

MySQL Restrictions and Limitations

Abstract

This is the MySQL Restrictions and Limitations extract from the MySQL 5.7 Reference Manual.

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](#).

Licensing information—MySQL 5.7. This product may include third-party software, used under license. If you are using a *Commercial* release of MySQL 5.7, see [this document](#) for licensing information, including licensing information relating to third-party software that may be included in this Commercial release. If you are using a *Community* release of MySQL 5.7, see [this document](#) for licensing information, including licensing information relating to third-party software that may be included in this Community release.

Licensing information—MySQL Cluster. This product may include third-party software, used under license. If you are using a *Community* release of MySQL Cluster NDB 7.5, see [this document](#) for licensing information, including licensing information relating to third-party software that may be included in this Community release.

Document generated on: 2016-08-17 (revision: 48585)

Table of Contents

Preface and Legal Notices	v
1 Restrictions and Limits	1
2 Restrictions on Stored Programs	3
3 Restrictions on Server-Side Cursors	7
4 Restrictions on Subqueries	9
5 Restrictions on Views	11
6 Restrictions on XA Transactions	13
7 Restrictions on Character Sets	15
8 Limits in MySQL	17
8.1 Limits on Joins	17
8.2 Limits on Number of Databases and Tables	17
8.3 Limits on Table Size	17
8.4 Limits on Table Column Count and Row Size	18
8.5 Limits Imposed by .frm File Structure	20
8.6 Windows Platform Limitations	20
9 Restrictions and Limitations on Partitioning	23
9.1 Partitioning Keys, Primary Keys, and Unique Keys	29
9.2 Partitioning Limitations Relating to Storage Engines	32
9.3 Partitioning Limitations Relating to Functions	33
9.4 Partitioning and Locking	34
10 MySQL Differences from Standard SQL	37
10.1 SELECT INTO TABLE Differences	37
10.2 UPDATE Differences	37
10.3 Foreign Key Differences	37
10.4 '--' as the Start of a Comment	38
11 Known Issues in MySQL	41

Preface and Legal Notices

This is the MySQL Restrictions and Limitations extract from the MySQL 5.7 Reference Manual.

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.

Chapter 1 Restrictions and Limits

The discussion here describes restrictions that apply to the use of MySQL features such as subqueries or views.

Chapter 2 Restrictions on Stored Programs

These restrictions apply to the features described in [Stored Programs and Views](#).

Some of the restrictions noted here apply to all stored routines; that is, both to stored procedures and stored functions. There are also some [restrictions specific to stored functions](#) but not to stored procedures.

The restrictions for stored functions also apply to triggers. There are also some [restrictions specific to triggers](#).

The restrictions for stored procedures also apply to the `DO` clause of Event Scheduler event definitions. There are also some [restrictions specific to events](#).

SQL Statements Not Permitted in Stored Routines

Stored routines cannot contain arbitrary SQL statements. The following statements are not permitted:

- The locking statements `LOCK TABLES` and `UNLOCK TABLES`.
- `ALTER VIEW`.
- `LOAD DATA` and `LOAD TABLE`.
- SQL prepared statements (`PREPARE`, `EXECUTE`, `DEALLOCATE PREPARE`) can be used in stored procedures, but not stored functions or triggers. Thus, stored functions and triggers cannot use dynamic SQL (where you construct statements as strings and then execute them).
- Generally, statements not permitted in SQL prepared statements are also not permitted in stored programs. For a list of statements supported as prepared statements, see [SQL Syntax for Prepared Statements](#). Exceptions are `SIGNAL`, `RESIGNAL`, and `GET DIAGNOSTICS`, which are not permissible as prepared statements but are permitted in stored programs.
- Because local variables are in scope only during stored program execution, references to them are not permitted in prepared statements created within a stored program. Prepared statement scope is the current session, not the stored program, so the statement could be executed after the program ends, at which point the variables would no longer be in scope. For example, `SELECT ... INTO local_var` cannot be used as a prepared statement. This restriction also applies to stored procedure and function parameters. See [PREPARE Syntax](#).
- Within all stored programs (stored procedures and functions, triggers, and events), the parser treats `BEGIN [WORK]` as the beginning of a `BEGIN ... END` block. To begin a transaction in this context, use `START TRANSACTION` instead.

Restrictions for Stored Functions

The following additional statements or operations are not permitted within stored functions. They are permitted within stored procedures, except stored procedures that are invoked from within a stored function or trigger. For example, if you use `FLUSH` in a stored procedure, that stored procedure cannot be called from a stored function or trigger.

- Statements that perform explicit or implicit commit or rollback. Support for these statements is not required by the SQL standard, which states that each DBMS vendor may decide whether to permit them.
- Statements that return a result set. This includes `SELECT` statements that do not have an `INTO var_list` clause and other statements such as `SHOW`, `EXPLAIN`, and `CHECK TABLE`. A function can process a result set either with `SELECT ... INTO var_list` or by using a cursor and `FETCH` statements. See [SELECT ... INTO Syntax](#), and [Cursors](#).

- [FLUSH](#) statements.
- Stored functions cannot be used recursively.
- A stored function or trigger cannot modify a table that is already being used (for reading or writing) by the statement that invoked the function or trigger.
- If you refer to a temporary table multiple times in a stored function under different aliases, a [Can't reopen table: 'tbl_name'](#) error occurs, even if the references occur in different statements within the function.
- [HANDLER ... READ](#) statements that invoke stored functions can cause replication errors and are disallowed.

Restrictions for Triggers

For triggers, the following additional restrictions apply:

- Triggers are not activated by foreign key actions.
- When using row-based replication, triggers on the slave are not activated by statements originating on the master. The triggers on the slave are activated when using statement-based replication. For more information, see [Replication and Triggers](#).
- The [RETURN](#) statement is not permitted in triggers, which cannot return a value. To exit a trigger immediately, use the [LEAVE](#) statement.
- Triggers are not permitted on tables in the [mysql](#) database.
- The trigger cache does not detect when metadata of the underlying objects has changed. If a trigger uses a table and the table has changed since the trigger was loaded into the cache, the trigger operates using the outdated metadata.

Name Conflicts within Stored Routines

The same identifier might be used for a routine parameter, a local variable, and a table column. Also, the same local variable name can be used in nested blocks. For example:

```
CREATE PROCEDURE p (i INT)
BEGIN
  DECLARE i INT DEFAULT 0;
  SELECT i FROM t;
  BEGIN
    DECLARE i INT DEFAULT 1;
    SELECT i FROM t;
  END;
END;
```

In such cases, the identifier is ambiguous and the following precedence rules apply:

- A local variable takes precedence over a routine parameter or table column.
- A routine parameter takes precedence over a table column.
- A local variable in an inner block takes precedence over a local variable in an outer block.

The behavior that variables take precedence over table columns is nonstandard.

Replication Considerations

Use of stored routines can cause replication problems. This issue is discussed further in [Binary Logging of Stored Programs](#).

The `--replicate-wild-do-table=db_name.tbl_name` option applies to tables, views, and triggers. It does not apply to stored procedures and functions, or events. To filter statements operating on the latter objects, use one or more of the `--replicate-*-db` options.

Debugging Considerations

There are no stored routine debugging facilities.

Unsupported Syntax from the SQL:2003 Standard

The MySQL stored routine syntax is based on the SQL:2003 standard. The following items from that standard are not currently supported:

- [UNDO](#) handlers
- [FOR](#) loops

Concurrency Considerations

To prevent problems of interaction between sessions, when a client issues a statement, the server uses a snapshot of routines and triggers available for execution of the statement. That is, the server calculates a list of procedures, functions, and triggers that may be used during execution of the statement, loads them, and then proceeds to execute the statement. While the statement executes, it does not see changes to routines performed by other sessions.

For maximum concurrency, stored functions should minimize their side-effects; in particular, updating a table within a stored function can reduce concurrent operations on that table. A stored function acquires table locks before executing, to avoid inconsistency in the binary log due to mismatch of the order in which statements execute and when they appear in the log. When statement-based binary logging is used, statements that invoke a function are recorded rather than the statements executed within the function. Consequently, stored functions that update the same underlying tables do not execute in parallel. In contrast, stored procedures do not acquire table-level locks. All statements executed within stored procedures are written to the binary log, even for statement-based binary logging. See [Binary Logging of Stored Programs](#).

Event Scheduler Restrictions

The following limitations are specific to the Event Scheduler:

- Event names are handled in case-insensitive fashion. For example, you cannot have two events in the same database with the names [anEvent](#) and [AnEvent](#).
- An event may not be created, altered, or dropped by a stored routine, trigger, or another event. An event also may not create, alter, or drop stored routines or triggers. (Bug #16409, Bug #18896)
- DDL statements on events are prohibited while a [LOCK TABLES](#) statement is in effect.
- Event timings using the intervals [YEAR](#), [QUARTER](#), [MONTH](#), and [YEAR_MONTH](#) are resolved in months; those using any other interval are resolved in seconds. There is no way to cause events scheduled to occur at the same second to execute in a given order. In addition—due to rounding, the nature of threaded applications, and the fact that a nonzero length of time is required to create events and to signal their execution—events may be delayed by as much as 1 or 2 seconds. However, the time shown in the [INFORMATION_SCHEMA.EVENTS](#) table's [LAST_EXECUTED](#) column or the [mysql.event](#) table's [last_executed](#) column is always accurate to within one second of the actual event execution time. (See also Bug #16522.)
- Each execution of the statements contained in the body of an event takes place in a new connection; thus, these statements has no effect in a given user session on the server's statement counts such as [Com_select](#) and [Com_insert](#) that are displayed by [SHOW STATUS](#). However, such counts are updated in the global scope. (Bug #16422)

- Events do not support times later than the end of the Unix Epoch; this is approximately the beginning of the year 2038. Such dates are specifically not permitted by the Event Scheduler. (Bug #16396)
- References to stored functions, user-defined functions, and tables in the **ON SCHEDULE** clauses of **CREATE EVENT** and **ALTER EVENT** statements are not supported. These sorts of references are not permitted. (See Bug #22830 for more information.)

Stored routines and triggers in MySQL Cluster. Stored procedures, stored functions, and triggers are all supported by tables using the **NDB** storage engine; however, it is important to keep in mind that they do *not* propagate automatically between MySQL Servers acting as Cluster SQL nodes. This is because of the following:

- Stored routine definitions are kept in tables in the **mysql** system database using the **MyISAM** storage engine, and so do not participate in clustering.
- The **.TRN** and **.TRG** files containing trigger definitions are not read by the **NDB** storage engine, and are not copied between Cluster nodes.

Any stored routine or trigger that interacts with MySQL Cluster tables must be re-created by running the appropriate **CREATE PROCEDURE**, **CREATE FUNCTION**, or **CREATE TRIGGER** statements on each MySQL Server that participates in the cluster where you wish to use the stored routine or trigger. Similarly, any changes to existing stored routines or triggers must be carried out explicitly on all Cluster SQL nodes, using the appropriate **ALTER** or **DROP** statements on each MySQL Server accessing the cluster.

