB Cluster 7.5.1 \(5.7.11-ndb-7.5.1\)](#)
- [Changes in MySQL NDB Cluster 7.5.0 \(5.7.10-ndb-7.5.0\)](#)

Changes in MySQL NDB Cluster 7.5.4 (5.7.16-ndb-7.5.4)

Functionality Added or Changed

- **Important Change; Packaging:** Naming and organization of the RPMs provided for MySQL Cluster have been changed to align with those released for the MySQL server. All MySQL Cluster RPM package names are now prefixed with `mysql-cluster`. Data nodes are now installed using the `data-node` package; management nodes are now installed from the `management-server` package; and SQL nodes require the `server` and `common` packages. **Important:** SQL nodes must use the `mysql-cluster` version of these RPMs; the versions released for the standard MySQL server do not provide support for the NDB storage engine. All client programs, including both the `mysql` client and the `ndb_mgm` management client, are now included in the `client` RPM.

For more information, see [Installing NDB Cluster from RPM](#).

- Added the `--print-sql-log` option for the `ndb_restore` program included with the MySQL Cluster distribution. This option causes the program to log SQL statements to `stdout`.

Note that each table being restored in this fashion must have an explicitly defined primary key; the hidden primary key implemented by the NDB storage engine is *not* sufficient for this purpose. (Bug #13511949)

- For fully replicated tables, `ndb_desc` shows only nodes holding main fragment replicas for partitions; nodes with copy fragment replicas only are ignored. To make this information available in the `mysql` client, several new tables have been introduced in the `ndbinfo` information database. These tables are listed here, with brief descriptions:
 - `dict_obj_info` provides the names and types of database (DICT) objects in NDB, such as tables and indexes, as well as information about parent objects where applicable
 - `table_distribution_status` provides NDB table distribution status information
 - `table_fragments` provides information about the distribution of NDB table fragments
 - `table_info` provides information about logging, checkpointing, storage, and other options in force for each NDB table

- `table_replicas` provides information about fragment replicas

For more information, see [The ndbinfo NDB Cluster Information Database](#). (Bug #81762, Bug #23547643)

Bugs Fixed

- **Important Change:** The default value of the `--ndb-default-column-format` server option has been changed from `DYNAMIC` to `FIXED`. This has been done for backwards compatibility. Only the default has been changed; setting this option to `DYNAMIC` continues to cause `DYNAMIC` to be used for `ROW_FORMAT` and `COLUMN_FORMAT` unless overridden. (Bug #24487363)
- **Important Change:** Event buffer status reporting has been improved by altering the semantics for calculating lag or slippage. Rather than defining this lag as the number of epochs behind, lag is now taken as the number of epochs completely buffered in the event buffer, but not yet consumed by the binlog injector thread. As part of this work, the default value for the `ndb_report_thresh_binlog_epoch_slip` system variable has been increased from 3 to 10. For more information, see the description of this variable in the documentation, as well as [Event Buffer Reporting in the Cluster Log](#). (Bug #22916457)

References: See also: Bug #22901309.

- **NDB Cluster APIs:** The term “fragment count type” has been superseded by “partition balance”. This change affects `NDB_TABLE` options for NDB tables as well as in the NDB API. In `NDB_TABLE` table option syntax, the `FRAGMENT_COUNT_TYPE` keyword is replaced with `PARTITION_BALANCE`. In the NDB API, the `Table` methods `getFragmentCountType()` and `setFragmentCountType()` have been renamed to `getPartitionBalance()` and `setPartitionBalance()`, respectively; `getFragmentCountTypeString()` is renamed to `getPartitionBalanceString()`. In addition, `Object::FragmentCountType` has been renamed to `PartitionBalance`, and the names of its enumerated values have been updated to be consistent with the new nomenclature.

For more information on how these changes affect NDB API applications, see the indicated `Table` and `Object` member descriptions. For more information on the SQL-level changes made as part of this fix, see [Setting NDB_TABLE Options in Table Comments](#). (Bug #81761, Bug #23547525)

References: See also: Bug #83147, Bug #24733331.

- If more than 4096 seconds elapsed while calculating an internal `NdbDuration::microSec()` value, this could cause an assert warning that the calculation would overflow. We fix this to avoid any overflow or precision loss when converting from the internal “tick” format to microseconds and nanoseconds, by performing the calculation in two parts corresponding to seconds and fractions of a second. (Bug #24695026)
- The serial commit protocol—which commits each operation at each replica individually and serially, and is used by the `DBTC` kernel block (see [The DBTC Block](#)) for takeover and when a transaction is judged to have timed out during the `COMMIT` or `COMPLETE` phase—had no support for `LATE_COMMIT`, which is required for the `READ_BACKUP` and `FULLY_REPLICATED` protocols. (Bug #24681305)
- In some cases, `ALTER TABLE ... REORGANIZE PARTITION` could lead to an unplanned shutdown of the cluster. This was due to the fact that, for fully replicated tables, the log part ID was assumed to be the same as the partition ID. This worked when `FOR_RA_BY_LDM` was used, but not necessarily for the other partition balancing types. (Bug #24610551)
- Using `ALGORITHM=INPLACE` when changing any of a table's `NDB_TABLE` properties (see [Setting NDB_TABLE Options in Table Comments](#)) caused the server to fail. (Bug #24584741)

- Following successive `ALTER TABLE` statements updating `NDB_TABLE` properties (see [Setting NDB_TABLE Options in Table Comments](#)), the current values were not always shown by `SHOW CREATE TABLE` or `ndb_desc`. (Bug #24584690)
- Case-insensitivity of keywords such as `FULLY_REPLICATED` in `NDB_TABLE` comments was not honored. (Bug #24577931)
- An `ALTER TABLE` statement attempting to set both `FULLY_REPLICATED` and `PARTITION_BALANCE` (see [Setting NDB_TABLE Options in Table Comments](#)) failed with a garbled error message. (Bug #24577894)
- A number of dependencies between the binlog injector thread and the `NDB` utility thread—a recurring source of synchronization and other problems—were removed. The principal changes are listed here:
 - Moved the setup of binlog injector structures from the utility thread to the injector thread itself.
 - Removed sharing of some utility and injector thread structures between these threads.
 - Moved stopping of the utility thread from the injector thread into a common block in which other such threads are stopped.
 - Removed a number of hacks required by the previous design.
 - Removed some injector mutex locking and injector condition signaling which were made obsolete by the changes already listed.

(Bug #24496910)

References: See also: Bug #22204186.

- A late commit `ACK` signal used for `FULLY_REPLICATED` or `READ_BACKUP` tables caused the associated `ApiConnectionRecord` to have an invalid state. (Bug #24459817)

References: See also: Bug #24444861.

- Added missing error information for a failure occurring when tables on disk became full. (Bug #24425373)
- When `ndbmt` crashed, the resulting error log incorrectly specified the name of the trace for thread 0, appending the nonexistent suffix `_t0` to the file name. (Bug #24353408)
- Passing a nonexistent node ID to `CREATE NODEGROUP` led to random data node failures. (Bug #23748958)
- `DROP TABLE` followed by a node shutdown and subsequent master takeover—and with the containing local checkpoint not yet complete prior to the takeover—caused the LCP to be ignored, and in some cases, the data node to fail. (Bug #23735996)

References: See also: Bug #23288252.

- Removed an invalid assertion to the effect that all cascading child scans are closed at the time API connection records are released following an abort of the main transaction. The assertion was invalid because closing of scans in such cases is by design asynchronous with respect to the main transaction, which means that subscans may well take some time to close after the main transaction is closed. (Bug #23709284)
- Although arguments to the `DUMP` command are 32-bit integers, `ndb_mgmd` used a buffer of only 10 bytes when processing them. (Bug #23708039)

- The `READ_BACKUP` setting was not honored when performing scans on `BLOB` tables. (Bug #23703536)
- Setting `FULLY_REPLICATED=1` (see [Setting NDB_TABLE Options in Table Comments](#)) did not propagate to the internal `BLOB` part tables used for `BLOB` and `TEXT` columns. (Bug #23703343)
- The `READ_BACKUP` setting was not applied to unique indexes. (Bug #23702848)
- In `ReadCommitted` mode, `DBSPJ` read primary fragment replicas for tables with `READ_BACKUP` (see [Setting NDB_TABLE Options in Table Comments](#)), even when a local fragment was available. (Bug #23633848)
- `ALL_REPORT MemoryUsage` produced incorrect output when fully replicated tables were in use. (Bug #23539805)
- Ordered indexes did not inherit `READ_BACKUP` (see [Setting NDB_TABLE Options in Table Comments](#)) from an indexed table, which meant that ordered index scans continued to be routed to only to primary fragment replicas and never to backup fragment replicas.

Now `DBDICT` sets this property on ordered indexes from the table property when it distributes this information to instances of `DBTC` and `DBSPJ`. (Bug #23522027)

- Updates to a table containing a virtual column could cause the binary logging thread to fail. (Bug #23514050)
- A number of potential buffer overflow issues were found and fixed in the `NDB` codebase. (Bug #23152979)
- During an online upgrade from a MySQL Cluster NDB 7.3 release to an NDB 7.4 (or later) release, the failures of several data nodes running the lower version during local checkpoints (LCPs), and just prior to upgrading these nodes, led to additional node failures following the upgrade. This was due to lingering elements of the `EMPTY_LCP` protocol initiated by the older nodes as part of an LCP-plus-restart sequence, and which is no longer used in NDB 7.4 and later due to LCP optimizations implemented in those versions. (Bug #23129433)
- A `SIGNAL_DROPPED_REP` handler invoked in response to long message buffer exhaustion was defined in the `SPJ` kernel block, but not actually used. This meant that the default handler from `SimulatedBlock` was used instead in such cases, which shut down the data node. (Bug #23048816)

References: See also: Bug #23251145, Bug #23251423.

