

本文是作者在ACMUG 2016 MySQL年会上的演讲内容，版权归作者所有。

中国MySQL用户组（China MySQL User Group）简称ACMUG。
ACMUG是覆盖中国MySQL技术爱好者的一个技术社区，是Oracle User Group Community和MairaDB Foundation共同认可的MySQL技术社区。

我们关注MySQL，MariaDB，以及其他一切周边的开源数据库和开源工具，我们交流使用经验，推广开源技术，为开源贡献力量。

我们是开放社区，欢迎任何关注MySQL及其相关技术的人加入，我愿意跟其他任何技术组织和团体保持沟通和展开合作。

我们期望在我们的活动中大家都能以开心的、轻松的姿态交流技术，分享技术，形成一个良性循环，从而每个人都可以有一份收获。

ACMUG的口号：开源，开放，开心

关注ACMUG公众号，参与社区活动，交流开源技术，分享学习心得，一起共同进步。



The Best Practice on AliCloud RDS for MySQL

Bo Wang (襄洛)
Senior Development Engineer, AliCloud.

Agenda

- Product Model
- Data Migration
- Database Monitoring
- Data Security

Agenda

Product Model

Basic Architecture

High Availability Service

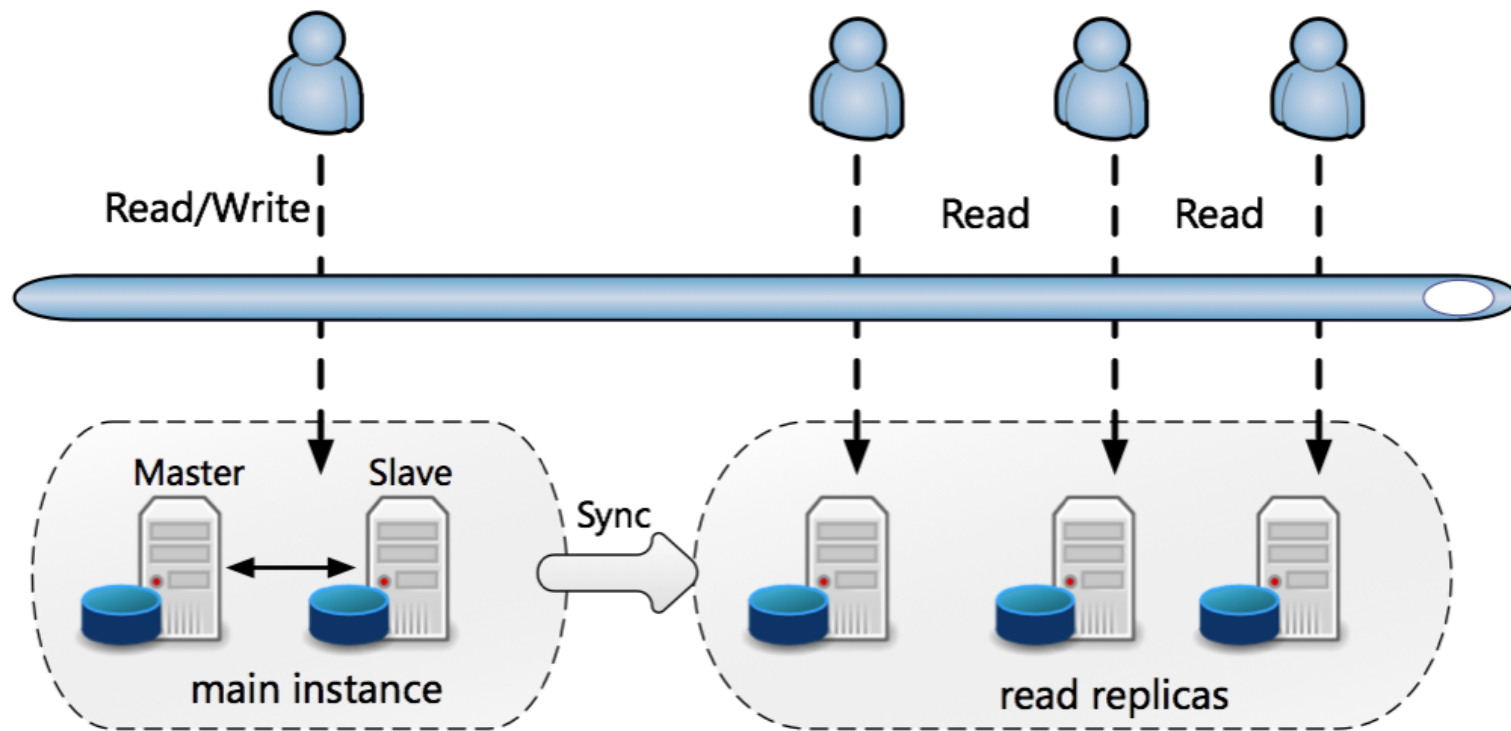
Read/write Separation

Backup & Recover

Elastic Upgrade

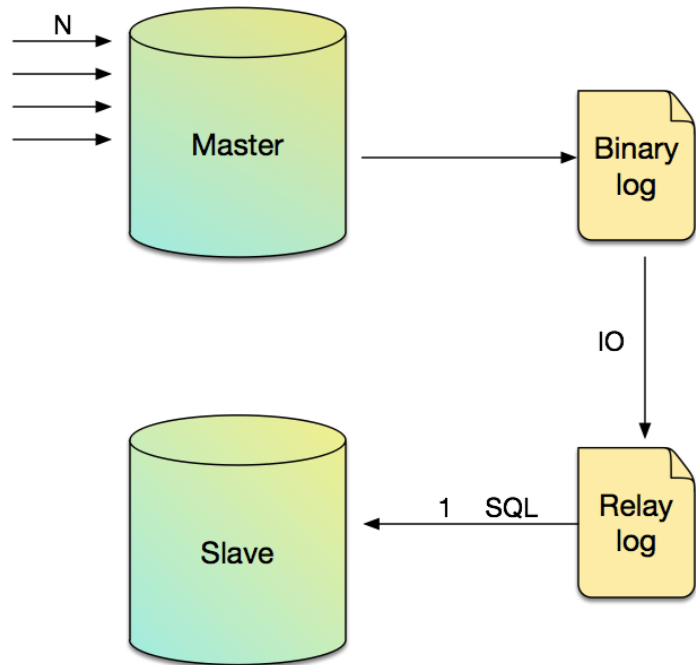
Monitoring & Notification

Performance Diagnostics



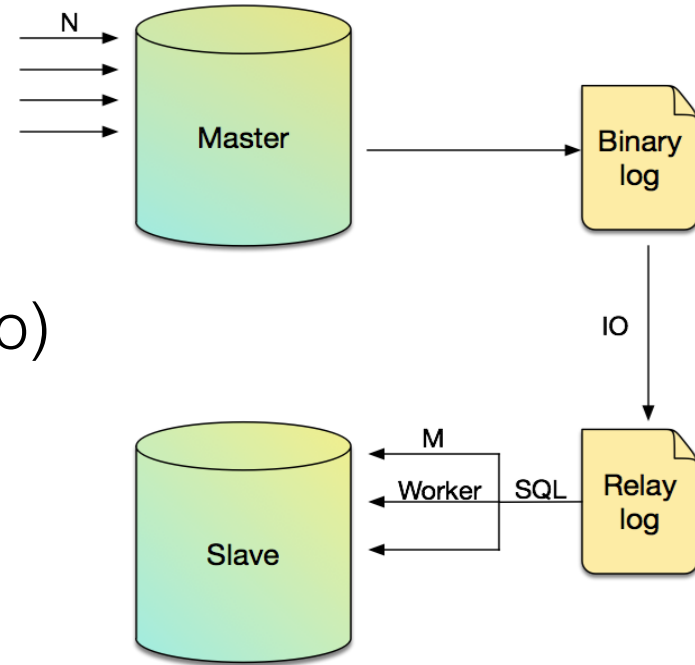
Basic Architecture - Replication

1. Multi Thread Slave



N:1

- logical replication(No)
- parallelism(Yes)



N:M

slave_parallel_workers
slave_pr_mode [SCHEMA|TABLE|TRX]

Basic Architecture - Replication

2. Implicit Key

```
mysql> SHOW CREATE TABLE t1\G
***** 1. row *****
      Table: t1
Create Table: CREATE TABLE `t1` (
  `id` int(11) DEFAULT NULL,
  `name` varchar(30) DEFAULT NULL,
  `__#alibaba_rds_row_id#__` bigint(20) NOT NULL AUTO_INCREMENT COMMENT 'Implicit Primary Key by RDS',