Warning

Do *not* attempt to work around the issue described in the first item mentioned previously by converting any **mysql** database tables to use the **NDB** storage engine. *Altering the system tables in the **mysql** database is not supported* and is very likely to produce undesirable results.

Chapter 3 Restrictions on Server-Side Cursors

Server-side cursors are implemented in the C API using the `mysql_stmt_attr_set()` function. The same implementation is used for cursors in stored routines. A server-side cursor enables a result set to be generated on the server side, but not transferred to the client except for those rows that the client requests. For example, if a client executes a query but is only interested in the first row, the remaining rows are not transferred.

In MySQL, a server-side cursor is materialized into an internal temporary table. Initially, this is a `MEMORY` table, but is converted to a `MyISAM` table when its size exceeds the minimum value of the `max_heap_table_size` and `tmp_table_size` system variables. The same restrictions apply to internal temporary tables created to hold the result set for a cursor as for other uses of internal temporary tables. See [Internal Temporary Table Use in MySQL](#). One limitation of the implementation is that for a large result set, retrieving its rows through a cursor might be slow.

Cursors are read only; you cannot use a cursor to update rows.

`UPDATE WHERE CURRENT OF` and `DELETE WHERE CURRENT OF` are not implemented, because updatable cursors are not supported.

Cursors are nonholdable (not held open after a commit).

Cursors are asensitive.

Cursors are nonscrollable.

Cursors are not named. The statement handler acts as the cursor ID.

You can have open only a single cursor per prepared statement. If you need several cursors, you must prepare several statements.

You cannot use a cursor for a statement that generates a result set if the statement is not supported in prepared mode. This includes statements such as `CHECK TABLE`, `HANDLER READ`, and `SHOW BINLOG EVENTS`.

Chapter 4 Restrictions on Subqueries

- In general, you cannot modify a table and select from the same table in a subquery. For example, this limitation applies to statements of the following forms:

```
DELETE FROM t WHERE ... (SELECT ... FROM t ...);
UPDATE t ... WHERE col = (SELECT ... FROM t ...);
{INSERT|REPLACE} INTO t (SELECT ... FROM t ...);
```

Exception: The preceding prohibition does not apply if you are using a subquery for the modified table in the **FROM** clause. Example:

```
UPDATE t ... WHERE col = (SELECT * FROM (SELECT ... FROM t...) AS _t ...);
```

Here the result from the subquery in the **FROM** clause is stored as a temporary table, so the relevant rows in **t** have already been selected by the time the update to **t** takes place.

- Row comparison operations are only partially supported:
 - For *expr* **[NOT] IN** *subquery*, *expr* can be an *n*-tuple (specified using row constructor syntax) and the subquery can return rows of *n*-tuples. The permitted syntax is therefore more specifically expressed as *row_constructor* **[NOT] IN** *table_subquery*
 - For *expr op* **{ALL|ANY|SOME}** *subquery*, *expr* must be a scalar value and the subquery must be a column subquery; it cannot return multiple-column rows.

In other words, for a subquery that returns rows of *n*-tuples, this is supported:

```
(expr_1, ..., expr_n) [NOT] IN table_subquery
```

But this is not supported:

```
(expr_1, ..., expr_n) op {ALL|ANY|SOME} subquery
```

The reason for supporting row comparisons for **IN** but not for the others is that **IN** is implemented by rewriting it as a sequence of **=** comparisons and **AND** operations. This approach cannot be used for **ALL**, **ANY**, or **SOME**.

- Subqueries in the **FROM** clause cannot be correlated subqueries. They are materialized in whole (evaluated to produce a result set) during query execution, so they cannot be evaluated per row of the outer query. The optimizer delays materialization until the result is needed, which may permit materialization to be avoided. See [Optimizing Derived Tables and View References](#).
- MySQL does not support **LIMIT** in subqueries for certain subquery operators:

```
mysql> SELECT * FROM t1
-> WHERE s1 IN (SELECT s2 FROM t2 ORDER BY s1 LIMIT 1);
ERROR 1235 (42000): This version of MySQL doesn't yet support
'LIMIT & IN/ALL/ANY/SOME subquery'
```

- MySQL permits a subquery to refer to a stored function that has data-modifying side effects such as inserting rows into a table. For example, if **f()** inserts rows, the following query can modify data:

```
SELECT ... WHERE x IN (SELECT f() ...);
```

This behavior is an extension to the SQL standard. In MySQL, it can produce indeterminate results because **f()** might be executed a different number of times for different executions of a given query depending on how the optimizer chooses to handle it.

For statement-based or mixed-format replication, one implication of this indeterminism is that such a query can produce different results on the master and its slaves.

Chapter 5 Restrictions on Views

View processing is not optimized:

- It is not possible to create an index on a view.
- Indexes can be used for views processed using the merge algorithm. However, a view that is processed with the temptable algorithm is unable to take advantage of indexes on its underlying tables (although indexes can be used during generation of the temporary tables).

Before MySQL 5.7.7, subqueries cannot be used in the `FROM` clause of a view.

There is a general principle that you cannot modify a table and select from the same table in a subquery. See [Chapter 4, Restrictions on Subqueries](#).

The same principle also applies if you select from a view that selects from the table, if the view selects from the table in a subquery and the view is evaluated using the merge algorithm. Example:

```
CREATE VIEW v1 AS
SELECT * FROM t2 WHERE EXISTS (SELECT 1 FROM t1 WHERE t1.a = t2.a);
UPDATE t1, v2 SET t1.a = 1 WHERE t1.b = v2.b;
```

If the view is evaluated using a temporary table, you *can* select from the table in the view subquery and still modify that table in the outer query. In this case the view will be stored in a temporary table and thus you are not really selecting from the table in a subquery and modifying it “at the same time.” (This is another reason you might wish to force MySQL to use the temptable algorithm by specifying `ALGORITHM = TEMPTABLE` in the view definition.)

You can use `DROP TABLE` or `ALTER TABLE` to drop or alter a table that is used in a view definition. No warning results from the `DROP` or `ALTER` operation, even though this invalidates the view. Instead, an error occurs later, when the view is used. `CHECK TABLE` can be used to check for views that have been invalidated by `DROP` or `ALTER` operations.

With regard to view updatability, the overall goal for views is that if any view is theoretically updatable, it should be updatable in practice. This includes views that have `UNION` in their definition. Not all views that are theoretically updatable can be updated. The initial view implementation was deliberately written this way to get usable, updatable views into MySQL as quickly as possible. Many theoretically updatable views can be updated now, but limitations still exist:

- Updatable views with subqueries anywhere other than in the `WHERE` clause. Some views that have subqueries in the `SELECT` list may be updatable.
- You cannot use `UPDATE` to update more than one underlying table of a view that is defined as a join.
- You cannot use `DELETE` to update a view that is defined as a join.

There exists a shortcoming with the current implementation of views. If a user is granted the basic privileges necessary to create a view (the `CREATE VIEW` and `SELECT` privileges), that user will be unable to call `SHOW CREATE VIEW` on that object unless the user is also granted the `SHOW VIEW` privilege.

That shortcoming can lead to problems backing up a database with `mysqldump`, which may fail due to insufficient privileges. This problem is described in Bug #22062.

The workaround to the problem is for the administrator to manually grant the `SHOW VIEW` privilege to users who are granted `CREATE VIEW`, since MySQL doesn't grant it implicitly when views are created.

Views do not have indexes, so index hints do not apply. Use of index hints when selecting from a view is not permitted.

`SHOW CREATE VIEW` displays view definitions using an `AS alias_name` clause for each column. If a column is created from an expression, the default alias is the expression text, which can be quite long. Aliases for column names in `CREATE VIEW` statements are checked against the maximum column length of 64 characters (not the maximum alias length of 256 characters). As a result, views created from the output of `SHOW CREATE VIEW` fail if any column alias exceeds 64 characters. This can cause problems in the following circumstances for views with too-long aliases:

- View definitions fail to replicate to newer slaves that enforce the column-length restriction.
- Dump files created with `mysqldump` cannot be loaded into servers that enforce the column-length restriction.

A workaround for either problem is to modify each problematic view definition to use aliases that provide shorter column names. Then the view will replicate properly, and can be dumped and reloaded without causing an error. To modify the definition, drop and create the view again with `DROP VIEW` and `CREATE VIEW`, or replace the definition with `CREATE OR REPLACE VIEW`.

For problems that occur when reloading view definitions in dump files, another workaround is to edit the dump file to modify its `CREATE VIEW` statements. However, this does not change the original view definitions, which may cause problems for subsequent dump operations.

Chapter 6 Restrictions on XA Transactions

XA transaction support is limited to the [InnoDB](#) storage engine.

For “external XA,” a MySQL server acts as a Resource Manager and client programs act as Transaction Managers. For “Internal XA”, storage engines within a MySQL server act as RMs, and the server itself acts as a TM. Internal XA support is limited by the capabilities of individual storage engines. Internal XA is required for handling XA transactions that involve more than one storage engine. The implementation of internal XA requires that a storage engine support two-phase commit at the table handler level, and currently this is true only for [InnoDB](#).

For [XA START](#), the [JOIN](#) and [RESUME](#) clauses are not supported.

For [XA END](#), the [SUSPEND \[FOR MIGRATE\]](#) clause is not supported.

The requirement that the [bqual](#) part of the [xid](#) value be different for each XA transaction within a global transaction is a limitation of the current MySQL XA implementation. It is not part of the XA specification.

Prior to MySQL 5.7.7, XA transactions were not compatible with replication. This was because an XA transaction that was in [PREPARED](#) state would be rolled back on clean server shutdown or client disconnect. Similarly, an XA transaction that was in [PREPARED](#) state would still exist in [PREPARED](#) state in case the server was shutdown abnormally and then started again, but the contents of the transaction could not be written to the binary log. In both of these situations the XA transaction could not be replicated correctly.

In MySQL 5.7.7 and later, there is a change in behavior and an XA transaction is written to the binary log in two parts. When [XA PREPARE](#) is issued, the first part of the transaction up to [XA PREPARE](#) is written using an initial GTID. A [XA_prepare_log_event](#) is used to identify such transactions in the binary log. When [XA COMMIT](#) or [XA ROLLBACK](#) is issued, a second part of the transaction containing only the [XA COMMIT](#) or [XA ROLLBACK](#) statement is written using a second GTID. Note that the initial part of the transaction, identified by [XA_prepare_log_event](#), is not necessarily followed by its [XA COMMIT](#) or [XA ROLLBACK](#), which can cause interleaved binary logging of any two XA transactions. The two parts of the XA transaction can even appear in different binary log files. This means that an XA transaction in [PREPARED](#) state is now persistent until an explicit [XA COMMIT](#) or [XA ROLLBACK](#) statement is issued, ensuring that XA transactions are compatible with replication.

