What's new in AliSQL

Alibaba's fork of MySQL

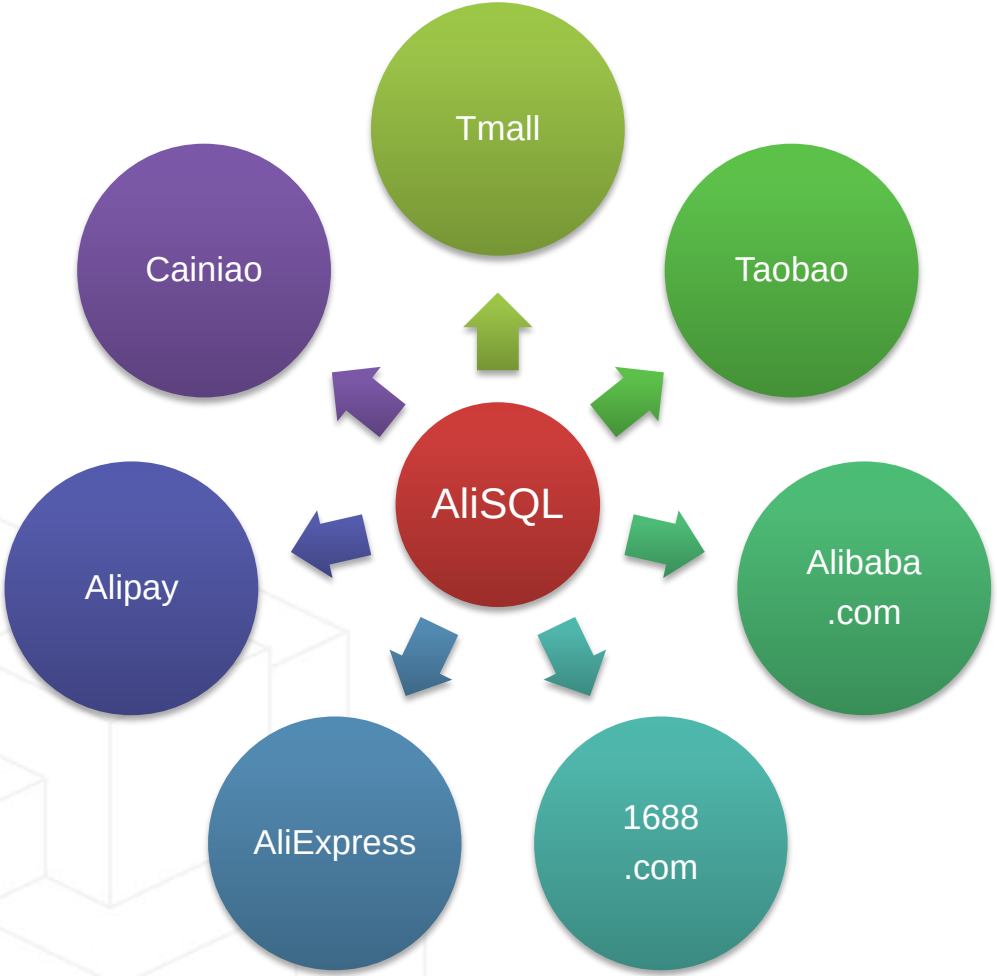
Zhang Yingqiang
Database Kernel Expert @ Alibaba Group

