InnoDB Troubleshooting

Sveta Smirnova
Principal Technical Services Engineer
August, 11, 2016

Table of Contents

- What you need to know about InnoDB?
- Data corruption
- Surprising data
- Performance
- Instruments

What you need to know about InnoDB?

Overview

- Transactional engine
  - Any statement is part of transaction!
  - Even single SELECT
    - Use READ ONLY transactions in 5.6+
- Supports MVCC
  - Stores old versions of data until all transactions which could request it are closed
- ACID-complaint
- Physical layout

Overview

- Transactional engine
- Supports MVCC
- ACID-complaint
- Physical layout
  - System (shared) tablespace
  - Redo log files
  - Individual tablespaces

InnoDB structures

- System tablespace
  - InnoDB data dictionary
    - Metadata for InnoDB-related objects
  - Doublewrite buffer
    - Stores write page after it was flushed from the Buffer Pool, but before it is written into its proper position
    - Used for recovery when crash happens in the middle of a page write
  - Change buffer
    - Cache of changes to secondary indexes which are not in the Buffer Pool
  - Undo log (rollback segment)
    - Copies of data, modified by active transactions
    - If another transaction needs to see old data it is retrieved from this storage area
    - Can be stored in separate files in 5.6+

InnoDB structures

- System tablespace
- Redo log files
  - Encoded change requests, produced by transactions
  - Writes are recorded in circular fashion
  - Writes what did not finish updating the data files are replicated from the redo log files after unexpected shutdown
- Buffer pool
  - In-memory cache of table and index data
  - Old data get purged using a variation of LRU algorithm

InnoDB structures

- System tablespace
- Redo log files
- Buffer pool
- Adaptive hash index (AHI)
  - A hash index, based on a prefix of the index key
  - If table fits in memory can speed up queries by enabling direct lookup to any element

InnoDB structures

- System tablespace
- Redo log files
- Buffer pool
- Adaptive hash index (AHI)
  - For certain queries maintaining AHI structure can be not doable
    - Multiple concurrent joins
    - LIKE with wildcards ('%')
  - Before 5.7.8 was protected by a single latch which could be become a point of contention under heavy workloads

InnoDB structures

- System tablespace
- Redo log files
- Buffer pool
- Adaptive hash index (AHI)
- File-Per-Table Tablespaces
  - Contain table's index and data
  - Support various formats
    - DYNAMIC
    - COMPRESSED
  - Can be stored on different disks

Data corruption

Self-healing

- InnoDB can fix most of corruption issues itself
  - Check during startup
    - Redo logs
    - Doublewrite buffer
- If does not: innodb_force_recovery
  - Start from 1
  - Make sure if you have a copy of the datadir before using ≥ 4
  - After server started
    - Dump corrupted table
    - Drop it
    - Restart server with innodb_force_recovery=0
    - Reload the table

Self-healing

- InnoDB can fix most of corruption issues itself
- If does not: innodb_force_recovery
- Unless this is hardware corruption

Corrupted data

- Crashes if finds corrupted table

```
2016-08-07T01:47:00.373528Z 2 [ERROR] InnoDB: Database page corruption on disk or a failed file
read of page [page id: space=23, page number=178]. You may have to recover from a backup.
2016-08-07T01:47:00.373541Z 2 [Note] InnoDB: Page dump in ascii and hex (16384 bytes):
 len 16384; hex 00051eba0000014403e031343738366331322d356334302d313165362d616335382d30
...
InnoDB: End of page dump
2016-08-07T01:47:00.627116Z 2 [Note] InnoDB: Uncompressed page, stored checksum in field1 335546,
calculated checksums for field1: crc32 3863909191/4269622915, innodb 58208022, none 3735928559,
stored checksum in field2 0, calculated checksums for field2: crc32 3863909191/4269622915,
innodb 1818597513, none 3735928559,  page LSN 825371957 1664364589, low 4 bytes of LSN at
page end 0, page number (if stored to page already) 324, space id (if created with >= MySQL-4.1.1
and stored already) 808596530
2016-08-07T01:47:00.627140Z 2 [Note] InnoDB: It is also possible that your operating system has
corrupted its own file cache and rebooting your computer removes the error. If the corrupt page
is an index page. You can also try to fix the corruption by dumping, dropping, and reimporting
the corrupt table. You can use CHECK TABLE to scan your table for corruption. Please refer to
http://dev.mysql.com/doc/.../forcing-innodb-recovery.html for information about forcing recovery.
```