The following restrictions exist for using XA transactions in MySQL 5.7.7 and later:

- XA is not fully resilient to an unexpected halt with respect to the binary log (on the master). If there is an unexpected halt before [XA PREPARE](#), between [XA PREPARE](#) and [XA COMMIT](#) (or [XA ROLLBACK](#)), or after [XA COMMIT](#) (or [XA ROLLBACK](#)), the server and binary log are correctly recovered and taken to a consistent state. However, if there is an unexpected halt in the middle of the execution of one of these statements, the server may not be able to recover to a correct state, leaving the server and the binary log in an inconsistent state.
- XA does not work with [relay-log-info-repository=TABLE](#).
- XA does not work with replication filters or binary log filters. Filters are permitted as long as they do not render any XA transactions empty. Filters that filter out XA transactions may cause the slave to stop with an error.
- In case GTIDs are enabled and the slave does not use either [log-bin=OFF](#) or does not use [log-slave-updates](#), XA transactions are not crash-safe with respect to GTIDs on the slave. If the slave stops unexpectedly while applying an [XA PREPARE](#) or [XA COMMIT](#), then after recovery [@@GLOBAL.GTID_EXECUTED](#) may not correctly describe the transactions that have been applied on the slave.

Chapter 7 Restrictions on Character Sets

- Identifiers are stored in `mysql` database tables (`user`, `db`, and so forth) using `utf8`, but identifiers can contain only characters in the Basic Multilingual Plane (BMP). Supplementary characters are not permitted in identifiers.
- The `ucs2`, `utf16`, `utf16le`, and `utf32` character sets have the following restrictions:
 - They cannot be used as a client character set, which means that they do not work for `SET NAMES` or `SET CHARACTER SET`. (See [Connection Character Sets and Collations](#).)
 - It is currently not possible to use `LOAD DATA INFILE` to load data files that use these character sets.
 - `FULLTEXT` indexes cannot be created on a column that uses any of these character sets. However, you can perform `IN BOOLEAN MODE` searches on the column without an index.
 - The use of `ENCRYPT()` with these character sets is not recommended because the underlying system call expects a string terminated by a zero byte.
- The `REGEXP` and `RLIKE` operators work in byte-wise fashion, so they are not multibyte safe and may produce unexpected results with multibyte character sets. In addition, these operators compare characters by their byte values and accented characters may not compare as equal even if a given collation treats them as equal.

Chapter 8 Limits in MySQL

Table of Contents

8.1 Limits on Joins	17
8.2 Limits on Number of Databases and Tables	17
8.3 Limits on Table Size	17
8.4 Limits on Table Column Count and Row Size	18
8.5 Limits Imposed by .frm File Structure	20
8.6 Windows Platform Limitations	20

This section lists current limits in MySQL 5.7.

8.1 Limits on Joins

The maximum number of tables that can be referenced in a single join is 61. This includes a join handled by merging derived tables (subqueries) and views in the **FROM** clause into the outer query block (see [Optimizing Derived Tables and View References](#)). It also applies to the number of tables that can be referenced in the definition of a view.

8.2 Limits on Number of Databases and Tables

MySQL has no limit on the number of databases. The underlying file system may have a limit on the number of directories.

MySQL has no limit on the number of tables. The underlying file system may have a limit on the number of files that represent tables. Individual storage engines may impose engine-specific constraints. **InnoDB** permits up to 4 billion tables.

8.3 Limits on Table Size

The effective maximum table size for MySQL databases is usually determined by operating system constraints on file sizes, not by MySQL internal limits. The following table lists some examples of operating system file-size limits. This is only a rough guide and is not intended to be definitive. For the most up-to-date information, be sure to check the documentation specific to your operating system.

Operating System	File-size Limit
Win32 w/ FAT/FAT32	2GB/4GB
Win32 w/ NTFS	2TB (possibly larger)
Linux 2.2-Intel 32-bit	2GB (LFS: 4GB)
Linux 2.4+	(using ext3 file system) 4TB
Solaris 9/10	16TB
OS X w/ HFS+	2TB

Windows users, please note that FAT and VFAT (FAT32) are *not* considered suitable for production use with MySQL. Use NTFS instead.

On Linux 2.2, you can get **MyISAM** tables larger than 2GB in size by using the Large File Support (LFS) patch for the ext2 file system. Most current Linux distributions are based on kernel 2.4 or higher and include all the required LFS patches. On Linux 2.4, patches also exist for ReiserFS to get support for big files (up to 2TB). With JFS and XFS, petabyte and larger files are possible on Linux.

For a detailed overview about LFS in Linux, have a look at Andreas Jaeger's *Large File Support in Linux* page at http://www.suse.de/~aj/linux_lfs.html.

If you do encounter a full-table error, there are several reasons why it might have occurred:

- The disk might be full.
- The **InnoDB** storage engine maintains **InnoDB** tables within a tablespace that can be created from several files. This enables a table to exceed the maximum individual file size. The tablespace can include raw disk partitions, which permits extremely large tables. The maximum tablespace size is 64TB.

If you are using **InnoDB** tables and run out of room in the **InnoDB** tablespace. In this case, the solution is to extend the **InnoDB** tablespace. See [Changing the Number or Size of InnoDB Redo Log Files](#).

- You are using **MyISAM** tables on an operating system that supports files only up to 2GB in size and you have hit this limit for the data file or index file.
- You are using a **MyISAM** table and the space required for the table exceeds what is permitted by the internal pointer size. **MyISAM** permits data and index files to grow up to 256TB by default, but this limit can be changed up to the maximum permissible size of 65,536TB ($256^7 - 1$ bytes).

If you need a **MyISAM** table that is larger than the default limit and your operating system supports large files, the **CREATE TABLE** statement supports **AVG_ROW_LENGTH** and **MAX_ROWS** options. See [CREATE TABLE Syntax](#). The server uses these options to determine how large a table to permit.

If the pointer size is too small for an existing table, you can change the options with **ALTER TABLE** to increase a table's maximum permissible size. See [ALTER TABLE Syntax](#).

```
ALTER TABLE tbl_name MAX_ROWS=1000000000 AVG_ROW_LENGTH=nnn;
```

You have to specify **AVG_ROW_LENGTH** only for tables with **BLOB** or **TEXT** columns; in this case, MySQL can't optimize the space required based only on the number of rows.

To change the default size limit for **MyISAM** tables, set the **myisam_data_pointer_size**, which sets the number of bytes used for internal row pointers. The value is used to set the pointer size for new tables if you do not specify the **MAX_ROWS** option. The value of **myisam_data_pointer_size** can be from 2 to 7. A value of 4 permits tables up to 4GB; a value of 6 permits tables up to 256TB.

You can check the maximum data and index sizes by using this statement:

```
SHOW TABLE STATUS FROM db_name LIKE 'tbl_name';
```

You also can use **myisamchk -dv /path/to/table-index-file**. See [SHOW Syntax](#), or [myisamchk — MyISAM Table-Maintenance Utility](#).

Other ways to work around file-size limits for **MyISAM** tables are as follows:

- If your large table is read only, you can use **myisampack** to compress it. **myisampack** usually compresses a table by at least 50%, so you can have, in effect, much bigger tables. **myisampack** also can merge multiple tables into a single table. See [myisampack — Generate Compressed, Read-Only MyISAM Tables](#).
- MySQL includes a **MERGE** library that enables you to handle a collection of **MyISAM** tables that have identical structure as a single **MERGE** table. See [The MERGE Storage Engine](#).
- You are using the **MEMORY (HEAP)** storage engine; in this case you need to increase the value of the **max_heap_table_size** system variable. See [Server System Variables](#).

8.4 Limits on Table Column Count and Row Size

There is a hard limit of 4096 columns per table, but the effective maximum may be less for a given table. The exact limit depends on several interacting factors.

- Every table (regardless of storage engine) has a maximum row size of 65,535 bytes. Storage engines may place additional constraints on this limit, reducing the effective maximum row size.

The maximum row size constrains the number (and possibly size) of columns because the total length of all columns cannot exceed this size. For example, `utf8` characters require up to three bytes per character, so for a `CHAR(255) CHARACTER SET utf8` column, the server must allocate $255 \times 3 = 765$ bytes per value. Consequently, a table cannot contain more than $65,535 / 765 = 85$ such columns.

Storage for variable-length columns includes length bytes, which are assessed against the row size. For example, a `VARCHAR(255) CHARACTER SET utf8` column takes two bytes to store the length of the value, so each value can take up to 767 bytes.

`BLOB` and `TEXT` columns count from one to four plus eight bytes each toward the row-size limit because their contents are stored separately from the rest of the row.

Declaring columns `NULL` can reduce the maximum number of columns permitted. For `MyISAM` tables, `NULL` columns require additional space in the row to record whether their values are `NULL`. Each `NULL` column takes one bit extra, rounded up to the nearest byte. The maximum row length in bytes can be calculated as follows:

```
row length = 1
             + (sum of column lengths)
             + (number of NULL columns + delete_flag + 7)/8
             + (number of variable-length columns)
```

`delete_flag` is 1 for tables with static row format. Static tables use a bit in the row record for a flag that indicates whether the row has been deleted. `delete_flag` is 0 for dynamic tables because the flag is stored in the dynamic row header. For information about `MyISAM` table formats, see [MyISAM Table Storage Formats](#).

For `InnoDB` tables, storage size is the same for `NULL` and `NOT NULL` columns, so the preceding calculations do not apply.

The following statement to create table `t1` succeeds because the columns require 32,765 + 2 bytes and 32,766 + 2 bytes, which falls within the maximum row size of 65,535 bytes:

```
mysql> CREATE TABLE t1
-> (c1 VARCHAR(32765) NOT NULL, c2 VARCHAR(32766) NOT NULL)
-> ENGINE = MyISAM CHARACTER SET latin1;
Query OK, 0 rows affected (0.02 sec)
```

The following statement to create table `t2` fails because the columns are `NULL` and `MyISAM` requires additional space that causes the row size to exceed 65,535 bytes:

```
mysql> CREATE TABLE t2
-> (c1 VARCHAR(32765) NULL, c2 VARCHAR(32766) NULL)
-> ENGINE = MyISAM CHARACTER SET latin1;
ERROR 1118 (42000): Row size too large. The maximum row size for the
used table type, not counting BLOBs, is 65535. You have to change some
columns to TEXT or BLOBs
```

The following statement to create table `t3` fails because, although the column length is within the maximum length of 65,535 bytes, two additional bytes are required to record the length, which causes the row size to exceed 65,535 bytes:

```
mysql> CREATE TABLE t3
-> (c1 VARCHAR(65535) NOT NULL)
-> ENGINE = MyISAM CHARACTER SET latin1;
ERROR 1118 (42000): Row size too large. The maximum row size for the
used table type, not counting BLOBs, is 65535. You have to change some
```

columns to TEXT or BLOBs

Reducing the column length to 65,533 or less permits the statement to succeed.