Alibaba Group 阿里巴巴集团

阿里技术保障 Ali Infrastructure Service

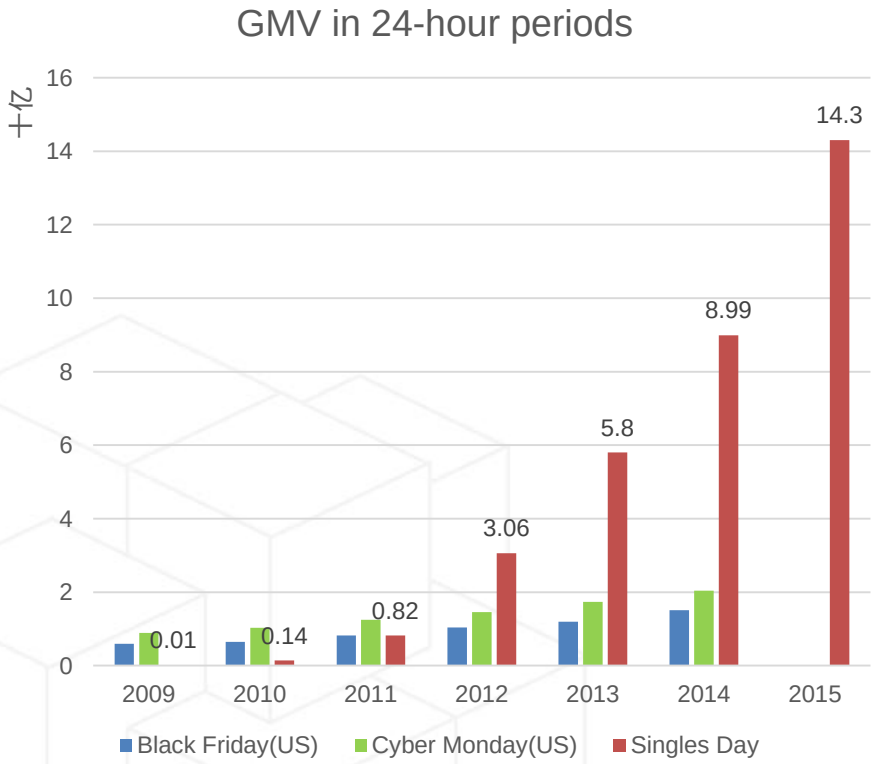
Agenda

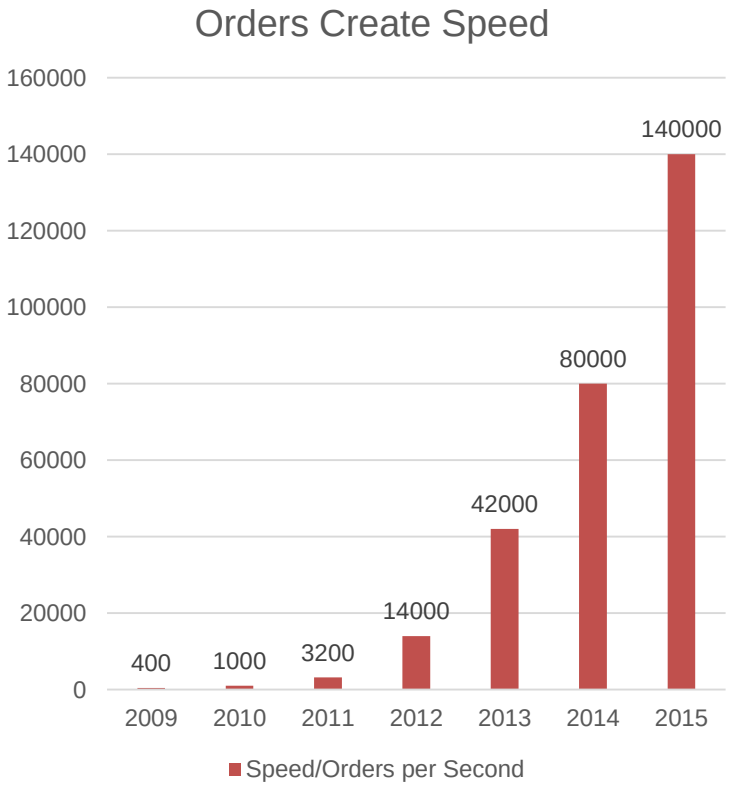
- ✓ Evolution of AliSQL
- ✓ Performance
- ✓ Robustness
- ✓ Security & Others



AliSQL are widely used by almost every Business Group in Alibaba Group and Ant Financial (Alipay).

The extreme business pressure for AliSQL





2011

- AliSQL 5.1

- Bugfix for DDL
- Eliminate race condition.

…

2012~

- AliSQL 5.5

- Parallel Replication
- Optimize Hot SKU v1

…

2014~

- AliSQL 5.6

- Enhanced Thread Pool
- SQL firewall

…

- 40+ new bugs have been found & fixed
- All have been reported to the community

BugFix

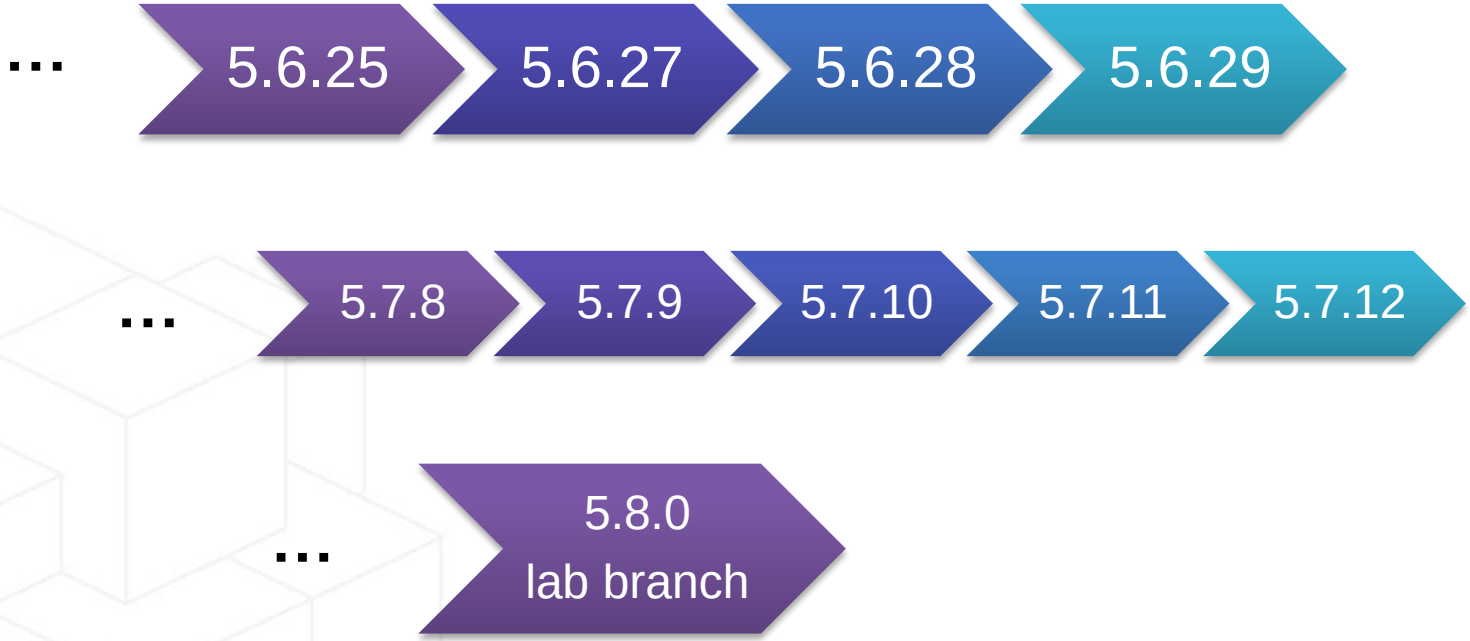
- 41 new Features have been added

New Feature

- 27 bottlenecks have been optimized

Performance enhance

The recently released Oracle/MySQL versions which contain AliSQL's contributions.