  KEY `__#alibaba_rds_row_id#__` (`__#alibaba_rds_row_id#__`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8
```

Basic Architecture - Replication

3. Minor Log

```
# at 67969776
#160930 11:15:24 server id 1732686928 end_log_pos 67969824 CRC32 0x5e77b4a9 GTID [commit=yes]
SET @@SESSION.GTID_NEXT= 'ebfc443c-8518-11e6-b6c3-7cd30ac3345e:53562'/*!*/;
# at 67969824
#160930 11:15:24 server id 1732686928 end_log_pos 67969895 CRC32 0x44f27e8d Query thread_id=1
SET TIMESTAMP=1475205324/*!*/;
BEGIN
/*!*/;
# at 67969895
#160930 11:15:24 server id 1732686928 end_log_pos 67969952 CRC32 0x80ab78c9 Table_map: `sb1`
# at 67969952
#160930 11:15:24 server id 1732686928 end_log_pos 67970366 CRC32 0x71aea2e0 Update_rows: table_id=1
BINLOG '
zNjtVxNQtEZnOQAAAKAJDQQAAMAAAAAAEAA3NiMQAic2J0ZXN0MTgABAMD/v4E7mj+tADJJeKuA
zNjtVxhQtEZnngEAAD41DQQAAMAAAAAAEABP//8NgOAAADkFgAAdwA1MjExODMwMzM0OC02MTky
NzE1OTY4OC05MjQ0NzNmM2NDk5NS05ODc2NzEzNjM2MS0xODc0MDI4MDkzMzY0YzZlODQyMjY0x
```

binlog of main instance

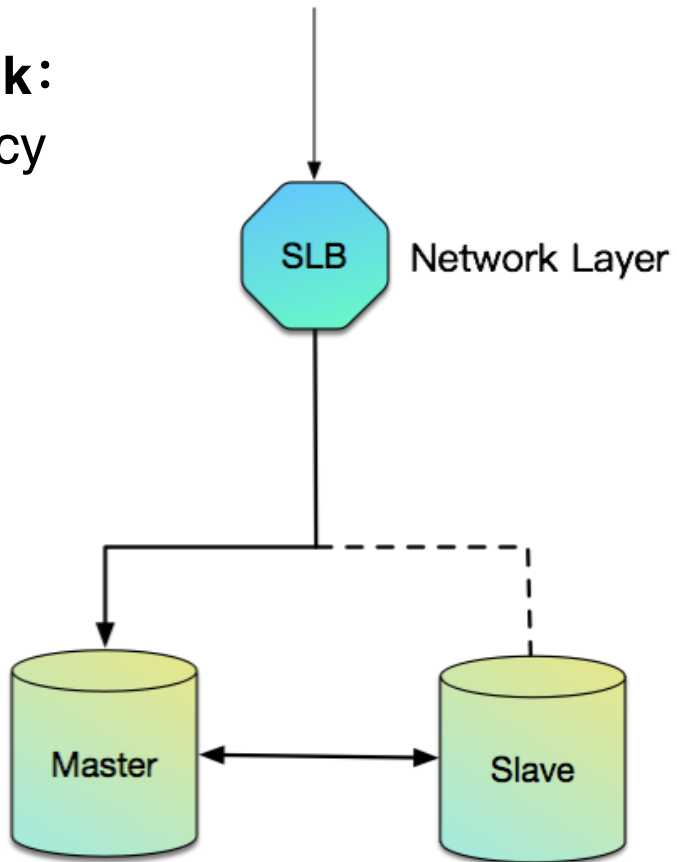
```
# at 11877398
#160930 11:15:24 server id 1732686928 end_log_pos 11877446 CRC32 0x984ab4e1 GTID |
SET @@SESSION.GTID_NEXT= 'ebfc443c-8518-11e6-b6c3-7cd30ac3345e:53562'/*!*/;
# at 11877446
#160930 11:15:24 server id 1732686928 end_log_pos 11877509 CRC32 0x63327628 Query
SET TIMESTAMP=1475205324/*!*/;
BEGIN
/*!*/;
# at 11877509
#160930 11:15:24 server id 1732686928 end_log_pos 11877573 CRC32 0x3cf807c5 Query
SET TIMESTAMP=1475205324/*!*/;
COMMIT
/*!*/;
```

