
Table of Contents

Introduction	1.1
	1.2
TiDB	
1 TiDB	2.1
2	2.2
3	2.3
3.1 Key-Value	2.3.1
3.2	2.3.2
3.3 SQL	2.3.3
4	2.4
4.1	2.4.1
4.2	2.4.2
5 TiDB MySQL	2.5
6 TiDB	2.6
6.1	2.6.1
6.2	2.6.2
6.3 4.0	2.6.3
7 TiDB DDL	2.7
7.1	2.7.1
7.2 DDL	2.7.2
7.3 Sequence	2.7.3
7.4 AutoRandom	2.7.4
8 Titan	2.8
8.1 Titan	2.8.1
8.2 TiDB Titan	2.8.2
9 TiFlash HTAP	2.9
9.1 TiDB HTAP	2.9.1
9.2 TiFlash	2.9.2
9.3 TiFlash	2.9.3
10 TiDB	2.10
10.1	2.10.1
10.2 RBAC	2.10.2
10.3	2.10.3
11 TiSpark	2.11
11.1 TiSpark	2.11.1
11.2 TiSpark	2.11.2
11.3 TiSpark on TiFlash	2.11.3

11.4 TiSpark	2.11.4
<hr/>	
1 &	3.1
1.1 TiUP & TiOps	3.1.1
 1.1.1 TiUP	3.1.1.1
 1.1.2 TiUP	3.1.1.2
 1.1.3 TiOps	3.1.1.3
 1.1.4 TiOps	3.1.1.4
1.2 TiDB on Kubernetes	3.1.2
 1.2.1 TiDB-Operator	3.1.2.1
 1.2.2 TiDB-Operator Kind	3.1.2.2
 1.2.3 TiDB-Operator	3.1.2.3
 1.2.3.1	3.1.2.3.1
 1.2.3.1.1 AWS EKS TiDB	3.1.2.3.1.1
 1.2.3.1.2 GCP GKE TiDB	3.1.2.3.1.2
 1.2.3.1.3 TiDB	3.1.2.3.1.3
 1.2.3.1.4 TiDB	3.1.2.3.1.4
 1.2.3.2	3.1.2.3.2
 1.2.3.2.1	3.1.2.3.2.1
 1.2.3.2.2 PV	3.1.2.3.2.2
 1.2.3.2.3 TiDB Operator	3.1.2.3.2.3
 1.2.3.2.4 TiDB	3.1.2.3.2.4
 1.2.3.2.5	3.1.2.3.2.5
 1.2.3.2.6 TiDB	3.1.2.3.2.6
 1.2.4 Kubernetes TiDB	3.1.2.4
 1.2.5 Kubernetes BR TiDB	3.1.2.5
 1.2.6 Kubernetes Lightning	3.1.2.6
 1.2.7 Kubernetes TiDB	3.1.2.7
 1.2.8 TiDB-Operator	3.1.2.8
1.3	3.1.3
 1.3.1 TiOps	3.1.3.1
 1.3.2 TiDB-Operator	3.1.3.2
1.4	3.1.4
 1.4.1 TiOps	3.1.4.1
 1.4.2 TiDB-Operator	3.1.4.2
1.5	3.1.5
2 TiDB	3.2
 2.1 4.0 CDC	3.2.1
 2.1.1 CDC	3.2.1.1
 2.1.2 CDC	3.2.1.2

2.1.3 CDC	3.2.1.3
2.2 TiDB Lightning	3.2.2
2.2.1 Lightning	3.2.2.1
2.2.2 Lightning	3.2.2.2
2.3 4.0 BR	3.2.3
2.3.1 BR	3.2.3.1
2.3.2 BR	3.2.3.2
2.4 4.0 Dumpling	3.2.4
2.4.1 Dumpling	3.2.4.1
2.4.2 Dumpling	3.2.4.2

TiDB Troubleshooting

1 SQL	4.1
1.1 TiDB	4.1.1
1.2	4.1.2
1.3 SQL Plan Management	4.1.3
1.4	4.1.4
1.5 SQL	4.1.5
2 TiDB Dashboard	4.2
2.1	4.2.1
2.2 SQL	4.2.2
2.3	4.2.3
2.4	4.2.4
2.5 CPU	4.2.5
3	4.3
3.1	4.3.1
3.2	4.3.2
3.3	4.3.3
3.4	4.3.4
3.5 SQL	4.3.5
3.6 Processlist	4.3.6
3.7 Statement Summary	4.3.7
4 TiDB	4.4
4.1	4.4.1
4.2 TiDB	4.4.2
4.4 Prometheus	4.4.3
5	4.5
5.1 GC	4.5.1
5.2 Recover/Flashback	4.5.2
5.3	4.5.3
6 TiDB-Operator Troubleshooting	4.6

TiDB

1		5.1
2		5.2
3		5.3
	3.1 Sysbench	5.3.1
	3.2 TPC-C	5.3.2
4		5.4
	4.1 binlog	5.4.1
	4.2 Raft	5.4.2
	4.3	5.4.3
	4.4 AWS AZ TiDB	5.4.4
5		5.5
	5.1 MySQL TiDB DM	5.5.1
	5.1.1 DM MySQL TiDB	5.5.1.1
	5.1.2 DM MySQL TiDB	5.5.1.2
	5.2 Oracle TiDB OGG	5.5.2
	5.3 SqlServer TiDB	5.5.3
	5.4 SqlServer TiDBDATAAX	5.5.4
	5.5 DB2 TiDB CDC	5.5.5
	5.6 TiDB TiDBDATAAX	5.5.6
	5.7 Mongodb TiDB	5.5.7
6		5.6
	6.1	5.6.1
	6.1.1	5.6.1.1
	6.1.2 TiDB	5.6.1.2
	6.1.3	5.6.1.3
	6.1.4	5.6.1.4
	6.1.5 hint	5.6.1.5
	6.2 SQL	5.6.2
	6.3 TiDB + TiSpark	5.6.3
	6.4	5.6.4
	6.4.1 TiDB	5.6.4.1
	6.4.2 TiDB	5.6.4.2
	6.4.3 TiDB	5.6.4.3
	6.4.4 TiDB	5.6.4.4
	6.5 TiDB	5.6.5
	6.6 TiDB HBaseESDruid	5.6.6
	6.7 TiDB SaikuGrafana	5.6.7
7		5.7
	7.1 Oncall	5.7.1
	7.2	5.7.2

7.3 TiKV is busy	5.7.3
7.4 TiDB OOM	5.7.4
7.5 TiKV	5.7.5
8 TiDB	5.8
8.1 TiDB	5.8.1
8.2 TiKV	5.8.2
8.2.1 TiKV	5.8.2.1
8.2.2 Region	5.8.2.2
8.2.3	5.8.2.3
8.3	5.8.3
8.3.1 TiDB	5.8.3.1
8.3.2	5.8.3.2

TiDB

1 TiDB	6.1
1.1 TiDB	6.1.1
1.2 TiDB	6.1.2
1.3 TiDB	6.1.3
1.4 TiDB	6.1.4
2 TiDB	6.2
2.1	6.2.1
2.1.1 TiDB Devcon	6.2.1.1
2.1.2 TiDB TechDay	6.2.1.2
2.1.3 Infra Meetup	6.2.1.3
2.1.4 TiDB Hackathon	6.2.1.4
2.1.5 TUG	6.2.1.5
2.2 TUGTiDB User Group	6.2.2
2.3 Talent Plan	6.2.3
2.4 Challenge Program	6.2.4
2.5 PingCAP Incubator	6.2.5
2.6 PingCAP University	6.2.6
2.7 AskTUG	6.2.7
2.8 Contributor Map	6.2.8

Gitbook

Read it: [TiDB In Action: based on 4.0](#)

Download PDF: [TiDB In Action: based on 4.0](#)

- [Project](#)
- [TODO](#)
- [Pull Request](#)

- [markdown](#).
- [Github](#)

```
res/{doc-path}/  
• {doc-path}  
  
• session1/chapter1/tidb-intro.md  
• /res/session1/chapter1/tidb-intro/  
• markdown ![1.png](/res/session1/chapter1/tidb-intro/1.png)
```

TiDB in Action

TiDB in Action

-
- TiDB

-
-
-

- -
 -

-
-
-
-
-
-
-

--

-
-
-
-
-
-

markdown

```
<!--TODO: /-->
```

License

Shield:  License CC BY-SA 4.0

This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](#).



1 TiDB

TiDB TiDB NewSQL TiDB

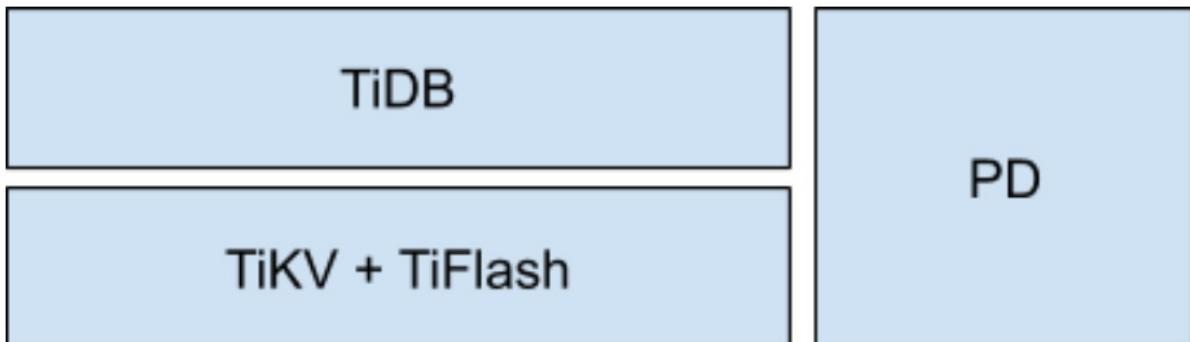
- SQL MySQL MySQL MySQL
- ACID
-

TiDB 4.0 TiDB

TiDB Google Spanner F1 TiDB



TiDB



1. TiDB (tidb-server, <https://github.com/pingcap/tidb>): SQL MySQL endpoint SQL TiDB TiDB TiDB tidb-server SQL TiKV
2. TiKV (tikv-server, <https://github.com/pingcap/tikv>): KV NoSQL TiDB TiKV NoSQL TiKV API KV SI Snapshot Isolation TiDB SQL TiDB SQL SQL TiKV API TiKV TiKV 3 TiFlash TiKV TiFlash
3. Placement Driver (pd-server PD <https://github.com/pingcap/pd>): TiDB TiKV Dashboard IDPD PD TiKV TiKV PD 3

2

TiDB TiKV

2.1 Key-Value Pairs ()

TiKV Key-Value TiKV

1. Map C++ std::map Key-Value Pairs
 2. Map Key-Value pair Key Seek Key Next Key Key-Value

SQL

TiKV KV SQL Table

SOL TiKV Key-Value

2.2 RocksDB

TiKV TiKV RocksDB RocksDB RocksDB Facebook KV TiKV RocksDB Key-Value Map

2.3 Raft

TiKV

TiKV Raft Raft Multi Paxos

Raft Raft Raft

1. Leader
 2. Leader
 - 3.

TiKV Raft Raft Raft Raft



RocksDB TiKV Raft Raft RocksDB RaftTiKV Key-Value Raft

2.4 Region

Region TiKV KV Map Raft Raft KV

- Hash Key Hash Hash
- Range Key Range Key

TiKV Key-Value Key Region Region TiKV 96MB Region [StartKeyEndKey)



Region SQL SQL KV Region TiKV

- Region Region
- Region Raft

Key Region Region TiDB PD Region Region PD Region Key Key Region
Region Key PD

TiKV Region Region TiKV ReplicaReplica Raft Region Replica Raft Group Replica Group Leader
Replica Follower Leader Leader Follower Region



Region KeyValue

2.5 MVCC

MVCCTiKV Key Value TiKV MVCC Key MVCC TiKV

```
Key1 -> Value
Key2 -> Value
.....
KeyN -> Value
```

MVCC TiKV Key

```
Key1_Version3 -> Value
Key1_Version2 -> Value
Key1_Version1 -> Value
.....
Key2_Version4 -> Value
Key2_Version3 -> Value
Key2_Version2 -> Value
Key2_Version1 -> Value
.....
KeyN_Version2 -> Value
KeyN_Version1 -> Value
.....
```

Key Key-Value Key Key + Version Value Key Version MVCC Key Key_Version RocksDB

SeekPrefix(Key_Version) API Key_Version

2.6 ACID

TiKV Google BigTable [Percolator](#) TiKV

TiKV API

```
tx = tikv.Begin()
    tx.Set(Key1, Value1)
    tx.Set(Key2, Value2)
    tx.Set(Key3, Value3)
tx.Commit()
```

3 Set TiKV TiDB SQL KV MySQL

NewSQL TiDB TiKV TiKV TiKV (Key, Value) TiDB TiDB SQL TiKV TiDB TiKV
TiFlash TiFlash [TiFlash HTAP](#))

3.1 Key-Value

TiDB (Key, Value)

1.
2.

3.1.1 Key-Value

(Key, Value) KeyOLTP Key ID ID OLAP Key

1. TiDB ID TableID ID
2. TiDB ID RowID ID IDTiDB TiDB ID

(Key, Value)

Key: tablePrefix{TableID}_recordPrefixSep{RowID}
Value: [col1, col2, col3, col4]

tablePrefix recordPrefixSep Key

3.1.2 Key-Value

TiDB TiDB ID IndexID

RowID (Key, Value)

Key: tablePrefix{tableID}_indexPrefixSep{indexID}_indexedColumnsValue
Value: RowID

RowID (Key, Value)

Key: tablePrefix{TableID}_indexPrefixSep{IndexID}_indexedColumnsValue_{RowID}
Value: null

3.1.3

tablePrefix recordPrefixSep indexPrefixSep Key

```
tablePrefix      = []byte{'t'}
recordPrefixSep = []byte{'r'}
indexPrefixSep  = []byte{'i'}
```

```
Key Key TiKV Key TiKV
      RowID TiKV Key indexedColumnsValue
      Key
```

3.1.4 Key-Value

TiDB Key-Value TiDB

```
CREATE TABLE User {
    ID int,
    Name varchar(20),
    Role varchar(20),
    Age int,
    PRIMARY KEY (ID),
    KEY idxAge (Age)
};
```

3

```
1, "TiDB", "SQL Layer", 10
2, "TiKV", "KV Engine", 20
3, "PD", "Manager", 30
```

(Key, Value) | int | RowID | TableID | 10 TiKV

```
t10_r1 --> ["TiDB", "SQL Layer", 10]
t10_r2 --> ["TiKV", "KV Engine", 20]
t10_r3 --> ["PD", "Manager", 30]
```

idxAge | IndexID | 1 TiKV

```
t10_i1_10_1 --> null
t10_i1_20_2 --> null
t10_i1_30_3 --> null
```

TiDB Key-Value

3.2

(Key, Value) TiDB Database Table TiDB TiKV

Database / Table ID ID Key-Value ID Key m_ KeyValue

TiDB (Key, Value) DDL 1TiDB pd-server etcd Key "/tidb/ddl/global_schema_version"Value int64 TiDB
Google F1 Online Schema etcd

3.3 SQL

TiDB SQL tidb-server Google

F1 SQL Key-Value Key-Value TiKV TiKV

3.3.1 SQL

Key-Value SQL KV KV

```
select count(*) from user where name = "TiDB"
1. Key Range RowID [0, MaxInt64] 0 MaxInt64 Key [StartKey, EndKey)
2. Key Range Key Range TiKV
3. name = "TiDB"
4. Count(*) Count(*)
```