Performance

commit_on_success hints

Group prepare

Enhanced Thread Pool

Select from Update

Hot SKU Optimization

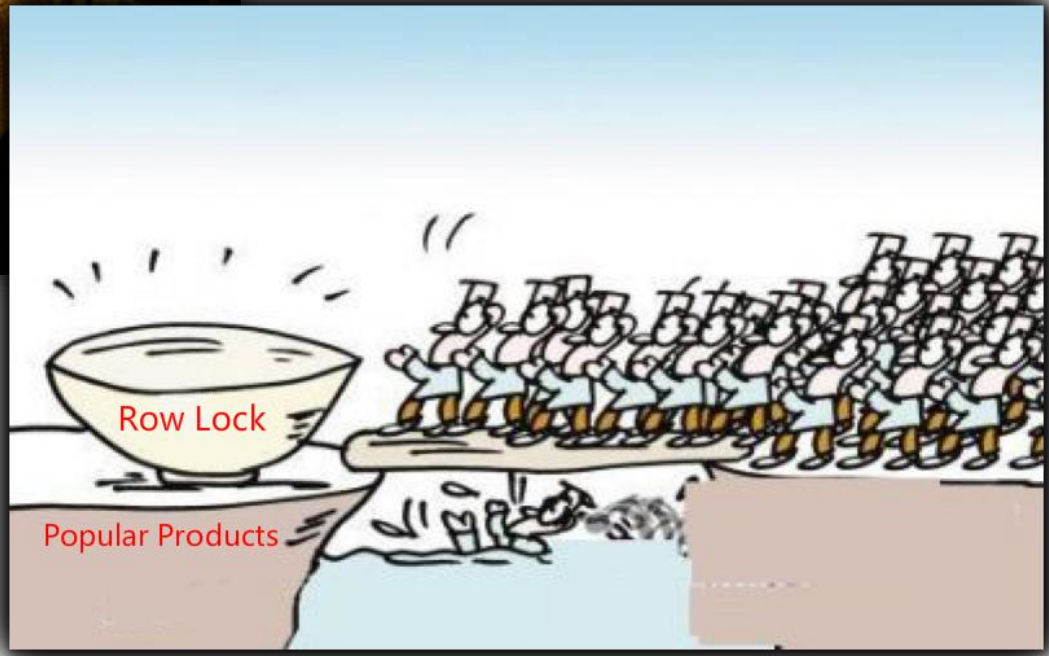
Dual redo log buffers

Table-level Parallel Replication

PK access optimization

Table Lock optimization for select

Enhanced Semi-Sync Replication



panic purchase

1st Generation

- new syntaxes

2nd Generation

- InnoDB Strict Concurrency
- Queue on PK

3rd Generation

(In development)

- Row Cache
- New InnoDB row Lock Type
- Group Update
- Associated Transaction

1st Generation: new syntaxes

Alibaba Group 阿里巴巴集团 | 阿里技术保障 Ali Infrastructure Service

Transaction model:
① begin;
② insert normal row;
③ update hot row;
④ select hot row;
⑤ commit;

1st step

Transaction model:
① begin;
② insert normal row;
③ select * from update hot row;
④ commit;

2nd step

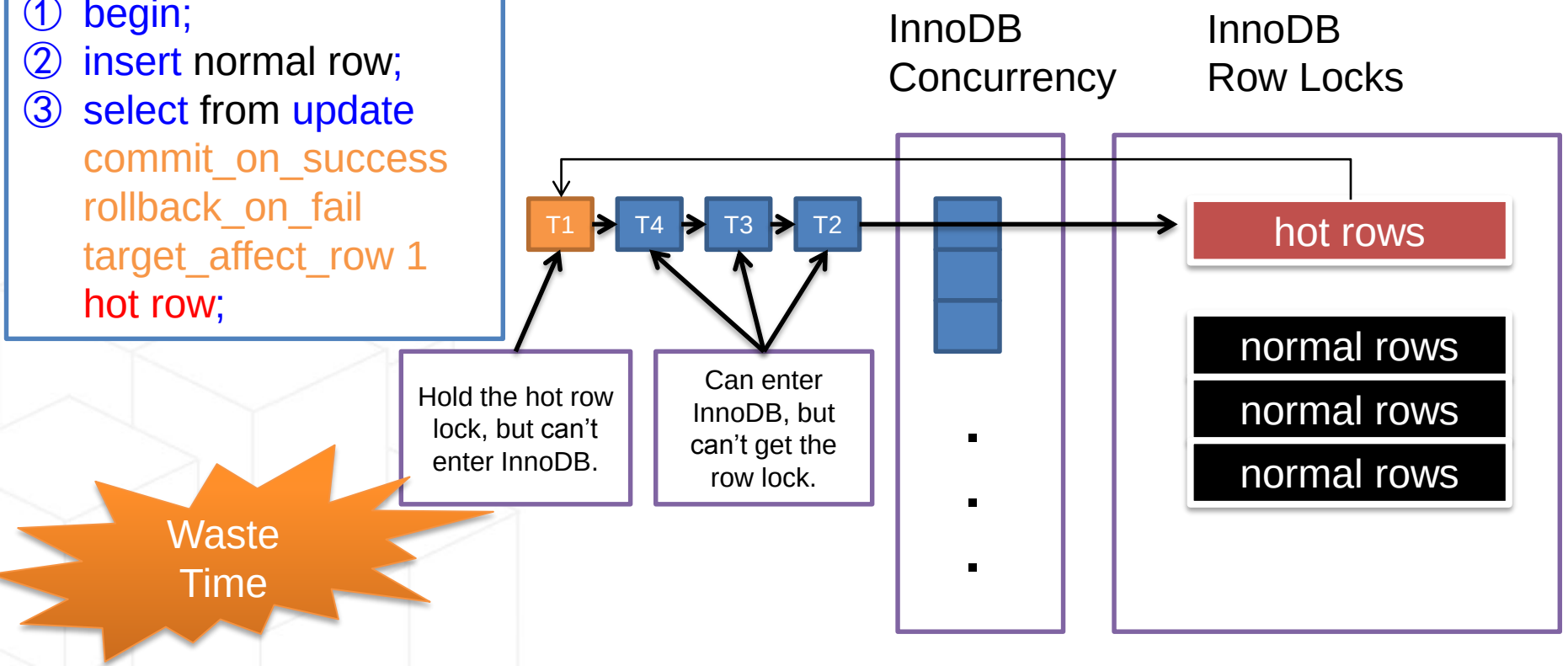
Transaction model:
① begin;
② insert normal row;
③ select * from update commit_on_success rollback_on_fail target_affect_row 1 hot row;

2nd Generation: InnoDB Strict Concurrency

Alibaba Group 阿里巴巴集团 | 阿里技术保障 Ali Infrastructure Service

Transaction model:

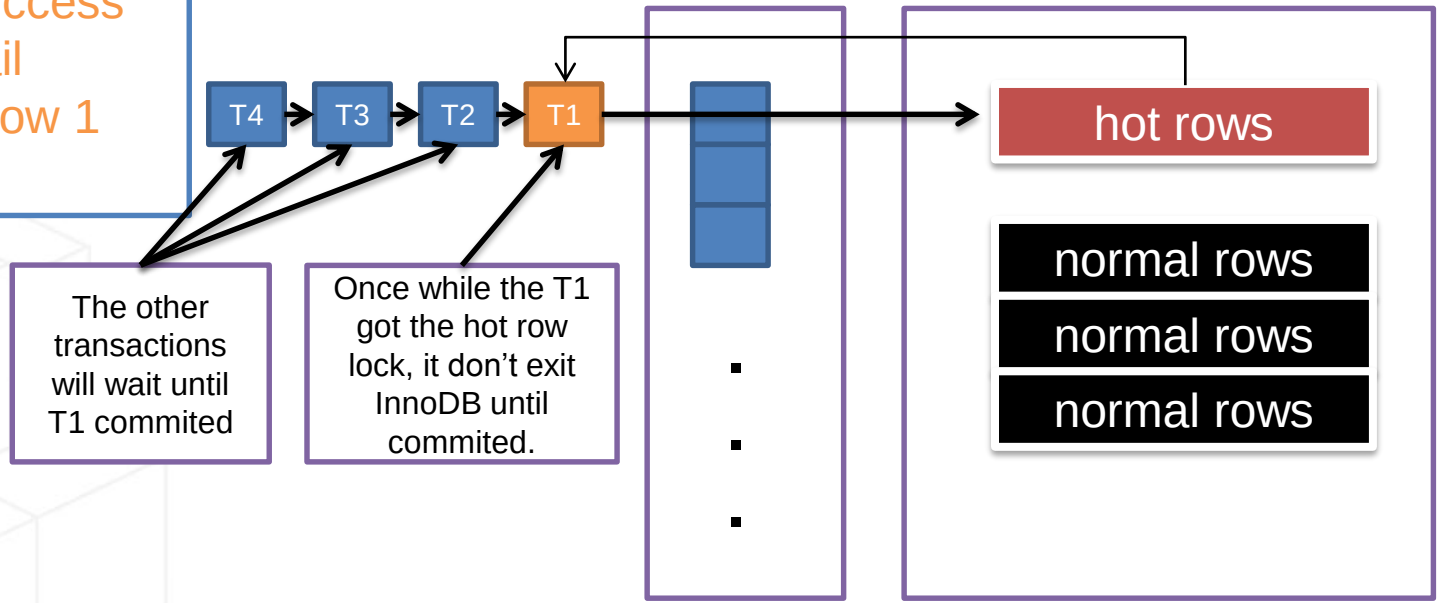
① begin;
② insert normal row;
③ select from update
commit_on_success
rollback_on_fail
target_affect_row 1
hot row;



2nd Generation: InnoDB Strict Concurrency

Transaction model:
① begin;
② insert normal row;
③ select from update
commit_on_success
rollback_on_fail
target_affect_row 1
hot row;

InnoDB Strict Concurrency



3rd Generation (In development)