binlog of replica

Basic Architecture - Data Link

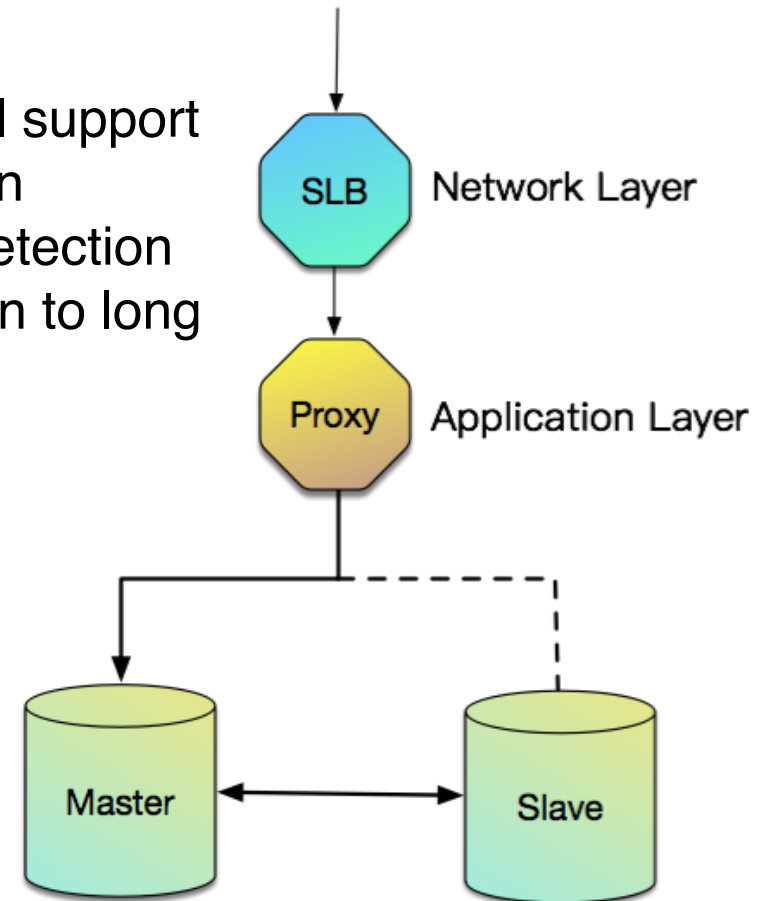
1. Data link choice

Basic Link:
Low latency



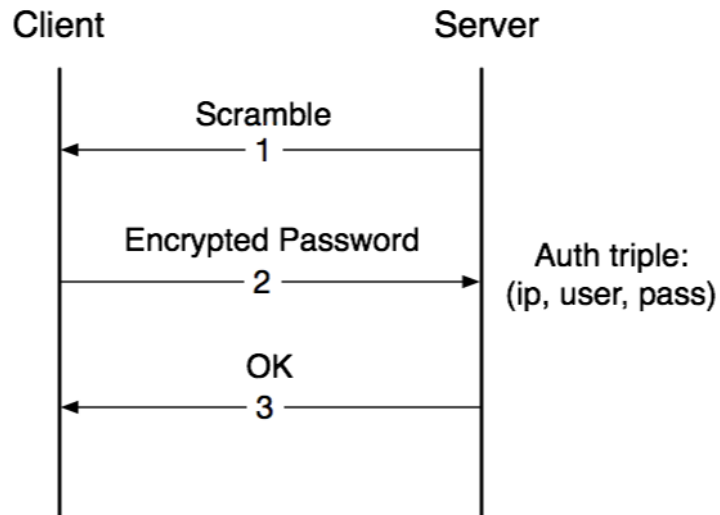
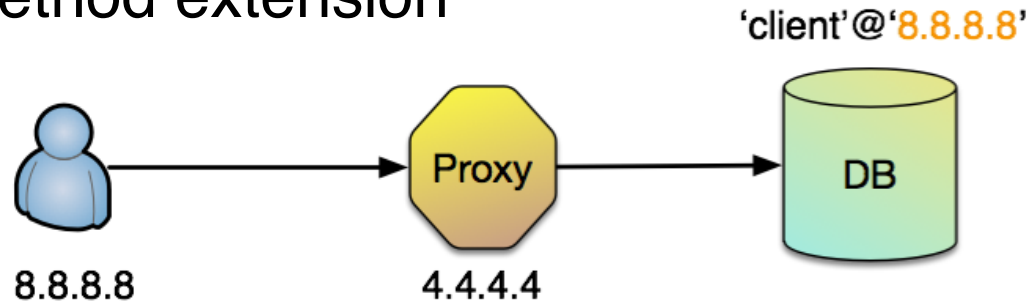
Security Link:

MySQL protocol support
Keep connection
SQL injection detection
Short connection to long
+20% latency