Work Work

1. KV TiKV RPC
- 2.
- 3.

3.3.2 SQL

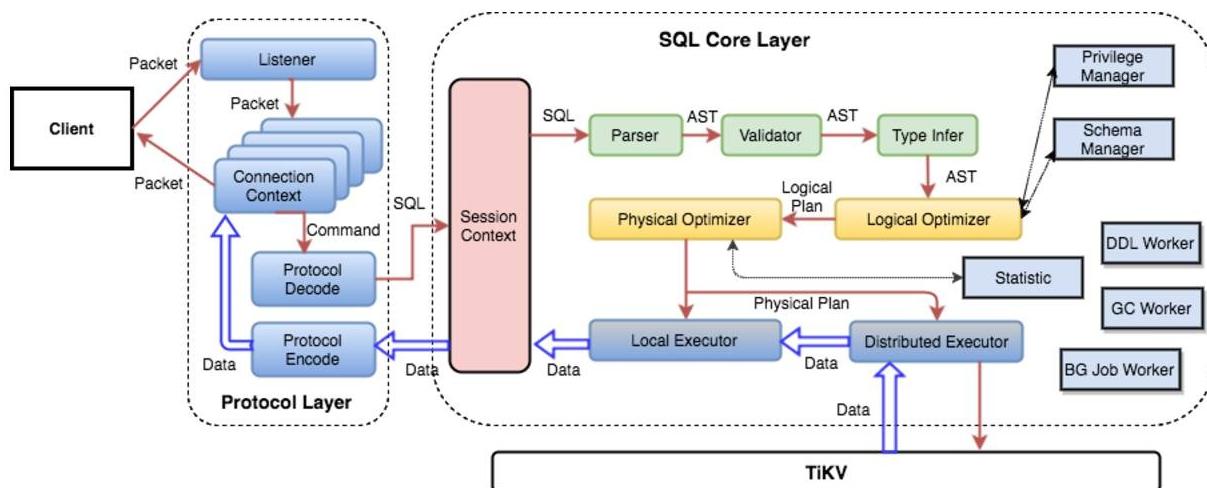
RPC SQL

Count(*) Count(*) SQL Count(*)



3.3.3 SQL

SQL TiDB SQL



SQL Load Balancer tidb-server tidb-server MySQL Protocol Packet SQL TiKV tidb-server TiKV
tidb-server

4

TiKV TiDB KV SQL PDPlacement Driver TiKV

4.1

4.1.1

TiKV TiDB KV Region Region Replica Replica TiKV Leader /Follower Leader raft log

- Region Replica TiKV
- TiKV Raft Group Replica
- TiKV ?
- Raft Group N Raft Group Replica Replica
- / Leader Leader
- Region Region
- IO CPU

Raft Group Region PD

4.1.2

-
-
-
-

- Leader
-
-
-
- /

Raft Group PD

4.1.3

- Replica
- Replica
- Leader Raft Group Replica transfer

Raft AddReplica RemoveReplica TransferLeader

4.1.4

TiKV Region TiKV PD

TiKV PD

TiKV Store PD PD Store Store Store

-
-
- Region
- /
- / Snapshot Replica Snapshot
-
- labels Tag

Raft Group Leader PD

Raft Group Leader PD Region

- Leader
- Followers
- Replica
- /

PD PD Store PD 30 Store Store Region PD PD Store PD Store Region

4.1.5

PD

Region Replica

PD Region Leader Region Replica Add/Remove Replica Replica

- Region Replica
- Replica Region Replica Replica
- max-replicas

Raft Group Replica

PD Replica Replica

-
- TiKV
- TiKV IDC

Region Replica labels PD location-labels label Replica Region Replica

Store

region

Leader Store

Raft Leader Leader PD Leader

Store

Store Region Leader Key /PD

Store

Store Capacity Store PD

CPU IO PD PD

4.1.6

PD Store Leader Region Leader PD Region Region Leader Region Leader Region Leader

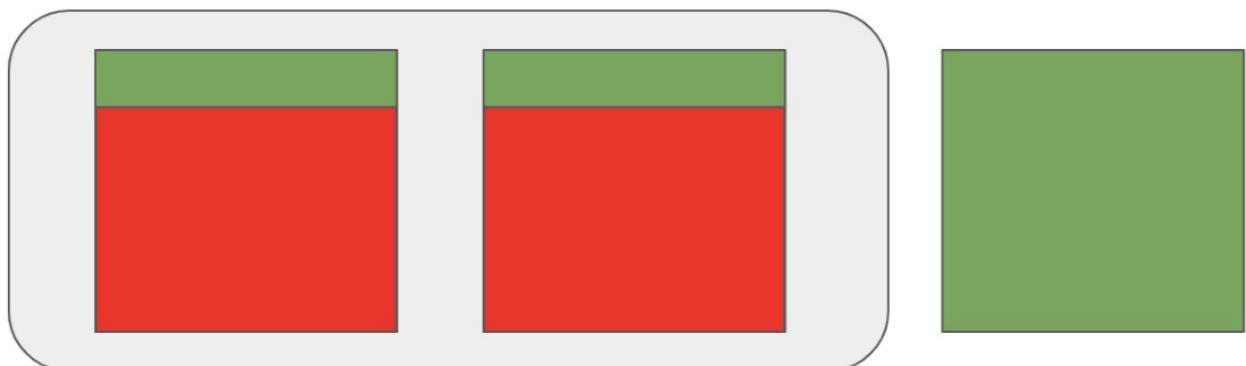
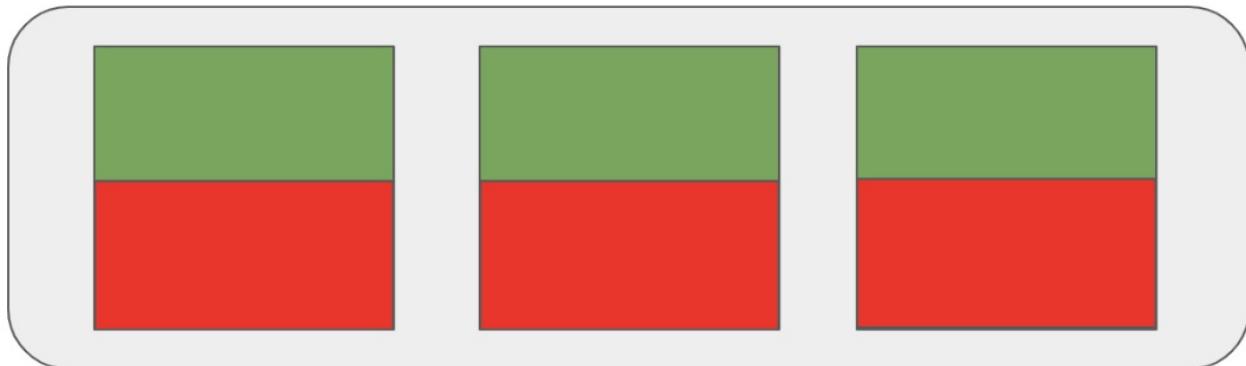
4.1.7

4.2

Elastic Schedule TiDB 4.0 TiDB **Adaptive Capacity** TiDB workload TiDB DBaaS

4.2.1

TiDB IDC



TiDB TiDB Operator 4.0

4.2.2

Auto-Scale

Aurora CPU TiDB Server TiKV Server QPSTiKV Server

TiDB TiDB Operator PD Auto-Scale TiDB Operator TiDB / TiKV metrics API TiDB / TiKV numbers TiDB Operator PD API Auto-Scaling TidbCluster.Spec.Replicas Auto-Scaling TiDB Operator Autoscaler API Autoscaler Controller Autoscaler API

```

apiVersion: pingcap.com/v1alpha1
kind: TidbClusterAutoScaler
metadata:
  name: autoscaler
  namespace: ela-demo
spec:
  cluster:
    name: ela-scheduling
    namespace: ela-demo
  metricsUrl: http://monitor-prometheus.el-a-demo.svc:9090
  tidb:
    minReplicas: 8
    maxReplicas: 8
    scaleOutIntervalSeconds: 100
    scaleInIntervalSeconds: 100
    metricsTimeDuration: "1m"
    metrics:
      - type: "Resource"
        resource:
          name: "cpu"
          target:
            type: "Utilization"
            averageUtilization: 90
  tikv:
    minReplicas: 3
    maxReplicas: 5
    scaleOutIntervalSeconds: 100
    scaleInIntervalSeconds: 100
    metricsTimeDuration: "1m"
    metrics:
      - type: "Resource"
        resource:
          name: "cpu"
          target:
            type: "Utilization"
            averageUtilization: 70

```

- minReplicas
- maxReplicas
- scaleOutIntervalSeconds scale-out
- scaleInIntervalSeconds scale-in

TiDB LB TiDB TiKV PD Region

4.2.3

TiDB Operator TiKV TiKV PD

- 1.
- 2.
- 3.
- 4.

Region 2 4

- | | | |
|--|-----|----------------------|
| 1. (Load Base Splitting) TiDB Region 4.0 Region CRDB
key Region | QPS | Region QPS 20 key ke |
| 2. (Isolate Frequently Access Region TiKV Range TiDB workload non-uniform shard_bits
TiDB workload PD | | |

4.2.4

TiDB 4.0 TiDB Operator DBaaS TiDB 4.0 TiDB workload TiDB Follower Read TiFlash
TiDB HTAP “”

5 TiDB MySQL

TiDB NewSQL SQL ACID TiDB MySQL MySQL MySQL MySQL TiDB

MySQL TiDB MySQL TiDB

TiDB MySQL

5.1

TiDB MySQL

	SQL	

4.0 TiDB MySQL

	MySQL	TiDB
JSON		MySQL 8.0
		asciilatin1binaryutf8utf8mb4
/		alter-primary-key
CREATE TABLE tblName AS SELECT stmt		
CREATE TEMPORARY TABLE		TiDB TEMPORARY
DML affected rows		
AutoRandom		
Sequence		

5.2

(1)

TiDB

```
tidb> SHOW CHARACTER SET;
+-----+-----+-----+-----+
| Charset | Description | Default collation | Maxlen |
+-----+-----+-----+-----+
| utf8   | UTF-8 Unicode | utf8_bin          |      3 |
| utf8mb4 | UTF-8 Unicode | utf8mb4_bin        |      4 |
| ascii   | US ASCII       | ascii_bin          |      1 |
| latin1  | Latin1          | latin1_bin         |      1 |
| binary  | binary          | binary             |      1 |
+-----+-----+-----+-----+
5 rows in set (0.00 sec)
```

TiDB utf8mb4 MySQL 5.7 latin1 MySQL 8.0 utf8mb4 utf8 utf8mb4 TiDB UTF8 incorrect utf8
 value MySQL 8.0 MySQL 5.7 UTF8 TiDB TiDB "tidb_skip_utf8_check" UTF8 TiDB
 Collation utf8 Collation utf8_bin TiDB Collation MySQL

	TiDB Collation	MySQL 5.7 Collation	MySQL 8.0 Collation
utf8	utf8_bin	utf8_general_ci	utf8_general_ci
utf8mb4	utf8mb4_bin	utf8mb4_general_ci	utf8mb4_0900_ai_ci
ascii	ascii_bin	ascii_general_ci	ascii_general_ci
latin1	latin1_bin	latin1_swedish_ci	latin1_swedish_ci
binary	binary	binary	binary

4.0 TiDB Collation Collation MySQL Collation PADDING

```
tidb> create table t(a varchar(20) charset utf8mb4 collate utf8mb4_general_ci primary key);
Query OK, 0 rows affected
tidb> insert into t values ('A');
Query OK, 1 row affected
tidb> insert into t values ('a');
Query OK, 1 row affected // MySQL utf8mb4_general_ci Duplicate entry 'a'.
tidb> insert into t1 values ('a ');
Query OK, 1 row affected // MySQL Duplicate entry 'a '
```

TiDB 4.0 Collation MySQL Collation new_collation_enabled_on_first_bootstrap Collation mysql . tidb
 new_collation_enabled Collation

```
tidb> select VARIABLE_VALUE from mysql.tidb where VARIABLE_NAME='new_collation_enabled';
+-----+
| VARIABLE_VALUE |
+-----+
| True           |
+-----+
1 row in set (0.00 sec)
```

Collation TiDB utf8mb4_general_bin utf8_general_bin PADDING utf8mb4_general_ci utf8_general_ci
 Collation MySQL

(2)

MySQL system_time_zone MySQL TZ --timezone

TiDB TiDB TiDB TiDB TZ mysql . tidb

```
tidb> select VARIABLE_VALUE from mysql.tidb where VARIABLE_NAME='system_tz';
+-----+
| VARIABLE_VALUE |
+-----+
| Asia/Shanghai |
+-----+
1 row in set (0.00 sec)
```

system_time_zone	system_tz
tidb> select @@system_time_zone;	
+-----+	
@@system_time_zone	
+-----+	
Asia/Shanghai	
+-----+	
1 row in set (0.00 sec)	

TiDB	time_zone
tidb> set @@global.time_zone='UTC';	
Query OK, 0 rows affected (0.00 sec)	

6.1

ACID TiDB Percolator

6.1.1

AtomicityConsistencyIsolationDurability ACID

6.1.1.1

SQL-92 4 (READ UNCOMMITTED) (READ COMMITTED) (REPEATABLE READ)
(SERIALIZABLE)

Isolation Level	Dirty Write	Dirty Read	Fuzzy Read	Phantom
READ UNCOMMITTED	Not Possible	Possible	Possible	Possible
READ COMMITTED	Not Possible	Not possible	Possible	Possible
REPEATABLE READ	Not Possible	Not possible	Not possible	Possible
SERIALIZABLE	Not Possible	Not possible	Not possible	Not possible

6.1.1.2

- OCC
- PCC

6.1.2 TiDB

TiDB Google [Percolator](#) ACID Snapshot Isolation TiDB TiKV

6.1.2.1 Snapshot Isolation

Percolator MVCC TiDB [PD](#) TSO

- start timestamp commit timestamp
- start timestamp
- timestamp

6.1.2.2 2PC

TiDB (Two-Phase Commit Prewrite Commit

- Prewrite Key lock TiDB Key Primary Key Primary Key
- Commit Prewrite Primary Key Secondary Keys

Percolator Coordinator

6.1.2.3



Percolator

- 1.
2. TiDB PD tso start timestamp
- 3.
4. Commit
5. TiDB
 - i. TiDB Key Primary Key
 - ii. TiDB TiKV Prewrite TiKV Prewrite
 - iii. TiDB Prewrite Prewrite
 - iv. TiDB PD commit timestamp
 - v. TiDB Primary Key TiKV TiKV Commit Prewrite
6. TiDB
7. TiDB

6.1.3

6.1.3.1

4 RTTRound trip time

- PD 2 Timestamp
- Prewrite Commit

query `autocommit = 1`

```
UPDATE my_table SET a='new_value' WHERE id = 1;
UPDATE my_table SET a='newer_value' WHERE id = 2;
UPDATE my_table SET a='newest_value' WHERE id = 3;
```

```
START TRANSACTION;
UPDATE my_table SET a='new_value' WHERE id = 1;
UPDATE my_table SET a='newer_value' WHERE id = 2;
UPDATE my_table SET a='newest_value' WHERE id = 3;
COMMIT;
```

INSERT

6.1.3.2

TiDB

- OOM (Out of Memory)
-
-

TiDB 100500

6.1.3.3

(1) TiKV Prewrite TiDB TiDB

- TiDB TiDB TiKV
- TiKV Prewrite TiDB TiDB TiDB TiKV

TiDB

`txn-local-latches :`

- enable
 -
 - false
- capacity
 - Hash slot 2 slot 32 Bytes
 - 1024000

capacity Key Key Hash Hash **capacity** Hash

- capacity
- capacity

TiKV TiKV hash

- scheduler-concurrency
 - scheduler Key Key hash
 - 2048000

TiKV latch



Scheduler latch wait duration

(2) 2

```
ERROR 8005 (HY000) : Write Conflict, txnStartTS is stale
```

MySQL SQL MySQL TiDB TiDB SQL

`tidb_disable_txn_auto_retry = off`

`tidb_retry_limit`

- `tidb_disable_txn_auto_retry 1`
- `tidb_retry_limit`

set

- session

```
set @@tidb_disable_txn_auto_retry = 0;
set @@tidb_retry_limit = 10;
```

-

```
set @@global.tidb_disable_txn_auto_retry = 0;
set @@global.tidb_retry_limit = 10;
```

`tidb_retry_limit 0 TiDB (try again later)`

(3) TiDB

1. start timestamp
2. SQL
- 3.