- When a data node has insufficient redo buffer during a system restart, it does not participate in the restart until after the other nodes have started. After this, it performs a takeover of its fragments from the nodes in its node group that have already started; during this time, the cluster is already running and user activity is possible, including DML and DDL operations.

During a system restart, table creation is handled differently in the `DIH` kernel block than normally, as this creation actually consists of reloading table definition data from disk on the master node. Thus, `DIH` assumed that any table creation that occurred before all nodes had restarted must be related to the restart and thus always on the master node. However, during the takeover, table creation can occur on non-master nodes due to user activity; when this happened, the cluster underwent a forced shutdown.

Now an extra check is made during system restarts to detect in such cases whether the executing node is the master node, and use that information to determine whether the table creation is part of the restart proper, or is taking place during a subsequent takeover. (Bug #23028418)

- `ndb_restore` set the `MAX_ROWS` attribute for a table for which it had not been set prior to taking the backup. (Bug #22904640)

- Whenever data nodes are added to or dropped from the cluster, the `NDB` kernel's Event API is notified of this using a `SUB_GCP_COMPLETE_REP` signal with either the `ADD` (add) flag or `SUB` (drop) flag set, as well as the number of nodes to add or drop; this allows `NDB` to maintain a correct count of `SUB_GCP_COMPLETE_REP` signals pending for every incomplete bucket. In addition to handling the bucket for the epoch associated with the addition or removal, it must also compensate for any later incomplete buckets associated with later epochs. Although it was possible to complete such buckets out of order, there was no handling of these, leading a stall in to event reception.

This fix adds detection and handling of such out of order bucket completion. (Bug #20402364)

References: See also: Bug #82424, Bug #24399450.

- When performing online reorganization of tables, unique indexes were not included in the reorganization. (Bug #13714258)
- Under very high loads, neither the receive thread nor any user thread had sufficient capacity to handle poll ownership properly. This meant that, as the load and the number of active thread increased, it became more difficult to sustain throughput. This fix attempts to increase the priority of the receive thread and retains poll ownership if successful.

This fix requires sufficient permissions to be enabled. On Linux systems, this means ensuring that either the data node binary or the user it runs as has permission to change the nice level. (Bug #83217, Bug #24761073)

- Local reads of unique index and blob tables did not work correctly for fully replicated tables using more than one node group. (Bug #83016, Bug #24675602)
- The effects of an `ALTER TABLE` statement changing a table to use `READ_BACKUP` were not preserved after a restart of the cluster. (Bug #82812, Bug #24570439)
- Using `FOR_RP_BY_NODE` or `FOR_RP_BY_LDM` for `PARTITION_BALANCE` did not work with fully replicated tables. (Bug #82801, Bug #24565265)
- Changes to `READ_BACKUP` settings were not propagated to internal blob tables. (Bug #82788, Bug #24558232)
- The count displayed by the `c_exec` column in the `ndbinfo.threadstat` table was incomplete. (Bug #82635, Bug #24482218)
- The default `PARTITION_BALANCE` setting for `NDB` tables created with `READ_BACKUP=1` (see [Setting NDB_TABLE Options in Table Comments](#)) has been changed from `FOR_RA_BY_LDM` to `FOR_RP_BY_LDM`. (Bug #82634, Bug #24482114)
- The internal function `ndbcluster_binlog_wait()`, which provides a way to make sure that all events originating from a given thread arrive in the binary log, is used by `SHOW BINLOG EVENTS` as well as when resetting the binary log. This function waits on an injector condition while the latest global epoch handled by `NDB` is more recent than the epoch last committed in this session, which implies that this condition must be signalled whenever the binary log thread completes and updates a new latest global epoch. Inspection of the code revealed that this condition signalling was missing, and that, instead of being awakened whenever a new latest global epoch completes (~100ms), client threads waited for the maximum timeout (1 second).

This fix adds the missing injector condition signalling, while also changing it to a condition broadcast to make sure that all client threads are alerted. (Bug #82630, Bug #24481551)

- Fully replicated internal foreign key or unique index triggers could fire multiple times, which led to aborted transactions for an insert or a delete operation. This happened due to redundant deferred

constraint triggers firing during pre-commit. Now in such cases, we ensure that only triggers specific to unique indexes are fired in this stage. (Bug #82570, Bug #24454378)

- Backups potentially could fail when using fully replicated tables due to their high usage (and subsequent exhaustion) of internal trigger resources. To compensate for this, the amount of memory reserved in the [NDB](#) kernel for internal triggers has been increased, and is now based in part on the maximum number of tables. (Bug #82569, Bug #24454262)

References: See also: Bug #23539733.

- In the [DBTC](#) function `executeFullyReplicatedTrigger()` in the [NDB](#) kernel, an incorrect check of state led in some cases to failure handling when no failure had actually occurred. (Bug #82568, Bug #24454093)

References: See also: Bug #23539733.

- When returning from [LQHKEYREQ](#) with failure in [LQHKEYREF](#) in an internal trigger operation, no check was made as to whether the trigger was fully replicated, so that those triggers that were fully replicated were never handled. (Bug #82566, Bug #24453949)

References: See also: Bug #23539733.

- When [READ_BACKUP](#) had not previously been set, then was set to 1 as part of an `ALTER TABLE ... ALGORITHM=INPLACE` statement, the change was not propagated to internal unique index tables or [BLOB](#) tables. (Bug #82491, Bug #24424459)
- Distribution of MySQL privileges was incomplete due to the failure of the `mysql_cluster_move_privileges()` procedure to convert the `mysql.proxies_priv` table to [NDB](#). The root cause of this was an `ALTER TABLE ... ENGINE NDB` statement which sometimes failed when this table contained illegal [TIMESTAMP](#) values. (Bug #82464, Bug #24430209)
- The internal variable `m_max_warning_level` was not initialized in `storage/ndb/src/kernel/blocks/thrman.cpp`. This sometimes led to node failures during a restart when the uninitialized value was treated as 0. (Bug #82053, Bug #23717703)
- Usually, when performing a system restart, all nodes are restored from redo logs and local checkpoints (LCPs), but in some cases some node might require a copy phase before it is finished with the system restart. When this happens, the node in question waits for all other nodes to start up completely before performing the copy phase. Notwithstanding the fact that it is thus possible to begin a local checkpoint before reaching start phase 4 in the [DBDIH](#) block, LCP status was initialized to [IDLE](#) in all cases, which could lead to a node failure. Now, when performing this variant of a system restart, the LCP status is no longer initialized. (Bug #82050, Bug #23717479)
- After adding a new node group online and executing `ALTER TABLE ... ALGORITHM=INPLACE REORGANIZE PARTITION`, partition IDs were not set correctly for new fragments.

In a related change done as part of fixing this issue, `ndb_desc -p` now displays rows relating to partitions in order of partition ID. (Bug #82037, Bug #23710999)

- When executing `STOP BACKUP` it is possible sometimes that a few bytes are written to the backup data file before the backup process actually terminates. When using [ODIRECT](#), this resulted in the wrong error code being returned. Now in such cases, nothing is written to `O_DIRECT` files unless the alignment is correct. (Bug #82017, Bug #23701911)
- When transaction coordinator (TC) connection records were used up, it was possible to handle scans only for local checkpoints and backups, so that operations coming from the [DBUTIL](#) block—used for `ALTER TABLE ... REORGANIZE PARTITION` and other operations that reorganize metadata—were unnecessarily blocked. In addition, such operations were not always retried when TC records were

exhausted. To fix this issue, a number of operation records are now earmarked for `DBUTIL` usage, as well as for LCP and backup usage so that these operations are also not negatively impacted by operations coming from `DBUTIL`.

For more information, see [The DBUTIL Block](#). (Bug #81992, Bug #23642198)

- Operations performing multiple updates of the same row within the same transaction could sometimes lead to corruption of lengths of page entries. (Bug #81938, Bug #23619031)
- During a node restart, a fragment can be restored using information obtained from local checkpoints (LCPs); up to 2 restorable LCPs are retained at any given time. When an LCP is reported to the `DIH` kernel block as completed, but the node fails before the last global checkpoint index written into this LCP has actually completed, the latest LCP is not restorable. Although it should be possible to use the older LCP, it was instead assumed that no LCP existed for the fragment, which slowed the restart process. Now in such cases, the older, restorable LCP is used, which should help decrease long node restart times. (Bug #81894, Bug #23602217)
- Optimized node selection (`ndb_optimized_node_selection` setting) was not respected by `ndb_data_node_neighbour` when this was enabled. (Bug #81778, Bug #23555834)
- `NDB` no longer retries a global schema lock if this has failed due to a timeout (default 3000ms) and there is the potential for this lock request to participate in a metadata lock-global schema lock deadlock. Now in such cases it selects itself as a “victim”, and returns the decision to the requestor of the metadata lock, which then handles the request as a failed lock request (preferable to remaining deadlocked indefinitely), or, where a deadlock handler exists, retries the metadata lock-global schema lock. (Bug #81775, Bug #23553267)
- Two issues were found in the implementation of hash maps—used by `NDB` for mapping a table row's hash value to a partition—for fully replicated tables:
 1. Hash maps were selected based on the number of fragments rather than the number of partitions. This was previously undetected due to the fact that, for other kinds of tables, these values are always the same.
 2. The hash map was employed as a partition-to-partition map, using the table row's hash value modulus the partition count as input.