Corrupted data

- Crashes if finds corrupted table
- Use Percona Server option
  - innodb_corrupt_table_action

Read Error Log!

- Seriously!
- Crash

```
InnoDB: We intentionally generate a memory trap.
InnoDB: Submit a detailed bug report to http://bugs.mysql.com.
InnoDB: If you get repeated assertion failures or crashes, even
InnoDB: immediately after the mysqld startup, there may be
InnoDB: corruption in the InnoDB tablespace. Please refer to
InnoDB: http://dev.mysql.com/doc/refman/5.7/en/forcing-innodb-recovery.html
InnoDB: about forcing recovery.
```

- Missed tables

Read Error Log!

- Seriously!
- Crash
- Missed tables

```
2016-08-07T02:05:04.718199Z 2 [ERROR] InnoDB: Failed to find tablespace for table ‘test‘.‘joinit‘
in the cache. Attempting to load the tablespace with space id 23
2016-08-07T02:05:04.718245Z 2 [ERROR] InnoDB: Operating system error number 2 in a file operation.
2016-08-07T02:05:04.718260Z 2 [ERROR] InnoDB: The error means the system cannot find the path
specified.
2016-08-07T02:05:04.718272Z 2 [ERROR] InnoDB: Cannot open datafile for read-only:
 ’./test/joinit.ibd’ OS error: 71
 ...
2016-08-07T02:05:04.718322Z 2 [ERROR] InnoDB: Could not find a valid tablespace file
for ‘test/joinit‘. Please refer to
http://dev.mysql.com/doc/.../innodb-troubleshooting-datadict.html for how to resolve the issue.
2016-08-07T02:05:04.718605Z 2 [Warning] InnoDB: Cannot calculate statistics for table
‘test‘.‘joinit‘ because the .ibd file is missing. Please refer to
http://dev.mysql.com/doc/.../innodb-troubleshooting.html for how to resolve the issue.
```

Surprising data

Transaction isolation levels

- You need to understand transaction isolation levels
- Which snapshots are created
- Which tables are locked

Different auto-increment locks

- Order of AI values is always consecutive
- Can contain gaps
- Affect replication
  - 0 (”traditional”) and 2 (”interleaved”) not safe for SBR
- Affect performance
  - Less safe modes are faster

Deadlocks

- Happen when two transactions block rows which each other wants to update
- Not avoidable
- Details are in the ENGINE INNODB STATUS

```
LATEST DETECTED DEADLOCK
------------------------
2016-08-07 06:29:36 0x7fa7d155d700
*** (1) TRANSACTION:
TRANSACTION 1295, ACTIVE 3 sec starting index read
mysql tables in use 1, locked 1
LOCK WAIT 2 lock struct(s), heap size 1160, 1 row lock(s)
MySQL thread id 4, OS thread handle 140358748063488, query id 35 localhost 127.0.0.1 root updating
DELETE FROM t WHERE i = 1
*** (1) WAITING FOR THIS LOCK TO BE GRANTED:
RECORD LOCKS space id 0 page no 311 n bits 72 index GEN_CLUST_INDEX of table `test`.`t` trx id 1295
...
```