SQL SQL start timestamp start timestamp

`SELECT FOR UPDATE` `SELECT FOR UPDATE`

TiDB

6.1.3.4 GC

TiDB MVCCGC

GC 10 GC TiDB

`safe point 10 TiDB`

`safe point`

TiDB GC mysql.tidb SQL

```
select VARIABLE_NAME, VARIABLE_VALUE from mysql.tidb;
```

GC

```
update mysql.tidb set VARIABLE_VALUE="24h" where VARIABLE_NAME="tikv_gc_life_time";
```

GC TiKV gc.max-write-bytes-per-sec GC worker 0

6.1.4

```
MySQL [test]> insert into tran select id,star from eyesight.t_request;
Query OK, 3327 rows affected (0.17 sec)
Records: 3327 Duplicates: 0 Warnings: 0
MySQL [test]>
```

BEGIN COMMIT

```
Database changed
MySQL [test]> set autocommit=0;
Query OK, 0 rows affected (0.00 sec)

MySQL [test]> begin;
Query OK, 0 rows affected (0.00 sec)

MySQL [test]> update tran set b=8888;
Query OK, 3327 rows affected (0.12 sec)
Rows matched: 3327 Changed: 3327 Warnings: 0

MySQL [test]> commit;
Query OK, 0 rows affected (2.31 sec)

MySQL [test]> []
```

```
Database changed
MySQL [test]> set autocommit=0;
Query OK, 0 rows affected (0.00 sec)

MySQL [test]> begin;
Query OK, 0 rows affected (0.00 sec)

MySQL [test]> update tran set b=9999;
Query OK, 3327 rows affected (0.04 sec)
Rows matched: 3327 Changed: 3327 Warnings: 0

MySQL [test]> commit;
Query OK, 0 rows affected (1.17 sec)

MySQL [test]> []
```

TiDB tidb.log startTs(415144181973647361) startTs(415144181960540213) 415144181973647361 ()

```
[2020/03/08 14:36:17.121 +08:00] [WARN] [session.go:419] [sql] [label=internal] [error="kv:9007]Write conflict, txnStartTS=415144181973647361, conflictStartTS=415144181960540213, conflictCommitTS=415144182052290562, key={tableID=19, indexID=1, indexValues={415144166153256976, 114421, }} primary={tableID=19, indexID=1, indexValues={415144166153256976, 114421, }} [try again later]" [txn="Txn{state=invalid}"]
[2020/03/08 14:36:17.121 +08:00] [WARN] [session.go:611] [retrying] [schemaVersion=1813] [retryCnt=0] [queryNum=1] [sql="update mysql.stats_meta set version = 415144181973647361, count = count + 52587, modify_count = modify_count + 52587 where table_id = 1807"]
[2020/03/08 14:36:17.125 +08:00] [WARN] [session.go:632] ["transaction association"] ["retrying txnStartTS=415144182301327413] ["original txnStartTS=415144181973647361] startTs
[2020/03/08 14:36:17.279 +08:00] [INFO] [2pc.go:1039] ["2PC clean up done"] [txnStartTS=415144181973647361]
```

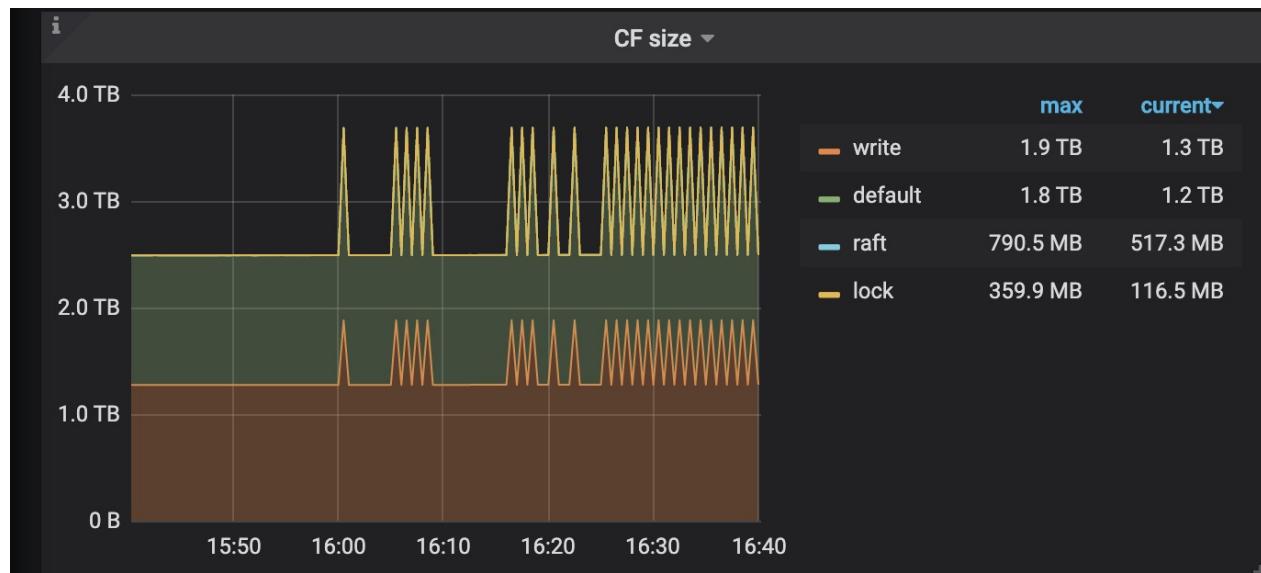
commit (2.311.17) SELECT

```
MySQL [test]> select * from tran limit 2;
+-----+-----+
| a     | b     |
+-----+-----+
| 100007 | 8888 |
| 100036 | 8888 |
+-----+-----+
2 rows in set (0.00 sec)
MySQL [test]>
```

6.1.5

6.1.5.1 tikv--server

lock



6.1.5.2 tidb-transaction



6.2

TiDB TiDB v3.0 TiDB

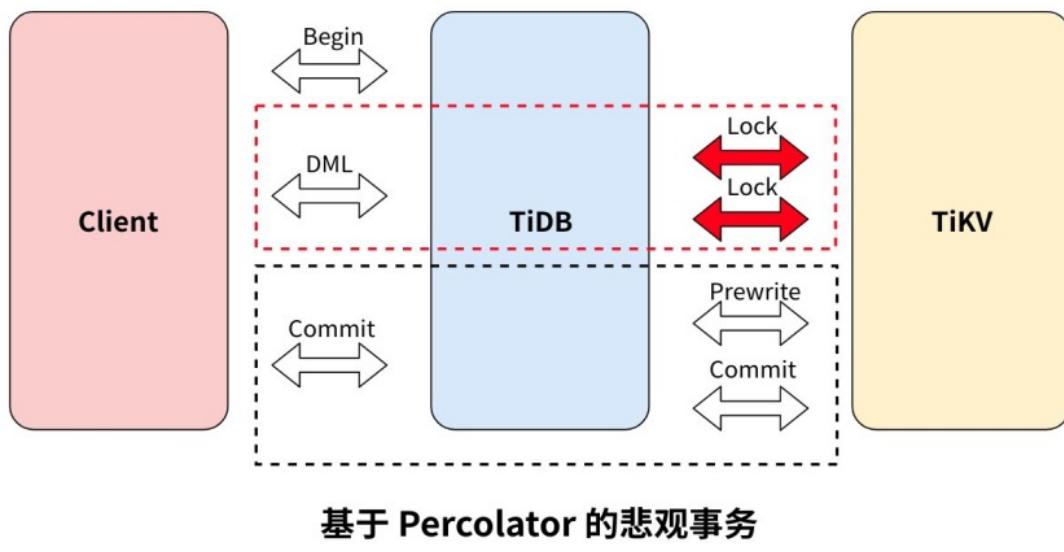
6.2.1

- v3.0.8 TiDB
- TiDB

6.2.2 Percolator

Percolator Prewrite Acquire Pessimistic Lock Prewrite

- DML TiKV raft
- Write Conflictkey
- Key Prewrite Percolator
- Percolator Prewrite Write Conflict



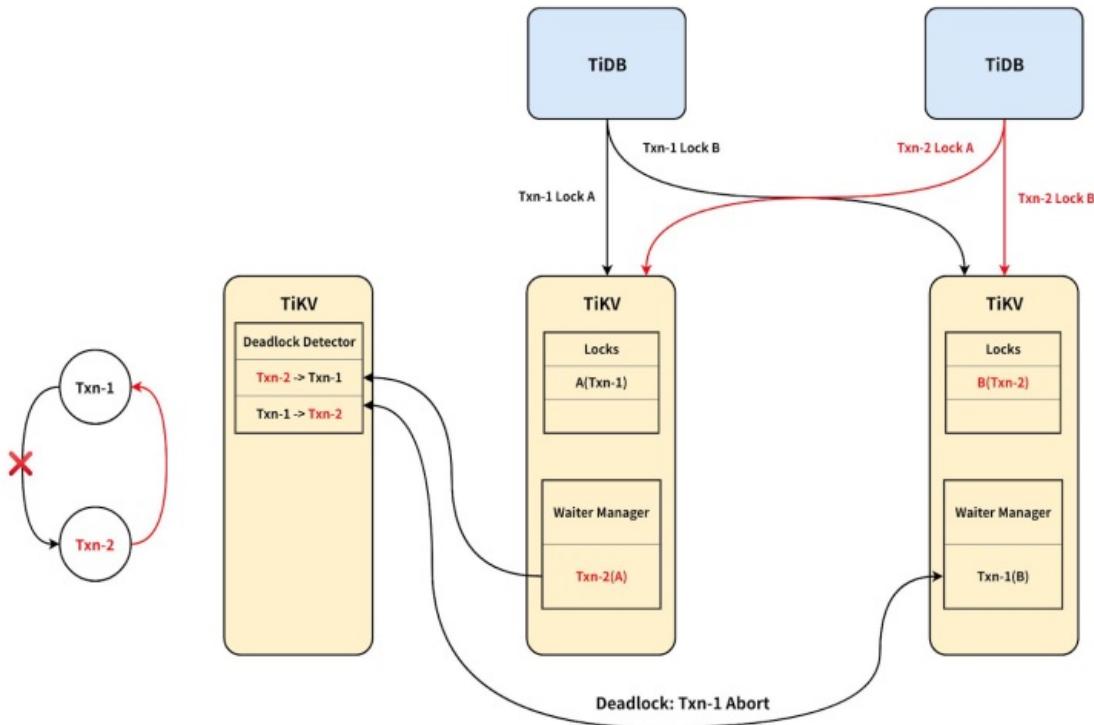
6.2.2.1

TiKV Waiter Manager Waiter Manager TiKV start timestamp

6.2.2.2

Waiter Manager TiDB

- TiKV leader
- leader



Raft leader leader

6.2.3

6.2.3.1

TiDB

-
-

6.2.3.2

v3.0.8 TiDB :

- BEGIN PESSIMISTIC ;
BEGIN /*!90000 PESSIMISTIC */; MySQL
- set @@tidb_txn_mode = 'pessimistic'; session autocommit
- set @@global.tidb_txn_mode = 'pessimistic'; session autocommit
set @@global.tidb_txn_mode = '';

6.2.3.3 Batch DML

DML TiKV DML DML DML

INSERT TiKV

```
BEGIN;
INSERT INTO my_table VALUES (1);
INSERT INTO my_table VALUES (2);
INSERT INTO my_table VALUES (3);
COMMIT;
```

INSERT

```
BEGIN;
INSERT INTO my_table VALUES (1), (2), (3);
COMMIT;
```

6.2.3.4

TiDB 2

MySQL Repeatable Read

TiDB MySQL InnoDB

1. TiDB range WHERE DML

SELECT FOR UPDATE

INSERT

InnoDB gap lock range INSERT statement based binlog READ COMMITTED gap lock TiDB gap lock

2. TiDB SELECT LOCK IN SHARE MODE

3. DDL

MySQL DDL TiDB DDL

ERROR 1105 (HY000): Information schema is changed. [try again later]

4. START TRANSACTION WITH CONSISTENT SNAPSHOT MySQL TiDB

5. autocommit

DML autocommit TiDB

select for update

SET SESSION TRANSACTION ISOLATION LEVEL READ COMMITTED; Oracle Read Committed

Oracle TiDB

6.3 4.0

6.3.1

DBA TiDB transaction too large

PingCAP FAQ

4.3.3 Transaction too large

Raft Raft

- SQL 5000
- KV entry 6MB
- KV entry 30w
- KV entry 100MB

Google Cloud Spanner TiDB

- MySQL/Oracle TiDB statement count 5001 exceeds the transaction limitation
- kv 30W 10W ERROR 8004 (HY000): transaction too large, len:300001

asktug4.0 TiDB

6.3.2

4.0

- Prewrite Prewrite
- Prewrite

4.0

6.3.2.1 Min Commit Timestamp

TiDB Snapshot Isolation start timestamp commit timestamp start timestamp

TiDB Primary Lock minCommitTs commit timestamp start timestamp Primary Lock

6.3.2.2 Time to live(TTL)

Percolator Prewrite TiDB TiDB TTL

v4.0 TTL v4.0 TTL Manager Primary Lock TTL

6.3.3

TiDB

```
[performance]
txn-total-size-limit = 104857600
```

10737418240(10G) v3.x v4.0

(1) 3.0.5

```

mysql> insert into t1 (name, age) select name, age from t1;
Query OK, 131072 rows affected (1.86 sec)
Records: 131072  Duplicates: 0  Warnings: 0

mysql> select count(1) from t1;
+-----+
| count(1) |
+-----+
| 262144 |
+-----+
1 row in set (0.14 sec)

mysql> insert into t1 (name, age) select name, age from t1;
ERROR 8004 (HY000): transaction too large, len:300001

```

(2) 4.0

```

MySQL [test]> select count(1) from t1;
+-----+
| count(1) |
+-----+
| 262144 |
+-----+
1 row in set (0.20 sec)

MySQL [test]> insert into t1 (name, age) select name, age from t1;
Query OK, 262144 rows affected (9.20 sec)
Records: 262144  Duplicates: 0  Warnings: 0

MySQL [test]> select count(1) from t1;
+-----+
| count(1) |
+-----+
| 524288 |
+-----+
1 row in set (0.52 sec)

MySQL [test]> create table t2 like t1;
Query OK, 0 rows affected (0.11 sec)

MySQL [test]> insert into t2 select * from t1;
Query OK, 524288 rows affected (17.61 sec)
Records: 524288  Duplicates: 0  Warnings: 0

```

4.0 tidb_batch_insert**6.3.4**

TiDB GC 10min	gc_life_time	v4.0	safepoint	start timestamp
• kv 6MB	Blob	TiDB		
• 10GB	10GB	10GB		
• 3~4	10GB	30~40GB		

7 TiDB DDL

DDL TiDB Google F1 schema

TiDB DDL 4.0

1. 3
2. DDL DDL DDL
3. Sequence 3
4. AutoRandom

7.1

7.1.1

(schema) case

7.1.2

TiDB

TiDB Region Region 96MBRegion Region Raft Group Raft Group Leader TiDB
PD Follower-ReadLeader



Region Region Region TiDB

7.1.3

1.