- Individual storage engines might impose additional restrictions that limit table column count. Examples:
 - **InnoDB** permits up to 1000 columns.
 - **InnoDB** restricts row size to slightly less than half of a database page for 4KB, 8KB, 16KB, and 32KB page sizes. For a page size of 64KB, **InnoDB** restricts row size to about 16000 bytes. Row size restrictions differ for variable-length columns (**VARBINARY**, **VARCHAR**, **BLOB**, and **TEXT**). For more information, see [Limits on InnoDB Tables](#).
 - Different **InnoDB** storage formats (**COMPRESSED**, **REDUNDANT**) use different amounts of page header and trailer data, which affects the amount of storage available for rows.

8.5 Limits Imposed by .frm File Structure

Each table has an **.frm** file that contains the table definition. The server uses the following expression to check some of the table information stored in the file against an upper limit of 64KB:

```
if (info_length+(ulong) create_fields.elements*FCOMP+288+
    n_length+int_length+com_length > 65535L || int_count > 255)
```

The portion of the information stored in the **.frm** file that is checked against the expression cannot grow beyond the 64KB limit, so if the table definition reaches this size, no more columns can be added.

The relevant factors in the expression are:

- **info_length** is space needed for “screens.” This is related to MySQL’s Unireg heritage.
- **create_fields.elements** is the number of columns.
- **FCOMP** is 17.
- **n_length** is the total length of all column names, including one byte per name as a separator.
- **int_length** is related to the list of values for **ENUM** and **SET** columns. In this context, “int” does not mean “integer.” It means “interval,” a term that refers collectively to **ENUM** and **SET** columns.
- **int_count** is the number of unique **ENUM** and **SET** definitions.
- **com_length** is the total length of column comments.

The expression just described has several implications for permitted table definitions:

- Using long column names can reduce the maximum number of columns, as can the inclusion of **ENUM** or **SET** columns, or use of column comments.
- A table can have no more than 255 unique **ENUM** and **SET** definitions. Columns with identical element lists are considered the same against this limit. For example, if a table contains these two columns, they count as one (not two) toward this limit because the definitions are identical:

```
e1 ENUM('a','b','c')
e2 ENUM('a','b','c')
```

- The sum of the length of element names in the unique **ENUM** and **SET** definitions counts toward the 64KB limit, so although the theoretical limit on number of elements in a given **ENUM** column is 65,535, the practical limit is less than 3000.

8.6 Windows Platform Limitations

The following limitations apply to use of MySQL on the Windows platform:

- **Process memory**

On Windows 32-bit platforms, it is not possible by default to use more than 2GB of RAM within a single process, including MySQL. This is because the physical address limit on Windows 32-bit is 4GB and the default setting within Windows is to split the virtual address space between kernel (2GB) and user/applications (2GB).

Some versions of Windows have a boot time setting to enable larger applications by reducing the kernel application. Alternatively, to use more than 2GB, use a 64-bit version of Windows.

- **File system aliases**

When using **MyISAM** tables, you cannot use aliases within Windows link to the data files on another volume and then link back to the main MySQL **datadir** location.

This facility is often used to move the data and index files to a RAID or other fast solution, while retaining the main **.frm** files in the default data directory configured with the **datadir** option.

- **Limited number of ports**

Windows systems have about 4,000 ports available for client connections, and after a connection on a port closes, it takes two to four minutes before the port can be reused. In situations where clients connect to and disconnect from the server at a high rate, it is possible for all available ports to be used up before closed ports become available again. If this happens, the MySQL server appears to be unresponsive even though it is running. Ports may be used by other applications running on the machine as well, in which case the number of ports available to MySQL is lower.

For more information about this problem, see <http://support.microsoft.com/default.aspx?scid=kb;en-us;196271>.

- **DATA DIRECTORY and INDEX DIRECTORY**

The **DATA DIRECTORY** option for **CREATE TABLE** is supported on Windows only for **InnoDB** tables, as described in [Creating a File-Per-Table Tablespace Outside the Data Directory](#). For **MyISAM** and other storage engines, the **DATA DIRECTORY** and **INDEX DIRECTORY** options for **CREATE TABLE** are ignored on Windows and any other platforms with a nonfunctional **realpath()** call.

- **DROP DATABASE**

You cannot drop a database that is in use by another session.

- **Case-insensitive names**

File names are not case sensitive on Windows, so MySQL database and table names are also not case sensitive on Windows. The only restriction is that database and table names must be specified using the same case throughout a given statement. See [Identifier Case Sensitivity](#).

- **Directory and file names**

On Windows, MySQL Server supports only directory and file names that are compatible with the current ANSI code pages. For example, the following Japanese directory name will not work in the Western locale (code page 1252):

```
datadir="C:/私たちのプロジェクトのデータ"
```

The same limitation applies to directory and file names referred to in SQL statements, such as the data file path name in **LOAD DATA INFILE**.

- **The “\” path name separator character**

Path name components in Windows are separated by the “\” character, which is also the escape character in MySQL. If you are using `LOAD DATA INFILE` or `SELECT ... INTO OUTFILE`, use Unix-style file names with “/” characters:

```
mysql> LOAD DATA INFILE 'C:/tmp/skr.txt' INTO TABLE skr;  
mysql> SELECT * INTO OUTFILE 'C:/tmp/skr.txt' FROM skr;
```

Alternatively, you must double the “\” character:

```
mysql> LOAD DATA INFILE 'C:\\tmp\\skr.txt' INTO TABLE skr;  
mysql> SELECT * INTO OUTFILE 'C:\\tmp\\skr.txt' FROM skr;
```

- **Problems with pipes**

Pipes do not work reliably from the Windows command-line prompt. If the pipe includes the character `^Z / CHAR(24)`, Windows thinks that it has encountered end-of-file and aborts the program.

This is mainly a problem when you try to apply a binary log as follows:

```
C:\> mysqlbinlog binary_log_file | mysql --user=root
```

If you have a problem applying the log and suspect that it is because of a `^Z / CHAR(24)` character, you can use the following workaround:

```
C:\> mysqlbinlog binary_log_file --result-file=/tmp/bin.sql  
C:\> mysql --user=root --execute "source /tmp/bin.sql"
```

The latter command also can be used to reliably read in any SQL file that may contain binary data.

Chapter 9 Restrictions and Limitations on Partitioning

Table of Contents

9.1 Partitioning Keys, Primary Keys, and Unique Keys	29
9.2 Partitioning Limitations Relating to Storage Engines	32
9.3 Partitioning Limitations Relating to Functions	33
9.4 Partitioning and Locking	34

This section discusses current restrictions and limitations on MySQL partitioning support.

Prohibited constructs. The following constructs are not permitted in partitioning expressions:

- Stored procedures, stored functions, UDFs, or plugins.
- Declared variables or user variables.

For a list of SQL functions which are permitted in partitioning expressions, see [Section 9.3, “Partitioning Limitations Relating to Functions”](#).

Arithmetic and logical operators. Use of the arithmetic operators `+`, `-`, and `*` is permitted in partitioning expressions. However, the result must be an integer value or `NULL` (except in the case of `[LINEAR] KEY` partitioning, as discussed elsewhere in this chapter; see [Partitioning Types](#), for more information).

The `DIV` operator is also supported, and the `/` operator is not permitted. (Bug #30188, Bug #33182)

The bit operators `|`, `&`, `^`, `<<`, `>>`, and `~` are not permitted in partitioning expressions.

HANDLER statements. Previously, the `HANDLER` statement was not supported with partitioned tables. This limitation is removed beginning with MySQL 5.7.1.

Server SQL mode. Tables employing user-defined partitioning do not preserve the SQL mode in effect at the time that they were created. As discussed in [Server SQL Modes](#), the results of many MySQL functions and operators may change according to the server SQL mode. Therefore, a change in the SQL mode at any time after the creation of partitioned tables may lead to major changes in the behavior of such tables, and could easily lead to corruption or loss of data. For these reasons, *it is strongly recommended that you never change the server SQL mode after creating partitioned tables*.

Examples. The following examples illustrate some changes in behavior of partitioned tables due to a change in the server SQL mode:

1. **Error handling.** Suppose that you create a partitioned table whose partitioning expression is one such as `column DIV 0` or `column MOD 0`, as shown here:

```
mysql> CREATE TABLE tn (c1 INT)
-> PARTITION BY LIST(1 DIV c1) (
-> PARTITION p0 VALUES IN (NULL),
-> PARTITION p1 VALUES IN (1)
-> );
Query OK, 0 rows affected (0.05 sec)
```

The default behavior for MySQL is to return `NULL` for the result of a division by zero, without producing any errors:

```
mysql> SELECT @@sql_mode;
+-----+
| @@sql_mode |
```

```

+-----+
|         |
+-----+
1 row in set (0.00 sec)
mysql> INSERT INTO tn VALUES (NULL), (0), (1);
Query OK, 3 rows affected (0.00 sec)
Records: 3  Duplicates: 0  Warnings: 0

```

However, changing the server SQL mode to treat division by zero as an error and to enforce strict error handling causes the same `INSERT` statement to fail, as shown here:

```

mysql> SET sql_mode='STRICT_ALL_TABLES,ERROR_FOR_DIVISION_BY_ZERO';
Query OK, 0 rows affected (0.00 sec)
mysql> INSERT INTO tn VALUES (NULL), (0), (1);
ERROR 1365 (22012): Division by 0

```

2. **Table accessibility.** Sometimes a change in the server SQL mode can make partitioned tables unusable. The following `CREATE TABLE` statement can be executed successfully only if the `NO_UNSIGNED_SUBTRACTION` mode is in effect:

```

mysql> SELECT @@sql_mode;
+-----+
| @@sql_mode |
+-----+
|             |
+-----+
1 row in set (0.00 sec)
mysql> CREATE TABLE tu (c1 BIGINT UNSIGNED)
-> PARTITION BY RANGE(c1 - 10) (
-> PARTITION p0 VALUES LESS THAN (-5),
-> PARTITION p1 VALUES LESS THAN (0),
-> PARTITION p2 VALUES LESS THAN (5),
-> PARTITION p3 VALUES LESS THAN (10),
-> PARTITION p4 VALUES LESS THAN (MAXVALUE)
-> );
ERROR 1563 (HY000): Partition constant is out of partition function domain
mysql> SET sql_mode='NO_UNSIGNED_SUBTRACTION';
Query OK, 0 rows affected (0.00 sec)
mysql> SELECT @@sql_mode;
+-----+
| @@sql_mode |
+-----+
| NO_UNSIGNED_SUBTRACTION |
+-----+
1 row in set (0.00 sec)
mysql> CREATE TABLE tu (c1 BIGINT UNSIGNED)
-> PARTITION BY RANGE(c1 - 10) (
-> PARTITION p0 VALUES LESS THAN (-5),
-> PARTITION p1 VALUES LESS THAN (0),
-> PARTITION p2 VALUES LESS THAN (5),
-> PARTITION p3 VALUES LESS THAN (10),
-> PARTITION p4 VALUES LESS THAN (MAXVALUE)
-> );
Query OK, 0 rows affected (0.05 sec)

```

If you remove the `NO_UNSIGNED_SUBTRACTION` server SQL mode after creating `tu`, you may no longer be able to access this table:

```

mysql> SET sql_mode='';
Query OK, 0 rows affected (0.00 sec)
mysql> SELECT * FROM tu;
ERROR 1563 (HY000): Partition constant is out of partition function domain
mysql> INSERT INTO tu VALUES (20);
ERROR 1563 (HY000): Partition constant is out of partition function domain

```

Server SQL modes also impact replication of partitioned tables. Differing SQL modes on master and slave can lead to partitioning expressions being evaluated differently; this can cause the distribution of

data among partitions to be different in the master's and slave's copies of a given table, and may even cause inserts into partitioned tables that succeed on the master to fail on the slave. For best results, you should always use the same server SQL mode on the master and on the slave.