comparison test

MySQL vs V1 vs V2 vs V3

commit_on_success hints

Group prepare

Enhanced Thread Pool

Select from Update

Hot SKU Optimization

Dual redo log buffers

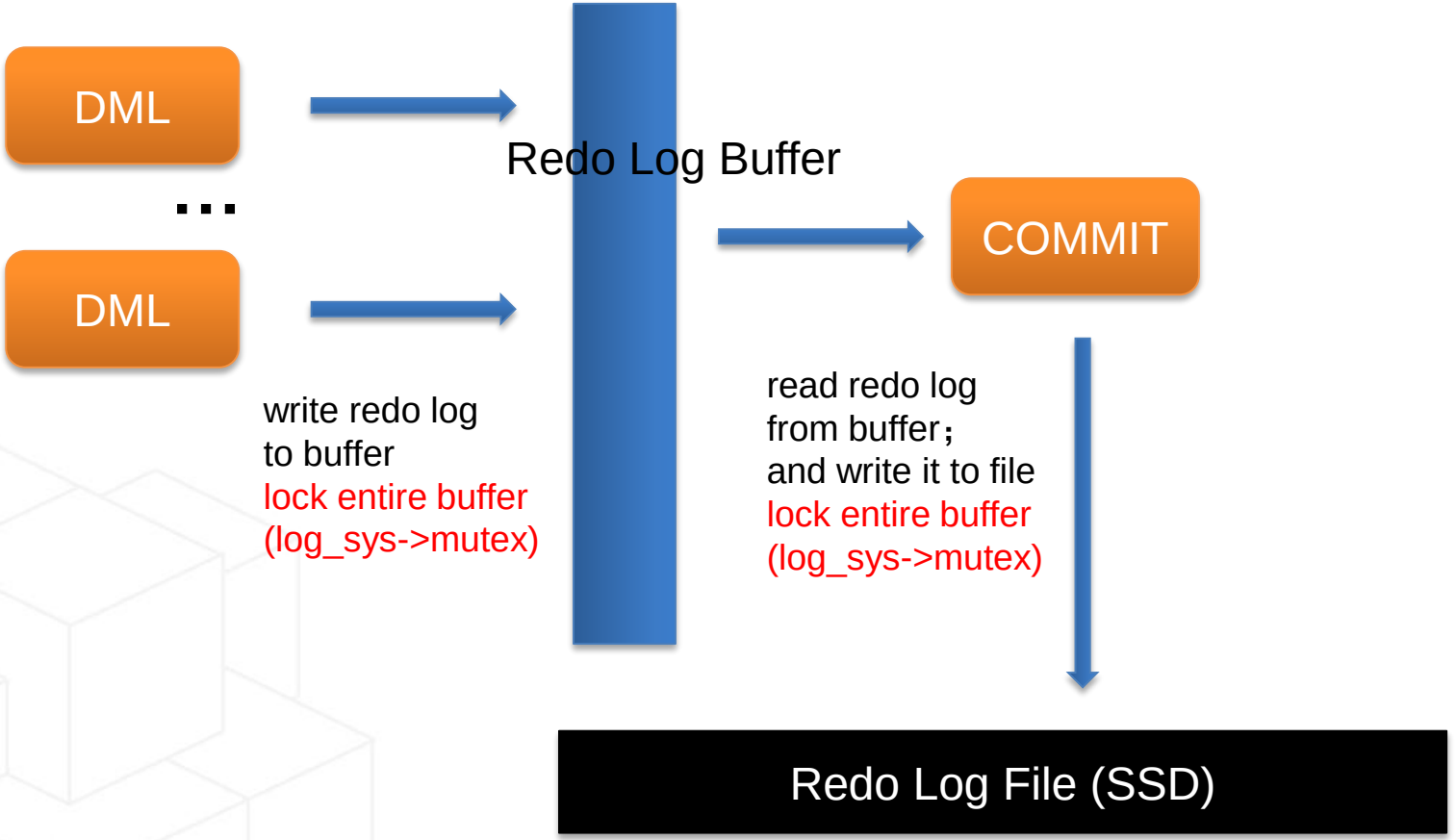
Table-level Parallel Replication

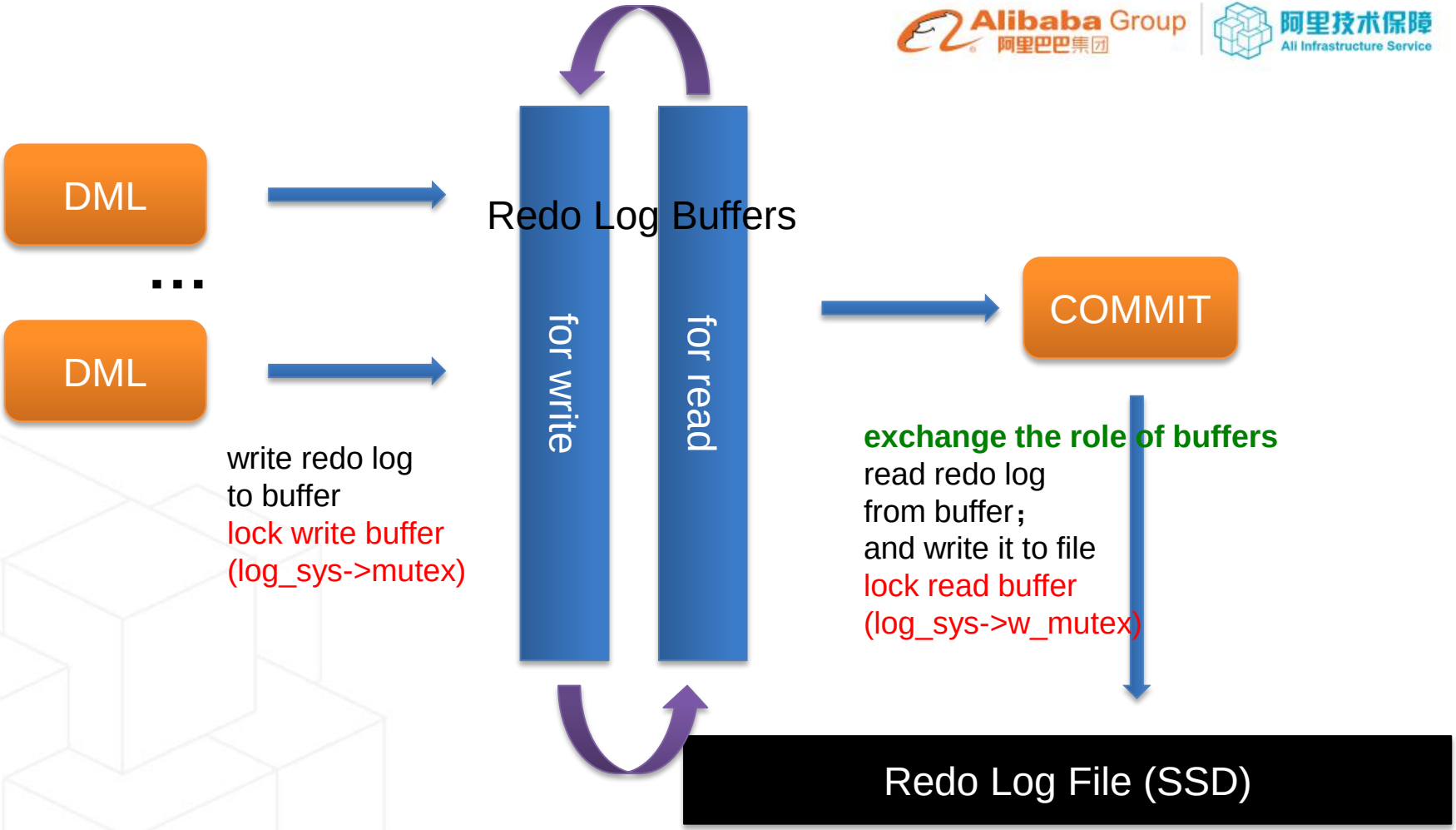
PK access optimization

Table Lock optimization for select

Enhanced Semi-Sync Replication

Race Condition of log_sys->mutex




DML
...
DML
Redo Log Buffers
for write
for read
COMMIT
write redo log
to buffer
lock write buffer
(log_sys->mutex)
exchange the role of buffers
read redo log
from buffer;
and write it to file
lock read buffer
(log_sys->w_mutex)
Redo Log File (SSD)