MySQL InnoDB DBA ID

TiDB Region TiDB Region Region PD Region Region Region TiKV Region TiDB Region
TiKV Region

TiDB 3.0

Grafana **Grafana > TiKV-Trouble-Shooting > Hot write**

store QPS CPU pd-ctl Region region_id

```
pd-ctl -u "http://{{pd-instance}}:2379" -d region topwrite
```

region_id

```
curl http://{{tidb-instance}}:10080/regions/{{region_id}}
```

Region Region Region TiDB

TiDB PK PK rowID_tidb_rowid SHARD_ROW_ID_BITS rowID Region RPC CPU

- SHARD_ROW_ID_BITS = 4 2^{416}
- SHARD_ROW_ID_BITS = 6 2^{664}
- SHARD_ROW_ID_BITS = 0 1

```
CREATE TABLE t (c int) SHARD_ROW_ID_BITS = 4;
ALTER TABLE t SHARD_ROW_ID_BITS = 4;
```

SHARD_ROW_ID_BITS Region

TiDB 3.1.0 AUTO_RANDOM AUTO_INCREMENT TiDB ID AUTO_RANDOM

2.

TiDB case

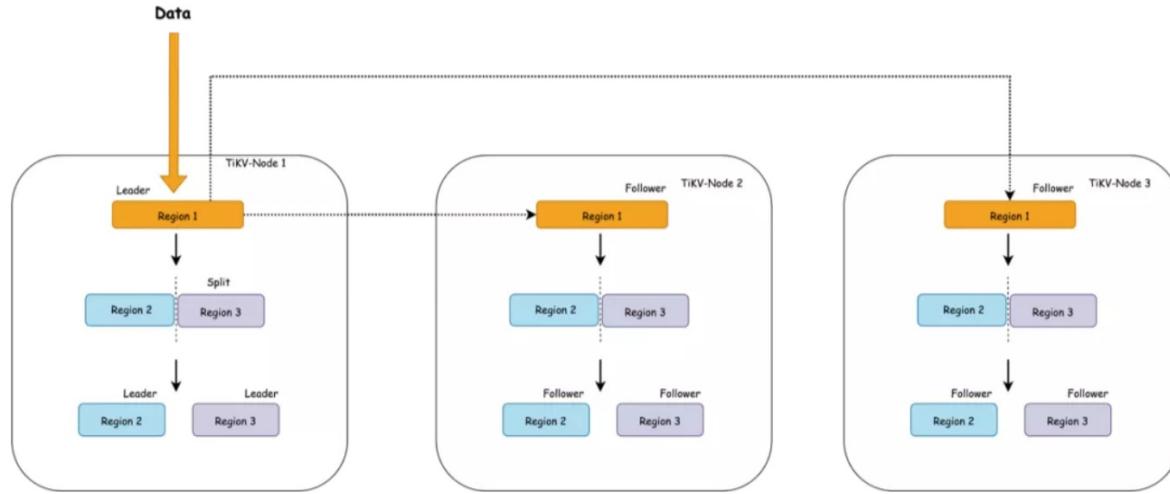
```
CREATE TABLE IF NOT EXISTS TEST_HOTSPOT(
    id      BIGINT PRIMARY KEY,
    age     INT,
    user_name VARCHAR(32),
    email   VARCHAR(128)
)
```

id id TiDB

TiKV Region:

```
[CommonPrefix + TableID, CommonPrefix + TableID + 1)
```

Region



TiKV Region Leader Store Store Region Leader Region 23 Region 1 TiKV-Node 1

PD Node 1 Leader Node TiKV Region Node PD Region

Region

Region

shard_row_id_bits Region pre_split_regions pre_split_regions shard_row_id_bits $2^{(pre_split_regions)}$ Region

```
CREATE TABLE t (a INT, b INT, INDEX idx1(a)) SHARD_ROW_ID_BITS = 4 PRE_SPLIT_REGIONS = 2;
```

Split Region Region

Split Region TiDB Region Region TiKV

- BETWEEN lower_value AND upper_value REGIONS region_num Region region_num Region

```
SPLIT TABLE table_name [INDEX index_name] BETWEEN (lower_value) AND (upper_value) REGIONS region_num
```

- BY value_list... Region

```
SPLIT TABLE table_name [INDEX index_name] BY (value_list) [, (value_list)] ...
```

Region Merge split Region split split-merge-interval Region Region Merge Region Merge

TiDB 4.0 TiDB 4.0 PD Load Based Splitting Region QPS Region

3.

delete TiDB KV RocksDB delete

1. TiDB TiKV
2. delete
3. TiKV
4. RocksDB Compaction

TiDB Delete Ranges

1. TiDB gc_delete_range
2. Delete Ranges gc_delete_range safe point

TiDB TiKV TiKV delete TiKV RocksDB compaction

4.

case

-
- TiDB MySQL DECIMAL
-
- case TiDB 70 TP99 130ms 7 TP99 60ms

7.1.4

case GC

7.2 DDL

7.2.1 TiDB DDL

DDL SQL DDL SQL DDL DDL

TiDB DDL SQL SQL DBA TiDB

7.2.2 DDL

TiDB MySQL ADMIN DDL DDL admin

- schema versionowner DDL

```
ADMIN SHOW DDL;
```

- DDL (DDL) NUM (10) DDL

```
ADMIN SHOW DDL JOBS [NUM] [WHERE where_condition];
```

- DDL 5 DDL

```
ADMIN SHOW DDL JOBS 5;
```

- test DDL 5 DDL

```
ADMIN SHOW DDL JOBS 5 WHERE state != 'synced' AND db_name = 'test';
```

- JOB_ID DDL

```
ADMIN SHOW DDL JOB QUERIES job_id [, job_id] ...;
```

- DDL

```
ADMIN CANCEL DDL JOBS job_id [, job_id] ...;
```

- JOB_ID RECOVER TABLE table_name

```
RECOVER TABLE BY JOB ddl_job_id;
```

DDL ADD INDEX mysql-client ctrl+c DDL DDL

```
> admin@sbttest02:04:32>alter table sbtest1 add key idx_c_pad(c, pad);
^CCtrl-C -- sending "KILL QUERY 15768094" to server ...
Ctrl-C -- query aborted.
^C^CCtrl-C -- exit!
Aborted
```

```
ADMIN SHOW DDL      ctrl+c    DDL
```

```
> admin@(none)02:04:56>ADMIN SHOW DDL\G
***** 1. row ****
SCHEMA_VER: 6765
OWNER_ID: 828a4567-91a5-4070-b55f-f90702e80e7a
OWNER_ADDRESS: 10.40.216.9:4000
RUNNING_JOBS: ID:8406, Type:add index, State:running, SchemaState:write reorganization, SchemaID:7299, TableID:7380, RowCount:196608, ArgLen:0, start time: 2020-03-07 14:04:53.321 +0800 CST, Err:<nil>, ErrCount:0, SnapshotVersion:415121039339290634
SELF_ID: 828a4567-91a5-4070-b55f-f90702e80e7a
QUERY: alter table sbtest1 add key idx_c_pad(c, pad)
1 row in set (0.02 sec)
```

TiDB HTTP owner TiDB TiDB `ddl_id lease`

```
# curl http://{TiDBIP}:10080/info/all
# $curl http://127.0.0.1:10080/info/all
{
  "servers_num": 2,
  "owner_id": "29a65ec0-d931-4f9e-a212-338eaeffab96",
  "is_all_server_version_consistent": true,
  "all_servers_info": {
    "29a65ec0-d931-4f9e-a212-338eaeffab96": {
      "version": "5.7.25-TiDB-v4.0.0-alpha-669-g8f2a09a52-dirty",
      "git_hash": "8f2a09a52fdcaf9d9bfd775d2c6023f363dc121e",
      "ddl_id": "29a65ec0-d931-4f9e-a212-338eaeffab96",
      "ip": "",
      "listening_port": 4000,
      "status_port": 10080,
      "lease": "45s",
      "binlog_status": "Off"
    },
    "cd13c9eb-c3ee-4887-af9b-e64f3162d92c": {
      "version": "5.7.25-TiDB-v4.0.0-alpha-669-g8f2a09a52-dirty",
      "git_hash": "8f2a09a52fdcaf9d9bfd775d2c6023f363dc121e",
      "ddl_id": "cd13c9eb-c3ee-4887-af9b-e64f3162d92c",
      "ip": "",
      "listening_port": 4001,
      "status_port": 10081,
      "lease": "45s",
      "binlog_status": "Off"
    }
  }
}
```

7.2.3 DDL

- `tidb_ddl_reorg_worker_cnt`

		GLOBAL
		4
		re-organize

- `tidb_ddl_reorg_batch_size`

		GLOBAL
		256
		32
		10240
		re-organize

- `tidb_ddl_reorg_priority`

		GLOBAL SESSION
		PRIORITY_LOW
		re-organize

- `tidb_ddl_error_count_limit`

		GLOBAL
		512
		DDL DDL

TiDB DDL DDL DDL TiKV TiKV ADD INDEX

TiDB `tidb_ddl_reorg_worker_cnt` `tidb_ddl_reorg_batch_size` DDL DDL DDL

- `tidb_ddl_reorg_priority` SQL
- `tidb_ddl_error_count_limit` TiKV DDL

7.2.4 DDL

TiDB-Server SQL DDL TiDB-Server TiDB TiDB-Server DDL

TiDB-Server DDL TiDB-Server DDL DDL Job TiKV

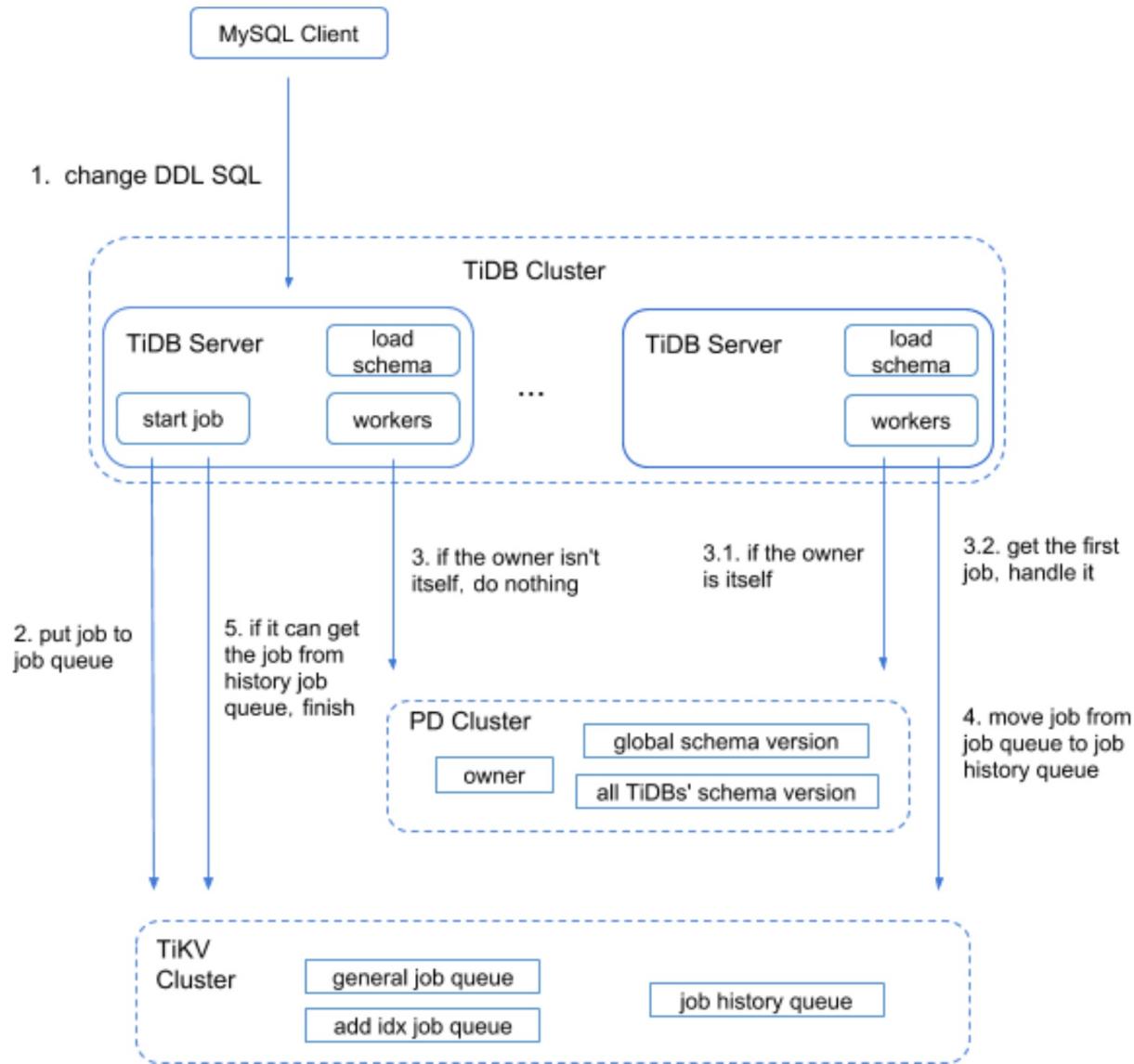
TiDB-Server DDL worker TiDB-Server Owner DDL TiDB-Server DDL DDL Owner Owner
Owner

Owner worker DDL ADD INDEX DDL DDL TiDB ADD INDEX DDL ADD INDEX DDL Owner
worker DDL DDL DDL

Owner worker ADD INDEX DDL `tidb_ddl_reorg_worker_cnt` `tidb_ddl_reorg_batch_size` DDL

DDL DDL job history queue DDL

DDL DDL



7.2.5 DDL

TiDB Online, Asynchronous Schema Change in F1 schema schema

ADD INDEX DDL

schema		
schema version 1	absent	
schema version 2	delete only	schema /
schema version 3	write only	
schema version 4	write reorganization	
schema version 5	public	

TiDB ADD COLUMN DDL TIDB default schema "" TiKV null "" null "" null TiKV

DDL DDL

DDL DROP INDEX DROP TABLE DROP DATABASE TRUNCATE TABLE DDL TiDB gc_delete_range GC

DDL schema

DDL TiDB DDL architecture

7.3 Sequence

Sequence SQL 2003 “Sequence” MySQL Sequence TiDB Sequence MariaDB Oracle IBM Db2

- Create Sequence

```
CREATE [TEMPORARY] SEQUENCE [IF NOT EXISTS] sequence_name
[ INCREMENT [ BY | = ] INCREMENT ]
[ MINVALUE [=] minvalue | NO MINVALUE | NOMINVALUE ]
[ MAXVALUE [=] maxvalue | NO MAXVALUE | NOMAXVALUE ]
[ START [ WITH | = ] start ]
[ CACHE [=] cache | NOCACHE | NO CACHE]
[ CYCLE | NOCYCLE | NO CYCLE]
[ ORDER | NOORDER | NO ORDER]
[table_options]
```

- Show Create Sequence

```
SHOW CREATE SEQUENCE sequence_name
```

- Drop Sequence

```
DROP [TEMPORARY] SEQUENCE [IF NOT EXISTS] sequence_name
```

-

```
SELECT NEXT VALUE FOR sequence_name;
SELECT NEXTVAL(sequence_name);
```

- /

```
SELECT PREVIOUS VALUE FOR sequence_name;
SELECT LASTVAL(sequence_name);
```

-

```
SELECT SETVAL(sequence_name, 100)
```

7.3.1

TiDB Sequence

-

Sequence Leaf - ID uid-generator TiDB Sequence

(1) Sequence

```
CREATE SEQUENCE seq_for_unique START WITH 1 INCREMENT BY 1 CACHE 1000 NOCYCLE;
```