Deadlocks

- Happen when two transactions block rows which each other wants to update
- Not avoidable
- Details are in the ENGINE INNODB STATUS
- In the error log: innodb_print_all_deadlocks

```
2016-08-07T03:29:36.938929Z 5 [Note] InnoDB: Transactions deadlock detected, dumping detailed ...
2016-08-07T03:29:36.938958Z 5 [Note] InnoDB:
*** (1) TRANSACTION:
...
RECORD LOCKS space id 0 page no 311 n bits 72 index GEN_CLUST_INDEX of table `test`.`t`
trx id 1296 lock_mode X waiting
2016-08-07T03:29:36.939398Z 5 [Note] InnoDB: *** WE ROLL BACK TRANSACTION (1)
```

Foreign key issues

- ENGINE INNODB STATUS

```
------------------------
LATEST FOREIGN KEY ERROR
------------------------
2016-08-07 06:38:12 0x7fa7d155d700 Error in foreign key constraint of table test/child:
FOREIGN KEY (parent_id)
        REFERENCES parent(id)
        ON DELETE CASCADE
) ENGINE=INNODB:
Cannot find an index in the referenced table where the
referenced columns appear as the first columns, or column types
in the table and the referenced table do not match for constraint.
Note that the internal storage type of ENUM and SET changed in
tables created with >= InnoDB-4.1.12, and such columns in old tables
cannot be referenced by such columns in new tables.
Please refer to ...innodb-foreign-key-constraints.html for correct foreign key definition.
```

Foreign key issues

- ENGINE INNODB STATUS

```
2016-08-07 06:40:38 0x7fa7d155d700 Transaction:
TRANSACTION 1329, ACTIVE 0 sec inserting
mysql tables in use 1, locked 1
3 lock struct(s), heap size 1160, 1 row lock(s), undo log entries 1
MySQL thread id 6, OS thread handle 140358748329728, query id 51 localhost 127.0.0.1 root update
insert into child values(1,4)
Foreign key constraint fails for table `test`.`child`:
,
  CONSTRAINT `child_ibfk_1` FOREIGN KEY (`parent_id`) REFERENCES `parent` (`id`) ON DELETE CASCADE
Trying to add in child table, in index par_ind tuple:
DATA TUPLE: 2 fields;
 0: len 4; hex 80000004; asc     ;;
 1: len 6; hex 000000000203; asc       ;;

But in parent table `test`.`parent`, in index PRIMARY,
the closest match we can find is record:
...
```

Foreign key issues

- ENGINE INNODB STATUS
- Foreign key require
  - Same datatype
  - Key in the parent table
- Error message itself is not detailed!

```
mysql> CREATE TABLE child (
    ->     id INT,
    ->     parent_id INT,
    ->     INDEX par_ind (parent_id),
    ->     FOREIGN KEY (parent_id)
    ->         REFERENCES parent(id)
    ->         ON DELETE CASCADE
    -> ) ENGINE=INNODB;
ERROR 1215 (HY000): Cannot add foreign key constraint
```

Foreign key issues

- ENGINE INNODB STATUS
- Foreign key require
  - Same datatype
  - Key in the parent table
- Error message itself is not detailed!

```
mysql> insert into child values(1,4);
ERROR 1452 (23000): Cannot add or update a child row: a foreign key constraint fails
(‘test‘.‘child‘, CONSTRAINT ‘child_ibfk_1‘ FOREIGN KEY (‘parent_id‘)
REFERENCES ‘parent‘ (‘id‘) ON DELETE CASCADE)
```