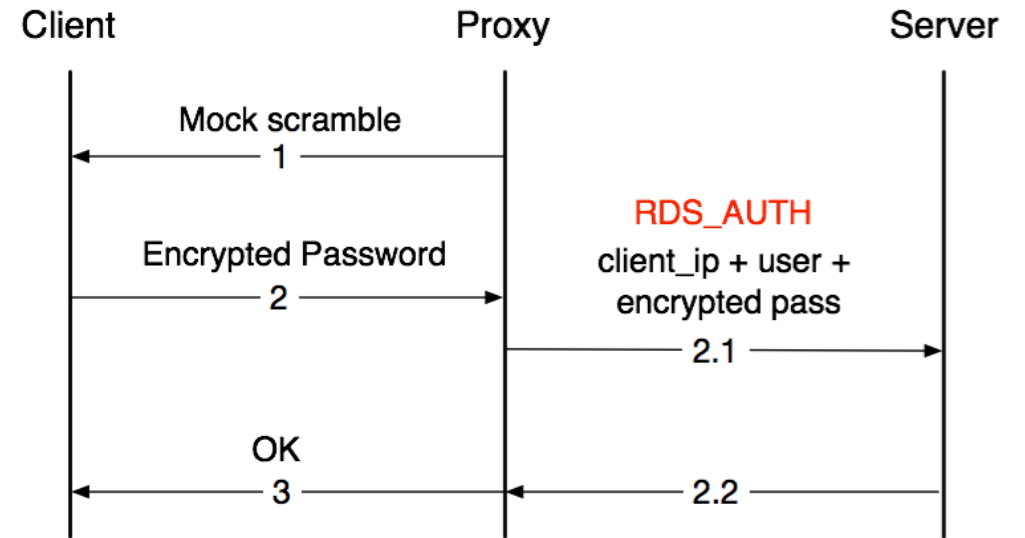


Basic Architecture - Data Link

2. Authentication method extension



original authentication method



original authentication method + RDS_AUTH

Basic Architecture - Data Link

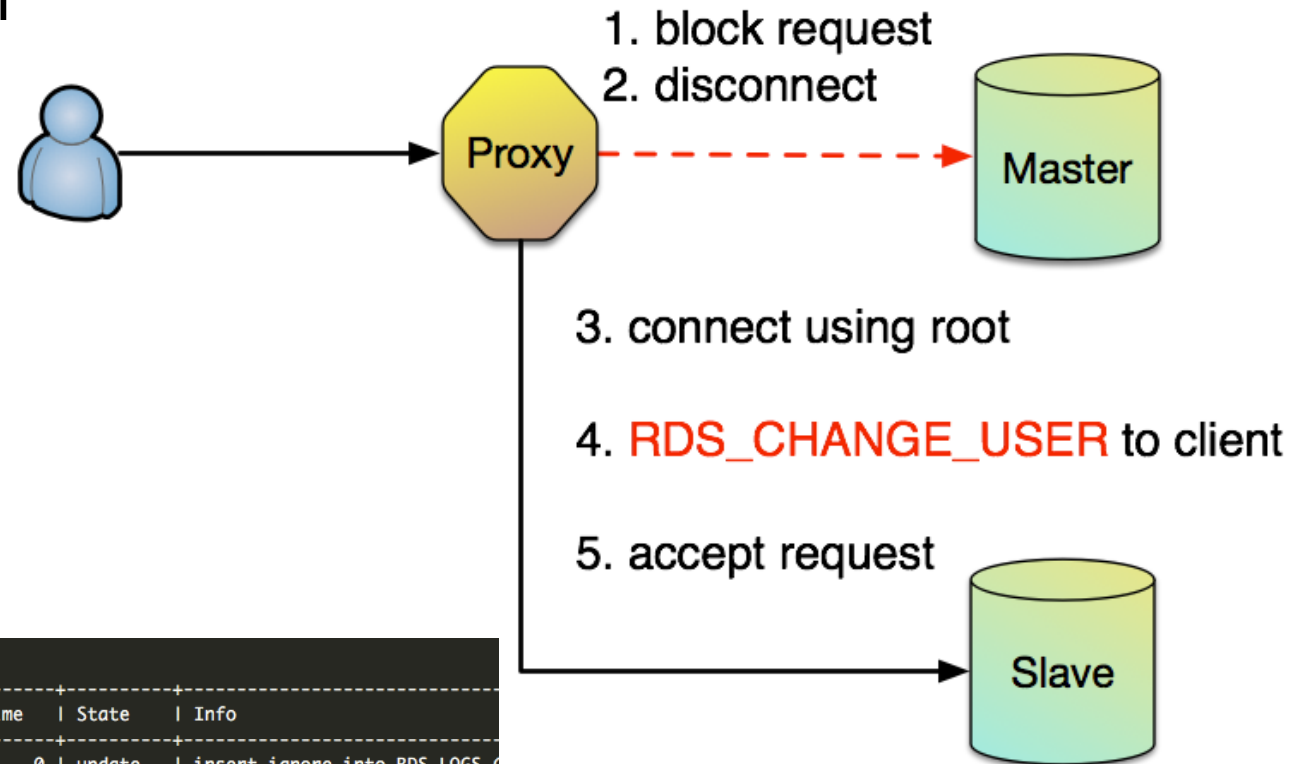
3. Switch over and keep connection

- RDS_CHANGER_USER command without user password

- rds set connection

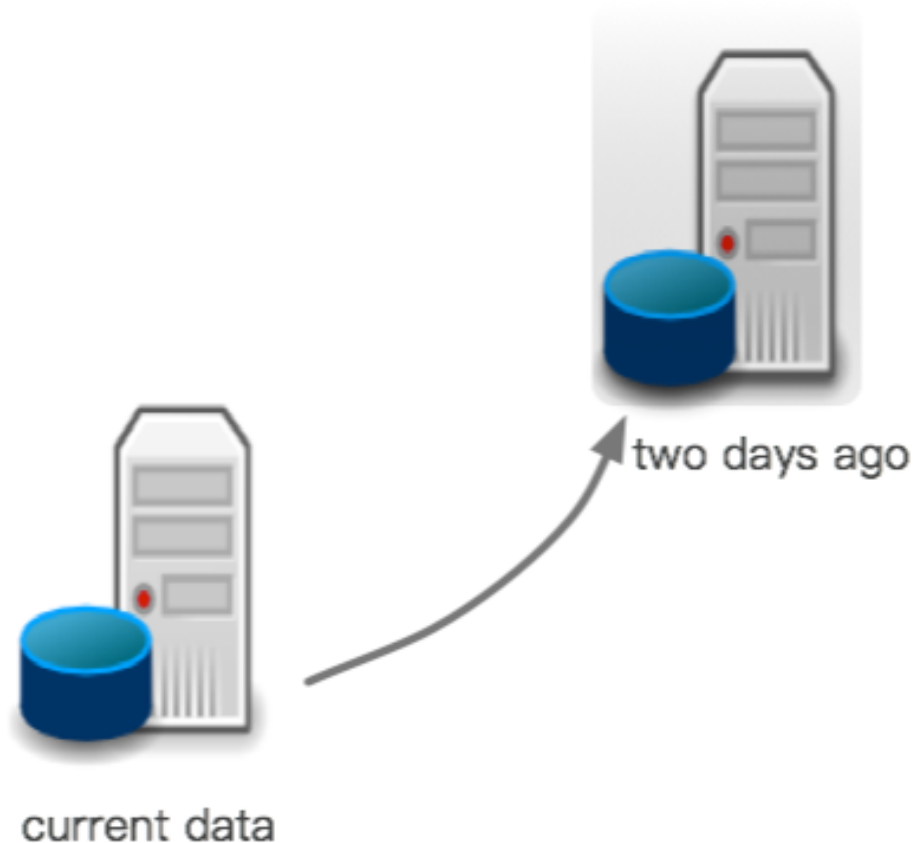
```
mysql> SHOW PROCESSLIST;
```

Id	User	Host	db	Command	Time	State	Info
113002	dn_tlognfs1	10.151.204.164:22285	tlognfs0	Query	0	update	insert ignore into RDS_LOGS C
113007	dn_tlognfs1	10.151.204.133:24139	tlognfs0	Query	0	update	insert ignore into RDS_LOGS C
113013	dn_tlognfs1	10.151.204.132:55170	tlognfs0	Sleep	0		NULL
113017	dn_tlognfs1	10.151.204.132:55171	tlognfs0	Query	0	update	insert ignore into RDS_LOGS C
113021	dn_tlognfs1	10.151.204.132:55172	tlognfs0	Query	0	update	insert ignore into RDS_LOGS C
113030	dn_tlognfs1	10.151.204.164:22286	tlognfs0	Query	0	update	insert ignore into RDS_LOGS C
113034	dn_tlognfs1	10.151.204.133:24140	tlognfs0	Query	0	update	insert ignore into RDS_LOGS C
113036	dn_tlognfs1	10.151.204.133:24141	tlognfs0	Query	0	update	insert ignore into RDS_LOGS C
113038	dn_tlognfs1	10.151.204.164:22287	tlognfs	Sleep	3494		NULL