(2) TiDB Sequence

TiDB

```

Atidb[test]> SELECT NEXT VALUE FOR seq_for_unique;
+-----+
| NEXT VALUE FOR seq_for_unique |
+-----+
|          1 |
+-----+
1 row in set (0.00 sec)

Btidb[test]> SELECT NEXT VALUE FOR seq_for_unique;
+-----+
| NEXT VALUE FOR seq_for_unique |
+-----+
|          2 |
+-----+
1 row in set (0.00 sec)

```

(3) TiDB TiDB

```

Atidb[test]> SELECT NEXT VALUE FOR seq_for_unique;
+-----+
| NEXT VALUE FOR seq_for_unique |
+-----+
|          1 |
+-----+
1 row in set (0.00 sec)

Btidb[test]> SELECT NEXT VALUE FOR seq_for_unique;
+-----+
| NEXT VALUE FOR seq_for_unique |
+-----+
|         1001 |
+-----+
1 row in set (0.00 sec)

```

•

MySQL auto_increment TiDB Sequence

(1) Sequence

```

CREATE SEQUENCE seq_for_autooid START WITH 1 INCREMENT BY 2 CACHE 1000 NOCYCLE;
CREATE SEQUENCE seq_for_logid START WITH 100 INCREMENT BY 1 CACHE 1000 NOCYCLE;

```

(2) default nextval(seq_name)

```

CREATE TABLE `user` (
  `userid` varchar(32) NOT NULL,
  `autooid` int(11) DEFAULT 'nextval(`test`.`seq_for_autooid`)',
  `logid` int(11) DEFAULT 'nextval(`test`.`seq_for_logid`)',
  PRIMARY KEY (`userid`)
)

```

(3)

```

INSERT INTO user (userid) VALUES ('usera');
INSERT INTO user (userid) VALUES ('userb');
INSERT INTO user (userid) VALUES ('userc');

```

(4) user autooid logid userid

```
tidb[test]> select * from user;
+-----+-----+-----+
| userid | autoid | logid |
+-----+-----+-----+
| usera |      1 |   100 |
| userb |      3 |   101 |
| userc |      5 |   102 |
+-----+-----+-----+
3 rows in set (0.01 sec)
```

•

20 Sequence update Sequence

(1)

```
tidb[test]> CREATE TABLE t( a int, name varchar(32));
Query OK, 0 rows affected (0.01 sec)
```

(2) Sequence

```
tidb[test]> CREATE SEQUENCE test;
Query OK, 0 rows affected (0.00 sec)
```

(3) 1

```
for i in $(seq 1 10000)
do
    echo "insert into t values($((RANDOM%1000)), 'user${i}');" >> user.sql
done
```

```
tidb[test]> select count(*) from t;
```

```
+-----+
```

```
| count(*) |
```

```
+-----+
```

```
| 10000 |
```

```
+-----+
```

```
1 row in set (0.05 sec)
```

```
tidb[test]> select * from t;
```

```
+-----+-----+
```

```
| a | name |
```

```
+-----+-----+
```

```
| 355 | user1 |
```

```
| 729 | user2 |
```

```
| 684 | user3 |
```

```
| 815 | user4 |
```

```
| 39 | user5 |
```

```
| 294 | user6 |
```

```
| 407 | user7 |
```

```
| 767 | user8 |
```

```
| 246 | user9 |
```

```
| 755 | user10 |
```

```
| 496 | user11 |
```

```
...
```

(4)

```
tidb[test]> update t set a=nextval(test);
Query OK, 10000 rows affected (0.20 sec)
Rows matched: 10000  Changed: 10000  Warnings: 0
```

(5), a

```
tidb[test]> select * from t;
+-----+-----+
| a    | name   |
+-----+-----+
| 1    | user1  |
| 2    | user2  |
| 3    | user3  |
| 4    | user4  |
| 5    | user5  |
| 6    | user6  |
| 7    | user7  |
| 8    | user8  |
| 9    | user9  |
| 10   | user10 |
| 11   | user11 |
| 12   | user12 |
| 13   | user13 |
| 14   | user14 |
| 15   | user15 |
| 16   | user16 |
| 17   | user17 |
| 18   | user18 |
| 19   | user19 |
| 20   | user20 |
...
...
```

7.3.2

Sequence Sequence Order/No Order Cache/No Cache Sequence Sequence

7.4 AutoRandom

AutoRandom TiDB 4.0 AutoIncrement ID

7.4.1 AutoRandom

```
TiDB      _tidb_rowid _tidb_rowid  Key TiKV TiKV Region
TiDB      _tidb_rowid ""          AUTO_INCREMENT     SHARD_ROW_ID_BITS
          AUTO_INCREMENT   AUTO_INCREMENT  AUTO_RANDOM TiDB ID
```

AutoRandom

- TiDB
- TiDB PRE_SPLIT_REGION
- AutoIncrement TiDB Server Binlog

7.4.2 AutoRandom

AUTO_RANDOM TiDB TiDB 4.0

```
column_definition:
  data_type [NOT NULL | NULL] [DEFAULT default_value]
    [AUTO_INCREMENT | AUTO_RANDOM [(length)]]
    [UNIQUE [KEY] | [PRIMARY] KEY]
    [COMMENT 'string']
    [reference_definition]
```

AutoRandom AUTO_RANDOM Shard Bits 5

7.4.3 AutoRandom

AUTO_RANDOM TiDB experimental allow-auto-random = true allow-auto-random

```
tidb> create table t (a int primary key auto_random);
```

INSERT INTO t(b) values... INSERT

```

tidb> insert into t values (), ();
Query OK, 2 rows affected (0.00 sec)
Records: 2  Duplicates: 0  Warnings: 0

tidb> select * from t;
+-----+
| a      |
+-----+
| 201326593 |
| 201326594 |
+-----+
2 rows in set (0.00 sec)

tidb> insert into t values (), ();
Query OK, 2 rows affected (0.01 sec)
Records: 2  Duplicates: 0  Warnings: 0

tidb> select * from t;
+-----+
| a      |
+-----+
| 201326593 |
| 201326594 |
| 2080374787 |
| 2080374788 |
+-----+
4 rows in set (0.00 sec)

tidb> select last_insert_id();
+-----+
| last_insert_id() |
+-----+
|      2080374787 |
+-----+
1 row in set (0.00 sec)

```

- INSERT a TiDB ID
- INSERT AutoIncrement TiDB
- INSERT AutoRandom ID LAST_INSERT_ID() LAST_INSERT_ID() ID

7.4.5 AutoRandom

AutoRandom TiDB

- alter-primary-key AutoIncrement + SHARD_ROW_ID_BITS
SHARD_ROW_ID_BITS
- UUID()
UUID() TiDB

AUTO_RANDOM

8 Titan

TiKV TiKV RocksDB TiKV RocksDB LSM-tree RocksDB IO CPU SSD SSD

Titan PingCAP RocksDB key-value value key LSM-tree value RocksDB compaction IO CPU SSD

8.1 Titan

Titan USENIX FAST 2016 WiscKey WiscKey SSD SSD value LSM-tree

value Titan RocksDB RocksDB Titan SSD SSD IO Titan

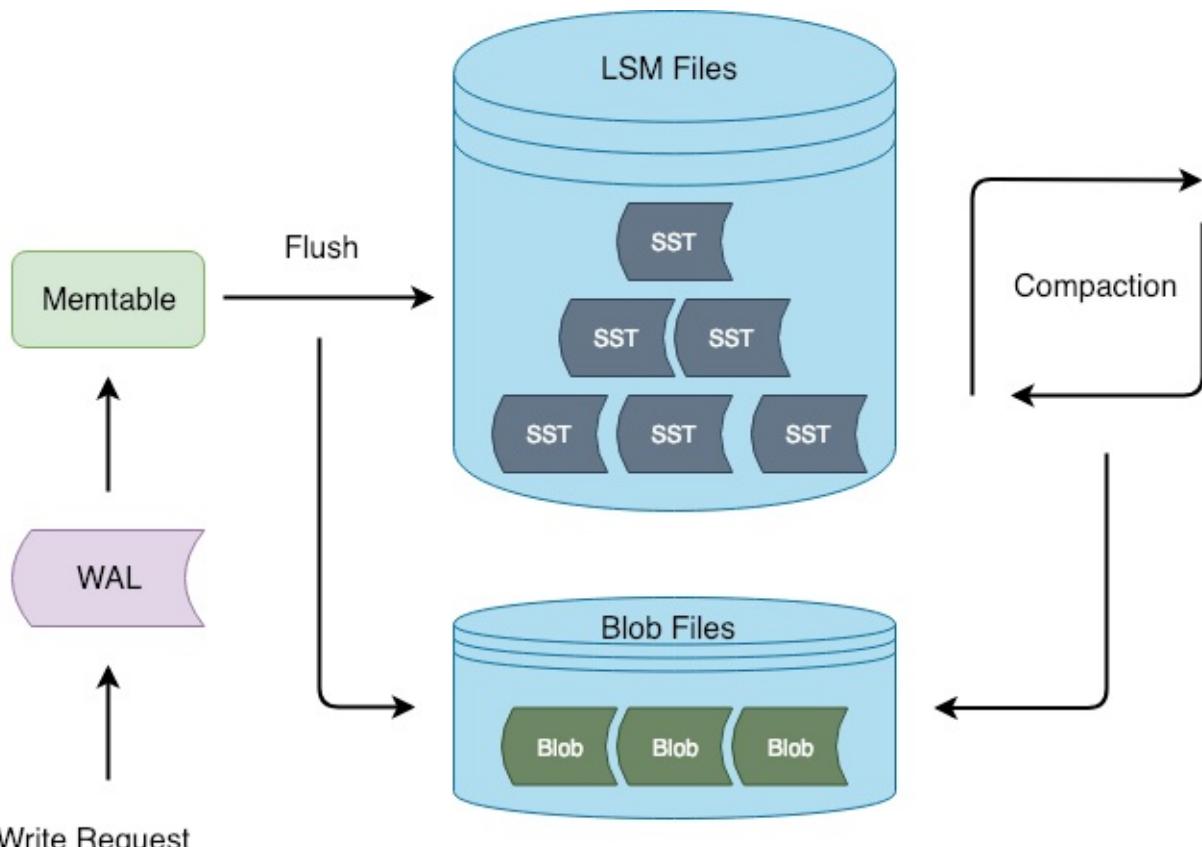
8.1.1

TiKV TiKV RocksDB TiKV Titan RocksDB RocksDB TiKV Titan TiKV Titan

- value LSM-tree value
- RocksDB Titan
- 100% TiKV RocksDB
- RocksDB Titan RocksDB

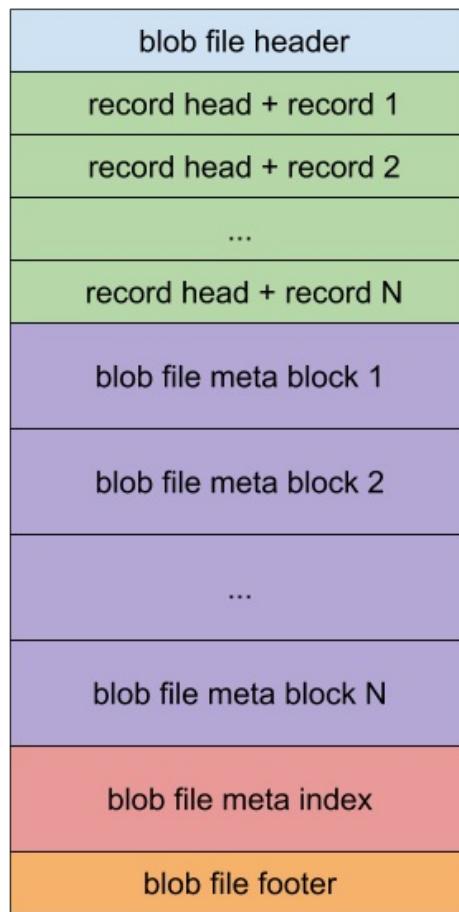
8.1.2

Titan RocksDB Flush Compaction value LSM-tree BlobFile RocksDB Titan BlobFileTitanTableBuilder Garbage CollectionGC



8.1.2.1 BlobFile

BlobFile LSM-tree KV headerrecord meta blockmeta index footer record key-value meta block meta index meta block



prefetch BlobFile key-value key LSM-tree value blob record key key Titan GC key GC key
 value Titan value BlobFile [SnappyLZ4](#) [Zstd](#) [Titan LZ4](#)

8.1.2.2 TitanTableBuilder

RocksDB TableBuilder table Titan TitanTableBuilder RocksDB table value LSM-tree



RocksDB SST TitanTableBuilder value value BlobFile value Titan value RocksDB SST value
 BlobFile value RocksDB SST value RocksDB TableBuilder RocksDB Compaction BlobFile value SST
 Titan RocksDB

* KV LSM-tree Badger Titan VLog WAL LSM-tree max level 5 10 LSM-tree 1 + 1 + 10 + 10 + 10 =
 42 BlobFile LSM-tree 1 : 42 WAL RocksDB Titan

8.1.2.3 Garbage Collection

RocksDB LSM-tree Compaction Titan Garbage Collection (GC) Titan GC GCLevel-Merge GC

1. GC

Blob GC GC GC GC

Titan GC RocksDB compaction RocksDB compaction BlobFile value Titan RocksDB compaction
 GC compaction SST BlobFile BlobFileSizeProperties BlobFile



inputs compaction SST BlobFileSizeProperties BlobFile ID : outputs SST BlobFile discardable size ID

BlobFileSizeProperties RocksDB TablePropertiesCollector SST Titan BlobFileSizeCollector SST BlobFile

SST

key - index {001:0000:512}
key - index {002:0000:1024}
key - value
key - index {001:2873:1024}
key - index {003:1752:256}

BlobFileSizeProperties

001:1536
002:1024
003:256



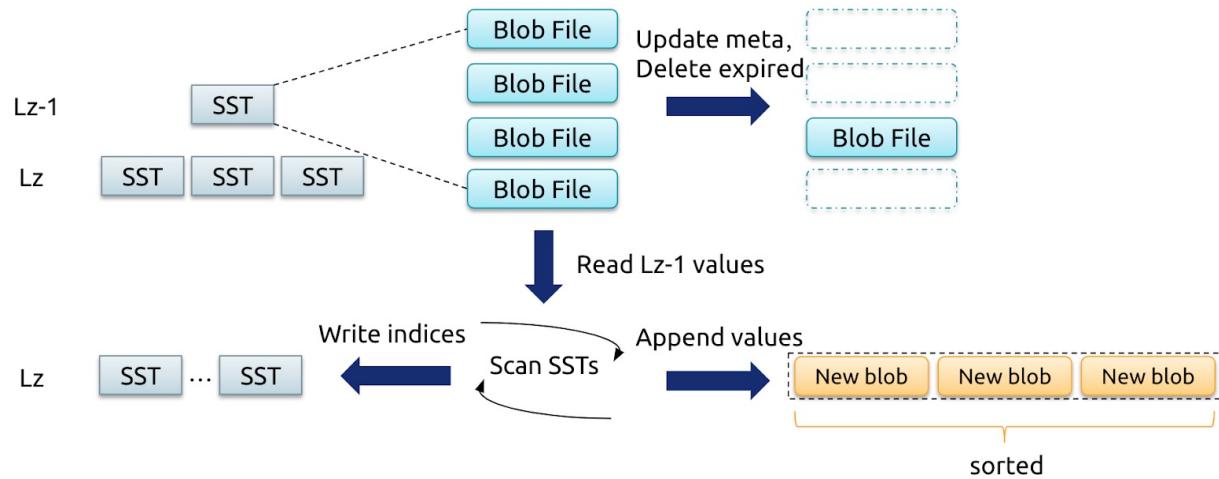
BlobFileSizeCollector SST KV SST BlobFile

GC BlobFileTitan BlobFile discardable size RocksDB compaction BlobFile discardable size BlobFile
SST BlobFile BlobFile GC GC BlobFile discardable size GC