Performance

Two main issues

- Concurrency
- IO

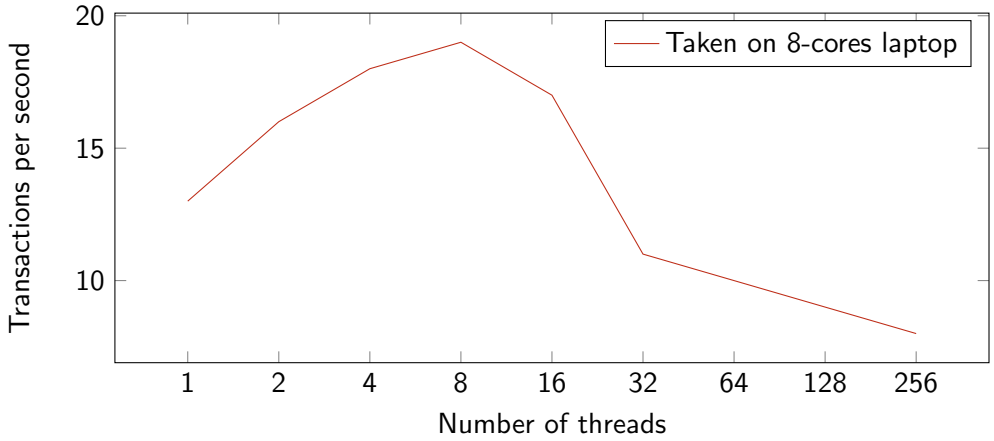
Concurrency

- More threads - more job can be performed at the same time
- InnoDB thread is not the same as connection thread
  - Connection threads may be sleeping while InnoDB threads are doing their job
  - If there are more running connection threads (Running_threads) than InnoDB threads the engine needs to prioritize

Concurrency

- More threads - more job can be performed at the same time
- InnoDB thread is not the same as connection thread
- Limited by number of CPU cores
- Unreasonable number of parallel threads can cause contentions
  - Performance decrease
  - Server hangs!

Unreasonable number of threads is not good



How to tune InnoDB parallelism

- Must-care options
  - innodb_thread_concurrency
  - innodb_concurrency_tickets
  - innodb_read_io_threads
  - innodb_write_io_threads
  - innodb_commit_concurrency

How to tune InnoDB parallelism

- Must-care options
- Other options, affecting CPU
  - innodb_log_checksum_algorithm
  - innodb_log_checksums
  - innodb_compression_level
  - innodb_thread_sleep_delay
  - innodb_page_cleaners
  - innodb_purge_threads

IO

- Less IO
  - Better performance
  - Higher risk of data corruption
- You don’t always need decrease IO
  - Fast disks
  - Application writes a lot

InnoDB IO tuning

- Must-care options
  - innodb_flush_method
  - innodb_log_file_size
  - innodb_io_capacity
  - innodb_read_io_threads
  - innodb_write_io_threads

InnoDB IO tuning

- Must-care options
- Less IO with risk of data loss
  - innodb_doublewrite
  - innodb_flush_log_at_trx_commit
  - innodb_flush_log_at_timeout
  - innodb_support_xa
  - Modify on read-only slave only!

InnoDB IO tuning

- Other IO-related options
  - innodb_adaptive_flushing
  - innodb_change_buffering
  - innodb_flush_neighbors
  - innodb_flush_sync
  - innodb_log_compressed_pages
  - innodb_max_dirty_pages_pct
  - innodb_max_purge_lag
  - innodb_page_cleaners
  - innodb_purge_threads

Instruments

Status variables

- Innodb_buffer_pool_*
- Innodb_data_*
- Innodb_dblwr_*
- Innodb_log_*
- Innodb_lsn_*
- Innodb_master_thread_[active|idle]_loops
- Innodb_max_trx_id
- ...

Status variables: usage tips

- Increments
  - Record current value

```
mysql> show global status like 'Innodb_rows_read'\G
*************************** 1. row ***************************
Variable_name: Innodb_rows_read
        Value: 111954358499
1 row in set (0,00 sec)
```

  - Wait 30 seconds - 5 minutes
  - Record updated value

```
mysql> show global status like 'Innodb_rows_read'\G
*************************** 1. row ***************************
Variable_name: Innodb_rows_read
        Value: 111960191111 - about 5M rows in measured time
1 row in set (0,00 sec)
```