113039	dn_tlognfs1	10.151.204.164:22288	tlognfs0	Query	0	update	insert ignore into RDS_LOGS C



Clone instance

1. generate multiple instance in batch
e.g. business scale out
2. point in time recover
e.g. important data deleted by mistake



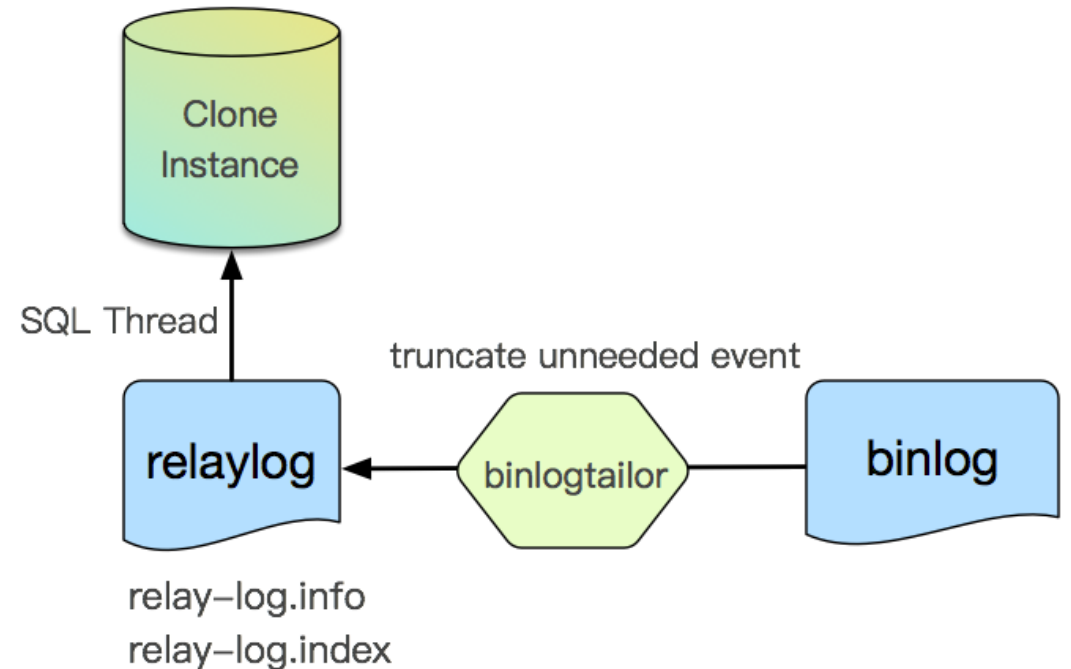
Clone instance

Full Backup + binlog incremental

1. usual way, mysqlbinlog tool decode binlog to executable SQL
`mysqlbinlog mysql-bin.000001 | mysql`

- single thread, slow
- error_code in Query_log_event
- event size >1G

2. improved way, treat binlog as relaylog, applied by SQL thread(MTS)



Clone instance

Flashback

- binlog ordered with transaction
- rows_log_event contains Before Image and After Image
- DDL not supported

```
### INSERT INTO `test`.`xp`
### SET
###   @1=1 /* INT meta=0 nullable=0 is_null=0 */
###   @2='pre' /* VARSTRING(100) meta=100 nullable=1 is_null=0 */

### UPDATE `test`.`xp`
### WHERE
###   @1=1 /* INT meta=0 nullable=0 is_null=0 */
###   @2='pre' /* VARSTRING(100) meta=100 nullable=1 is_null=0 */
### SET
###   @1=1 /* INT meta=0 nullable=0 is_null=0 */
###   @2='post' /* VARSTRING(100) meta=100 nullable=1 is_null=0 */

### DELETE FROM `test`.`xp`
### WHERE
###   @1=1 /* INT meta=0 nullable=0 is_null=0 */
###   @2='post' /* VARSTRING(100) meta=100 nullable=1 is_null=0 */
```