GC Titan BlobFile record record key LSM-tree blob index blob record BlobFile BlobFile blob index LSM-tree snapshot GC Titan snapshot BlobFile GC RocksDB sequence number snapshot sequence number

1. Level-Merge

GC GC LSM-tree record key record index GC LSM-tree Titan Level-Merge LSM-tree compaction SST
BlobFile BlobFile GC LSM-tree BlobFile Scan



Level-Merge RocksDB level z-1 level z SST compaction KV SST BlobFile value BlobFile SST blob index
compaction key BlobFile BlobFile GC LSM-tree 99% Level-Merge Titan LSM-tree BlobFile Level-Merge

8.2 TiDB Titan

Titan

- Value value value Titan Titan 1KB value value 512B value Titan
- Titan RocksDB Titan RocksDB 40%
- Titan Titan RocksDB block Titan RocksDB Titan RocksDB

8.2.1 Titan

Titan RocksDB RocksDB TiKV Titan TiKV TiKV

```
[rocksdb.titan]
enabled = true
```

Titan Titan RocksDB compaction key-value Titan TiKV Details - Titan kv - blob file size Titan

Titan tikv-ctl compaction tikv-ctl

RocksDB Titan RocksDB Titan Titan TiKV Titan rocksdb.titan.enabled = false TiKV TiKV log “You have disabled titan when its data directory is not empty” “Titan”

Titan

TiKV Titan Titan RocksDB Titan leader PD Titan leader count pd-ctl store weight Titan leader weight leader count

8.2.2

```
[rocksdb.titan]
max-background-gc1
```

Titan GC TiKV Details - Thread CPU - RocksDB CPU Titan GC Titan GC

```
[rocksdb.defaultcf.titan]
min-blob-size 1kb
```

value value value RocksDB Titan blob file value value RocksDB value value Titan RocksDB compaction

```
[rocksdb.defaultcf.titan]
blob-file-compressionlz4
```

Titan value Titan value

```
[rocksdb.defaultcf.titan]
blob-cache-size0
```

Titan value Titan OOM RocksDB block cache storage.block-cache.capacity store size blob file size blob-cache-size * 50% block cache block cache RocksDB blob cache

```
[rocksdb.defaultcf.titan]
discardable-ratio0.5
```

blob file key Titan GC Titan

= 1 / discardable_ratio

= 1 / (1 - discarable_ratio)

Titan GC Titan GC IO CPU

```
[rocksdb]
rate-bytes-per-second
```

Titan RocksDB compaction IO RocksDB compaction IO CPU Titan RocksDB compaction Titan GC IO
RocksDB compaction Titan GC IO / CPU IO

8.2.3 Titan

rocksdb.defaultcf.titan.blob-run-mode Titanblob-run-mode

- “kNormal” Titan
- “kReadonly” value RocksDB
- “kFallback” value RocksDB RocksDB compaction Titan blob file value RocksDB

Titan blob-run-mode = “kFallback” tikv-ctl compaction blob file size 0 rocksdb.titan.enabled = false TiKV

Titan

8.2.4 Level Merge

TiKV 4.0 Titan Titan GC level mergeLevel merge

```
[rocksdb.defaultcf.titan]
level-merge = true
```

level merge

- Titan
- Titan GC
- Titan

level mege Titan RocksDB

Level merge Titan Titan TiKV Level merge

9 TiFlash HTAP

—— OLTP OLAP ETL “” T+N

HTAP HTAP OLTP OLAP 2014 Garnter / (HTAP) “” “” “”

TiDB NewSQL HTAP TiDB HTAP HTAP TiFlash TiFlash

- [9.1 TiDB HTAP](#)
- [9.2 TiFlash](#)
- [9.3 TiFlash](#)

9.1 TiDB HTAP

HTAP Hybrid Transactional / Analytical Processing 2014 Gartner TP AP HTAP

HTAP TiDB

1.

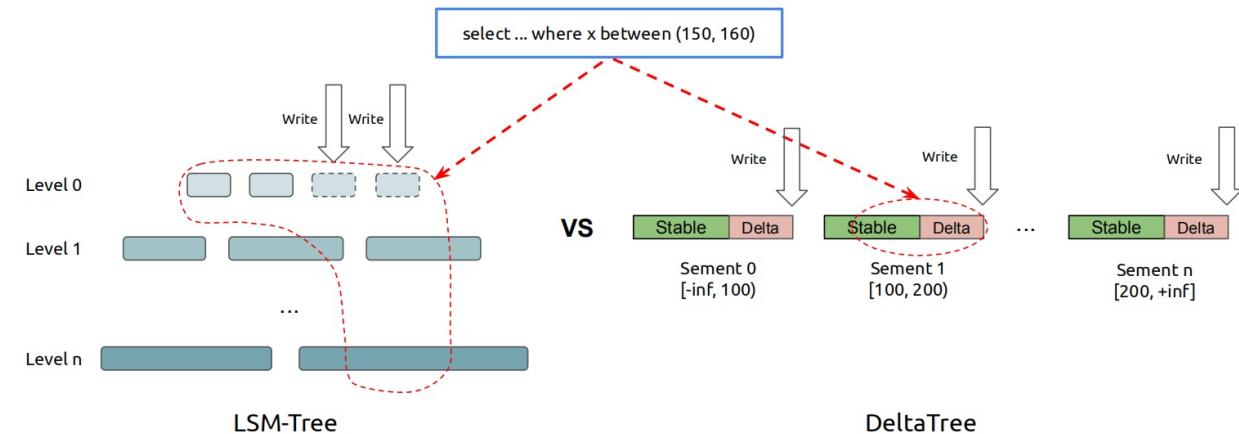
-
-

TiKV TiFlash

-
- ClickHouse SIMD
- TiKV
- TiDB

2. Delta Tree

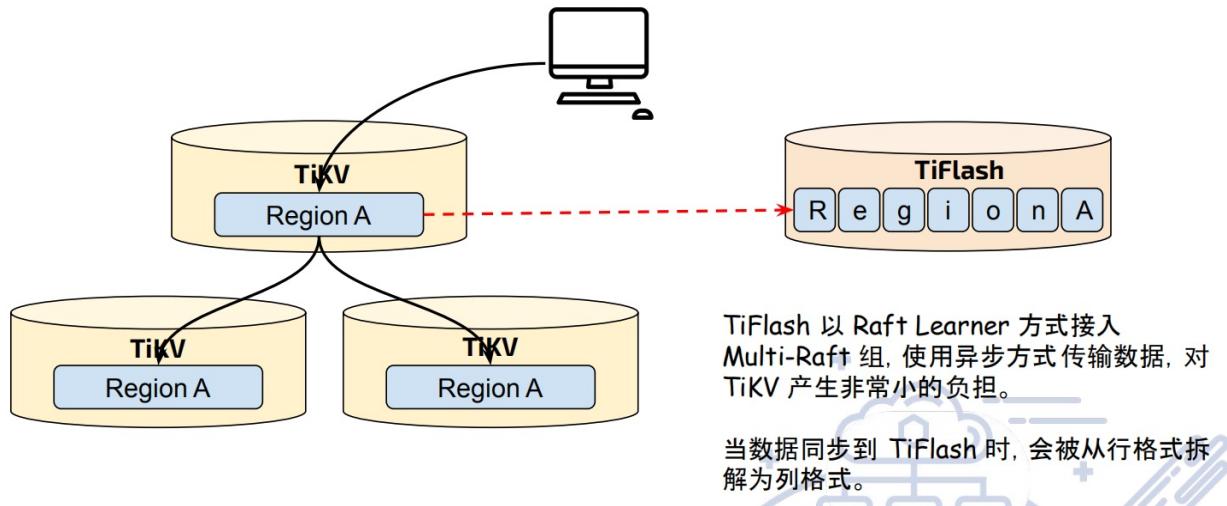
TiFlash Delta Main TiFlash Delta Main LSM Delta Tree LSM LSM TiFlash



3.

TiFlash TiDB Multi-Raft Raft Learner TiFlash TiKV TiFlash TiKV TiFlash TiKV

Raft Learner - Sync



4.

TiFlash TiKV TiDB TiSpark BI

TiDB TiDB



5.

ETL TiDB

6.

TiFlash TiKV TiKV TiDB HTAP

9.2 TiFlash

TiFlash Schema ClickHouse TiFlash SISnapshot Isolation TiFlash

9.2.1

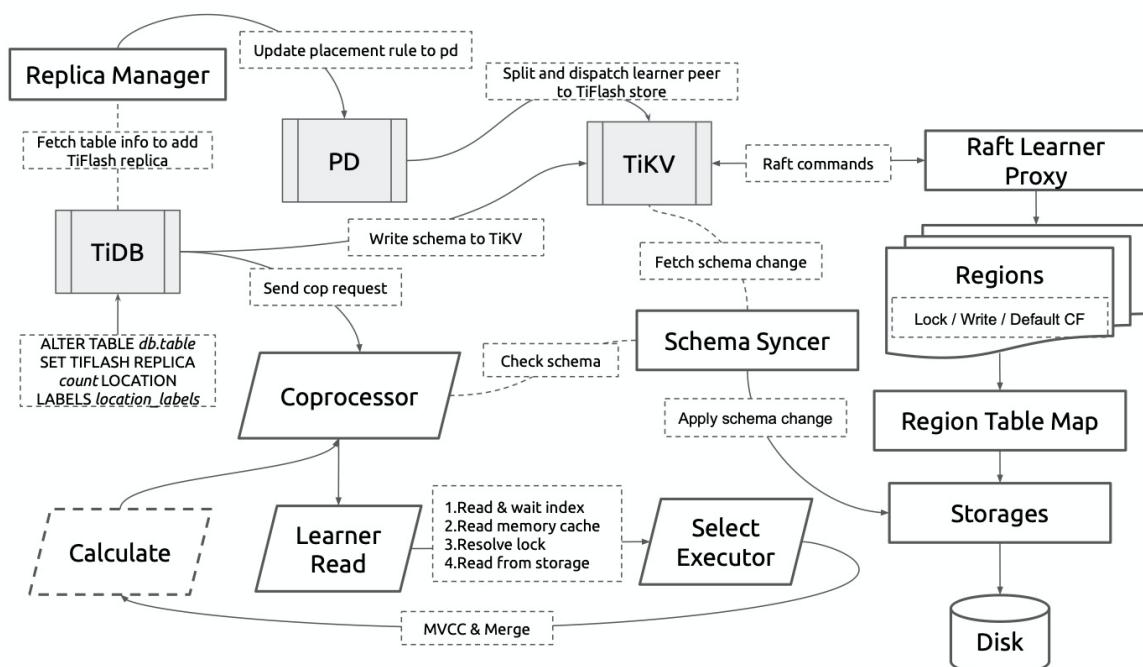


TiFlash TiDB OLAP TiDB TiFlash MPP TiDB TiFlash

TiFlash Raft Learner Store TiKV TiKV Key Region Region Peer Raft Group Raft Group Peer Leader Follower Learner TiKV Peer TiFlash Peer Learner TiFlash TiKV PD Region TiDB TiSpark TiKV Coprocessor TiFlash

TiDB Table SchemaIndexRecord TiKV Table TiDB Table Schema Key-Value Region Multi-Raft TiFlash TiFlash Schema

9.2.2



TiFlash

1. Replica Manager

Index TiFlash Table TiFlash PD Region



TiFlash PD ETCD Replica Manager PD TiDB TiDB TiFlash DDL PD Placement Rule PD TiKV Key
Region Learner Peer TiFlash Table TiFlash Replica Manager TiFlash Table TiDB

2. Raft Learner Proxy

Raft Learner TiFlash Region Learner Peer TiKV TiKV Raft Proxy Proxy TiFlash Region TiFlash
LearnerTiFLash TiFlash Raft Schema TiFlash Raft WAL
TiDB Percolator TiKV 3 CF Column Family WriteDefaultLockTiFlash Region TiFlash CF Region
Table Schema

3. Schema Syncer

TiFlash TiKV Table Schema TiDB Schema

4. Learner Read / Coprocessor



CH TiSpark CHSpark TiDB TiFlash Learner Read Raft Group Index Raft Region Version Conf Version
Split/Merge/ChangePeer

Region Region PD TiKV Timestamp Snapshot Read PD Table Schema Version TiDB

1. Schema
2. Region Leader Peer Index Learner Applied Index
3. Version Conf Version Lock CF
4. Region
5. Timestamp Snapshot Read
6. Coprocessor ClickHouse

9.2.3 TiFlash OLAP

OLAP

-
-
-

TiFlash

(1) I/O

- I/O
-
- I/O

(2) CPU

- CPU Cache
-

9.2.4 TiKV TiFlash

TiFlash TiKV TiFlash

9.2.5

TiFlash TiDB HTAP

9.3 TiFlash

TiDB TiSpark TiFlash TiDB OLAP TiSpark OLAP

9.3.1 TiFlash

TiFlash TiKV MySQL TiDB DDL TiFlash

```
ALTER TABLE table_name SET TIFLASH REPLICA count;
```

- count 0 TiFlash
- DDL

1 2 TiFlash

```
ALTER TABLE `tpch50`.`lineitem` SET TIFLASH REPLICA 2;
```

2

```
ALTER TABLE `tpch50`.`lineitem` SET TIFLASH REPLICA 0;
```

SQL WHERE WHERE TiFlash

```
SELECT * FROM information_schema.tiflash_replica
WHERE TABLE_SCHEMA = '<db_name>' and TABLE_NAME = '<table_name>';
```

AVAILABLE TiFlash

t DDL TiFlash TiFlash

```
CREATE TABLE table_name LIKE t;
```

9.3.2 TiDB TiFlash

TiDB TiFlash TiFlash engine CBO (Cost Based Optimization)

1. CBO

```
TiFlash TiDB CBO TiFlash TiFlash      explain analyze
mysql> explain analyze select * from t;
+-----+-----+-----+-----+
| id   | count | task    | operator info          | execution info
|      |       |         | memory                 | time:54.230814ms, loops:2,
+-----+-----+-----+-----+
| TableReader_7 | 10000.00 | root     | data:TableScan_6
rows:1, rpc time:54.147725ms | 163 Bytes |
| \TableScan_6  | 10000.00 | cop[tiflash] | table:t, range:[-inf,+inf], keep order:false, stats:pseudo | time:40ms, loops:1, rows:1
|               |           | N/A                  |
+-----+-----+-----+-----+
2 rows in set (0.06 sec)
```

2. Engine

Engine engine engine tikv tiflash 2

(1) SESSION GLOBAL

```
set@@session.tidb_isolation_read_engines=" engine list";
```

```
set SESSION tidb_isolation_read_engines = " engine list";
```

TiDB

```
set SESSION tidb_isolation_read_engines = "tikv,tiflash";
```

TiFlash

```
set SESSION tidb_isolation_read_engines = "tiflash";
```

(2) TiDB INSTANCE INSTANCE "tikv,tiflash" SESSION "tikv" tikv TiDB

```
[isolation-read]
engines = ["tikv", "tiflash"]
```

"tikv,tiflash" tikv tiflash

engine engine tikv engine tiflash tiflash engine

Engine engine

3. hint

hint TiDB TiFlash CBO engine

```
SELECT /*+ read_from_storage(tiflash[t]) */ * FROM t;
```

hint tiflash tiflash

10 TiDB

TiDB RBAC RBAC TLS

- [10.1](#)
- [10.2 RBAC](#)
- [10.3](#)

