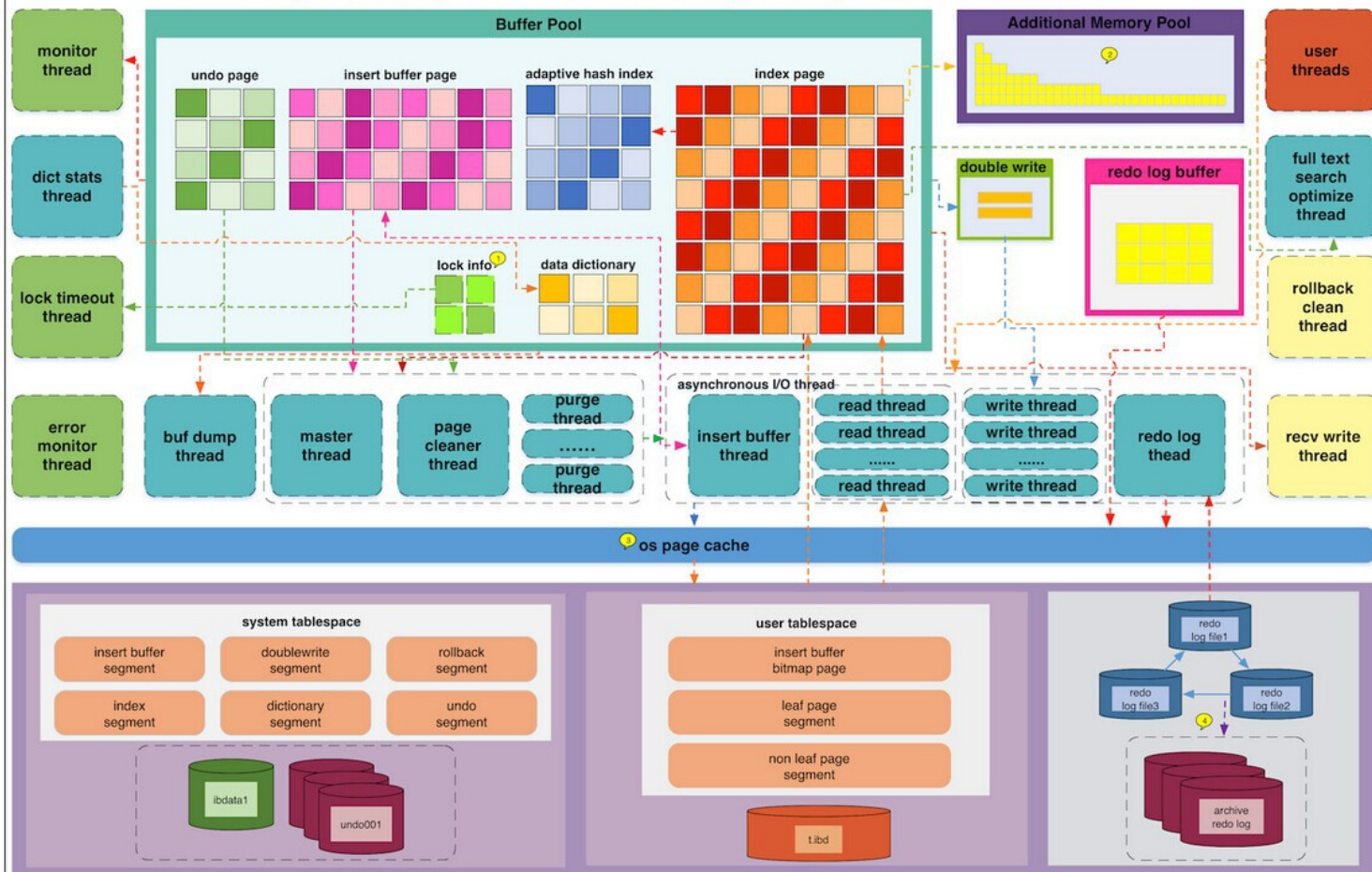


图解 INNODB 存储引擎的体系结构

MySQL 5.6 | InnoDB Storage Engine Architecture



Thread 后台线程

Master Thread
主线程

4 个 I/O
Thread

1 个锁 lock
监控线程

1 个错误
监控线程

4 个 I/O
Thread

Insert buffer thread
插入缓冲线程

log thread
日志线程

read thread
读线程

write thread
写线程

IO thread 的数量可以通过以下参数来调整：
innodb_read_io_threads
innodb_write_io_threads

Master Thread 主线程


Loop
主循环

Background Loop
后台循环

Flush Loop
刷新循环

Suspend Loop
暂停循环

每秒
1s

 sleep 1s

Fulsh
log buffer
to disk

if last_one_second_io
<
5% * innodb_io_capacity

Merge
5% * innodb_io_capacity
insert buffer

if buf_get_modified_ratio_pct
>
innodb_max_dirty_pages_pct

Flush
100% * innodb_io_capacity dirty page
buffer pool

Flush
desired amount dirty page
buffer pool

if no user activity

Goto backgroup loop

十秒
10s

Fulsh
log buffer
to disk

Full purge

Merge
5% * innodb_io_capacity
insert buffer

if last_ten_second_io
<
innodb_io_capacity

Flush
100% * innodb_io_capacity dirty page
buffer pool

if buf_get_modified_ratio_pct
>
innodb_max_dirty_pages_pct

Flush
100% * innodb_io_capacity dirty page
buffer pool

Flush
10% * innodb_io_capacity dirty page
buffer pool

Fuzzy checkpoint

Goto loop

Full purge

Merge
10% * innodb_io_capacity
insert buffer

if no user activity:
Goto backgroup loop
Else:
Goto flush loop

Flush
100% * innodb_io_capacity dirty
page
buffer pool

if buf_get_modified_ratio_pct
>
innodb_max_dirty_pages_pct
Goto flush loop

Goto suspend Loop

Wait event

Goto loop

每秒会做很多事情，包括：

- 1) 睡觉
- 2) 将日志缓冲刷新到磁盘，即使事务未提交
- 3) 如果【前 1s 的 I/O 次数】小于【磁盘 I/O 吞吐量的 5%（默认 200*5%=10 个页）】，则合并插入缓冲（数量为磁盘 I/O 吞吐量 *5% 即 10 个页）
- 4) 如果【脏页比例】大于【阈值（默认为 75，脏页比例为百分之七十五）】，则刷新缓冲池中脏页到磁盘，否则只刷新合适的脏页数量（innodb_adaptive_flushing 参数决定）
- 5) 当前没有用户活动，切换到后台循环

每十秒会做的事情，包括：

- 1) 将日志缓冲刷新到磁盘，即使事务未提交
- 2) 删除无用的 undo 页
- 3) 合并插入缓冲，缓冲数量为【磁盘 I/O 吞吐量的 5%，如果默认 200，则合并插入缓冲 10 个】