TokuDB Engine

- High compress ratio (Case: 2.7T - > 400G)
 - Delay writes (put msg)
 - Transactional ACID + MVCC
 - Secondary Clustering Index
 - Replace into, Read Free Replication
-
- TokuDB community not active like InnoDB
 - Sharp checkpoint, lock contention, etc.

```
CREATE TABLE ... ENGINE=TokuDB  
ROW_FORMAT=TOKUDB_SNAPPY;
```

- adjust tokudb_buffer_pool_ratio
- Convert Archive to TokuDB

Feature	Archive	TokuDB
DML	only INSERT	yes
Transactions	no	yes
ACID	no	yes
Indexes	no	yes
Online DDL	no	yes

Agenda

Data Migration

mysqldump

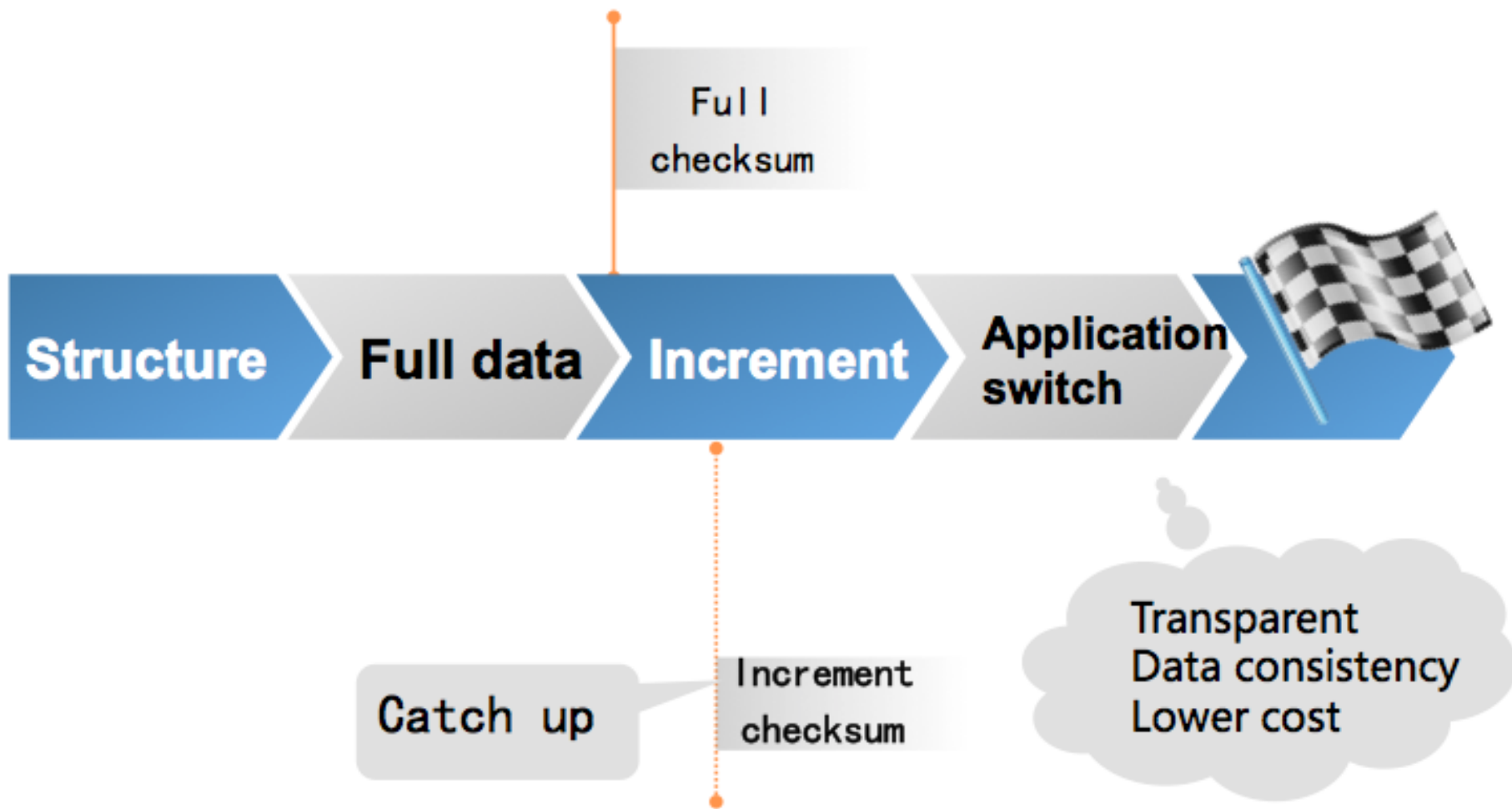
1. dump from local source db

```
mysqldump -h localhost -u userName -p --opt --default-character-set=utf8  
--hex-blob dbName --skip-triggers > /tmp/dbName.sql
```

2. load into remote RDS db

```
mysql -h intranet4example.mysql.rds.aliyuncs.com -u userName -p  
dbName < /tmp/dbName.sql
```

Data Transfer Service(DTS)



Agenda

Database Monitoring

Basic monitoring

Process info: CPU, Mem, Disk usage, IOPS, etc.

MySQL metrics: Connections, TPS/QPS, COMDML, Innodb_rows, Tokudb_rows, replica(SBM), etc.

Threshold notification: Disk, IOPS, Connections, CPU, etc.

Table and index statistics

```
mysql> SELECT * FROM INFORMATION_SCHEMA.TABLE_STATISTICS;
```

TABLE_SCHEMA	TABLE_NAME	ROWS_READ	ROWS_CHANGED	ROWS_CHANGED_X_INDEXES	ROWS_INSERTED	ROWS_DELETED	ROWS_UPDATED
test	sbtest1	1076141	17846	35692	2536	2536	12774

```
mysql> SELECT * FROM INFORMATION_SCHEMA.INDEX_STATISTICS;
```

TABLE_SCHEMA	TABLE_NAME	INDEX_NAME	ROWS_READ
test	sbtest1	k_1	66
test	sbtest1	PRIMARY	1076075

Table statistics: read and write of business, scale out by more instance, or more replica.

Index statistics: frequency of each index, drop unused index.