10.1

TiDB

TiDB MySQL TiDB MySQL

TiDB

10.1.1

-
-
-

10.1.2

- mysql.user
- mysql.tables_priv
- mysql.db

TiDB

1.mysql.user

mysql.user

```

CREATE TABLE `user` (
  `Host` char(64) NOT NULL,
  `User` char(32) NOT NULL,
  `authentication_string` text DEFAULT NULL,
  `Select_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Insert_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Update_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Delete_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Create_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Drop_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Process_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Grant_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `References_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Alter_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Show_db_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Super_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Create_tmp_table_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Lock_tables_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Execute_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Create_view_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Show_view_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Create_routine_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Alter_routine_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Index_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Create_user_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Event_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Trigger_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Create_role_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Drop_role_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Account_locked` enum('N','Y') NOT NULL DEFAULT 'N',
  `Shutdown_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Reload_priv` enum('N','Y') DEFAULT 'N',
  `File_priv` enum('N','Y') DEFAULT 'N',
  PRIMARY KEY (`Host`,`User`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_bin

```

mysql.user Host User Host user ip Host 192.168.1.7 root mysql.user
root @ 192.168.1.* User root ip 192.168.1.7 Host

root @ 172.16.*

root @ 192.168.1.*

2.mysql.table, mysql.db

mysql.table mysql.db

```

CREATE TABLE `tables_priv` (
  `Host` char(60) NOT NULL,
  `DB` char(64) NOT NULL,
  `User` char(32) NOT NULL,
  `Table_name` char(64) NOT NULL,
  `Grantor` char(77) DEFAULT NULL,
  `Timestamp` timestamp DEFAULT CURRENT_TIMESTAMP,
  `Table_priv` set('Select','Insert','Update','Delete','Create','Drop','Grant','Index','Alter','Create View','Show View','Trigger','References') DEFAULT NULL,
  `Column_priv` set('Select','Insert','Update') DEFAULT NULL,
  PRIMARY KEY (`Host`,`DB`,`User`,`Table_name`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_bin

```

```

CREATE TABLE `db` (
  `Host` char(60) NOT NULL,
  `DB` char(64) NOT NULL,
  `User` char(32) NOT NULL,
  `Select_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Insert_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Update_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Delete_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Create_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Drop_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Grant_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `References_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Index_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Alter_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Create_tmp_table_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Lock_tables_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Create_view_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Show_view_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Create_routine_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Alter_routine_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Execute_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Event_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  `Trigger_priv` enum('N','Y') NOT NULL DEFAULT 'N',
  PRIMARY KEY (`Host`,`DB`,`User`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_bin

```

mysql.tables_priv mysql.db

bitmask User, Host

GRANT, REVOKE TiDB sql INSERT, UPDATE, DELETE

10.1.3

developer 192.168.0.* 'test_user'

```
root> CREATE USER 'developer'@'192.168.0.%' IDENTIFIED BY 'test_user';
```

developer read_table write_table

```
root> GRANT SELECT ON app.read_table TO 'developer'@'192.168.0.%';
root> GRANT INSERT, UPDATE ON app.write_table TO 'developer'@'192.168.0.%';
```

developer

```
root> SHOW GRANTS FOR 'developer'@'192.168.0.%';
GRANT USAGE ON *.* TO 'developer'@'192.168.0.%'
GRANT Select ON app.read_table TO 'developer'@'192.168.0.%'
GRANT Insert,Update,Delete ON app.write_table TO 'developer'@'192.168.0.%'
```

developer

```
developer> SEECT * FROM app.read_table;
Empty set (0.01 sec)
developer> INSERT INTO write_table VALUES (1),(2),(3);
Query OK, 3 rows affected (0.00 sec)
Records: 3  Duplicates: 0  Warnings: 0
```

developer write_table TiDB

```
developer> INSERT INTO read_table VALUES (1),(2),(3);
ERROR 1142 (42000): INSERT command denied to user 'developer'@'192.168.0.%' for table 'read_table'
```

```
GRANT_OPTION GRANT,REVOKE developer read_table
```

```
root> REVOKE SELECT ON app.read_table from 'developer'@'192.168.0.%';
```

```
CREATE USER
```

```
root> ALTER USER 'developer'@'192.168.0.%' IDENTIFIED BY 'password';
root> DROP USER 'developer'@'192.168.0.%'
```

10.1.4

TiDB TiDB RBAC

10.2 RBAC

TiDB -- RBAC

Role-based access control RBAC

MAC (Mandatory access control) DAC (Discretionary Access Control) RBAC

TiDB (RBAC) MySQL 8.0 RBAC MySQL RBAC

10.2.1 RBAC

- • • •

10.2.2 RBAC

- TiDB
 - session ActiveRole SET ROLE default_roles ActiveRole
 - mysql.user Account_Locked Y.

- mysql.role_edges r1 test

```
+-----+-----+-----+-----+
| FROM_HOST | FROM_USER | TO_HOST | TO_USER | WITH_ADMIN_OPTION |
+-----+-----+-----+-----+
| %         | r1       | %         | test     | N
+-----+-----+-----+-----+
```

- mysql.default_roles

```
+-----+-----+-----+
| HOST | USER | DEFAULT_ROLE_HOST | DEFAULT_ROLE_USER |
+-----+-----+-----+
| %    | test |             | r_1           |
+-----+-----+-----+
```

10.2.3 RBAC

- r_1r_2

```
CREATE ROLE `r_1`@`%`, `r_2`@`%`;
```

- r_1

```
GRANT SELECT ON db_1.* TO 'r_1'@'%';
```

- r_1 test@'%'

```
grant r_1 to test@'%';
```

•

```
SET DEFAULT ROLE 'r_1';
```

- session

```
SET ROLE 'r_1';
```

•

```
SELECT CURRENT_ROLE();
```

•

```
TiDB > SHOW GRANTS FOR 'test'@'%' USING 'r_1';
+-----+
| Grants for test@%          |
+-----+
| GRANT USAGE ON *.* TO 'test'@'%'      |
| GRANT Select ON test.* TO 'test'@'%'   |
| GRANT 'r_1'@'%' TO 'test'@'%'          |
+-----+
```

```
REVOKE 'r_1' FROM 'test'@'%', 'root'@'%';
```

10.2.4

bi_user

```

# reader
root@127.0.0.1:(none)>create role reader@'%';
Query OK, 0 rows affected (0.012 sec)

# reader mysql.role_edges
root@127.0.0.1:mysql>grant select on mysql.role_edges to reader'%';
Query OK, 0 rows affected (0.017 sec)

# bi_user
root@127.0.0.1:(none)>create user bi_user@'%';
Query OK, 0 rows affected (0.011 sec)

# reader bi_user
root@127.0.0.1:mysql>grant reader to bi_user'%';
Query OK, 0 rows affected (0.014 sec)

# bi_user
bi_user@127.0.0.1:(none)>show databases;
+-----+
| Database      |
+-----+
| INFORMATION_SCHEMA |
+-----+
1 row in set (0.000 sec)

# bi_user
bi_user@127.0.0.1:(none)>SELECT CURRENT_ROLE();
+-----+
| CURRENT_ROLE() |
+-----+
|          |
+-----+
1 row in set (0.000 sec)

# session bi_user reader
bi_user@127.0.0.1:(none)>set role reader;
Query OK, 0 rows affected (0.000 sec)

# bi_user
bi_user@127.0.0.1:(none)>SELECT CURRENT_ROLE();
+-----+
| CURRENT_ROLE() |
+-----+
| `reader`@`%`   |
+-----+
1 row in set (0.000 sec)

# bi_user mysql
bi_user@127.0.0.1:mysql>select * from role_edges;
+-----+-----+-----+-----+
| FROM_HOST | FROM_USER | TO_HOST | TO_USER | WITH_ADMIN_OPTION |
+-----+-----+-----+-----+
| %        | reader    | %        | bi_user  | N           |
+-----+-----+-----+-----+
1 row in set (0.000 sec)

# bi_user delete
bi_user@127.0.0.1:mysql>delete from role_edges;
ERROR 1105 (HY000): privilege check fail

# bi_user
bi_user@127.0.0.1:mysql>select * from user;
ERROR 1142 (42000): SELECT command denied to user 'bi_user'@'127.0.0.1' for table 'user'

# bi_user
bi_user@127.0.0.1:(none)>use mysql
ERROR 1044 (42000): Access denied for user 'bi_user'@'%' to database 'mysql'

```

10.2.5

RBAC RBAC TiDB ——

10.3

TiDB 3.0.8 TiDB TiDB MySQL

10.3.1

TiDB TiDB TiDB

-
-
-

10.3.2

TiDB X.509 CA()

- CA CA
- TiDB CA CA
-

TiDB subject issuer cipher mysql.global_priv

TiDB

10.3.3

1. CA

[OpenSSL](#) OpenSSL

```
sudo apt-get install openssl
```

CA CA

```
sudo openssl genrsa 2048 > ca-key.pem
```

CA CA

```
sudo openssl req -new -x509 -nodes -days 365000 -key ca-key.pem -out ca-cert.pem
```

```
Country Name (2 letter code) [AU]:US
State or Province Name (full name) [Some-State]:California
Locality Name (eg, city) []:San Francisco
Organization Name (eg, company) [Internet Widgits Pty Ltd]:PingCAP Inc.
Organizational Unit Name (eg, section) []:TiDB
Common Name (e.g. server FQDN or YOUR name) []:TiDB admin
Email Address []:s@pingcap.com
```

CA CA

2.

```
sudo openssl req -newkey rsa:2048 -days 365000 -nodes -keyout server-key.pem -out server-req.pem
```

```
Country Name (2 letter code) [AU]:US
State or Province Name (full name) [Some-State]:California
Locality Name (eg, city) []:San Francisco
Organization Name (eg, company) [Internet Widgits Pty Ltd]:PingCAP Inc.
Organizational Unit Name (eg, section) []:TiKV
Common Name (e.g. server FQDN or YOUR name) []:TiKV Test Server
Email Address []:k@pingcap.com

Please enter the following 'extra' attributes
to be sent with your certificate request
A challenge password []:
An optional company name []:
```

RAS

```
sudo openssl rsa -in server-key.pem -out server-key.pem
```

CA key

```
sudo openssl x509 -req -in server-req.pem -days 365000 -CA ca-cert.pem -CAkey ca-key.pem -set_serial 01 -out server-cert.pem
```

3.

```
sudo openssl req -newkey rsa:2048 -days 365000 -nodes -keyout client-key.pem -out client-req.pem
sudo openssl rsa -in client-key.pem -out client-key.pem
sudo openssl x509 -req -in client-req.pem -days 365000 -CA ca-cert.pem -CAkey ca-key.pem -set_serial 01 -out client-cert.pem
```

```
openssl verify -CAfile ca-cert.pem server-cert.pem client-cert.pem
```

```
server-cert.pem: OK
client-cert.pem: OK
```

4. TiDB

TiDB

TiDB [security] CA

```
[security]
ssl-cert ="path/to/server-cert.pem"
ssl-key ="path/to/server-key.pem"
ssl-ca="path/to/ca-cert.pem"
```

TiDB

```
[INFO] [server.go:264] ["secure connection is enabled"] ["client verification enabled=true"]
```

TiDB --ssl-ca

```
mysql -utest -h0.0.0.0 -P4000 --ssl-cert /path/to/client-cert.pem --ssl-key /path/to/client-key.pem --ssl-ca /path/to/ca-cert.pem
```

TiDB subjectissuercipher

- subject subject
- issuer CA subject CA

subject issuer CA

```
create user 'u1'@'%' require issuer '/C=US/ST=California/L=San Francisco/O=PingCAP Inc./OU=TiDB/CN=TiDB admin/emailAddress=s@pingcap.com' subject '/C=US/ST=California/L=San Francisco/O=PingCAP Inc./OU=TiDB/CN=tpch-user1/emailAddress=zz@pingcap.com' cipher 'TLS_AES_256_GCM_SHA384'
```

GRANT

```
> create user 'u1'@'%';
> grant all on *.* to 'u1'@'%' require issuer '/C=US/ST=California/L=San Francisco/O=PingCAP Inc./OU=TiDB/CN=TiDB admin/emailAddress=s@pingcap.com' subject '/C=US/ST=California/L=San Francisco/O=PingCAP Inc./OU=TiDB/CN=tpch-user1/emailAddress=zz@pingcap.com' cipher 'TLS_AES_256_GCM_SHA384';
```

TiDB

- SSL Server CA
- Issuer
- Subject

ERROR 1045 (28000): Access denied

10.3.4

CA CA CA

1. CA

CA ca-key.pem CA

```
mv ca-key.pem ca-key.old.pem
mv ca-cert.pem ca-cert.old.pem
```

CA

```
sudo openssl genrsa 2048 > ca-key.pem
```

CA

```
sudo openssl req -new -x509 -nodes -days 365000 -key ca-key.pem -out ca-cert.new.pem
```

CA

```
cat ca-cert.new.pem ca-cert.old.pem > ca-cert.pem
```

CA TiDB Server CA

2.

```
sudo openssl req -newkey rsa:2048 -days 365000 -nodes -keyout server-key.new.pem -out server-req.new.pem
sudo openssl rsa -in server-key.new.pem -out server-key.new.pem
```

CA

```
sudo openssl x509 -req -in server-req.new.pem -days 365000 -CA ca-cert.pem -CAkey ca-key.pem -set_serial 01 -out server-cert.new.pem
```

TiDB

3.

```
sudo openssl req -newkey rsa:2048 -days 365000 -nodes -keyout client-key.new.pem -out client-req.new.pem
sudo openssl rsa -in client-key.new.pem -out client-key.new.pem
```

require subject CA

```
sudo openssl x509 -req -in client-req.new.pem -days 365000 -CA ca-cert.pem -CAkey ca-key.pem -set_serial 01 -out client-cert.new.pem
```

TiDB

```
mysql -utest -h0.0.0.0 -P4000 --ssl-cert /path/to/client-cert.new.pem --ssl-key /path/to/client-key.new.pem --ssl-ca /path/to/ca-cert.pem
```

11 TiSpark

TiSpark PingCAP OLAP Spark SQL TiKV TiDB HTAP (Hybrid Transactional/Analytical Processing)

TiSpark TiSpark TiSpark PingCAP TiFlash TiSpark

- [11.1 TiSpark](#)
- [11.2 TiSpark](#)
- [11.3 TiSpark on TiFlash](#)
- [11.4 TiSpark](#)

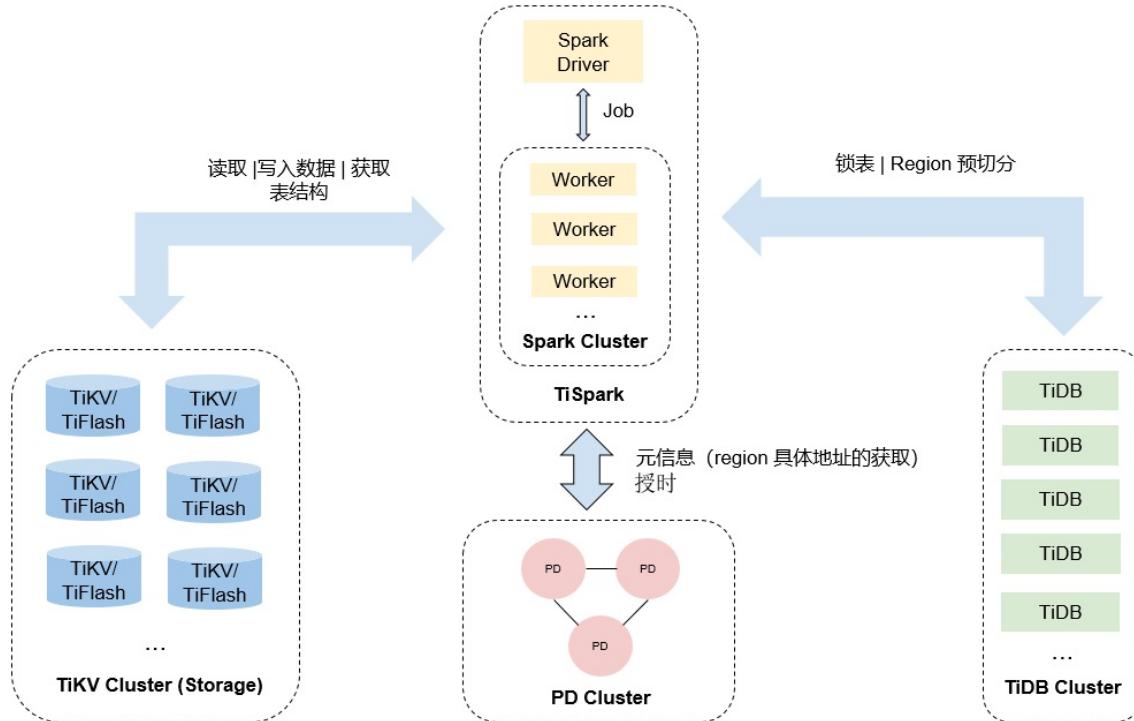
11.1 TiSpark

TiSpark PingCAP OLAP Spark TiKV TiFlash TiDB HTAP (Hybrid Transactional/Analytical Processing)
TiSpark TiKV Placement Driver (PD) Spark