This fix addresses both of the problems just described. (Bug #81757, Bug #23544220)

References: See also: Bug #81761, Bug #23547525, Bug #23553996.

- Using `mysqld` together with `--initialize` and `--ndbcluster` led to problems later when attempting to use `mysql_upgrade`. When running with `--initialize`, the server does not require `NDB` support, and having it enabled can lead to issues with `ndbinfo` tables. To prevent this from happening, using the `--initialize` option now causes `mysqld` to ignore the `--ndbcluster` option if the latter is also specified.

This issue affects upgrades from MySQL Cluster NDB 7.5.2 or 7.5.3 only. In cases where such upgrades fail for the reasons outlined previously, you can work around the issue by deleting all `.frm` files in the `data/ndbinfo` directory following a rolling restart of the entire cluster, then running `mysql_upgrade`. (Bug #81689, Bug #23518923)

References: See also: Bug #82724, Bug #24521927.

- While a `mysqld` was waiting to connect to the management server during initialization of the `NDB` handler, it was not possible to shut down the `mysqld`. If the `mysqld` was not able to make the connection, it could become stuck at this point. This was due to an internal wait condition in the utility

and index statistics threads that could go unmet indefinitely. This condition has been augmented with a maximum timeout of 1 second, which makes it more likely that these threads terminate themselves properly in such cases.

In addition, the connection thread waiting for the management server connection performed 2 sleeps in the case just described, instead of 1 sleep, as intended. (Bug #81585, Bug #23343673)

- `ALTER TABLE ... ALGORITHM=INPLACE` on a fully replicated table did not copy the associated trigger ID, leading to a failure in the `DBDICT` kernel block. (Bug #81544, Bug #23330359)
- The list of deferred tree node lookup requests created when preparing to abort a `DBSPJ` request were not cleared when this was complete, which could lead to deferred operations being started even after the `DBSPJ` request aborted. (Bug #81355, Bug #23251423)

References: See also: Bug #23048816.

- Error and abort handling in `Dbspj::execTRANSID_AI()` was implemented such that its `abort()` method was called before processing of the incoming signal was complete. Since this method sends signals to the LDM, this partly overwrote the contents of the signal which was later required by `execTRANSID_AI()`. This could result in aborted `DBSPJ` requests cleaning up their allocated resources too early, or not at all. (Bug #81353, Bug #23251145)

References: See also: Bug #23048816.

- The read backup feature added in MySQL Cluster NDB 7.5.2 that makes it possible to read from backup replicas was not used for reads with lock, or for reads of `BLOB` tables or unique key tables where locks were upgraded to reads with lock. Now the `TCKEYREQ` and `SCAN_TABREQ` signals use a flag to convey information about such locks making it possible to read from a backup replica when a read lock was upgraded due to being the read of the base table for a `BLOB` table, or due to being the read for a unique key. (Bug #80861, Bug #23001841)
- Primary replicas of partitioned tables were not distributed evenly among node groups and local data managers.

As part of the fix for this issue, the maximum number of node groups supported for a single MySQL Cluster, which was previously not determined, is now set at 48 (`MAX_NDB_NODE_GROUPS`). (Bug #80845, Bug #22996305)

- Several object constructors and similar functions in the `NDB` codebase did not always perform sanity checks when creating new instances. These checks are now performed under such circumstances. (Bug #77408, Bug #21286722)
- An internal call to `malloc()` was not checked for `NULL`. The function call was replaced with a direct write. (Bug #77375, Bug #21271194)
- **NDB Cluster APIs:** Reuse of transaction IDs could occur when `Ndb` objects were created and deleted concurrently. As part of this fix, the NDB API methods `lock_ndb_objects()` and `unlock_ndb_objects` are now declared as `const`. (Bug #23709232)
- **NDB Cluster APIs:** When the management server was restarted while running an MGM API application that continuously monitored events, subsequent events were not reported to the application, with timeouts being returned indefinitely instead of an error.

This occurred because sockets for event listeners were not closed when restarting `mgmd`. This is fixed by ensuring that event listener sockets are closed when the management server shuts down, causing applications using functions such as `ndb_logevent_get_next()` to receive a read error following the restart. (Bug #19474782)

- **NDB Cluster APIs:** To process incoming signals, a thread which wants to act as a receiver must acquire polling rights from the transporter layer. This can be requested and assigned to a separate receiver thread, or each client thread can take the receiver role when it is waiting for a result.

When the thread acting as poll owner receives a sufficient amount of data, it releases locks on any other clients taken while delivering signals to them. This could make them runnable again, and the operating system scheduler could decide that it was time to wake them up, which happened at the expense of the poll owner threads, which were in turn excluded from the CPU while still holding polling rights on it. After this fix, polling rights are released by a thread before unlocking and signalling other threads. This makes polling rights available for other threads that are actively executing on this CPU.

This change increases concurrency when polling receiver data, which should also reduce latency for clients waiting to be woken up. (Bug #83129, Bug #24716756)

- **NDB Cluster APIs:** `libndbclient` and `libmysqlclient` exported conflicting symbols, resulting in a segmentation fault in debug builds on Linux. To fix this issue, the conflicting symbols in `libndbclient.so` are no longer publicly visible. Due to this change, the version number for `libndbclient.so` has been raised from 6.0.0 to 6.1.0. (Bug #83093, Bug #24707521)

References: See also: Bug #80352, Bug #22722555.

- **NDB Cluster APIs:** When NDB schema object ownership checks are enabled by a given `NdbTransaction`, objects used by this transaction are checked to make sure that they belong to the `NdbDictionary` owned by this connection. An attempt to create a `NdbOperation`, `NdbScanOperation`, or `NdbIndexScanOperation` on a table or index not belonging to the same connection fails.

This fix corrects a resource leak which occurred when the operation object to be created was allocated before checking schema object ownership and subsequently not released when the object creation failed. (Bug #81949, Bug #23623978)

References: See also: Bug #81945, Bug #23623251.

- **NDB Cluster APIs:** NDB API objects are allocated in the context of an `Ndb` object, or of an `NdbTransaction` object which is itself owned by an `Ndb` object. When a given `Ndb` object is destroyed, all remaining `NdbTransaction` objects are terminated, and all NDB API objects related to this `Ndb` object should be released at this time as well. It was found, when there remained unclosed `NdbTransaction` objects when their parent `Ndb` object was destroyed, leaks of objects allocated from the `NdbTransaction` objects could occur. (However, the `NdbTransaction` objects themselves did not leak.)

While it is advisable (and, indeed, recommended) to close an `NdbTransaction` explicitly as soon as its lifetime ends, the destruction of the parent `Ndb` object should be sufficient to release whatever objects are dependent on it. Now in cases such as described previously, the `Ndb` destructor checks to ensure that all objects derived from a given `Ndb` instance are truly released. (Bug #81945, Bug #23623251)

- **NDB Cluster APIs:** In some of the NDB API example programs included with the MySQL Cluster distribution, `ndb_end()` was called prior to calling the `Ndb_cluster_connection` destructor. This caused a segmentation fault in debug builds on all platforms. The example programs affected have also been extensively revised and refactored. See [NDB API Examples](#), for more information. (Bug #80352, Bug #22722555)

References: See also: Bug #83093, Bug #24707521.

Changes in MySQL NDB Cluster 7.5.3 (5.7.13-ndb-7.5.3)

Functionality Added or Changed

- **Important Change:** It is now possible to set `READ_BACKUP` for an existing table online using an SQL statement such as `ALTER TABLE ... ALGORITHM=INPLACE, COMMENT="NDB_TABLE=READ_BACKUP=1"`. See [Setting NDB_TABLE Options in Table Comments](#), for further information about the `READ_BACKUP` option. (Bug #80858, Bug #23001617)

References: See also: Bug #18435416.

- Added three new tables to the `ndbinfo` information database to provide running information about locks and lock attempts in an active MySQL Cluster. These tables, with brief descriptions, are listed here:
 - `cluster_locks`: Current lock requests which are waiting for or holding locks; this information can be useful when investigating stalls and deadlocks. Analogous to `cluster_operations`.
 - `locks_per_fragment`: Counts of lock claim requests, and their outcomes per fragment, as well as total time spent waiting for locks successfully and unsuccessfully. Analogous to `operations_per_fragment` and `memory_per_fragment`.
 - `server_locks`: Subset of cluster transactions—those running on the local `mysqld`, showing a connection ID per transaction. Analogous to `server_operations`.

For more information, see [The ndbinfo NDB Cluster Information Database](#).

- The `NDB` storage engine now supports generated columns (see [CREATE TABLE and Generated Columns](#)) as well as indexes on stored generated columns.

Indexes on virtual generated columns of `NDB` tables are not supported.

`ALTER TABLE` statements that add stored generated columns to `NDB` tables cannot be performed online.

Bugs Fixed

- The `ndbinfo` `cpustat_1sec` and `cpustat_20sec` tables did not provide any history information. (Bug #23520271)
- During shutdown, the `mysqld` process could sometimes hang after logging `NDB Util: Stop ... NDB Util: Wakeup`. (Bug #23343739)

References: See also: Bug #21098142.

- During expansion or reduction of a hash table, allocating a new overflow page in the `DBACC` kernel block caused the data node to fail when it was out of index memory. This could sometimes occur after a large table had increased or decreased very rapidly in size. (Bug #23304519)

References: This issue is a regression of: Bug #13436216.