  - Compare

Status variables: usage tips

- Increments
- Ratios

```
mysql> show global status like 'Innodb_buffer_pool_pages%';
+----------------------------------+-----------+
| Variable_name                    | Value     |
+----------------------------------+-----------+
| Innodb_buffer_pool_pages_data    | 2696428   |
| Innodb_buffer_pool_pages_dirty   | 76082     |
| Innodb_buffer_pool_pages_flushed | 679700112 |
| Innodb_buffer_pool_pages_free    | 7713      | - 0.34 % free
| Innodb_buffer_pool_pages_misc    | 179435    |
| Innodb_buffer_pool_pages_total   | 2883576   |
+----------------------------------+-----------+
6 rows in set (0,00 sec)
```

Status variables: usage tips

- Increments
- Ratios
- Status

```
mysql> show global status like 'Innodb_buffer_pool%status';
+---------------------------------+-------------------------------------------------+
| Variable_name                   | Value                                           |
+---------------------------------+-------------------------------------------------+
| Innodb_buffer_pool_dump_status  | Dumping of buffer pool not started              |
| Innodb_buffer_pool_load_status  | Buffer pool(s) load completed at 160811  0:23:10 |
| Innodb_buffer_pool_resize_status |                                                 |
+---------------------------------+-------------------------------------------------+
3 rows in set (0,00 sec)
```

ENGINE INNODB STATUS

- Has information of what is happening inside InnoDB
- Always available
  - SHOW ENGINE INNODB STATUS
- Dump to file periodically
  - Before 5.6
    - CREATE TABLE innodb_monitor(...) ENGINE=INNODB;
    - Dump will be in the error log file
  - 5.6+
    - innodb-status-file
    - innodb_status_output
    - innodb_status_output_locks

ENGINE INNODB STATUS

Name	Description
BACKGROUND THREAD	Background thread activity
SEMAPHORES	Internal semaphores
TRANSACTIONS	Current transactions, see also Troubleshooting Locking Issues
FILE I/O	IO operations, all internal threads
INSERT BUFFER AND ADAPTIVE HASH INDEX	Insert buffer and adaptive hash index usage
LOG	Operations with InnoDB log files
BUFFER POOL AND MEMORY	InnoDB buffer pool and memory usage
ROW OPERATIONS	Row operations and statistics

BACKGROUND THREAD

```
-----------------
BACKGROUND THREAD
-----------------
srv_master_thread loops: 1914468 srv_active, 0 srv_shutdown, 7275 srv_idle
srv_master_thread log flush and writes: 1921714
...
-----------------
BACKGROUND THREAD
-----------------
srv_master_thread loops: 1914497 srv_active, 0 srv_shutdown, 7275 srv_idle
srv_master_thread log flush and writes: 1921743
```

SEMAPHORES

```
----------
SEMAPHORES
----------
OS WAIT ARRAY INFO: reservation count 898439525, signal count 12201032157
--Thread 140127188350720 has waited at btr0sea.c line 633 for 0.00 seconds the semaphore:
X-lock (wait_ex) on RW-latch at 0x7f724215d688 created in file btr0sea.c line 178
a writer (thread id 140127188350720) has reserved it in mode  wait exclusive
number of readers 1, waiters flag 1, lock_word: fffffffffffffff
Last time read locked in file btr0sea.c line 882
Last time write locked in file /build/buildd/mysql-5.5-5.5.40/storage/innobase/btr/btr0sea.c line 633
--Thread 140063303272192 has waited at row0sel.c line 3713 for 0.00 seconds the semaphore:
S-lock on RW-latch at 0x7f724215d688 created in file btr0sea.c line 178
a writer (thread id 140127188350720) has reserved it in mode  wait exclusive
number of readers 1, waiters flag 1, lock_word: fffffffffffffff
...
AHI not effective
```