Performance considerations. Some affects of partitioning operations on performance are given in the following list:

- **File system operations.** Partitioning and repartitioning operations (such as `ALTER TABLE` with `PARTITION BY ...`, `REORGANIZE PARTITIONS`, or `REMOVE PARTITIONING`) depend on file system operations for their implementation. This means that the speed of these operations is affected by such factors as file system type and characteristics, disk speed, swap space, file handling efficiency of the operating system, and MySQL server options and variables that relate to file handling. In particular, you should make sure that `large_files_support` is enabled and that `open_files_limit` is set properly. For partitioned tables using the `MyISAM` storage engine, increasing `myisam_max_sort_file_size` may improve performance; partitioning and repartitioning operations involving `InnoDB` tables may be made more efficient by enabling `innodb_file_per_table`.

See also [Maximum number of partitions](#).

- **MyISAM and partition file descriptor usage.** For a partitioned `MyISAM` table, MySQL uses 2 file descriptors for each partition, for each such table that is open. This means that you need many more file descriptors to perform operations on a partitioned `MyISAM` table than on a table which is identical to it except that the latter table is not partitioned, particularly when performing `ALTER TABLE` operations.

Assume a `MyISAM` table `t` with 100 partitions, such as the table created by this SQL statement:

```
CREATE TABLE t (c1 VARCHAR(50))
PARTITION BY KEY (c1) PARTITIONS 100
ENGINE=MYISAM;
```

Note

For brevity, we use `KEY` partitioning for the table shown in this example, but file descriptor usage as described here applies to all partitioned `MyISAM` tables, regardless of the type of partitioning that is employed. Partitioned tables using other storage engines such as `InnoDB` are not affected by this issue.

Now assume that you wish to repartition `t` so that it has 101 partitions, using the statement shown here:

```
ALTER TABLE t PARTITION BY KEY (c1) PARTITIONS 101;
```

To process this `ALTER TABLE` statement, MySQL uses 402 file descriptors—that is, two for each of the 100 original partitions, plus two for each of the 101 new partitions. This is because all partitions (old and new) must be opened concurrently during the reorganization of the table data. It is recommended that, if you expect to perform such operations, you should make sure that `--open-files-limit` is not set too low to accommodate them.

- **Table locks.** The process executing a partitioning operation on a table takes a write lock on the table. Reads from such tables are relatively unaffected; pending `INSERT` and `UPDATE` operations are performed as soon as the partitioning operation has completed.
- **Storage engine.** Partitioning operations, queries, and update operations generally tend to be faster with `MyISAM` tables than with `InnoDB` or `NDB` tables.
- **Indexes; partition pruning.** As with nonpartitioned tables, proper use of indexes can speed up queries on partitioned tables significantly. In addition, designing partitioned tables and queries

on these tables to take advantage of *partition pruning* can improve performance dramatically. See [Partition Pruning](#), for more information.

Previously, index condition pushdown was not supported for partitioned tables. This limitation was removed in MySQL 5.7.3. See [Index Condition Pushdown Optimization](#).

- **Performance with LOAD DATA.** In MySQL 5.7, [LOAD DATA](#) uses buffering to improve performance. You should be aware that the buffer uses 130 KB memory per partition to achieve this.

Maximum number of partitions.

The maximum possible number of partitions for a given table not using the [NDB](#) storage engine is 8192. This number includes subpartitions.

The maximum possible number of user-defined partitions for a table using the [NDB](#) storage engine is determined according to the version of the MySQL Cluster software being used, the number of data nodes, and other factors. See [NDB and user-defined partitioning](#), for more information.

If, when creating tables with a large number of partitions (but less than the maximum), you encounter an error message such as `Got error ... from storage engine: Out of resources when opening file`, you may be able to address the issue by increasing the value of the [open_files_limit](#) system variable. However, this is dependent on the operating system, and may not be possible or advisable on all platforms; see [File Not Found and Similar Errors](#), for more information. In some cases, using large numbers (hundreds) of partitions may also not be advisable due to other concerns, so using more partitions does not automatically lead to better results.

See also [File system operations](#).

Query cache not supported.

The query cache is not supported for partitioned tables, and is automatically disabled for queries involving partitioned tables. The query cache cannot be enabled for such queries.

Per-partition key caches.

In MySQL 5.7, key caches are supported for partitioned [MyISAM](#) tables, using the [CACHE INDEX](#) and [LOAD INDEX INTO CACHE](#) statements. Key caches may be defined for one, several, or all partitions, and indexes for one, several, or all partitions may be preloaded into key caches.

Foreign keys not supported for partitioned InnoDB tables.

Partitioned tables using the [InnoDB](#) storage engine do not support foreign keys. More specifically, this means that the following two statements are true:

1. No definition of an [InnoDB](#) table employing user-defined partitioning may contain foreign key references; no [InnoDB](#) table whose definition contains foreign key references may be partitioned.
2. No [InnoDB](#) table definition may contain a foreign key reference to a user-partitioned table; no [InnoDB](#) table with user-defined partitioning may contain columns referenced by foreign keys.

The scope of the restrictions just listed includes all tables that use the [InnoDB](#) storage engine. [CREATE TABLE](#) and [ALTER TABLE](#) statements that would result in tables violating these restrictions are not allowed.

ALTER TABLE ... ORDER BY. An [ALTER TABLE ... ORDER BY column](#) statement run against a partitioned table causes ordering of rows only within each partition.

Effects on REPLACE statements by modification of primary keys. It can be desirable in some cases (see [Section 9.1, “Partitioning Keys, Primary Keys, and Unique Keys”](#)) to modify a table's primary key. Be aware that, if your application uses [REPLACE](#) statements and you do this, the results of these statements can be drastically altered. See [REPLACE Syntax](#), for more information and an example.

FULLTEXT indexes.

Partitioned tables do not support [FULLTEXT](#) indexes or searches, even for partitioned tables employing the [InnoDB](#) or [MyISAM](#) storage engine.

Spatial columns. Columns with spatial data types such as [POINT](#) or [GEOMETRY](#) cannot be used in partitioned tables.

Temporary tables.

Temporary tables cannot be partitioned. (Bug #17497)

Log tables. It is not possible to partition the log tables; an [ALTER TABLE ... PARTITION BY ...](#) statement on such a table fails with an error.

Data type of partitioning key.

A partitioning key must be either an integer column or an expression that resolves to an integer. Expressions employing [ENUM](#) columns cannot be used. The column or expression value may also be [NULL](#). (See [How MySQL Partitioning Handles NULL](#).)

There are two exceptions to this restriction:

1. When partitioning by [\[LINEAR\] KEY](#), it is possible to use columns of any valid MySQL data type other than [TEXT](#) or [BLOB](#) as partitioning keys, because MySQL's internal key-hashing functions produce the correct data type from these types. For example, the following two [CREATE TABLE](#) statements are valid:

```
CREATE TABLE tkc (c1 CHAR)
PARTITION BY KEY(c1)
PARTITIONS 4;
CREATE TABLE tke
( c1 ENUM('red', 'orange', 'yellow', 'green', 'blue', 'indigo', 'violet') )
PARTITION BY LINEAR KEY(c1)
PARTITIONS 6;
```

2. When partitioning by [RANGE COLUMNS](#) or [LIST COLUMNS](#), it is possible to use string, [DATE](#), and [DATETIME](#) columns. For example, each of the following [CREATE TABLE](#) statements is valid:

```
CREATE TABLE rc (c1 INT, c2 DATE)
PARTITION BY RANGE COLUMNS(c2) (
    PARTITION p0 VALUES LESS THAN('1990-01-01'),
    PARTITION p1 VALUES LESS THAN('1995-01-01'),
    PARTITION p2 VALUES LESS THAN('2000-01-01'),
    PARTITION p3 VALUES LESS THAN('2005-01-01'),
    PARTITION p4 VALUES LESS THAN(MAXVALUE)
);
CREATE TABLE lc (c1 INT, c2 CHAR(1))
PARTITION BY LIST COLUMNS(c2) (
    PARTITION p0 VALUES IN('a', 'd', 'g', 'j', 'm', 'p', 's', 'v', 'y'),
    PARTITION p1 VALUES IN('b', 'e', 'h', 'k', 'n', 'q', 't', 'w', 'z'),
    PARTITION p2 VALUES IN('c', 'f', 'i', 'l', 'o', 'r', 'u', 'x', NULL)
);
```

Neither of the preceding exceptions applies to [BLOB](#) or [TEXT](#) column types.

Subqueries.

A partitioning key may not be a subquery, even if that subquery resolves to an integer value or [NULL](#).

Issues with subpartitions.

Subpartitions must use [HASH](#) or [KEY](#) partitioning. Only [RANGE](#) and [LIST](#) partitions may be subpartitioned; [HASH](#) and [KEY](#) partitions cannot be subpartitioned.