- When tracking time elapsed from sending a `SCAN_FRAGREQ` signal to receiving a corresponding `SCAN_FRAGCONF`, the assumption was made in the `DBTC` kernel block that a `SCAN_FRAGCONF` can occur only after sending a `SCAN_FRAGREQ` or `SCAN_NEXTREQ` signal, which is not always the case: It is actually possible that a local query handler can, immediately after sending a `SCAN_FRAGCONF`, send an additional `SCAN_FRAGCONF` signal upon reporting that the scan is closed. This problem is fixed by ensuring that the timer value is initialized each time before use. (Bug #81907, Bug #23605817, Bug #23306695)
- Table indexes were listed in the output of `ndb_desc` in a nondeterministic order that could vary between platforms. Now these indexes are ordered by ID in the output. (Bug #81763, Bug #23547742)
- Following a restart of the cluster, the first attempt to read from any of the `ndbinfo` `cpustat`, `cpustat_50ms`, `cpustat_1sec`, or `cpustat_20sec` tables generated a warning to the effect that

columns were missing from the table. Subsequently, the `thread_sleeping` and `spin_time` columns were found to be missing from each of these tables. (Bug #81681, Bug #23514557)

References: See also: Bug #23305078.

- Using a `ThreadConfig` parameter value with a trailing comma led to an assertion. (Bug #81588, Bug #23344374)
- **NDB Cluster APIs:** The scan lock takeover issued by `NdbScanOperation::lockCurrentTuple()` did not set the operation type for the takeover operation. (Bug #23314028)

Changes in MySQL NDB Cluster 7.5.2 (5.7.12-ndb-7.5.2)

Functionality Added or Changed

- **Performance:** A deficiency in event buffer memory allocation was identified as inefficient and possibly leading to undesirable results. This could happen when additional memory was allocated from the operating system to buffer received event data even when memory had already been allocated but remained unused. This is fixed by allocating the event buffer memory directly from the page allocation memory manager (`mmap()`), where such functionality is offered by the operating system, allowing for direct control over the memory such that it is in fact returned to the system when released.

This reimplemention avoids the tendencies of the existing one to approach worst-case memory usage, maintenance of data structures for a worst-case event buffer event count, and useless caching of free memory in unusable positions. This work should also help minimize the runtime costs of buffering events, minimize heap fragmentation, and avoid OS-specific problems due to excessive numbers of distinct memory mappings.

In addition, the relationship between epochs and internal `EventData` objects is now preserved throughout the event lifecycle, reception to consumption, thus removing the need for iterating, and keeping in synch, two different lists representing the epochs and their `EventData` objects.

As part of this work, better reporting on the relevant event buffer metrics is now provided in the cluster logs.

References: See also: Bug #21651536, Bug #21660947, Bug #21661297, Bug #21673318, Bug #21689380, Bug #21809959.

- `ndb_restore` now performs output logging for specific stages of its operation. (Bug #21097957)
- An improvement in the hash index implementation used by MySQL Cluster data nodes means that partitions may now contain more than 16 GB of data for fixed columns, and the maximum partition size for fixed column data is now increased to 128 TB. The previous limitation originated with the `DBACC` block in the `NDB` kernel using only 32-bit references to the fixed-size part of a row handled in the `DBTUP` block, even though 45-bit references were already in use elsewhere in the kernel outside the `DBACC` block; all such references in `DBACC` now use 45-bit pointers instead.

As part of this work, error messages returned by the `DBACC` kernel block that were overly generic have now been improved by making them more specific. (Bug #17465232)

- The `NDB` storage engine supports the MySQL `JSON` data type and MySQL `JSON` functions implemented in MySQL 5.7.8 and later. This support is subject to the limitation that a single `NDB` table can have at most 3 `JSON` columns.
- A number of changes and improvements were made to the handling of send threads by `NDB`. The changes are summarized with brief descriptions in the following list:

- Decreased resource requirements for send threads, making sure that a given configuration using send threads outperforms the same configuration without send threads.
- Made use of otherwise idle threads (other than receiver threads) as send threads without incurring extra CPU resources in low-load situations.
- Improved response time for write transactions.
- Provided for handling of bad configuration data in a more graceful manner.
- Made it possible to measure CPU usage with improved real-time reporting from data nodes on their resource usage.

As part of implementing the last item of those just listed, a number of new tables providing information about CPU and thread activity by node, thread ID, and thread type have been added to the `ndbinfo` information database. These tables are listed here:

- `cpustat`: Provides per-second, per-thread CPU statistics
- `cpustat_50ms`: Shows raw per-thread CPU statistics data, gathered every 50ms
- `cpustat_1sec`: Provides raw per-thread CPU statistics data, gathered each second
- `cpustat_20sec`: Displays raw per-thread CPU statistics data, gathered every 20 seconds
- `threads`: Shows names and descriptions of thread types

For more information, see [The ndbinfo NDB Cluster Information Database](#).

- Rewrote the implementation of the NDB storage engine's global schema lock functionality to make use of the metadata lock hooks implemented in the MySQL Server in MySQL 5.7.11.
- A number of improvements provide additional read scalability for NDB by making it possible to read tables locally. It is now possible to enable reads from any replica, rather than from the primary replica only. This is disabled by default to remain compatible with previous behavior, but can be enabled for a given SQL node using the `ndb_read_backup` system variable added in this release.

It also becomes possible to be more flexible about the assignment of partitions by setting a fragment count type. Possible count types are one per node, one per node group, one per Local Data Manager (LDM) per node (this is the previous assignment scheme and still the default), and one per LDM per node group. This setting can be controlled for individual tables by means of a `FRAGMENT_COUNT_TYPE` option embedded in an `NDB_TABLE` comment in `CREATE TABLE` or `ALTER TABLE`.

Tables using one of the two per-node-group settings for the fragment count type can also be fully replicated. This requires that the table's fragment count type is `ONE_PER_NODE_GROUP` or `ONE_PER_LDM_PER_NODE_GROUP`, and can be enabled using the option `FULLY_REPLICATED=1` within in an `NDB_TABLE` comment. The option can be enabled by default for all new NDB tables using the `ndb_fully_replicated` system variable added in this release.

Settings for table-level `READ_BACKUP` are also supported using the `COMMENT="NDB_TABLE=..."` syntax. It is also possible (and often preferable) to set multiple options in one comment within a single `CREATE TABLE` or `ALTER TABLE` statement. For more information and examples, see [Setting NDB_TABLE Options in Table Comments](#).

This release also introduces the `ndb_data_node_neighbour` system variable, which is intended to be employed with transaction hinting (and fully-replicated tables), and to provide the current SQL node with

the ID of a “nearby” data node to use. See the description of this variable in the documentation for more information.

References: See also: Bug #18435416.

- Made the following enhancements and additions to the `ThreadConfig` multithreaded data node (`ndbmt.d`) configuration parameter:
 - Added support for non-exclusive CPU locking on FreeBSD and Windows using `cpubind` and `cpuset`.
 - Added support for exclusive CPU locking on Solaris, using `cpubind_exclusive` and `cpuset_exclusive`, which are added in this release.
 - Added thread prioritization using `thread_prio`, which is added in this release. `thread_prio` is supported on Linux, FreeBSD, Windows, and Solaris, but the exact effects of this setting are platform-specific; see the documentation for details.
 - Added support for setting `realtime` on Windows platforms.

For more information, see the description of the `ThreadConfig` parameter in the online documentation.

- **NDB Cluster APIs:** Added the `Ndb::setEventBufferQueueEmptyEpoch()` method, which makes it possible to enable queuing of empty events (event type `TE_EMPTY`). (Bug #22157845)
- **NDB Cluster APIs:** `Ndb_free_list_t` is a template used in the implementation of the NDB API to create a list keeping released API objects such as `NdbTransaction`, `NdbOperation`, and `NdbRecAttr`. One drawback to this template is that released API objects are kept in the list for the lifetime of the owning `Ndb` object, such that a transient peak in the demand for any object causes an effective leak in memory that persists until this `Ndb` object itself has been released.

This work adds statistics to each `Ndb_free_list` instance which samples the usage of objects maintained by the list; now, when objects are released, they can be released into the free list, or deallocated, based on the collected usage statistics.

Bugs Fixed

- **Performance:** A performance problem was found in an internal polling method `do_poll()` where the polling client did not check whether it had itself been woken up before completing the poll. Subsequent analysis showed that it is sufficient that only some clients in the polling queue receive data. `do_poll()` can then signal these clients and give up its polling rights, even if the maximum specified wait time (10 ms) has not expired.

This change allows `do_poll()` to continue polling until either the maximum specified wait time has expired, or the polling client itself has been woken up (by receiving what it was waiting for). This avoids unnecessary thread switches between client threads and thus reduces the associated overhead by as much as 10% in the API client, resulting in a significant performance improvement when client threads perform their own polling. (Bug #81229, Bug #23202735)

- **Incompatible Change:** When the data nodes are only partially connected to the API nodes, a node used for a pushdown join may get its request from a transaction coordinator on a different node, without (yet) being connected to the API node itself. In such cases, the `NodeInfo` object for the requesting API node contained no valid info about the software version of the API node, which caused the `DBSPJ` block to assume (incorrectly) when aborting to assume that the API node used NDB version 7.2.4 or earlier, requiring the use of a backward compatibility mode to be used during query abort which sent a node failure error instead of the real error causing the abort.