FILE I/O

```
--------
FILE I/O
--------
I/O thread 0 state: waiting for completed aio requests (insert buffer thread)
I/O thread 1 state: waiting for completed aio requests (log thread)
I/O thread 2 state: complete io for buf page (read thread)
I/O thread 3 state: complete io for buf page (read thread)
I/O thread 4 state: complete io for buf page (read thread)
I/O thread 5 state: complete io for buf page (read thread)
I/O thread 6 state: waiting for completed aio requests (write thread)
I/O thread 7 state: waiting for completed aio requests (write thread)
I/O thread 8 state: waiting for completed aio requests (write thread)
I/O thread 9 state: waiting for completed aio requests (write thread)
Pending normal aio reads: 71 [5, 64, 2, 0] , aio writes: 0 [0, 0, 0, 0] ,
 ibuf aio reads: 0, log i/o's: 0, sync i/o's: 0
Pending flushes (fsync) log: 0; buffer pool: 0
817997725 OS file reads, 2452891797 OS file writes, 69035872 OS fsyncs
9 pending preads, 0 pending pwrites
1465.12 reads/s, 16361 avg bytes/read, 342.74 writes/s, 41.46 fsyncs/s
```

INSERT BUFFER AND ADAPTIVE HASH INDEX

```
-------------------------------------
INSERT BUFFER AND ADAPTIVE HASH INDEX
-------------------------------------
Ibuf: size 1, free list len 54033, seg size 54035, 1875378 merges
merged operations:
 insert 1466727, delete mark 1069988, delete 484172
discarded operations:
 insert 0, delete mark 0, delete 0
Hash table size 446244899, node heap has 81058 buffer(s)
1064000.00 hash searches/s, 2302000.00 non-hash searches/s

1064000 < 2302000 - AHI maybe not effective
```

LOG

```
---
LOG
---
Log sequence number 2212908859806
Log flushed up to   2212908846938
Last checkpoint at  2212902001954
1 pending log writes, 0 pending chkp writes
1020322391 log i/o's done, 145.87 log i/o's/second
```

BUFFER POOL AND MEMORY

```
----------------------
BUFFER POOL AND MEMORY
----------------------
Total memory allocated 61538828288; in additional pool allocated 0
Dictionary memory allocated 42219333
Buffer pool size   3670016
Free buffers       12
Database pages     3426027
Old database pages 1264665
Modified db pages  3187
Pending reads 0
Pending writes: LRU 0, flush list 1, single page 0
Pages made young 1419367465, not young 0
4.48 youngs/s, 0.00 non-youngs/s
Pages read 1250034217, created 7773872, written 248275353
1.65 reads/s, 0.61 creates/s, 53.95 writes/s
Buffer pool hit rate 1000 / 1000, young-making rate 0 / 1000 not 0 / 1000
Pages read ahead 0.00/s, evicted without access 0.00/s, Random read ahead 0.00/s
LRU len: 3426027, unzip_LRU len: 3564
```

ROW OPERATIONS

```
--------------
ROW OPERATIONS
--------------
2 queries inside InnoDB, 0 queries in queue
2 read views open inside InnoDB
Main thread process no. 1025, id 140064232015616, state: flushing buffer pool pages
Number of rows inserted 288238328, updated 551243613, deleted 119193909, read 2277604799383
31.26 inserts/s, 40.22 updates/s, 8.39 deletes/s, 203214.21 reads/s
```

Tables in Information Schema

- INNODB_BUFFER_*
  - INNODB_BUFFER_PAGE

```
mysql> select count(*), sum(NUMBER_RECORDS), if(IS_OLD='YES', 'old', 'new') as old_block
    -> from innodb_buffer_page where table_name='`sbtest`.`sbtest1`' group by old_block;
+----------+--------------------+-----------+
| count(*) | sum(NUMBER_RECORDS) | old_block |
+----------+--------------------+-----------+
|      611 |             135383 | new       |
|      345 |              26296 | old       |
+----------+--------------------+-----------+
2 rows in set (0.05 sec)
```

  - INNODB_BUFFER_PAGE_LRU
  - INNODB_BUFFER_POOL_STATS

Tables in Information Schema

- INNODB_BUFFER_*
- INNODB_CHANGED_PAGES - Percona Server only!
- INNODB_CMP*
  - INNODB_CMP
  - INNODB_CMPMEM
  - INNODB_CMPMEM_RESET
  - INNODB_CMP_PER_INDEX
  - INNODB_CMP_PER_INDEX_RESET
  - INNODB_CMP_RESET