SQL Mem/IO/Temp

```
Id: 39
User: xiangluo
Host: 30.40.13.11:56566
db: test
Command: Query
Time: 0
State: updating
Info: UPDATE sbtest1 SET c='64862974741'
Memory_used: 2175648
Memory_used_by_query: 8248
Logical_read: 12246
Physical_sync_read: 0
Physical_async_read: 0
Temp_user_table_size: 0
Temp_sort_table_size: 0
Temp_sort_file_size: 0
```

Memory statistics:

Memory_used

Memory_used_by_query

IO statistics:

Logical_read

Physical_sync_read

Physical_async_read

Temp space statistics:

Temp_user_table_size

Temp_sort_table_size

Temp_sort_file_size

SQL audit

SQL日志明细

SQL日志报告

SQL注入检测

慢SQL统计(按天)

慢SQL明细

主键统计

索引统计

存储引擎

大表优化

缺失索引

其它

日期时间:

2016-09-28 16:51:37



-

2016-09-29 16:51:37



关键字:

查询

SQL归档

SQL归档查询

SQL语句	执行时间	数据库用户名	客户端IP	消耗时间(微秒)	返回记录
SET NAMES 'utf8'	2016-09-29 16:51:26	wwb wwb	10.47.160.210	32	0

SQL and resources used of each request

- host, user, db, tid, sql
- start_time of trx/stmt, latency, lock wait time

- MEM / IO / Temp space
- error_code

Diagnose report

1. resource usage
2. dead lock, lock wait
3. long time transaction, uncommitted transaction
4. TOP SQL (latency, scan rows, update rows, frequency)

综合评分：64分



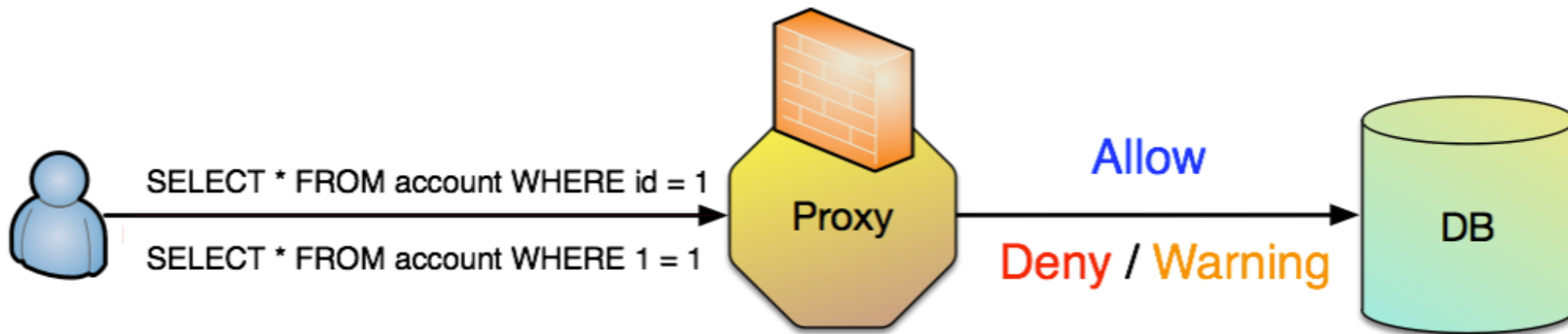
RDS DBA Service!