Now, whenever this situation occurs, it is assumed that, if the [NDB](#) software version is not yet available, the API node version is greater than 7.2.4. (Bug #23049170)

- **Important Change:** When started with the `--initialize` option, `mysqld` no longer enables the `NDBCLUSTER` storage engine plugin. This change was needed to prevent attempted initialization of system databases as distributed (rather than as specific to individual SQL nodes), which could result in a metadata lock deadlock. This fix also brings the behavior of `--initialize` in this regard into line with that of the discontinued `--bootstrap` option, which started a minimal `mysqld` instance without enabling `NDB`. (Bug #22758238)
- **Microsoft Windows:** `ndb_mgmd` failed to start on 32-bit Windows platforms, due to an issue with calling dynamically loaded functions; such issues were also likely to occur with other `NDB` programs using `ndb_init()`. It was found that all of the functions used are already supported in targeted versions of Windows, so this problem is fixed by removing the dynamic loading of these functions and using the versions provided by the Windows header files instead. (Bug #80876, Bug #23014820)
- **Microsoft Windows:** When building MySQL Cluster on Windows using more than one parallel build job it was sometimes possible for the build to fail because `host_info.exe` could not be installed. To fix this problem, the `install_mcc` target is now always built prior to the `host_info` target. (Bug #80051, Bug #22566368)
- **Microsoft Windows:** Performing `ANALYZE TABLE` on a table having one or more indexes caused `ndbmt` to fail with an `InvalidAttrInfo` error due to signal corruption. This issue occurred consistently on Windows, but could also be encountered on other platforms. (Bug #77716, Bug #21441297)
- Reserved send buffer for the loopback transporter, introduced in MySQL Cluster NDB 7.4.8 and used by API and management nodes for administrative signals, was calculated incorrectly. (Bug #23093656, Bug #22016081)

References: This issue is a regression of: Bug #21664515.

- During a node restart, re-creation of internal triggers used for verifying the referential integrity of foreign keys was not reliable, because it was possible that not all distributed TC and LDM instances agreed on all trigger identities. To fix this problem, an extra step is added to the node restart sequence, during which the trigger identities are determined by querying the current master node. (Bug #23068914)

References: See also: Bug #23221573.

- Following the forced shutdown of one of the 2 data nodes in a cluster where `NoOfReplicas=2`, the other data node shut down as well, due to arbitration failure. (Bug #23006431)
- Aborting a `CREATE LOGFILE GROUP` statement which had failed due to lack of shared global memory was not performed correctly, causing node failure. In addition, the transaction in which this occurred was not rolled back correctly, also causing any subsequent `CREATE LOGFILE GROUP` to fail. (Bug #22982618)
- The `ndbinfo.tc_time_track_stats` table uses histogram buckets to give a sense of the distribution of latencies. The sizes of these buckets were also reported as `HISTOGRAM BOUNDARY INFO` messages during data node startup; this printout was redundant and so has been removed. (Bug #22819868)
- Online upgrades from previous versions of MySQL Cluster to MySQL Cluster NDB 7.5 were not possible due to missing entries in the matrix used to test upgrade compatibility between versions. (Bug #22024947)
- A failure occurred in `DBTUP` in debug builds when variable-sized pages for a fragment totalled more than 4 GB. (Bug #21313546)

- Restoration of metadata with `ndb_restore -m` occasionally failed with the error message `Failed to create index...` when creating a unique index. While diagnosing this problem, it was found that the internal error `PREPARE_SEIZE_ERROR` (a temporary error) was reported as an unknown error. Now in such cases, `ndb_restore` retries the creation of the unique index, and `PREPARE_SEIZE_ERROR` is reported as NDB Error 748 `Busy during read of event table`. (Bug #21178339)

References: See also: Bug #22989944.

- `mysqld` did not shut down cleanly when executing `ndb_index_stat`. (Bug #21098142)

References: See also: Bug #23343739.

- The following improvements were made to the data node error logging mechanism:

- Increased the message slot size 499 bytes to 999 bytes to prevent log messages from overwriting one another or from being truncated.
- Added a `Trace file name` field to the output. This field contains the trace file name (without any path) and trace file number for the thread causing the trace.
- `ndbmt` trace files are also now shown in the error log.

(Bug #21082710)

- `DBDICT` and `GETTABINFOREQ` queue debugging were enhanced as follows:

- Monitoring by a data node of the progress of `GETTABINFOREQ` signals can be enabled by setting `DictTrace >= 2`.
- Added the `ApiVerbose` configuration parameter, which enables NDB API debug logging for an API node where it is set greater than or equal to 2.
- Added `DUMP` code 1229 which shows the current state of the `GETTABINFOREQ` queue. (See `DUMP 1229`.)

See also [The DBDICT Block](#). (Bug #20368450)

References: See also: Bug #20368354.

- When a write to the `ndb_binlog_index` table failed during a MySQL Server shutdown, `mysqld` killed the NDB binary logging thread. (Bug #81166, Bug #23142945)
- Memory associated with table descriptions was not freed by the internal table information method `NdbDictInterface::parseTableInfo()`. (Bug #81141, Bug #23130084)
- Improved memory usage by the internal `TransporterFacade` constructor when performing mutex array initialization. (Bug #81134, Bug #23127824)
- Fixed a memory leak that occurred when an error was raised in `ha_ndbcluster::get_metadata()` or one of the functions which this method calls. (Bug #81045, Bug #23089566)
- An internal function used to validate connections failed to update the connection count when creating a new `Ndb` object. This had the potential to create a new `Ndb` object for every operation validating the connection, which could have an impact on performance, particularly when performing schema operations. (Bug #80750, Bug #22932982)
- A table scan on an NDB table using neither an ordered index nor any Disk Data columns normally uses an `ACC` scan. If this happened while scanning an unique but unordered index which shrank (due to rows

being deleted) after the scan started and then grew again (rows inserted), a single row that had been neither deleted nor inserted could be scanned twice. (Bug #80733, Bug #22926938)

- On OS X, `ndb_config` failed when an empty string was used for the `--host` option. (Bug #80689, Bug #22908696)
- Starting a backup in the `ndb_mgm` client after creating a large number of tables caused a forced shutdown of the cluster. (Bug #80640, Bug #228849958)
- When an SQL node was started, and joined the schema distribution protocol, another SQL node, already waiting for a schema change to be distributed, timed out during that wait. This was because the code incorrectly assumed that the new SQL node would also acknowledge the schema distribution even though the new node joined too late to be a participant in it.

As part of this fix, printouts of schema distribution progress now always print the more significant part of a bitmask before the less significant; formatting of bitmasks in such printouts has also been improved. (Bug #80554, Bug #22842538)

- The `ndb_print_file` utility failed consistently on Solaris 9 for SPARC. (Bug #80096, Bug #22579581)
- The MySQL Cluster Auto-Installer failed to work in various ways on different platforms. (Bug #79853, Bug #22502247)
- The internal function `ndbrequire()`, which, like `assert()`, evaluates a given expression and terminates the process if the expression does not evaluate as true, now includes the failed expression in its output to the error logs. (Bug #77021, Bug #21128318)
- Trying to drop a table during an ongoing backup failed with the error message `Unknown table`; now, it fails with `Unable to alter table as backup is in progress`. (Bug #47301, Bug #11755512)

References: See also: Bug #44695, Bug #11753280.

- **NDB Disk Data:** The following improvements were made to logging during restarts by data nodes using MySQL Cluster Disk Data:
 - The total amount of undo log to be applied by the data node is now provided as the total number of pages present in the log. This is a worst case estimate.
 - Progress information is now provided at regular intervals (once for each 30000 records) as the undo log is applied. This information is supplied as the number of records and number of undo log pages applied so far during the current restart.

(Bug #22513381)

- **NDB Cluster APIs:** Deletion of Ndb objects used a disproportionately high amount of CPU. (Bug #22986823)
- **NDB Cluster APIs:** Executing a transaction with an `NdbIndexOperation` based on an obsolete unique index caused the data node process to fail. Now the index is checked in such cases, and if it cannot be used the transaction fails with an appropriate error. (Bug #79494, Bug #22299443)

Changes in MySQL NDB Cluster 7.5.1 (5.7.11-ndb-7.5.1)

Functionality Added or Changed

- **Important Change:** When creating or adding columns to NDB tables, the default value used for both the `COLUMN_FORMAT` and `ROW_FORMAT` options is now `DYNAMIC` rather than `FIXED`.

This change does not affect the row format or column format used by existing tables. New columns added to such tables use the new defaults, and existing columns are changed to use these as well, provided that the `ALTER TABLE` statement in question uses `ALGORITHM=COPY`. Note that this cannot be done implicitly if `mysqld` is run with `--ndb-allow-copying-alter-table=FALSE` (the default is `TRUE`).

A new MySQL server option `--ndb-default-column-format` is added for backwards compatibility; set this to `FIXED` to force the old defaults to be used for `COLUMN_FORMAT` and `ROW_FORMAT`.



Note

This behavior change is superseded in MySQL Cluster NDB 7.5.4.

- Scans have been improved by replacing the `DIH_SCAN_GET_NODES_REQ` signal, formerly used for communication between the `DBTC` and `DBDIH` kernel blocks in `NDB`, with the `DIGETNODESREQ` signal, which supports direct execution and allows for the elimination of `DIH_SCAN_GET_NODES_REF` and `DIH_SCAN_GET_NODES_CONF`, as well as for `DIH_SCAN_TAB_REQ` and `DIH_SCAN_TAB_COMPLETE_REP` signals to `EXECUTE_DIRECT`. This enables higher scalability of data nodes when used for scan operations by decreasing the use of CPU resources for scan operations by an estimated five percent in some cases. This should also improve response times, which could help prevent issues with overload of the main threads.