[SUBPARTITION BY KEY](#) requires that the subpartitioning column or columns be specified explicitly, unlike the case with [PARTITION BY KEY](#), where it can be omitted (in which case the table's primary key column is used by default). Consider the table created by this statement:

```
CREATE TABLE ts (
    id INT NOT NULL AUTO_INCREMENT PRIMARY KEY,
```

```
name VARCHAR(30)
);
```

You can create a table having the same columns, partitioned by [KEY](#), using a statement such as this one:

```
CREATE TABLE ts (
  id INT NOT NULL AUTO_INCREMENT PRIMARY KEY,
  name VARCHAR(30)
)
PARTITION BY KEY()
PARTITIONS 4;
```

The previous statement is treated as though it had been written like this, with the table's primary key column used as the partitioning column:

```
CREATE TABLE ts (
  id INT NOT NULL AUTO_INCREMENT PRIMARY KEY,
  name VARCHAR(30)
)
PARTITION BY KEY(id)
PARTITIONS 4;
```

However, the following statement that attempts to create a subpartitioned table using the default column as the subpartitioning column fails, and the column must be specified for the statement to succeed, as shown here:

```
mysql> CREATE TABLE ts (
->   id INT NOT NULL AUTO_INCREMENT PRIMARY KEY,
->   name VARCHAR(30)
-> )
-> PARTITION BY RANGE(id)
-> SUBPARTITION BY KEY()
-> SUBPARTITIONS 4
-> (
->   PARTITION p0 VALUES LESS THAN (100),
->   PARTITION p1 VALUES LESS THAN (MAXVALUE)
-> );
ERROR 1064 (42000): You have an error in your SQL syntax; check the manual that
corresponds to your MySQL server version for the right syntax to use near ')'
mysql> CREATE TABLE ts (
->   id INT NOT NULL AUTO_INCREMENT PRIMARY KEY,
->   name VARCHAR(30)
-> )
-> PARTITION BY RANGE(id)
-> SUBPARTITION BY KEY(id)
-> SUBPARTITIONS 4
-> (
->   PARTITION p0 VALUES LESS THAN (100),
->   PARTITION p1 VALUES LESS THAN (MAXVALUE)
-> );
Query OK, 0 rows affected (0.07 sec)
```

This is a known issue (see Bug #51470).

DATA DIRECTORY and INDEX DIRECTORY options. [DATA DIRECTORY](#) and [INDEX DIRECTORY](#) are subject to the following restrictions when used with partitioned tables:

- Table-level [DATA DIRECTORY](#) and [INDEX DIRECTORY](#) options are ignored (see Bug #32091).
- On Windows, the [DATA DIRECTORY](#) and [INDEX DIRECTORY](#) options are not supported for individual partitions or subpartitions of [MyISAM](#) tables. However, you can use [DATA DIRECTORY](#) for individual partitions or subpartitions of [InnoDB](#) tables.

Repairing and rebuilding partitioned tables. The statements [CHECK TABLE](#), [OPTIMIZE TABLE](#), [ANALYZE TABLE](#), and [REPAIR TABLE](#) are supported for partitioned tables.

In addition, you can use [ALTER TABLE ... REBUILD PARTITION](#) to rebuild one or more partitions of a partitioned table; [ALTER TABLE ... REORGANIZE PARTITION](#) also causes partitions to be rebuilt. See [ALTER TABLE Syntax](#), for more information about these two statements.

Starting in MySQL 5.7.2, [ANALYZE](#), [CHECK](#), [OPTIMIZE](#), [REPAIR](#), and [TRUNCATE](#) operations are supported with subpartitions. [REBUILD](#) was also accepted syntax prior to MySQL 5.7.5, although this had no effect. (Bug #19075411, Bug #73130) See also [ALTER TABLE Partition Operations](#).

[mysqlcheck](#), [myisamchk](#), and [myisampack](#) are not supported with partitioned tables.

FOR EXPORT option (FLUSH TABLES). The [FLUSH TABLES](#) statement's [FOR EXPORT](#) option is not supported for partitioned [InnoDB](#) tables in MySQL 5.7.4 and earlier. (Bug #16943907)

9.1 Partitioning Keys, Primary Keys, and Unique Keys

This section discusses the relationship of partitioning keys with primary keys and unique keys. The rule governing this relationship can be expressed as follows: All columns used in the partitioning expression for a partitioned table must be part of every unique key that the table may have.

In other words, *every unique key on the table must use every column in the table's partitioning expression*. (This also includes the table's primary key, since it is by definition a unique key. This particular case is discussed later in this section.) For example, each of the following table creation statements is invalid:

```
CREATE TABLE t1 (
  col1 INT NOT NULL,
  col2 DATE NOT NULL,
  col3 INT NOT NULL,
  col4 INT NOT NULL,
  UNIQUE KEY (col1, col2)
)
PARTITION BY HASH(col3)
PARTITIONS 4;
CREATE TABLE t2 (
  col1 INT NOT NULL,
  col2 DATE NOT NULL,
  col3 INT NOT NULL,
  col4 INT NOT NULL,
  UNIQUE KEY (col1),
  UNIQUE KEY (col3)
)
PARTITION BY HASH(col1 + col3)
PARTITIONS 4;
```

In each case, the proposed table would have at least one unique key that does not include all columns used in the partitioning expression.

Each of the following statements is valid, and represents one way in which the corresponding invalid table creation statement could be made to work:

```
CREATE TABLE t1 (
  col1 INT NOT NULL,
  col2 DATE NOT NULL,
  col3 INT NOT NULL,
  col4 INT NOT NULL,
  UNIQUE KEY (col1, col2, col3)
)
PARTITION BY HASH(col3)
PARTITIONS 4;
CREATE TABLE t2 (
  col1 INT NOT NULL,
```

```
col2 DATE NOT NULL,
col3 INT NOT NULL,
col4 INT NOT NULL,
UNIQUE KEY (col1, col3)
)
PARTITION BY HASH(col1 + col3)
PARTITIONS 4;
```

This example shows the error produced in such cases:

```
mysql> CREATE TABLE t3 (
->   col1 INT NOT NULL,
->   col2 DATE NOT NULL,
->   col3 INT NOT NULL,
->   col4 INT NOT NULL,
->   UNIQUE KEY (col1, col2),
->   UNIQUE KEY (col3)
-> )
-> PARTITION BY HASH(col1 + col3)
-> PARTITIONS 4;
ERROR 1491 (HY000): A PRIMARY KEY must include all columns in the table's partitioning function
```

The `CREATE TABLE` statement fails because both `col1` and `col3` are included in the proposed partitioning key, but neither of these columns is part of both of unique keys on the table. This shows one possible fix for the invalid table definition:

```
mysql> CREATE TABLE t3 (
->   col1 INT NOT NULL,
->   col2 DATE NOT NULL,
->   col3 INT NOT NULL,
->   col4 INT NOT NULL,
->   UNIQUE KEY (col1, col2, col3),
->   UNIQUE KEY (col3)
-> )
-> PARTITION BY HASH(col3)
-> PARTITIONS 4;
Query OK, 0 rows affected (0.05 sec)
```

In this case, the proposed partitioning key `col3` is part of both unique keys, and the table creation statement succeeds.

The following table cannot be partitioned at all, because there is no way to include in a partitioning key any columns that belong to both unique keys:

```
CREATE TABLE t4 (
  col1 INT NOT NULL,
  col2 INT NOT NULL,
  col3 INT NOT NULL,
  col4 INT NOT NULL,
  UNIQUE KEY (col1, col3),
  UNIQUE KEY (col2, col4)
);
```

Since every primary key is by definition a unique key, this restriction also includes the table's primary key, if it has one. For example, the next two statements are invalid:

```
CREATE TABLE t5 (
  col1 INT NOT NULL,
  col2 DATE NOT NULL,
  col3 INT NOT NULL,
  col4 INT NOT NULL,
  PRIMARY KEY(col1, col2)
)
PARTITION BY HASH(col3)
PARTITIONS 4;
CREATE TABLE t6 (
```

```

col1 INT NOT NULL,
col2 DATE NOT NULL,
col3 INT NOT NULL,
col4 INT NOT NULL,
PRIMARY KEY(col1, col3),
UNIQUE KEY(col2)
)
PARTITION BY HASH( YEAR(col2) )
PARTITIONS 4;

```

In both cases, the primary key does not include all columns referenced in the partitioning expression. However, both of the next two statements are valid:

```

CREATE TABLE t7 (
  col1 INT NOT NULL,
  col2 DATE NOT NULL,
  col3 INT NOT NULL,
  col4 INT NOT NULL,
  PRIMARY KEY(col1, col2)
)
PARTITION BY HASH(col1 + YEAR(col2))
PARTITIONS 4;
CREATE TABLE t8 (
  col1 INT NOT NULL,
  col2 DATE NOT NULL,
  col3 INT NOT NULL,
  col4 INT NOT NULL,
  PRIMARY KEY(col1, col2, col4),
  UNIQUE KEY(col2, col1)
)
PARTITION BY HASH(col1 + YEAR(col2))
PARTITIONS 4;

```

If a table has no unique keys—this includes having no primary key—then this restriction does not apply, and you may use any column or columns in the partitioning expression as long as the column type is compatible with the partitioning type.

For the same reason, you cannot later add a unique key to a partitioned table unless the key includes all columns used by the table's partitioning expression. Consider the partitioned table created as shown here:

```

mysql> CREATE TABLE t_no_pk (c1 INT, c2 INT)
-> PARTITION BY RANGE(c1) (
-> PARTITION p0 VALUES LESS THAN (10),
-> PARTITION p1 VALUES LESS THAN (20),
-> PARTITION p2 VALUES LESS THAN (30),
-> PARTITION p3 VALUES LESS THAN (40)
-> );
Query OK, 0 rows affected (0.12 sec)

```

It is possible to add a primary key to `t_no_pk` using either of these `ALTER TABLE` statements:

```

# possible PK
mysql> ALTER TABLE t_no_pk ADD PRIMARY KEY(c1);
Query OK, 0 rows affected (0.13 sec)
Records: 0 Duplicates: 0 Warnings: 0
# drop this PK
mysql> ALTER TABLE t_no_pk DROP PRIMARY KEY;
Query OK, 0 rows affected (0.10 sec)
Records: 0 Duplicates: 0 Warnings: 0
# use another possible PK
mysql> ALTER TABLE t_no_pk ADD PRIMARY KEY(c1, c2);
Query OK, 0 rows affected (0.12 sec)
Records: 0 Duplicates: 0 Warnings: 0
# drop this PK
mysql> ALTER TABLE t_no_pk DROP PRIMARY KEY;
Query OK, 0 rows affected (0.09 sec)

```

```
Records: 0 Duplicates: 0 Warnings: 0
```

However, the next statement fails, because `c1` is part of the partitioning key, but is not part of the proposed primary key:

```
# fails with error 1503
mysql> ALTER TABLE t_no_pk ADD PRIMARY KEY(c2);
ERROR 1503 (HY000): A PRIMARY KEY must include all columns in the table's partitioning function
```

Since `t_no_pk` has only `c1` in its partitioning expression, attempting to adding a unique key on `c2` alone fails. However, you can add a unique key that uses both `c1` and `c2`.