commit_on_success hints

Group prepare

Enhanced Thread Pool

Select from Update

Hot SKU Optimization

Dual redo log buffers

Table-level Parallel Replication

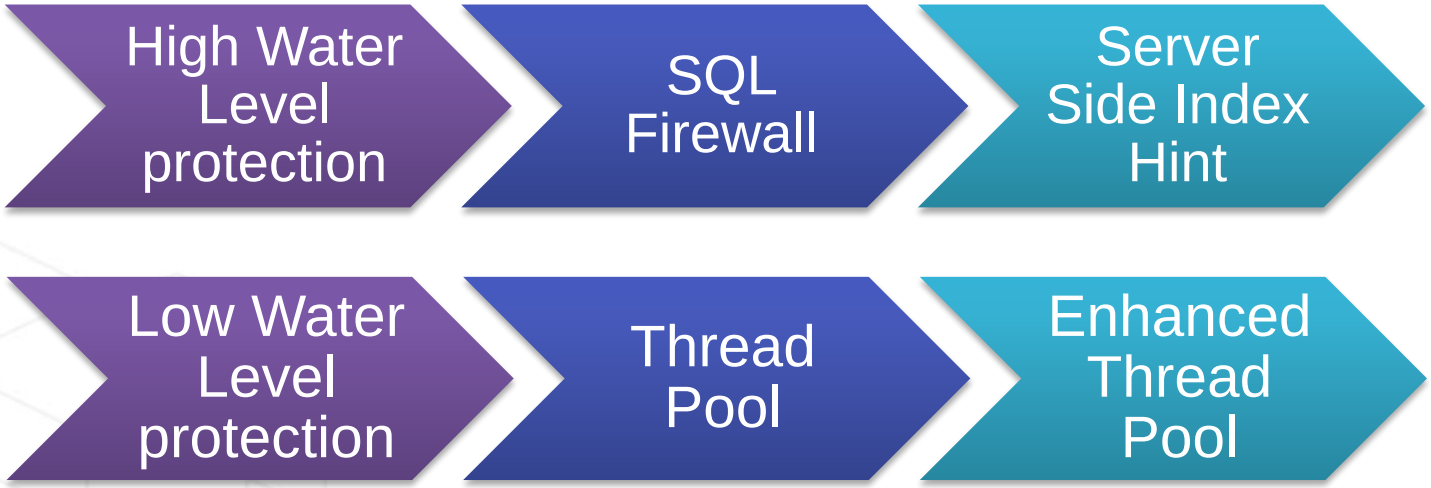
PK access optimization

Table Lock optimization for select

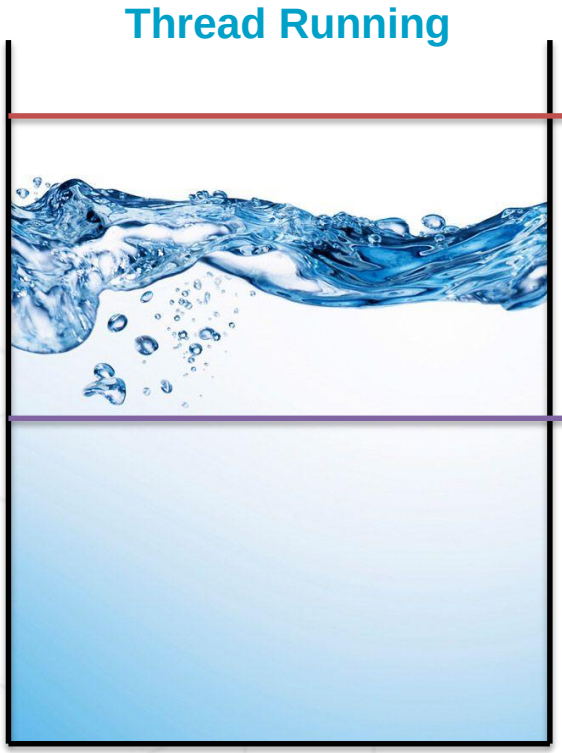
Enhanced Semi-Sync Replication

Version	AliSQL with TP				AliSQL without TP				MySQL 5.6.24 without TP			
Scenarios	L1	L2	L3	L4	L1	L2	L3	L4	L1	L2	L3	L4
QPS	33.0K	34.4K	44.3K	56.6K	34.0K	34.3K	46.7K	27.3K	16.1K	7.8K	NA	NA
TPS	11.0K	11.5K	14.8K	18.8K	11.4K	11.4K	15.4K	9.7K	5K	3K	NA	NA
RT	1.3ms	1.4ms	6.4ms	38.8ms	4.2ms	4.6ms	10.3ms	128.2ms	105.1ms	435.2ms	NA	NA
CS	23W	23W	26W	33W	30W	31W	33W	36W	40W	47W	NA	NA

Robustness


High Water Level protection
SQL Firewall
Server Side Index Hint
Low Water Level protection
Thread Pool
Enhanced Thread Pool

High/Low Water Level Protection



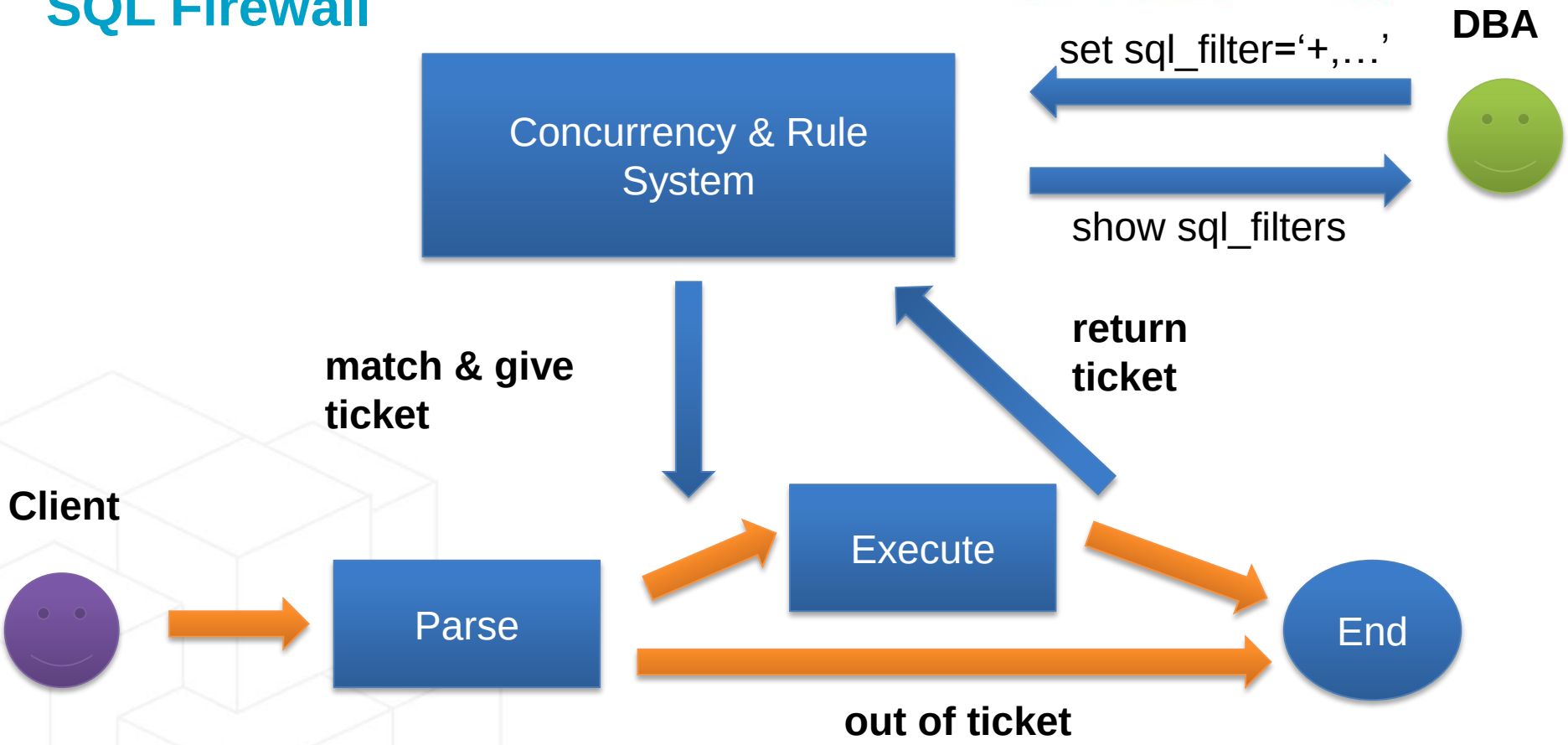
High Water Level Protection

Deny the query indiscriminately, while the thread running exceed the water level.