As part of these changes, scans made in the `BACKUP` kernel block have also been improved and made more efficient.

References: See also: Bug #80640, Bug #22884995.

- The following improvements have been made to event buffer reporting in the cluster log:
 - Each report now identifies the API node that sent it.
 - The fields shown in the report have been improved and expanded. Percentages are now better specified and used only when appropriate. For improved clarity, the `apply_epoch` and `latest_epoch` fields have been renamed to `latest_consumed_epoch` and `latest_buffered_epoch`, respectively. The ID of the `Ndb` object serving as the source of the report is now shown, as is the reason for making the report (as the `report_reason` field).
 - The frequency of unnecessary reporting has been reduced by limiting reports to those showing only significant changes in event buffer usage.
 - The MGM API adds a new `NDB_LE_EventBufferStatus2` event type to handle the additional information provided by the new event buffer reporting. The `NDB_LE_EventBufferStatus` event type used in older versions of MySQL Cluster is now deprecated, and will eventually be removed.

For more information, see [Event Buffer Reporting in the Cluster Log](#), as well as [The `ndb_logevent` Structure](#).

Bugs Fixed

- **Important Change:** The minimum value for the `BackupDataBufferSize` data node configuration parameter has been lowered from 2 MB to 512 KB. The default and maximum values for this parameter remain unchanged. (Bug #22749509)
- During node failure handling, the request structure used to drive the cleanup operation was not maintained correctly when the request was executed. This led to inconsistencies that were harmless

during normal operation, but these could lead to assertion failures during node failure handling, with subsequent failure of additional nodes. (Bug #22643129)

- The previous fix for a lack of mutex protection for the internal `TransporterFacade::deliver_signal()` function was found to be incomplete in some cases. (Bug #22615274)

References: This issue is a regression of: Bug #77225, Bug #21185585.

- Compilation of MySQL with Visual Studio 2015 failed in `ConfigInfo.cpp`, due to a change in Visual Studio's handling of spaces and concatenation. (Bug #22558836, Bug #80024)
- When setup of the binary log as an atomic operation on one SQL node failed, this could trigger a state in other SQL nodes in which they appeared to detect the SQL node participating in schema change distribution, whereas it had not yet completed binary log setup. This could in turn cause a deadlock on the global metadata lock when the SQL node still retrying binary log setup needed this lock, while another mysqld had taken the lock for itself as part of a schema change operation. In such cases, the second SQL node waited for the first one to act on its schema distribution changes, which it was not yet able to do. (Bug #22494024)
- Duplicate key errors could occur when `ndb_restore` was run on a backup containing a unique index. This was due to the fact that, during restoration of data, the database can pass through one or more inconsistent states prior to completion, such an inconsistent state possibly having duplicate values for a column which has a unique index. (If the restoration of data is preceded by a run with `--disable-indexes` and followed by one with `--rebuild-indexes`, these errors are avoided.)

Added a check for unique indexes in the backup which is performed only when restoring data, and which does not process tables that have explicitly been excluded. For each unique index found, a warning is now printed. (Bug #22329365)

- When setting up event logging for `ndb_mgmd` on Windows, MySQL Cluster tries to add a registry key to `HKEY_LOCAL_MACHINE`, which fails if the user does not have access to the registry. In such cases `ndb_mgmd` logged the error `Could neither create or open key`, which is not accurate and which can cause confusion for users who may not realize that file logging is available and being used. Now in such cases, `ndb_mgmd` logs a warning `Could not create or access the registry key needed for the application to log to the Windows EventLog. Run the application with sufficient privileges once to create the key, or add the key manually, or turn off logging for that application. An error (as opposed to a warning) is now reported in such cases only if there is no available output at all for ndb_mgmd event logging.` (Bug #20960839)
- `NdbDictionary` metadata operations had a hard-coded 7-day timeout, which proved to be excessive for short-lived operations such as retrieval of table definitions. This could lead to unnecessary hangs in user applications which were difficult to detect and handle correctly. To help address this issue, timeout behaviour is modified so that read-only or short-duration dictionary interactions have a 2-minute timeout, while schema transactions of potentially long duration retain the existing 7-day timeout.

Such timeouts are intended as a safety net: In the event of problems, these return control to users, who can then take corrective action. Any reproducible issue with `NdbDictionary` timeouts should be reported as a bug. (Bug #20368354)

- Optimization of signal sending by buffering and sending them periodically, or when the buffer became full, could cause `SUB_GCP_COMPLETE_ACK` signals to be excessively delayed. Such signals are sent for each node and epoch, with a minimum interval of `TimeBetweenEpochs`; if they are not received in time, the `SUMA` buffers can overflow as a result. The overflow caused API nodes to be disconnected, leading to current transactions being aborted due to node failure. This condition made it difficult for long

transactions (such as altering a very large table), to be completed. Now in such cases, the `ACK` signal is sent without being delayed. (Bug #18753341)

- MySQL Cluster did not compile correctly with Microsoft Visual Studio 2015, due to a change from previous versions in the VS implementation of the `_vsnprintf()` function. (Bug #80276, Bug #22670525)
- When setting CPU spin time, the value was needlessly cast to a boolean internally, so that setting it to any nonzero value yielded an effective value of 1. This issue, as well as the fix for it, apply both to setting the `SchedulerSpinTimer` parameter and to setting `spintime` as part of a `ThreadConfig` parameter value. (Bug #80237, Bug #22647476)
- Processing of local checkpoints was not handled correctly on Mac OS X, due to an uninitialized variable. (Bug #80236, Bug #22647462)
- A logic error in an `if` statement in `storage/ndb/src/kernel/blocks/dbacc/DbaccMain.cpp` rendered useless a check for determining whether `ZREAD_ERROR` should be returned when comparing operations. This was detected when compiling with `gcc` using `-Werror=logical-op`. (Bug #80155, Bug #22601798)

References: This issue is a regression of: Bug #21285604.

- Suppressed a `CMake` warning that was caused by use of an incorrectly quoted variable name. (Bug #80066, Bug #22572632)
- When using `CREATE INDEX` to add an index on either of two `NDB` tables sharing circular foreign keys, the query succeeded but a temporary table was left on disk, breaking the foreign key constraints. This issue was also observed when attempting to create an index on a table in the middle of a chain of foreign keys—that is, a table having both parent and child keys, but on different tables. The problem did not occur when using `ALTER TABLE` to perform the same index creation operation; and subsequent analysis revealed unintended differences in the way such operations were performed by `CREATE INDEX`.

To fix this problem, we now make sure that operations performed by a `CREATE INDEX` statement are always handled internally in the same way and at the same time that the same operations are handled when performed by `ALTER TABLE` or `DROP INDEX`. (Bug #79156, Bug #22173891)

- The `PortNumber` `SCI`, `SHM`, and `TCP` configuration parameters, which were deprecated in MySQL 5.1.3, have now been removed and are no longer accepted in configuration files.

This change does not affect the `PortNumber` management node configuration parameter, whose behavior remains unaltered. (Bug #77405, Bug #21280456)

Changes in MySQL NDB Cluster 7.5.0 (5.7.10-ndb-7.5.0)

Functionality Added or Changed

- **Important Change:** Previously, the `ndbinfo` information database included lookup tables that used the `MyISAM` storage engine. This dependency on `MyISAM` has now been removed. (Bug #20075747)
- **Important Change:** Previously, the `NDB` scheduler always optimized for speed against throughput in a predetermined manner (this was hard coded); this balance can now be set using the `SchedulerResponsiveness` data node configuration parameter. This parameter accepts an integer in the range of 0-10 inclusive, with 5 as the default. Higher values provide better response times relative to throughput. Lower values provide increased throughput, but impose longer response times. (Bug #78531, Bug #21889312)

- **Important Change:** A number of MySQL Cluster data node configuration parameters were deprecated in earlier versions of MySQL Cluster, and have been removed with this release. These parameters include `Id`, `NoOfDiskPagesToDiskDuringRestartTUP`, `NoOfDiskPagesToDiskDuringRestartACC`, `NoOfDiskPagesToDiskAfterRestartACC`, `NoOfDiskPagesToDiskAfterRestartTUP`, `ReservedSendBufferMemory`, `MaxNoOfIndexes`, and `Discless` (use `Diskless` instead), as well as `DiskCheckpointSpeed` and `DiskCheckpointSpeedInRestart`. The archaic and unused `ByteOrder` computer configuration parameter has also been removed, as well as the unused `MaxNoOfSavedEvents` management node configuration parameter. These parameters are no longer supported; most of them already did not have (or no longer had) any effect. Trying to use any of these parameters in a MySQL Cluster configuration file now results in an error.

For more information, see [What is New in MySQL NDB Cluster 7.5](#). (Bug #77404, Bug #21280428)

- **Important Change:** The `ExecuteOnComputer` configuration parameter for management, data, and API nodes is now deprecated, and is subject to removal in a future MySQL Cluster release. For all types of MySQL Cluster nodes, you should now use the `HostName` parameter exclusively for identifying hosts in the cluster configuration file.

This information is also now displayed in the output of `ndb_config --configinfo --xml`. (Bug #53052, Bug #11760628)

- **Important Change:** The `ndbinfo` database can now provide default and current information about MySQL Cluster node configuration parameters as a result of the following changes:
 1. The `config_params` table has been enhanced with additional columns providing information about each configuration parameter, including its type, default, and maximum and minimum values (where applicable).
 2. A new `config_values` table has been added. A row in this table shows the current value of a parameter on a given node.