These rules also apply to existing nonpartitioned tables that you wish to partition using `ALTER TABLE ... PARTITION BY`. Consider a table `np_pk` created as shown here:

```
mysql> CREATE TABLE np_pk (
->   id INT NOT NULL AUTO_INCREMENT,
->   name VARCHAR(50),
->   added DATE,
->   PRIMARY KEY (id)
-> );
Query OK, 0 rows affected (0.08 sec)
```

The following `ALTER TABLE` statement fails with an error, because the `added` column is not part of any unique key in the table:

```
mysql> ALTER TABLE np_pk
->   PARTITION BY HASH( TO_DAYS(added) )
->   PARTITIONS 4;
ERROR 1503 (HY000): A PRIMARY KEY must include all columns in the table's partitioning function
```

However, this statement using the `id` column for the partitioning column is valid, as shown here:

```
mysql> ALTER TABLE np_pk
->   PARTITION BY HASH(id)
->   PARTITIONS 4;
Query OK, 0 rows affected (0.11 sec)
Records: 0 Duplicates: 0 Warnings: 0
```

In the case of `np_pk`, the only column that may be used as part of a partitioning expression is `id`; if you wish to partition this table using any other column or columns in the partitioning expression, you must first modify the table, either by adding the desired column or columns to the primary key, or by dropping the primary key altogether.

9.2 Partitioning Limitations Relating to Storage Engines

The following limitations apply to the use of storage engines with user-defined partitioning of tables.

MERGE storage engine. User-defined partitioning and the `MERGE` storage engine are not compatible. Tables using the `MERGE` storage engine cannot be partitioned. Partitioned tables cannot be merged.

FEDERATED storage engine. Partitioning of `FEDERATED` tables is not supported; it is not possible to create partitioned `FEDERATED` tables.

CSV storage engine. Partitioned tables using the `CSV` storage engine are not supported; it is not possible to create partitioned `CSV` tables.

InnoDB storage engine. `InnoDB` foreign keys and MySQL partitioning are not compatible. Partitioned `InnoDB` tables cannot have foreign key references, nor can they have columns referenced by foreign keys. `InnoDB` tables which have or which are referenced by foreign keys cannot be partitioned.

In addition, `ALTER TABLE ... OPTIMIZE PARTITION` does not work correctly with partitioned tables that use the InnoDB storage engine. Use `ALTER TABLE ... REBUILD PARTITION` and `ALTER TABLE ... ANALYZE PARTITION`, instead, for such tables. For more information, see [ALTER TABLE Partition Operations](#).

User-defined partitioning and the NDB storage engine (MySQL Cluster). Partitioning by `KEY` (including `LINEAR KEY`) is the only type of partitioning supported for the NDB storage engine. It is not possible under normal circumstances in MySQL Cluster to create a MySQL Cluster table using any partitioning type other than `[LINEAR] KEY`, and attempting to do so fails with an error.

Exception (not for production): It is possible to override this restriction by setting the `new` system variable on MySQL Cluster SQL nodes to `ON`. If you choose to do this, you should be aware that tables using partitioning types other than `[LINEAR] KEY` are not supported in production. *In such cases, you can create and use tables with partitioning types other than `KEY` or `LINEAR KEY`, but you do this entirely at your own risk.*

The maximum number of partitions that can be defined for an NDB table depends on the number of data nodes and node groups in the cluster, the version of the MySQL Cluster software in use, and other factors. See [NDB and user-defined partitioning](#), for more information.

As of MySQL Cluster NDB 7.5.2, the maximum amount of fixed-size data that can be stored per partition in an NDB table is 128 TB. Previously, this was 16 GB.

`CREATE TABLE` and `ALTER TABLE` statements that would cause a user-partitioned NDB table not to meet either or both of the following two requirements are not permitted, and fail with an error:

1. The table must have an explicit primary key.
2. All columns listed in the table's partitioning expression must be part of the primary key.

Exception. If a user-partitioned NDB table is created using an empty column-list (that is, using `PARTITION BY KEY()` or `PARTITION BY LINEAR KEY()`), then no explicit primary key is required.

Upgrading partitioned tables. When performing an upgrade, tables which are partitioned by `KEY` and which use any storage engine other than NDB must be dumped and reloaded.

Same storage engine for all partitions. All partitions of a partitioned table must use the same storage engine and it must be the same storage engine used by the table as a whole. In addition, if one does not specify an engine on the table level, then one must do either of the following when creating or altering a partitioned table:

- Do *not* specify any engine for *any* partition or subpartition
- Specify the engine for *all* partitions or subpartitions

9.3 Partitioning Limitations Relating to Functions

This section discusses limitations in MySQL Partitioning relating specifically to functions used in partitioning expressions.

Only the MySQL functions shown in the following table are allowed in partitioning expressions.

<code>ABS()</code>	<code>CEILING()</code> (see CEILING() and FLOOR())	<code>DAY()</code>
<code>DAYOFMONTH()</code>	<code>DAYOFWEEK()</code>	<code>DAYOFYEAR()</code>
<code>DATEDIFF()</code>	<code>EXTRACT()</code> (see EXTRACT() function with <code>WEEK</code> specifier)	<code>FLOOR()</code> (see CEILING() and FLOOR())
<code>HOURL()</code>	<code>MICROSECOND()</code>	<code>MINUTE()</code>
<code>MOD()</code>	<code>MONTH()</code>	<code>QUARTER()</code>

SECOND()	TIME_TO_SEC()	TO_DAYS()
TO_SECONDS()	UNIX_TIMESTAMP() (with TIMESTAMP columns)	WEEKDAY()
YEAR()		YEARWEEK()

In MySQL 5.7, partition pruning is supported for the [TO_DAYS\(\)](#), [TO_SECONDS\(\)](#), [YEAR\(\)](#), and [UNIX_TIMESTAMP\(\)](#) functions. See [Partition Pruning](#), for more information.

CEILING() and FLOOR(). Each of these functions returns an integer only if it is passed an argument of an exact numeric type, such as one of the [INT](#) types or [DECIMAL](#). This means, for example, that the following [CREATE TABLE](#) statement fails with an error, as shown here:

```
mysql> CREATE TABLE t (c FLOAT) PARTITION BY LIST( FLOOR(c) )(
->     PARTITION p0 VALUES IN (1,3,5),
->     PARTITION p1 VALUES IN (2,4,6)
-> );
ERROR 1490 (HY000): The PARTITION function returns the wrong type
```

EXTRACT() function with WEEK specifier. The value returned by the [EXTRACT\(\)](#) function, when used as [EXTRACT\(WEEK FROM col\)](#), depends on the value of the [default_week_format](#) system variable. For this reason, [EXTRACT\(\)](#) is not permitted as a partitioning function when it specifies the unit as [WEEK](#). (Bug #54483)

See [Mathematical Functions](#), for more information about the return types of these functions, as well as [Numeric Types](#).

9.4 Partitioning and Locking

For storage engines such as [MyISAM](#) that actually execute table-level locks when executing DML or DDL statements, such a statement in older versions of MySQL (5.6.5 and earlier) that affected a partitioned table imposed a lock on the table as a whole; that is, all partitions were locked until the statement was finished. In MySQL 5.7, *partition lock pruning* eliminates unneeded locks in many cases, and most statements reading from or updating a partitioned [MyISAM](#) table cause only the effected partitions to be locked. For example, a [SELECT](#) from a partitioned [MyISAM](#) table locks only those partitions actually containing rows that satisfy the [SELECT](#) statement's [WHERE](#) condition are locked.

For statements effecting partitioned tables using storage engines such as [InnoDB](#), that employ row-level locking and do not actually perform (or need to perform) the locks prior to partition pruning, this is not an issue.

The next few paragraphs discuss the effects of partition lock pruning for various MySQL statements on tables using storage engines that employ table-level locks.

Effects on DML statements

[SELECT](#) statements (including those containing unions or joins) lock only those partitions that actually need to be read. This also applies to [SELECT ... PARTITION](#).

An [UPDATE](#) prunes locks only for tables on which no partitioning columns are updated.

[REPLACE](#) and [INSERT](#) lock only those partitions having rows to be inserted or replaced. However, if an [AUTO_INCREMENT](#) value is generated for any partitioning column then all partitions are locked.

[INSERT ... ON DUPLICATE KEY UPDATE](#) is pruned as long as no partitioning column is updated.

[INSERT ... SELECT](#) locks only those partitions in the source table that need to be read, although all partitions in the target table are locked.

Locks imposed by [LOAD DATA](#) statements on partitioned tables cannot be pruned.

The presence of `BEFORE INSERT` or `BEFORE UPDATE` triggers using any partitioning column of a partitioned table means that locks on `INSERT` and `UPDATE` statements updating this table cannot be pruned, since the trigger can alter its values: A `BEFORE INSERT` trigger on any of the table's partitioning columns means that locks set by `INSERT` or `REPLACE` cannot be pruned, since the `BEFORE INSERT` trigger may change a row's partitioning columns before the row is inserted, forcing the row into a different partition than it would be otherwise. A `BEFORE UPDATE` trigger on a partitioning column means that locks imposed by `UPDATE` or `INSERT ... ON DUPLICATE KEY UPDATE` cannot be pruned.

Affected DDL statements

`CREATE VIEW` does not cause any locks.

`ALTER TABLE ... EXCHANGE PARTITION` prunes locks; only the exchanged table and the exchanged partition are locked.

`ALTER TABLE ... TRUNCATE PARTITION` prunes locks; only the partitions to be emptied are locked.

In addition, `ALTER TABLE` statements take metadata locks on the table level.

Other statements

`LOCK TABLES` cannot prune partition locks.

`CALL stored_procedure(expr)` supports lock pruning, but evaluating *expr* does not.

`DO` and `SET` statements do not support partitioning lock pruning.

Chapter 10 MySQL Differences from Standard SQL

Table of Contents

10.1 SELECT INTO TABLE Differences	37
10.2 UPDATE Differences	37
10.3 Foreign Key Differences	37
10.4 '--' as the Start of a Comment	38

We try to make MySQL Server follow the ANSI SQL standard and the ODBC SQL standard, but MySQL Server performs operations differently in some cases:

- There are several differences between the MySQL and standard SQL privilege systems. For example, in MySQL, privileges for a table are not automatically revoked when you delete a table. You must explicitly issue a [REVOKE](#) statement to revoke privileges for a table. For more information, see [REVOKE Syntax](#).
- The [CAST\(\)](#) function does not support cast to [REAL](#) or [BIGINT](#). See [Cast Functions and Operators](#).

10.1 SELECT INTO TABLE Differences

MySQL Server doesn't support the [SELECT ... INTO TABLE](#) Sybase SQL extension. Instead, MySQL Server supports the [INSERT INTO ... SELECT](#) standard SQL syntax, which is basically the same thing. See [INSERT ... SELECT Syntax](#). For example:

```
INSERT INTO tbl_temp2 (fld_id)
  SELECT tbl_temp1.fld_order_id
  FROM tbl_temp1 WHERE tbl_temp1.fld_order_id > 100;
```

Alternatively, you can use [SELECT ... INTO OUTFILE](#) or [CREATE TABLE ... SELECT](#).