Tables in Information Schema

- INNODB_BUFFER_*
- INNODB_CHANGED_PAGES - Percona Server only!
- INNODB_CMP*
- INNODB_FT_*
  - INNODB_FT_BEING_DELETED
  - INNODB_FT_CONFIG
  - INNODB_FT_DEFAULT_STOPWORD
  - INNODB_FT_DELETED
  - INNODB_FT_INDEX_CACHE
  - INNODB_FT_INDEX_TABLE

Tables in Information Schema

- INNODB_BUFFER_*
- INNODB_CHANGED_PAGES - Percona Server only!
- INNODB_CMP*
- INNODB_FT_*
- INNODB_LOCKS*
  - INNODB_LOCKS - See Troubleshooting Locking Issues
  - INNODB_LOCK_WAITS - See Troubleshooting Locking Issues

INNODB_METRICS

- Contains information about currently enabled and disabled monitors
- To turn monitors ON or OFF:
  - set global innodb_monitor_enable = 'NAME'
  - set global innodb_monitor_disable = 'NAME'
  - set global innodb_monitor_enable = 'subsystem%'
  - set global innodb_monitor_disable = 'all'

INNODB_METRICS subsystems

Subsystem	Description
adaptive_hash_index	Adaptive hash index usage
buffer	InnoDB buffer pool metrics, see also INNODB_BUFFER_POOL_STATS
buffer_page_io	Information about page reads and writes, see also INNODB_BUFFER_PAGE and INNODB_BUFFER_PAGE_LRU
change_buffer	Change buffer metrics
compression	Stats for compressed pages, see also INNODB_CMP* tables
ddl	Statistics about current DDL operations
dml	How many rows was read, inserted, updated and deleted
file_system	file_num_open_files: number of currently opened files inside InnoDB

INNODB_METRICS subsystems - cont.

Subsystem	Description
icp	Index push-down condition stats
index	Index statistics
lock	Informatioin about deadlocks, table, row, rec locks and timeouts
metadata	Statistics of table handles usage
os	Reads, writes and fsyncs metrics
purge	Purge statistics
recovery	Redo log usage: LSN, etc.
server	Information about global InnoDB structures
transaction	Transactions perforrmance statistics

INNODB_METRICS example: writes vs fsyncs

```
mysql> set global innodb_monitor_enable = 'os%';
Query OK, 0 rows affected (0.00 sec)

-- Perform data writes
mysql> insert into t1 select md5(rand()) from t1;
Query OK, 4 rows affected (0.05 sec)
Records: 4  Duplicates: 0  Warnings: 0

mysql> select name, count, max_count from information_schema.innodb_metrics
    -> where name in ('os_data_writes', 'os_data_fsyncs');
+----------------+-------+-----------+
| name           | count | max_count |
+----------------+-------+-----------+
| os_data_writes |    66 |        66 |
| os_data_fsyncs |    40 |        40 |
+----------------+-------+-----------+
2 rows in set (0.00 sec)
```

Tables in Information Schema

- INNODB_SYS_*
  - INNODB_SYS_COLUMNS
  - INNODB_SYS_DATAFILES
  - INNODB_SYS_FIELDS
  - INNODB_SYS_FOREIGN
  - INNODB_SYS_FOREIGN_COLS
  - INNODB_SYS_INDEXES
  - INNODB_SYS_TABLES
  - INNODB_SYS_TABLESPACES
  - INNODB_SYS_TABLESTATS
  - INNODB_SYS_VIRTUAL

Tables in Information Schema

- INNODB_SYS_*
- INNODB_TEMP_TABLE_INFO
  - Does not contain optimized internal temporary tables used by InnoDB
- INNODB_TRX