- 4) 如果【前 10s 的 I/O 次数】小于【磁盘 I/O 吞吐量（默认 200*100%=200 个页）】，则刷新缓冲池中脏页到磁盘（数量为磁盘 I/O 吞吐量 *100% 即 200 个页）
- 5) 如果【脏页比例】大于【阈值（默认为 75，脏页比例为百分之七十五）】，则刷新缓冲池中脏页到磁盘（数量为磁盘 I/O 吞吐量 *100% 即 200 个页），否则脏页刷新数量为（磁盘 I/O 吞吐量 *10% 即 20 个页）
- 6) 产生一个检查点（模糊检查点）

Master Thread 主线程


Loop
主循环

Background Loop
后台循环

Flush Loop
刷新循环

Suspend Loop
暂停循环

每秒
1s

 sleep 1s

Fulsh
log buffer
to disk

if last_one_second_io
<
5% * innodb_io_capacity

Merge
5% * innodb_io_capacity
insert buffer

if buf_get_modified_ratio_pct
>
innodb_max_dirty_pages_pct

Flush
100% * innodb_io_capacity dirty page
buffer pool

Flush
desired amount dirty pape
buffer pool

if no user activity

Goto backgroup loop

十秒
10s

Fulsh
log buffer
to disk

Full purge

Merge
5% * innodb_io_capacity
insert buffer

if last_ten_second_io
<
innodb_io_capacity

Flush
100% * innodb_io_capacity dirty page
buffer pool

if buf_get_modified_ratio_pct
>
innodb_max_dirty_pages_pct

Flush
100% * innodb_io_capacity dirty page
buffer pool

Flush
10% * innodb_io_capacity dirty page
buffer pool

Fuzzy checkpoint

Goto loop

Full purge

Merge
10% * innodb_io_capacity
insert buffer

if no user activity:
Goto backgroup loop
Else:
Goto flush loop

Flush
100% * innodb_io_capacity dirty
page
buffer pool

if buf_get_modified_ratio_pct
>
innodb_max_dirty_pages_pct
Goto flush loop

Goto suspend Loop

Wait event

Goto loop

后台循环，包括：

- 1) 删除无用的 undo 页
- 2) 合并插入缓冲，缓冲数量为【磁盘 I/O 吞吐量的 10%，如果默认 200，则合并插入缓冲 20 个】
- 3) 当前有用户活动，切换到主循环
- 4) 当前没有用户活动，切换到刷新循环

刷新循环，包括：

- 1) 刷新缓冲池中脏页到磁盘（数量为磁盘 I/O 吞吐量 * 100% 即 200 个页）
- 2) 如果【脏页比例】大于【阈值（默认为 75，脏页比例为百分之七十五）】，则继续进入刷新循环，否则进入暂停循环

暂停循环，包括：

- 1) 等待事件发生，进入主循环

MEM 内存空间

Buffer pool
缓冲池

Redo log buffer
重做日志缓冲

Additional memory pool
额外的内存池

缓冲池的大小可以通过以下参数来调整：
innodb_buffer_pool_size
innodb_log_buffer_size
innodb_additional_mem_pool_size

Buffer pool
缓冲池

Undo page
数据页

Insert buffer
插入缓冲

Index page
索引页

Lock info
锁信息

Data dictionary
数据字典

Adaptive hash index
自适应哈希索引