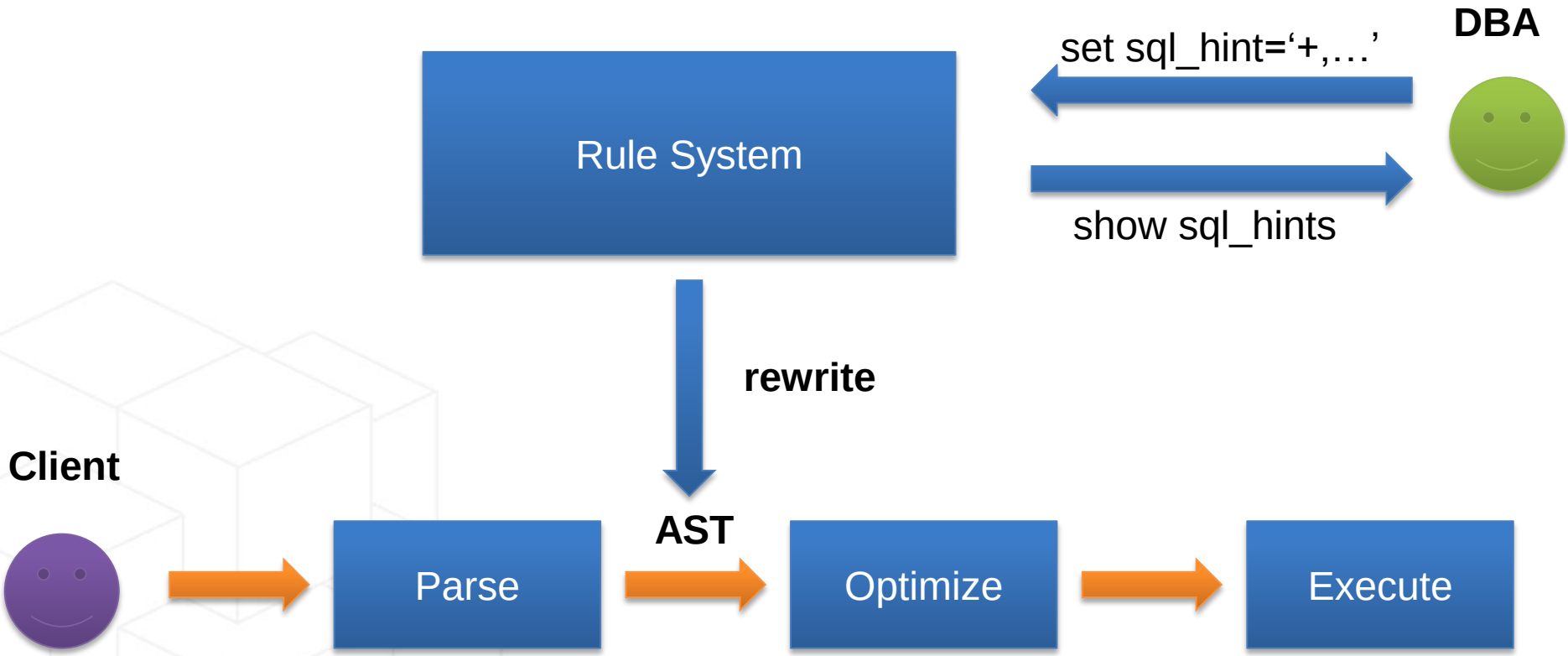
Low Water Level Protection

Buffer the query while the thread running exceed the low water level, deny the query which have been buffer a very long time.

SQL Firewall



Server side index hint



Other Protection Methods

Alibaba Group 阿里巴巴集团 | 阿里技术保障 Ali Infrastructure Service

Binlog speed limitation

Smart slave exec mode

Enrich Online DDL Type Support

Enrich Deadlock Information

more

Security & Others

Operation SQLs of Recycle Bin

- ✓ DROP TABLE [PURGE]
- ✓ FLASHBACK TABLE [TO BEFORE DROP] [RENAME TO <new_tablename>]
- ✓ PURGE TABLE <table_id>
- ✓ PURGE RECYCLEBIN