You can obtain values of MySQL Cluster configuration parameters by name using a join on these two tables such as the one shown here:

```
SELECT  p.param_name AS Name,
        v.node_id AS Node,
        p.param_type AS Type,
        p.param_default AS 'Default',
        v.config_value AS Current
FROM    config_params p
JOIN    config_values v
ON      p.param_number = v.config_param
WHERE   p.param_name IN ('NodeId', 'HostName', 'DataMemory', 'IndexMemory');
```

- Deprecated MySQL Cluster node configuration parameters are now indicated as such by `ndb_config --configinfo --xml`. For each parameter currently deprecated, the corresponding `<param/>` tag in the XML output now includes the attribute `deprecated="true"`. (Bug #21127135)
- Added the `--ndb-cluster-connection-pool-nodeids` option for `mysqld`, which can be used to specify a list of nodes by node ID for connection pooling. The number of node IDs in the list must equal the value set for `--ndb-cluster-connection-pool`. (Bug #19521789)
- Added the `PROMPT` command in the `ndb_mgm` client. This command has the syntax `PROMPT string`, which sets the client's prompt to `string`. Issuing the command without an argument causes the prompt to be reset to the default (`ndb_mgm>`). See [Commands in the NDB Cluster Management Client](#), for more information. (Bug #18421338)

- When the `--database` option has not been specified for `ndb_show_tables`, and no tables are found in the `TEST_DB` database, an appropriate warning message is now issued. (Bug #50633, Bug #11758430)
- The `NDB` storage engine now uses the improved records-per-key interface for index statistics introduced for the optimizer in MySQL 5.7. Some improvements due to this change are listed here:
 - The optimizer can now choose better execution plans for queries on `NDB` tables in many cases where a less optimal join index or table join order would previously have been chosen.
 - `EXPLAIN` now provides more accurate row estimates than previously.
 - Improved cardinality estimates can be obtained from `SHOW INDEX`.

Bugs Fixed

- **Incompatible Change; NDB Cluster APIs:** The `pollEvents2()` method now returns -1, indicating an error, whenever a negative value is used for the time argument. (Bug #20762291)
- **Important Change:** The options controlling behavior of `NDB` programs with regard to the number and timing of successive attempts to connect to a management server have changed as listed here:
 - The minimum value for the `--connect-retry-delay` option common to all `NDB` programs has been changed from 0 to 1; this means that all `NDB` programs now wait at least 1 second between successive connection attempts, and it is no longer possible to set a waiting time equal to 0.
 - The semantics for the `--connect-retries` option have changed slightly, such that the value of this option now sets the number of times an `NDB` program tries to connect to a management server. Setting this option to 0 now causes the program to attempt the connection indefinitely, until it either succeeds or is terminated by other means (such as `kill`).
 - In addition, the default for the `--connect-retries` option for the `ndb_mgm` client has been changed from 3 to 12, so that the minimum, maximum, and default values for this option when used with `ndb_mgm` are now exactly the same as for all other `NDB` programs.

The `ndb_mgm --try-reconnect` option, although deprecated in MySQL Cluster NDB 7.4, continues to be supported as a synonym for `ndb_mgm --connect-retries` to provide backwards compatibility. The default value for `--try-reconnect` has also been changed from 3 to 12, respectively, so that this option continues to behave in the exactly in the same way as `--connect-retries`.

(Bug #22116937)

- **Important Change:** In previous versions of MySQL Cluster, other DDL operations could not be part of `ALTER ONLINE TABLE ... RENAME ...` (This was disallowed by the fix for BUG#16021021.) MySQL Cluster NDB 7.5 makes the following changes:
 - Support for the `ONLINE` and `OFFLINE` keywords, which was deprecated in MySQL Cluster NDB 7.3, is now removed, and use of these now causes a syntax error; the `NDB` storage engine now accepts *only* `ALGORITHM = DEFAULT`, `ALGORITHM = COPY`, and `ALGORITHM = INPLACE` to specify whether the `ALTER` operation is copying or in-place, just as in the standard MySQL Server.
 - `NDB` now allows `ALTER TABLE ... ALGORITHM=COPYING RENAME`.
- (Bug #20804269, Bug #76543, Bug #20479917, Bug #75797)

References: See also: Bug #16021021.

- **Important Change; NDB Cluster APIs:** `Ndb::pollEvents()` is now compatible with the `TE_EMPTY`, `TE_INCONSISTENT`, and `TE_OUT_OF_MEMORY` event types introduced in MySQL Cluster NDB 7.4.3. For detailed information about this change, see the description of this method in the *MySQL Cluster API Developer Guide*. (Bug #20646496)
- **Important Change; NDB Cluster APIs:** Added the method `Ndb::isExpectingHigherQueuedEpochs()` to the NDB API to detect when additional, newer event epochs were detected by `pollEvents2()`.

The behavior of `Ndb::pollEvents()` has also been modified such that it now returns `NDB_FAILURE_GCI` (equal to `~(Uint64) 0`) when a cluster failure has been detected. (Bug #18753887)

- **Important Change; NDB Cluster APIs:** To release the memory used for dropped event operations, the event API formerly depended on `pollEvents()` and `nextEvent()` to consume all events possibly referring to the dropped events. This dependency between `dropEventOperation()` and the first two methods required the entire event buffer to be read before attempting to release event operation memory (that is, until successive calls to `pollEvents()` and `nextEvent()` returned no more events).

A related cleanup issue arose following the reset of the event buffer (when all event operations had previously been dropped), and the event buffer was truncated by the first `createEventOperation()` call subsequent to the reset.

To fix these problems, the event buffer is now cleared when the last event operation is dropped, rather than waiting for a subsequent create operation which might or might not occur. Memory taken up by dropped event operations is also now released when the event queue has been cleared, which removes the hidden requirement for consuming all events to free up memory. In addition, event operation memory is now released as soon as all events referring to the operation have been consumed, rather than waiting for the entire event buffer to be consumed. (Bug #78145, Bug #21661297)

- **Important Change; NDB Cluster APIs:** The MGM API error-handling functions `ndb_mgm_get_latest_error()`, `ndb_mgm_get_latest_error_msg()`, and `ndb_mgm_get_latest_error_desc()` each failed when used with a `NULL` handle. You should note that, although these functions are now null-safe, values returned in this case are arbitrary and not meaningful. (Bug #78130, Bug #21651706)
- **Important Change; NDB Cluster APIs:** The following NDB API methods were not actually implemented and have been removed from the sources:
 - `Datafile` methods: `getNode()`, `setNode()`, and `getFileNo()`
 - `Undofile` methods: `getNode()`, `setNode()`, and `getFileNo()`
 - `Table` methods: `getObjectType()` and `setObjectType()`
- A serious regression was inadvertently introduced in MySQL Cluster NDB 7.4.8 whereby local checkpoints and thus restarts often took much longer than expected. This occurred due to the fact that the setting for `MaxDiskWriteSpeedOwnRestart` was ignored during restarts and the value of `MaxDiskWriteSpeedOtherNodeRestart`, which is much lower by default than the default for `MaxDiskWriteSpeedOwnRestart`, was used instead. This issue affected restart times and performance only and did not have any impact on normal operations. (Bug #22582233)
- The epoch for the latest restorable checkpoint provided in the cluster log as part of its reporting for `EventBufferStatus` events (see [NDB Cluster : Messages in the Cluster Log](#)) was not well defined and thus unreliable; depending on various factors, the reported epoch could be the one currently being consumed, the one most recently consumed, or the next one queued for consumption.

This fix ensures that the latest restorable global checkpoint is always regarded as the one that was most recently completely consumed by the user, and thus that it was the latest restorable global checkpoint that existed at the time the report was generated. (Bug #22378288)

- Added the `--ndb-allow-copying-alter-table` option for `mysqld`. Setting this option (or the equivalent system variable `ndb_allow_copying_alter_table`) to `OFF` keeps `ALTER TABLE` statements from performing copying operations. The default value is `ON`. (Bug #22187649)

References: See also: Bug #17400320.

- Attempting to create an `NDB` table having greater than the maximum supported combined width for all `BIT` columns (4096) caused data node failure when these columns were defined with `COLUMN_FORMAT DYNAMIC`. (Bug #21889267)
- Creating a table with the maximum supported number of columns (512) all using `COLUMN_FORMAT DYNAMIC` led to data node failures. (Bug #21863798)
- In a MySQL Cluster with multiple LDM instances, all instances wrote to the node log, even inactive instances on other nodes. During restarts, this caused the log to be filled with messages from other nodes, such as the messages shown here:

```
2015-06-24 00:20:16 [ndbd] INFO      -- We are adjusting Max Disk Write Speed,
a restart is ongoing now
...
2015-06-24 01:08:02 [ndbd] INFO      -- We are adjusting Max Disk Write Speed,
no restarts ongoing anymore
```

Now this logging is performed only by the active LDM instance. (Bug #21362380)

- Backup block states were reported incorrectly during backups. (Bug #21360188)

References: See also: Bug #20204854, Bug #21372136.

- For a timeout in `GET_TABINFOREQ` while executing a `CREATE INDEX` statement, `mysqld` returned Error 4243 (`Index not found`) instead of the expected Error 4008 (`Receive from NDB failed`).