Agenda

Data Security

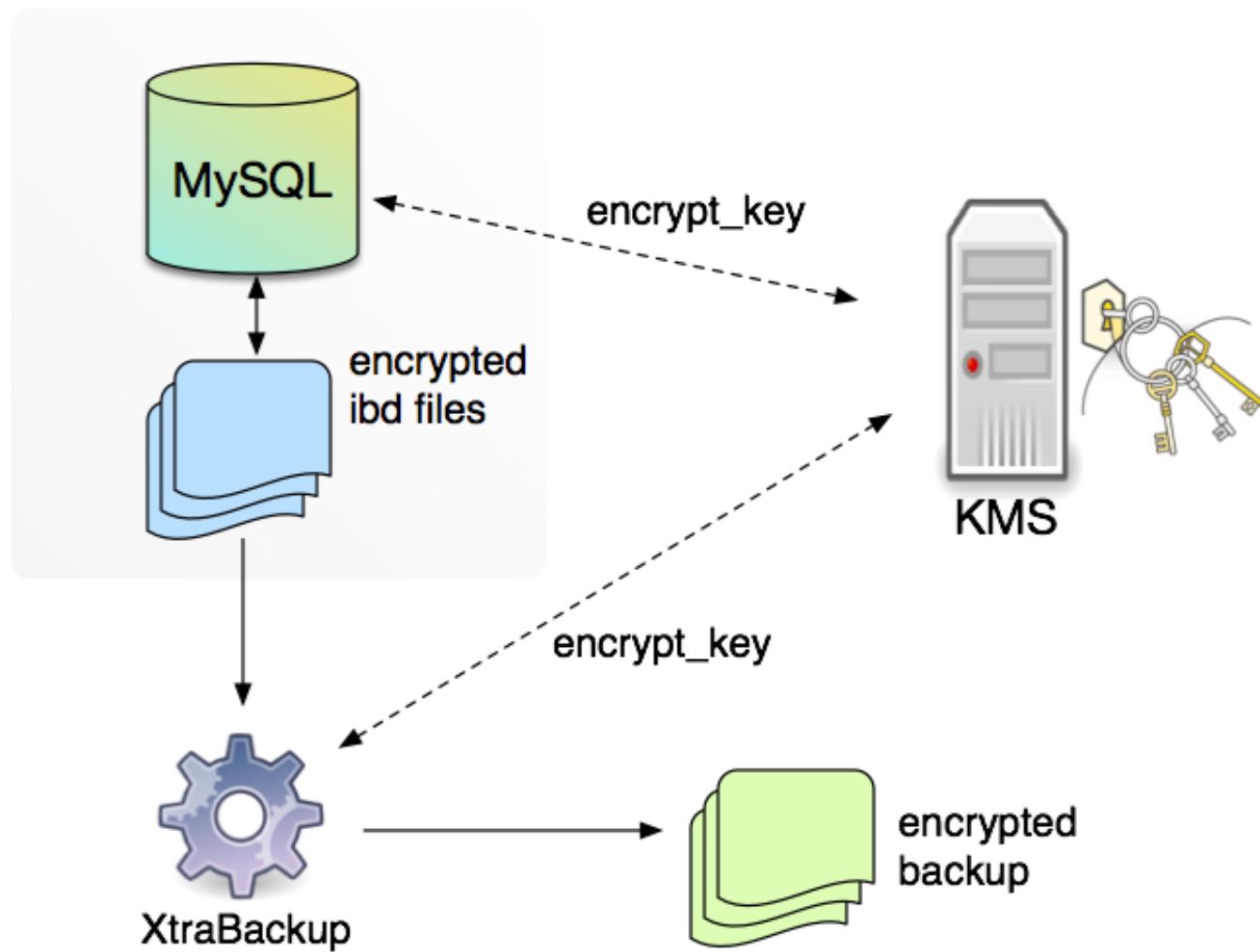
Data link security

1. IP whitelist
2. SSL support
3. SQL firewall



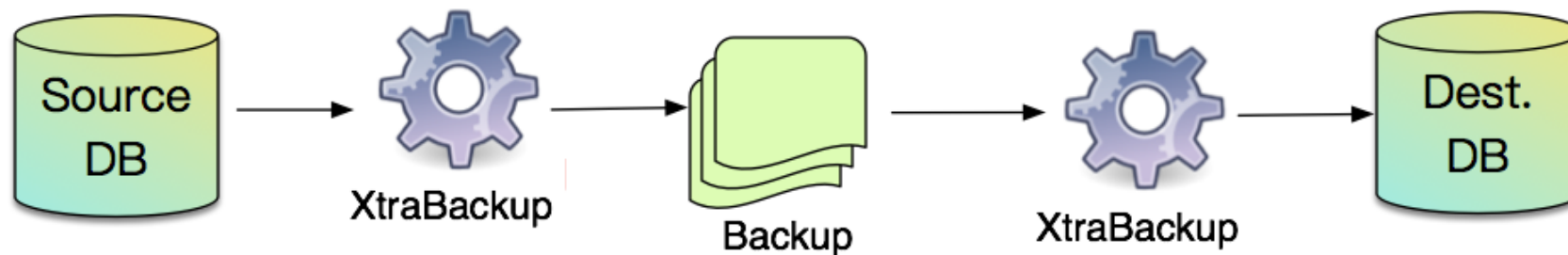
Transparent Data Encryption(TDE)

1. ibd file encrypted by block (AES)
2. backup encrypted
3. Key Management System
4. encrypt_key never persisted



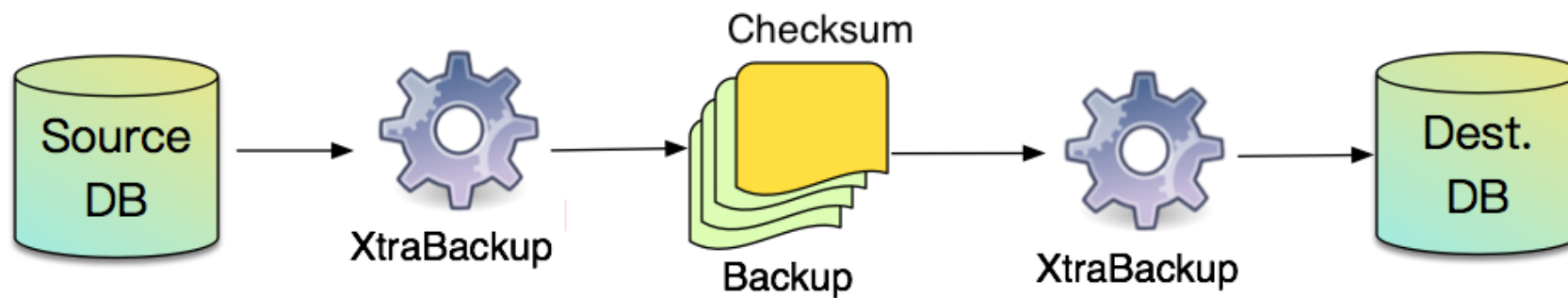
Backup checksum

backup set validation



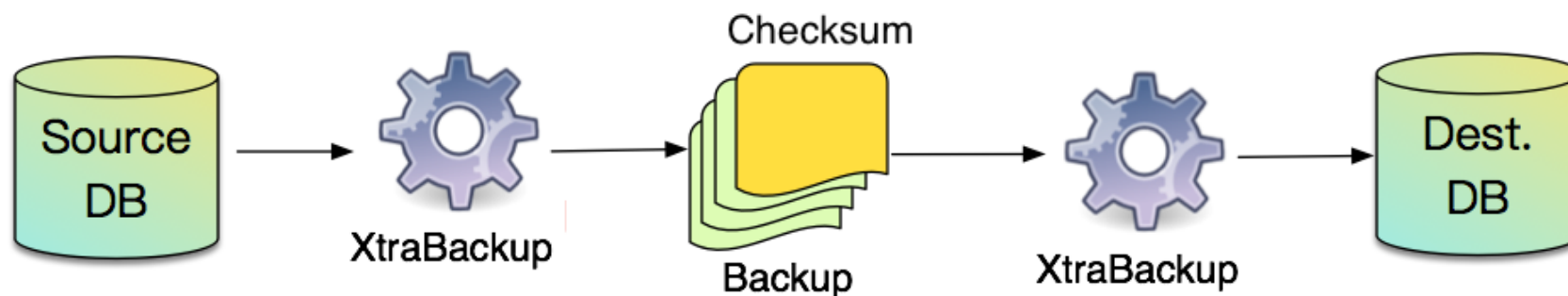
Backup checksum

backup set validation



Backup checksum

backup set validation



- IOPS threshold, SET SESSION `rds_sql_max_iops=100`;
- Buffer pool LRU, CHECKSUM `ENGINE_NO_CACHE` TABLE t1;

Contact us ^_^

1. AliSQL (Code)

<https://github.com/alibaba/AliSQL>

<https://github.com/alibaba/AliSQLBackup>

2. Database Kernel Monthly (technical article on database)

<http://mysql.taobao.org/monthly/>

3. @阿里丁奇

Thanks



Q&A

为了无法计算的价值 |  阿里云