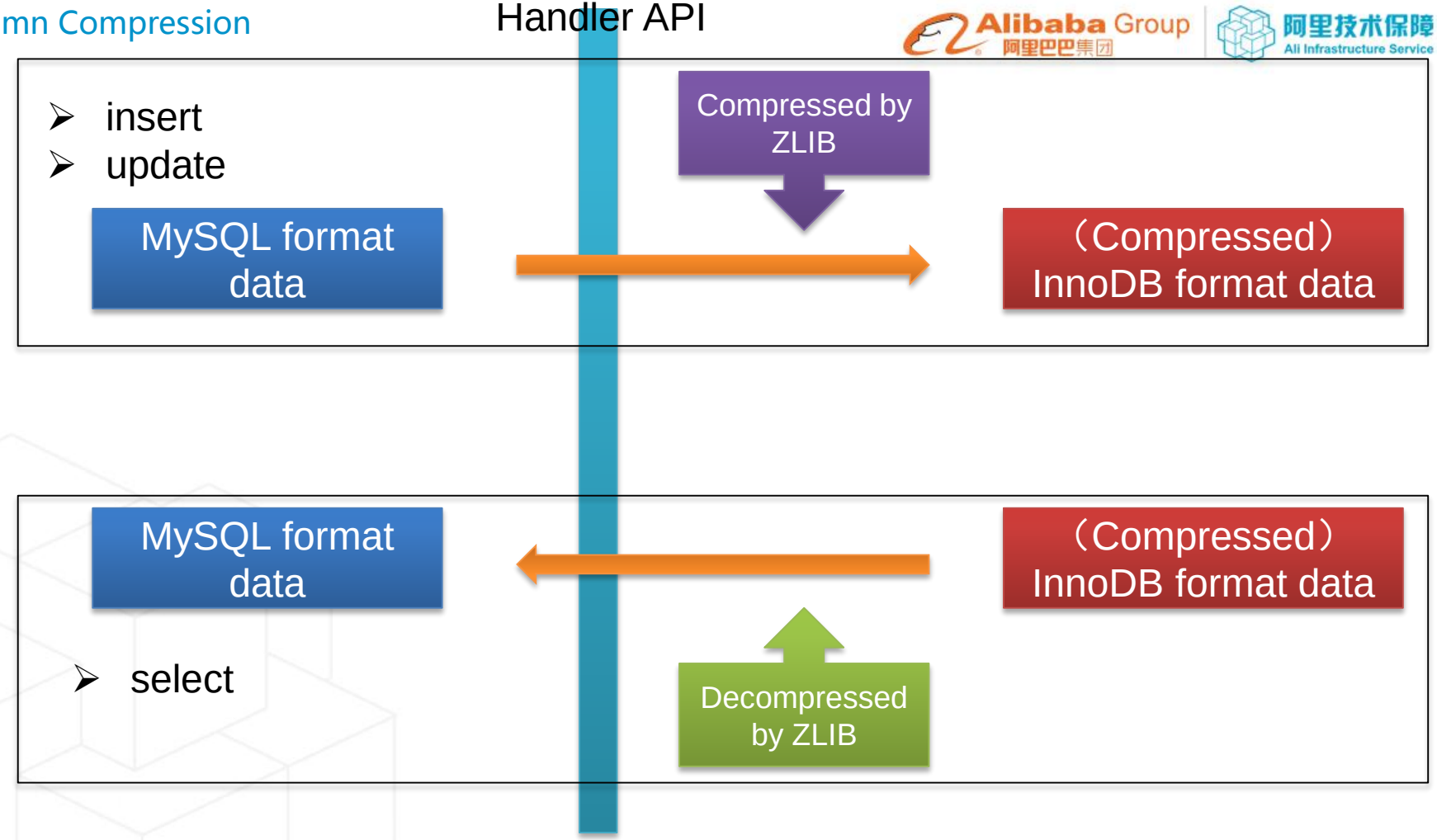
Process of drop tables

1. Judge if the variable recycle_bin is ON
2. Generate the autoincrement table id
3. Collect the other meta information (data size, drop time)
4. Write the meta info to the table mysql.recycle_bin_info
5. Rename the dropped table to recycle_bin.<table_id>

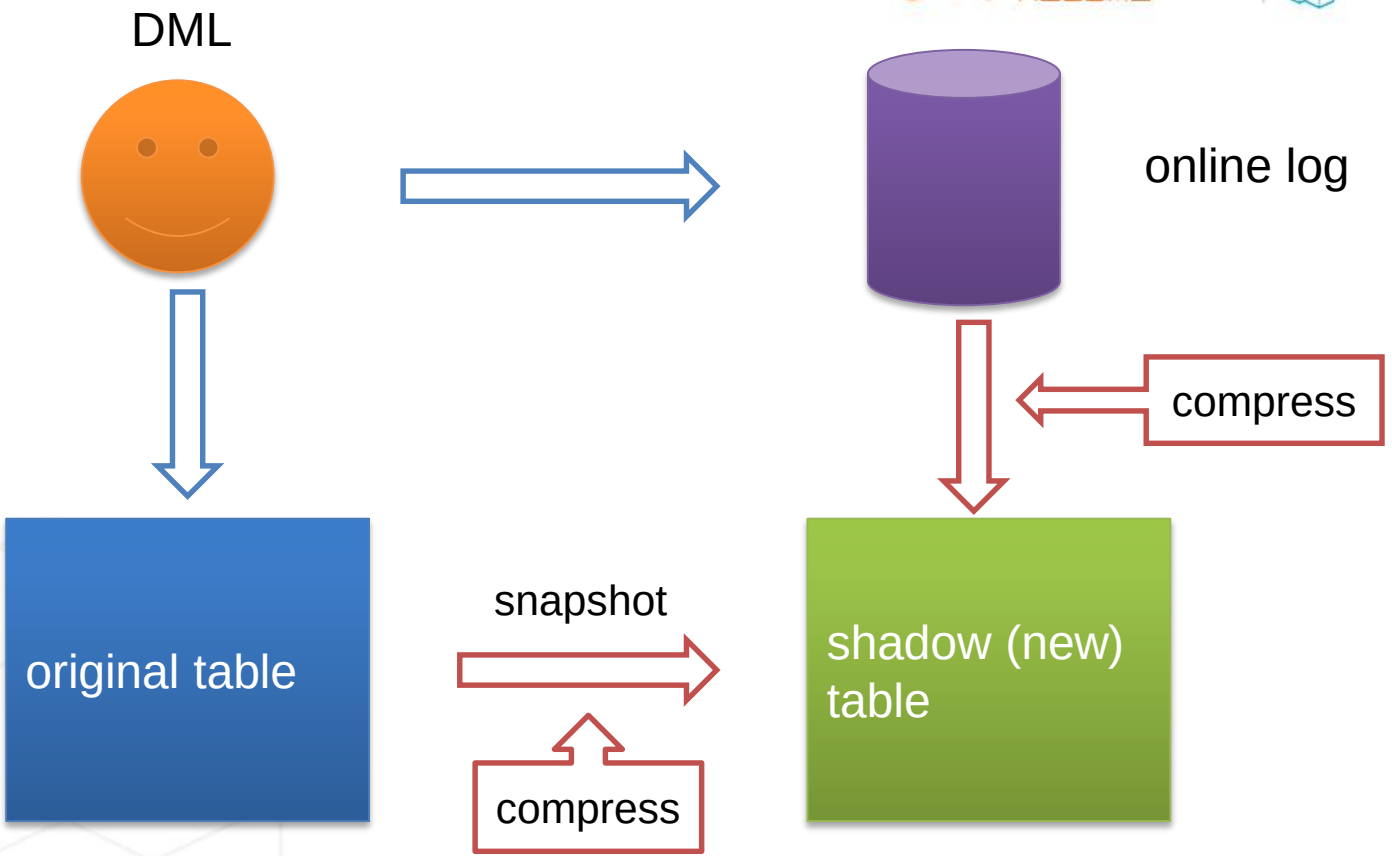
Process of flashback a table

1. Delete the corresponding row in the table mysql.recycle_bin_info
2. Rename the table from recycle_bin.<table_id> to original_tablename or new_tablename

Column Compression



Online Column Compression



cooperation

Q & A

Thank you


阿里技术保障
Ali Infrastructure Service