You can use [SELECT ... INTO](#) with user-defined variables. The same syntax can also be used inside stored routines using cursors and local variables. See [SELECT ... INTO Syntax](#).

10.2 UPDATE Differences

If you access a column from the table to be updated in an expression, [UPDATE](#) uses the current value of the column. The second assignment in the following statement sets [col2](#) to the current (updated) [col1](#) value, not the original [col1](#) value. The result is that [col1](#) and [col2](#) have the same value. This behavior differs from standard SQL.

```
UPDATE t1 SET col1 = col1 + 1, col2 = col1;
```

10.3 Foreign Key Differences

MySQL's implementation of foreign keys differs from the SQL standard in the following key respects:

- If there are several rows in the parent table that have the same referenced key value, [InnoDB](#) acts in foreign key checks as if the other parent rows with the same key value do not exist. For example, if you have defined a [RESTRICT](#) type constraint, and there is a child row with several parent rows, [InnoDB](#) does not permit the deletion of any of those parent rows.

[InnoDB](#) performs cascading operations through a depth-first algorithm, based on records in the indexes corresponding to the foreign key constraints.

- A **FOREIGN KEY** constraint that references a non-**UNIQUE** key is not standard SQL but rather an **InnoDB** extension.
- If **ON UPDATE CASCADE** or **ON UPDATE SET NULL** recurses to update the *same table* it has previously updated during the same cascade, it acts like **RESTRICT**. This means that you cannot use self-referential **ON UPDATE CASCADE** or **ON UPDATE SET NULL** operations. This is to prevent infinite loops resulting from cascaded updates. A self-referential **ON DELETE SET NULL**, on the other hand, is possible, as is a self-referential **ON DELETE CASCADE**. Cascading operations may not be nested more than 15 levels deep.
- In an SQL statement that inserts, deletes, or updates many rows, foreign key constraints (like unique constraints) are checked row-by-row. When performing foreign key checks, **InnoDB** sets shared row-level locks on child or parent records that it must examine. MySQL checks foreign key constraints immediately; the check is not deferred to transaction commit. According to the SQL standard, the default behavior should be deferred checking. That is, constraints are only checked after the *entire SQL statement* has been processed. This means that it is not possible to delete a row that refers to itself using a foreign key.

For information about how the **InnoDB** storage engine handles foreign keys, see [InnoDB and FOREIGN KEY Constraints](#).

10.4 '--' as the Start of a Comment

Standard SQL uses the C syntax `/* this is a comment */` for comments, and MySQL Server supports this syntax as well. MySQL also support extensions to this syntax that enable MySQL-specific SQL to be embedded in the comment, as described in [Comment Syntax](#).

Standard SQL uses “--” as a start-comment sequence. MySQL Server uses “#” as the start comment character. MySQL Server 3.23.3 and up also supports a variant of the “--” comment style. That is, the “--” start-comment sequence must be followed by a space (or by a control character such as a newline). The space is required to prevent problems with automatically generated SQL queries that use constructs such as the following, where we automatically insert the value of the payment for **payment**:

```
UPDATE account SET credit=credit-payment
```

Consider about what happens if **payment** has a negative value such as **-1**:

```
UPDATE account SET credit=credit--1
```

credit--1 is a valid expression in SQL, but “--” is interpreted as the start of a comment, part of the expression is discarded. The result is a statement that has a completely different meaning than intended:

```
UPDATE account SET credit=credit
```

The statement produces no change in value at all. This illustrates that permitting comments to start with “--” can have serious consequences.

Using our implementation requires a space following the “--” for it to be recognized as a start-comment sequence in MySQL Server 3.23.3 and higher. Therefore, **credit--1** is safe to use.

Another safe feature is that the **mysql** command-line client ignores lines that start with “--”.

The following information is relevant only if you are running a MySQL version earlier than 3.23.3:

If you have an SQL script in a text file that contains “--” comments, you should use the **replace** utility as follows to convert the comments to use “#” characters before executing the script:

```
shell> replace " --" " #" < text-file-with-funny-comments.sql \  
| mysql db_name
```

That is safer than executing the script in the usual way:

```
shell> mysql db_name < text-file-with-funny-comments.sql
```

You can also edit the script file “in place” to change the “--” comments to “#” comments:

```
shell> replace " --" " #" -- text-file-with-funny-comments.sql
```

Change them back with this command:

```
shell> replace " #" " --" -- text-file-with-funny-comments.sql
```

See [replace — A String-Replacement Utility](#).

Chapter 11 Known Issues in MySQL

This section lists known issues in recent versions of MySQL.

For information about platform-specific issues, see the installation and porting instructions in [General Installation Guidance](#), and [Debugging and Porting MySQL](#).

The following problems are known:

- Subquery optimization for `IN` is not as effective as for `=`.
- Even if you use `lower_case_table_names=2` (which enables MySQL to remember the case used for databases and table names), MySQL does not remember the case used for database names for the function `DATABASE()` or within the various logs (on case-insensitive systems).
- Dropping a `FOREIGN KEY` constraint does not work in replication because the constraint may have another name on the slave.
- `REPLACE` (and `LOAD DATA` with the `REPLACE` option) does not trigger `ON DELETE CASCADE`.
- `DISTINCT` with `ORDER BY` does not work inside `GROUP_CONCAT()` if you do not use all and only those columns that are in the `DISTINCT` list.
- When inserting a big integer value (between 2^{63} and $2^{64}-1$) into a decimal or string column, it is inserted as a negative value because the number is evaluated in a signed integer context.
- With statement-based binary logging, the master writes the executed queries to the binary log. This is a very fast, compact, and efficient logging method that works perfectly in most cases. However, it is possible for the data on the master and slave to become different if a query is designed in such a way that the data modification is nondeterministic (generally not a recommended practice, even outside of replication).

For example:

- `CREATE TABLE ... SELECT` or `INSERT ... SELECT` statements that insert zero or `NULL` values into an `AUTO_INCREMENT` column.
- `DELETE` if you are deleting rows from a table that has foreign keys with `ON DELETE CASCADE` properties.
- `REPLACE ... SELECT`, `INSERT IGNORE ... SELECT` if you have duplicate key values in the inserted data.

If and only if the preceding queries have no `ORDER BY` clause guaranteeing a deterministic order.

For example, for `INSERT ... SELECT` with no `ORDER BY`, the `SELECT` may return rows in a different order (which results in a row having different ranks, hence getting a different number in the `AUTO_INCREMENT` column), depending on the choices made by the optimizers on the master and slave.

A query is optimized differently on the master and slave only if:

- The table is stored using a different storage engine on the master than on the slave. (It is possible to use different storage engines on the master and slave. For example, you can use `InnoDB` on the master, but `MyISAM` on the slave if the slave has less available disk space.)
- MySQL buffer sizes (`key_buffer_size`, and so on) are different on the master and slave.
- The master and slave run different MySQL versions, and the optimizer code differs between these versions.

This problem may also affect database restoration using `mysqlbinlog|mysql`.

The easiest way to avoid this problem is to add an `ORDER BY` clause to the aforementioned nondeterministic queries to ensure that the rows are always stored or modified in the same order. Using row-based or mixed logging format also avoids the problem.

- Log file names are based on the server host name if you do not specify a file name with the startup option. To retain the same log file names if you change your host name to something else, you must explicitly use options such as `--log-bin=old_host_name-bin`. See [Server Command Options](#). Alternatively, rename the old files to reflect your host name change. If these are binary logs, you must edit the binary log index file and fix the binary log file names there as well. (The same is true for the relay logs on a slave server.)
- `mysqlbinlog` does not delete temporary files left after a `LOAD DATA INFILE` statement. See [mysqlbinlog — Utility for Processing Binary Log Files](#).
- `RENAME` does not work with `TEMPORARY` tables or tables used in a `MERGE` table.
- When using `SET CHARACTER SET`, you cannot use translated characters in database, table, and column names.
- You cannot use “_” or “%” with `ESCAPE` in `LIKE ... ESCAPE`.
- The server uses only the first `max_sort_length` bytes when comparing data values. This means that values cannot reliably be used in `GROUP BY`, `ORDER BY`, or `DISTINCT` if they differ only after the first `max_sort_length` bytes. To work around this, increase the variable value. The default value of `max_sort_length` is 1024 and can be changed at server startup time or at runtime.
- Numeric calculations are done with `BIGINT` or `DOUBLE` (both are normally 64 bits long). Which precision you get depends on the function. The general rule is that bit functions are performed with `BIGINT` precision, `IF()` and `ELT()` with `BIGINT` or `DOUBLE` precision, and the rest with `DOUBLE` precision. You should try to avoid using unsigned long long values if they resolve to be larger than 63 bits (9223372036854775807) for anything other than bit fields.
- You can have up to 255 `ENUM` and `SET` columns in one table.
- In `MIN()`, `MAX()`, and other aggregate functions, MySQL currently compares `ENUM` and `SET` columns by their string value rather than by the string's relative position in the set.
- In an `UPDATE` statement, columns are updated from left to right. If you refer to an updated column, you get the updated value instead of the original value. For example, the following statement increments `KEY` by 2, not 1:

```
mysql> UPDATE tbl_name SET KEY=KEY+1,KEY=KEY+1;
```

- You can refer to multiple temporary tables in the same query, but you cannot refer to any given temporary table more than once. For example, the following does not work:

```
mysql> SELECT * FROM temp_table, temp_table AS t2;  
ERROR 1137: Can't reopen table: 'temp_table'
```

- The optimizer may handle `DISTINCT` differently when you are using “hidden” columns in a join than when you are not. In a join, hidden columns are counted as part of the result (even if they are not shown), whereas in normal queries, hidden columns do not participate in the `DISTINCT` comparison.

An example of this is:

```
SELECT DISTINCT mp3id FROM band_downloads  
WHERE userid = 9 ORDER BY id DESC;
```

and

```
SELECT DISTINCT band_downloads.mp3id
FROM band_downloads,band_mp3
WHERE band_downloads.userid = 9
AND band_mp3.id = band_downloads.mp3id
ORDER BY band_downloads.id DESC;
```

In the second case, using MySQL Server 3.23.x, you may get two identical rows in the result set (because the values in the hidden `id` column may differ).

Note that this happens only for queries that do not have the `ORDER BY` columns in the result.

- If you execute a `PROCEDURE` on a query that returns an empty set, in some cases the `PROCEDURE` does not transform the columns.
- Creation of a table of type `MERGE` does not check whether the underlying tables are compatible types.
- If you use `ALTER TABLE` to add a `UNIQUE` index to a table used in a `MERGE` table and then add a normal index on the `MERGE` table, the key order is different for the tables if there was an old, non-`UNIQUE` key in the table. This is because `ALTER TABLE` puts `UNIQUE` indexes before normal indexes to be able to detect duplicate keys as early as possible.