The fix for this bug also fixes similar timeout issues for a number of other signals that are sent the `DBDICT` kernel block as part of DDL operations, including `ALTER_TAB_REQ`, `CREATE_INDX_REQ`, `DROP_FK_REQ`, `DROP_INDX_REQ`, `INDEX_STAT_REQ`, `DROP_FILE_REQ`, `CREATE_FILEGROUP_REQ`, `DROP_FILEGROUP_REQ`, `CREATE_EVENT`, `WAIT_GCP_REQ`, `DROP_TAB_REQ`, and `LIST_TABLES_REQ`, as well as several internal functions used in handling `NDB` schema operations. (Bug #21277472)

References: See also: Bug #20617891, Bug #20368354, Bug #19821115.

- Previously, multiple send threads could be invoked for handling sends to the same node; these threads then competed for the same send lock. While the send lock blocked the additional send threads, work threads could be passed to other nodes.

This issue is fixed by ensuring that new send threads are not activated while there is already an active send thread assigned to the same node. In addition, a node already having an active send thread assigned to it is no longer visible to other, already active, send threads; that is, such a node is longer added to the node list when a send thread is currently assigned to it. (Bug #20954804, Bug #76821)

- Queueing of pending operations when the redo log was overloaded (`DefaultOperationRedoProblemAction` API node configuration parameter) could lead to timeouts when data nodes ran out of redo log space (`P_TAIL_PROBLEM` errors). Now when the redo log is full, the node aborts requests instead of queuing them. (Bug #20782580)

References: See also: Bug #20481140.

- An `NDB` event buffer can be used with an `Ndb` object to subscribe to table-level row change event streams. Users subscribe to an existing event; this causes the data nodes to start sending event data signals (`SUB_TABLE_DATA`) and epoch completion signals (`SUB_GCP_COMPLETE`) to the `Ndb` object. `SUB_GCP_COMPLETE_REP` signals can arrive for execution in concurrent receiver thread before completion of the internal method call used to start a subscription.

Execution of `SUB_GCP_COMPLETE_REP` signals depends on the total number of `SUMA` buckets (sub data streams), but this may not yet have been set, leading to the present issue, when the counter used for tracking the `SUB_GCP_COMPLETE_REP` signals (`TOTAL_BUCKETS_INIT`) was found to be set to erroneous values. Now `TOTAL_BUCKETS_INIT` is tested to be sure it has been set correctly before it is used. (Bug #20575424, Bug #76255)

References: See also: Bug #20561446, Bug #21616263.

- `NDB` statistics queries could be delayed by the error delay set for `ndb_index_stat_option` (default 60 seconds) when the index that was queried had been marked with internal error. The same underlying issue could also cause `ANALYZE TABLE` to hang when executed against an `NDB` table having multiple indexes where an internal error occurred on one or more but not all indexes.

Now in such cases, any existing statistics are returned immediately, without waiting for any additional statistics to be discovered. (Bug #20553313, Bug #20707694, Bug #76325)

- Memory allocated when obtaining a list of tables or databases was not freed afterward. (Bug #20234681, Bug #74510)

References: See also: Bug #18592390, Bug #72322.

- Added the `BackupDiskWriteSpeedPct` data node parameter. Setting this parameter causes the data node to reserve a percentage of its maximum write speed (as determined by the value of `MaxDiskWriteSpeed`) for use in local checkpoints while performing a backup. `BackupDiskWriteSpeedPct` is interpreted as a percentage which can be set between 0 and 90 inclusive, with a default value of 50. (Bug #20204854)

References: See also: Bug #21372136.

- After restoring the database schema from backup using `ndb_restore`, auto-discovery of restored tables in transactions having multiple statements did not work correctly, resulting in `Deadlock found when trying to get lock; try restarting transaction` errors.

This issue was encountered both in the `mysql` client, as well as when such transactions were executed by application programs using Connector/J and possibly other MySQL APIs.

Prior to upgrading, this issue can be worked around by executing `SELECT TABLE_NAME, TABLE_SCHEMA FROM INFORMATION_SCHEMA.TABLES WHERE ENGINE = 'NDBCLUSTER'` on all SQL nodes following the restore operation, before executing any other statements. (Bug #18075170)

- Using `ndb_mgm STOP -f` to force a node shutdown even when it triggered a complete shutdown of the cluster, it was possible to lose data when a sufficient number of nodes were shut down, triggering a cluster shutdown, and the timing was such that `SUMA` handovers had been made to nodes already in the process of shutting down. (Bug #17772138)
- When using a sufficiently large value for `TransactionDeadlockDetectionTimeout` and the default value for `sort_buffer_size`, executing `SELECT * FROM ndbinfo.cluster_operations ORDER BY transid` with multiple concurrent conflicting or deadlocked transactions, each transaction

having several pending operations, caused the SQL node where the query was run to fail. (Bug #16731538, Bug #67596)

- The `ndbinfo.config_params` table is now read-only. (Bug #11762750, Bug #55383)
- **NDB** failed during a node restart due to the status of the current local checkpoint being set but not as active, even though it could have other states under such conditions. (Bug #78780, Bug #21973758)
- `ndbmtd` checked for signals being sent only after a full cycle in `run_job_buffers`, which is performed for all job buffer inputs. Now this is done as part of `run_job_buffers` itself, which avoids executing for extended periods of time without sending to other nodes or flushing signals to other threads. (Bug #78530, Bug #21889088)
- When attempting to enable index statistics, creation of the required system tables, events and event subscriptions often fails when multiple `mysqld` processes using index statistics are started concurrently in conjunction with starting, restarting, or stopping the cluster, or with node failure handling. This is normally recoverable, since the affected `mysqld` process or processes can (and do) retry these operations shortly thereafter. For this reason, such failures are no longer logged as warnings, but merely as informational events. (Bug #77760, Bug #21462846)
- It was possible to end up with a lock on the send buffer mutex when send buffers became a limiting resource, due either to insufficient send buffer resource configuration, problems with slow or failing communications such that all send buffers became exhausted, or slow receivers failing to consume what was sent. In this situation worker threads failed to allocate send buffer memory for signals, and attempted to force a send in order to free up space, while at the same time the send thread was busy trying to send to the same node or nodes. All of these threads competed for taking the send buffer mutex, which resulted in the lock already described, reported by the watchdog as `Stuck in Send`. This fix is made in two parts, listed here:
 1. The send thread no longer holds the global send thread mutex while getting the send buffer mutex; it now releases the global mutex prior to locking the send buffer mutex. This keeps worker threads from getting stuck in send in such cases.
 2. Locking of the send buffer mutex done by the send threads now uses a try-lock. If the try-lock fails, the node to make the send to is reinserted at the end of the list of send nodes in order to be retried later. This removes the `Stuck in Send` condition for the send threads.

(Bug #77081, Bug #21109605)

- **NDB Disk Data:** A unique index on a column of an **NDB** table is implemented with an associated internal ordered index, used for scanning. While dropping an index, this ordered index was dropped first, followed by the drop of the unique index itself. This meant that, when the drop was rejected due to (for example) a constraint violation, the statement was rejected but the associated ordered index remained deleted, so that any subsequent operation using a scan on this table failed. We fix this problem by causing the unique index to be removed first, before removing the ordered index; removal of the related ordered index is no longer performed when removal of a unique index fails. (Bug #78306, Bug #21777589)
- **NDB Cluster APIs:** The binary log injector did not work correctly with `TE_INCONSISTENT` event type handling by `Ndb::nextEvent()`. (Bug #22135541)

References: See also: Bug #20646496.

- **NDB Cluster APIs:** While executing `dropEvent()`, if the coordinator `DBDICT` failed after the subscription manager (`SUMA` block) had removed all subscriptions but before the coordinator had deleted the event from the system table, the dropped event remained in the table, causing any subsequent drop

or create event with the same name to fail with NDB error 1419 `Subscription already dropped` or error 746 `Event name already exists`. This occurred even when calling `dropEvent()` with a nonzero force argument.

Now in such cases, error 1419 is ignored, and `DBDICT` deletes the event from the table. (Bug #21554676)

- **NDB Cluster APIs:** Creation and destruction of `Ndb_cluster_connection` objects by multiple threads could make use of the same application lock, which in some cases led to failures in the global dictionary cache. To alleviate this problem, the creation and destruction of several internal NDB API objects have been serialized. (Bug #20636124)
- **NDB Cluster APIs:** When an `Ndb` object created prior to a failure of the cluster was reused, the event queue of this object could still contain data node events originating from before the failure. These events could reference “old” epochs (from before the failure occurred), which in turn could violate the assumption made by the `nextEvent()` method that epoch numbers always increase. This issue is addressed by explicitly clearing the event queue in such cases. (Bug #18411034)

References: See also: Bug #20888668.

- **NDB Cluster APIs:** `Ndb::pollEvents()` and `pollEvents2()` were slow to receive events, being dependent on other client threads or blocks to perform polling of transporters on their behalf. This fix allows a client thread to perform its own transporter polling when it has to wait in either of these methods.

Introduction of transporter polling also revealed a problem with missing mutex protection in the `ndbcluster_binlog` handler, which has been added as part of this fix. (Bug #79311, Bug #20957068, Bug #22224571)

- **NDB Cluster APIs:** After the initial restart of a node following a cluster failure, the cluster failure event added as part of the restart process was deleted when an event that existed prior to the restart was later deleted. This meant that, in such cases, an Event API client had no way of knowing that failure handling was needed. In addition, the GCI used for the final cleanup of deleted event operations, performed by `pollEvents()` and `nextEvent()` when these methods have consumed all available events, was lost. (Bug #78143, Bug #21660947)