INNODB_TRX

```
mysql> select id, state, info from information_schema.processlist where id=3\G
*************************** 1. row ***************************
   id: 3
state: updating
 info: update t1 set f1=md5(rand()) limit 2
1 row in set (0.00 sec)

mysql> select * from information_schema.innodb_trx where trx_mysql_thread_id=3\G
*************************** 1. row ***************************
                    trx_id: 1300
                 trx_state: LOCK WAIT
               trx_started: 2016-08-10 01:30:45
     trx_requested_lock_id: 1300:0:311:2
          trx_wait_started: 2016-08-10 01:30:45
                trx_weight: 2
       trx_mysql_thread_id: 3
                 trx_query: update t1 set f1=md5(rand()) limit 2
       trx_operation_state: starting index read
       ...
```

INNODB_TRX - cont.

```
         trx_tables_in_use: 1
         trx_tables_locked: 1
          trx_lock_structs: 2
     trx_lock_memory_bytes: 1160
           trx_rows_locked: 1
         trx_rows_modified: 0
   trx_concurrency_tickets: 0
       trx_isolation_level: REPEATABLE READ
         trx_unique_checks: 1
    trx_foreign_key_checks: 1
trx_last_foreign_key_error: NULL
 trx_adaptive_hash_latched: 0
 trx_adaptive_hash_timeout: 0
          trx_is_read_only: 0
trx_autocommit_non_locking: 0
1 row in set (0.01 sec)
```

InnoDB mutex status

- Shows currently used InnoDB mutexes

```
mysql> show engine innodb mutex;
+--------+----------------------------+----------+
| Type   | Name                       | Status   |
+--------+----------------------------+----------+
| InnoDB | rwlock: dict0dict.cc:1184  | waits=2  |
| InnoDB | rwlock: log0log.cc:907     | waits=10 |
| InnoDB | sum rwlock: buf0buf.cc:1453 | waits=3  |
+--------+----------------------------+----------+
3 rows in set (0.00 sec)
```

- More details can be found in Performance Schema

InnoDB mutex status

- Shows currently used InnoDB mutexes
- More details can be found in Performance Schema

```
mysql> select event_name, SOURCE, count(*) from performance_schema.events_waits_current
    -> where event_name like 'wait/synch/mutex/innodb%' group by event_name, SOURCE;
+-----------------------------------------------------+---------------------+----------+
| event_name                                          | SOURCE              | count(*) |
+-----------------------------------------------------+---------------------+----------+
| wait/synch/mutex/innodb/buf_pool_flush_state_mutex | buf0flu.cc:1986     |        2 |
| wait/synch/mutex/innodb/flush_list_mutex           | buf0buf.cc:408      |        1 |
| wait/synch/mutex/innodb/lock_wait_mutex            | lock0wait.cc:508    |        1 |
| wait/synch/mutex/innodb/log_sys_mutex              | log0log.ic:427      |        1 |
| wait/synch/mutex/innodb/recalc_pool_mutex          | dict0stats_bg.cc:164 |        1 |
| wait/synch/mutex/innodb/rw_lock_debug_mutex        | sync0debug.cc:1296  |        3 |
+-----------------------------------------------------+---------------------+----------+
6 rows in set (0.00 sec)
```

Performance Schema instrumentation

- wait/synch/mutex/innodb/
- wait/synch/sxlock/innodb
- wait/synch/cond/innodb/commit_cond
- wait/io/file/innodb
- stage/innodb/alter table
  - Watch ALTER TABLE in progress
- memory/innodb
  - Memory usage of InnoDB structures

Summary

Summary

- Every query is part of transaction
- Follow advices in the error log to fix corruption issues
- Mind transaction isolation levels and AI lock modes
- More threads can do more job unless compete for system resources
- IO subsystem is another one which affects performance

More information

- InnoDB Architecture
- Troubleshooting Locking Issues
- MySQL 5.5 Guide to InnoDB Status
- Jeremy Cole about InnoDB internals, structures, and behavior
- InnoDB Team blog
- InnoDB at the Percona Blog

Place for your questions

???

Thank you!

http://www.slideshare.net/SvetaSmirnova

https://twitter.com/svetsmirnova