TiSpark Spark Apache Spark Spark

11.1.1

TiSpark Spark SQL TiKV OLAP



- TiSpark TiKV PD Java Client TiSpark gRPC TiKV PD TiKV / TiFlash Key-Value Pair TiSpark SQL PD TiKV Region
- TiSpark TiDB Region
- TiSpark Driver
 - PD Client PD TiDB metadata TiDB metadata Spark metadata TiSpark TiDB
 - Spark SQL TiKV
 - PD Region
 - Region TiKV
 -
 - region
 -
- Spark Executor
 - TiKV
 - TiKV Spark SQL

11.1.2 TiKV Java Client

TiSpark TiKV Key-Value Pair TiSpark TiKV TiKV Java Client gRPC TiKV Server TiKV API

- TiKV Table Schema TiKV Schema Spark SQL Schema
- TiKV

- TiKV Key-Value Pair Key-Value Pair Spark SQL TiDB
- TiKV TiKV
- SQL TiSpark Java TiKV Client TiDB

11.1.3 TiKV TiSpark

Java TiKV Client TiSpark TiKV TiKV Spark Spark Spark Plan Spark SparkSessionExtensions
Spark SQL Plan Spark API

The screenshot shows the official Apache Spark documentation for the `SparkSessionExtensions` class. The title is "SparkSessionExtensions" under the "org.apache.spark.sql" package. It includes a "Related Doc: package sql" link and developer API/experimental status indicators.

class `SparkSessionExtensions` extends AnyRef

Holder for injection points to the `SparkSession`. We make NO guarantee about the stability regarding binary compatibility and source compatibility of methods here.

This current provides the following extension points:

- Analyzer Rules.
- Check Analysis Rules.
- Optimizer Rules.
- Planning Strategies.
- Customized Parser.
- (External) Catalog listeners.

The extensions can be used by calling `withExtension` on the `SparkSession.Builder`, for example:

```
SparkSession.builder()
  .master("...")
  .conf("...", true)
  .withExtensions { extensions =>
    extensions.injectResolutionRule { session =>
      ...
    }
    extensions.injectParser { (session, parser) =>
      ...
    }
  }
  .getOrCreate()
```

Note that none of the injected builders should assume that the `SparkSession` is fully initialized and should not touch the session's internals (e.g. the `SessionState`).

TiSpark

spark.sql.extensions org.apache.spark.sql.TiExtensions

TiExtensionsSpark SparkSessionExtensions TiSpark Spark

The screenshot shows a code editor with the file `TiExtensions.scala` open. The code defines a class `TiExtensions` that implements `SparkSessionExtensions`. It contains methods for injecting resolution rules, parsers, and planner strategies. The code editor interface shows the project structure on the left, with files like `TiAggregation.scala`, `TiContext.scala`, `TiStrategy.scala`, and others.

```
16   tiContextMap.put(sparkSession, tiContext)
17   tiContext
18 }
19 }
20
21 override def apply(e: SparkSessionExtensions): Unit = {
22   e.injectParser(TiParser(getOrCreateTiContext))
23   e.injectResolutionRule(TiDDLRule(getOrCreateTiContext))
24   e.injectResolutionRule(TiResolutionRule(getOrCreateTiContext))
25   e.injectPlannerStrategy(TiStrategy(getOrCreateTiContext))
26 }
27
28 }
```

Spark SQL Spark SQL Catalyst DataSource API TiSpark Spark Plan SQL SQL TiKV TiKV
Spark TiKV Spark TiKV



11.1.4

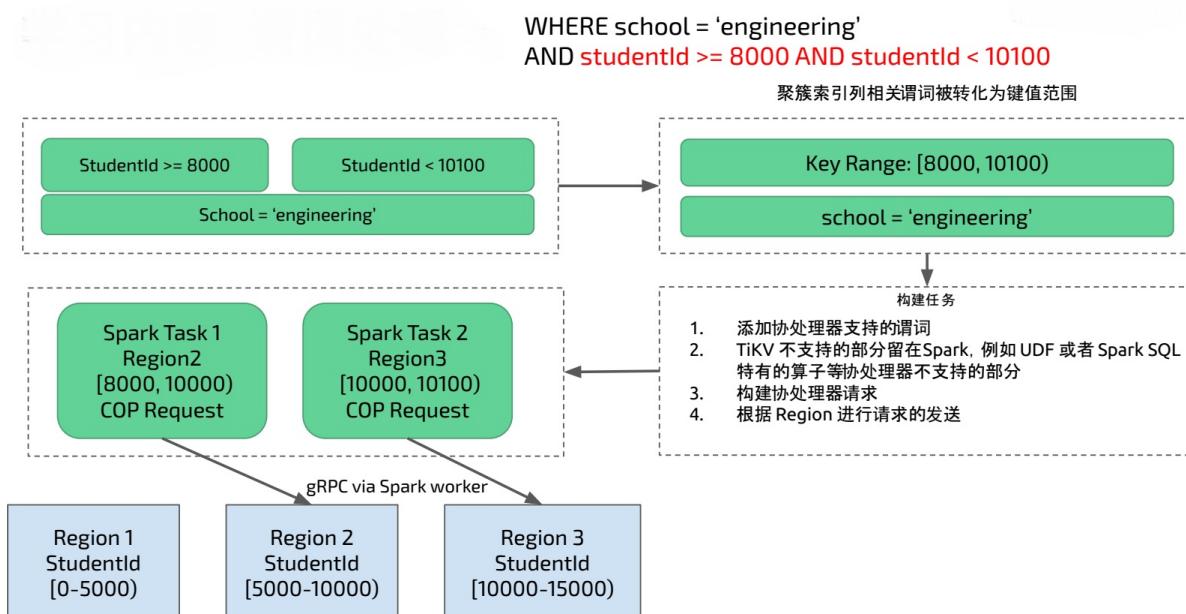
Spark SQL

Spark SQL SQL student TiDB studentID school

SELECT class, avg(score) FROM student

WHERE school = 'engineering' AND studentId >= 8000 AND studentId < 10100

GROUP BY class ;



studentID TiKV Region Region 2 studentId [5000 - 10000] Region 3 studentID [10000 - 15000] SQL TiKV
8000 10100 TiSpark Region 2 Region 3 TiSpark Spark Executor TiKV TiKV TiKV TiKV Spark
SQL

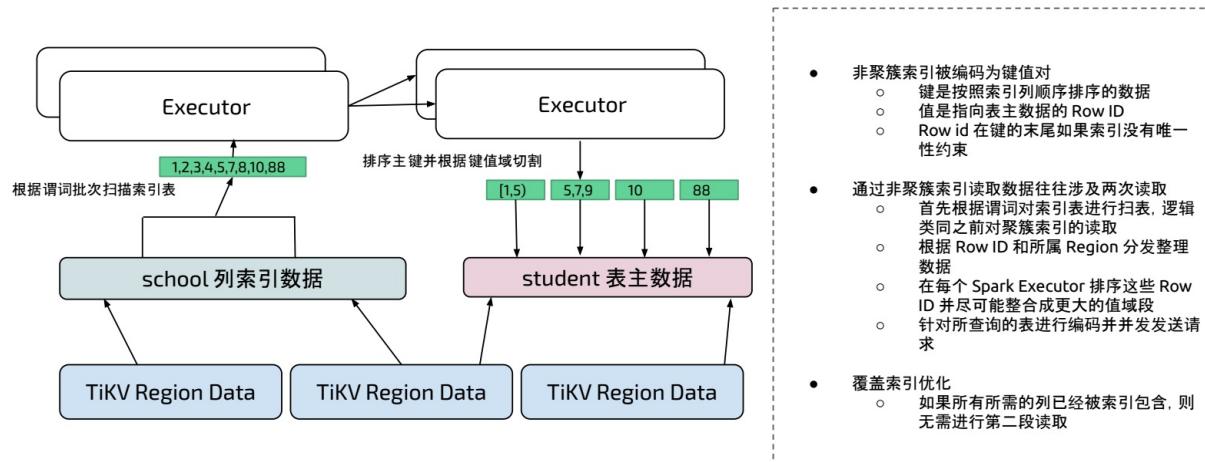
11.1.5

Row ID

- Row ID Region
- Spark Executor Row ID
-

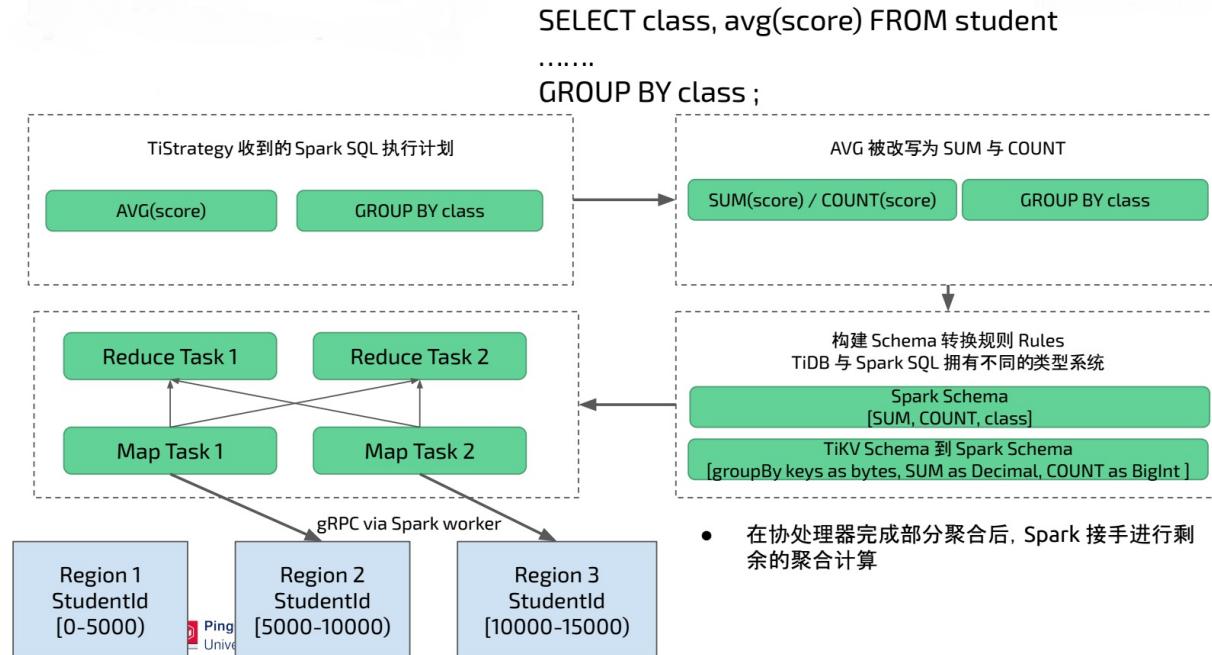
school 1,2,3,4,5,7,8,10,88 Row ID Spark Executor Row ID Row ID student TiKV Spark

WHERE school = 'engineering'
AND studentId >= 8000 AND studentId < 10100



11.1.6

TiStrategy TiSpark Spark SQL AVG(score) GROUP BY class TiSpark AVG SUM(score) / COUNT(score) TiSpark
TiDB TiKV TiDB group by bytes SUM Decimal Count Bigint TiDB Spark Spark Spark
Executor Map Task TiKV Reduce Task Map Task



11.1.7

TiSpark TiDB JDBC TiDB TiSpark TiKV Spark DataFrame TiSpark DataFrame TiKV TiKV Java Client TiKV

- DataFrame Region
- TiKV TiKV Percolator
 - Spark Driver TiKV
 - Spark Executor DataFrame TiKV Key-Value Pair gRPC DataFrame TiKV
 - Spark Driver Spark Executor Percolator
 - Spark Executor

11.1.8

TiSpark TiKV Java Client Spark Spark Plan TiKV Spark
TiKV TiSpark TiKV

TiKV TiSpark TiSpark TiKV TiSpark

11.2 TiSpark

TiSpark TiSpark

11.2.1 TiSpark

TiSpark Apache Spark Apache Spark 2.1 TiSpark
[Release](#) YARN Standalone Apache Spark Spark [TiSpark](#)
 TiSpark JAR

TiSpark

1. Spark Driver Executor TiSpark JAR
2. Spark spark-defaults.conf

```
spark.sql.extensions      org.apache.spark.sql.TiExtensions
spark.tispark.pd.addresses  pd-host1:port1,pd-host2:port2,pd-host3:port3
```

Standalone TiSpark

1. SPARK_HOME/spark-defaults.conf
2. Spark TiSpark JAR spark-shell

```
./spark-shell --jars /path/your-tispark.jar
```

11.2.2 TiSpark

TiSpark Spark

11.2.2.1 spark-shell

spark-shell Spark API

spark-shell lineitem

```
scala>spark.sql("use tpch")
tpch
scala>spark.sql("select count(*) from lineitem").show
lineitem

+-----+
| Count (1) |
+-----+
| 600000000 |
+-----+
```

Spark SQL sql

```

scala> spark.sql(
    """select
       |   l_returnflag,
       |   l_linenumber,
       |   sum(l_quantity) as sum_qty,
       |   sum(l_extendedprice) as sum_base_price,
       |   sum(l_extendedprice * (1 - l_discount)) as sum_disc_price,
       |   sum(l_extendedprice * (1 - l_discount) * (1 + l_tax)) as sum_charge,
       |   avg(l_quantity) as avg_qty,
       |   avg(l_extendedprice) as avg_price,
       |   avg(l_discount) as avg_disc,
       |   count(*) as count_order
      |from
      |   lineitem
     |where
     |   l_shipdate <= date '1998-12-01' - interval '90' day
    |group by
       |   l_returnflag,
       |   l_linenumber
    |order by
       |   l_returnflag,
       |   l_linenumber
    """.stripMargin).show
scala>
+-----+-----+-----+-----+
|l_returnflag|l_linenumber|sum_qty|sum_base_price|sum_disc_price|
+-----+-----+-----+-----+
|        A|          F| 380456.00| 532348211.65| 505822441.4861|
|        N|          F|    8971.00| 12384801.37| 11798257.2080|
|        N|          O| 742802.00| 1041502841.45| 989737518.6346|
|        R|          F| 381449.00| 534594445.35| 507996454.4067|
+-----+-----+-----+-----+

```

11.2.2.2 Spark SQL

TiSpark Spark SQL spark-sql sql

spark-sql lineitem

```

spark-sql> use tpch;
tpch
spark-sql> select count(*) from lineitem;
lineitem
2000
Time taken: 0.673 seconds, Fetched 1 row(s)

```

11.2.2.3 JDBC TiSpark

Thrift JDBC TiSpark

beeline JDBC TiSpark

```

beeline> !connect jdbc:hive2://localhost:10000
1: jdbc:hive2://localhost:10000> use testdb;
+-----+---+
| Result |
+-----+---+
+-----+---+
No rows selected (0.013 seconds)
select count(*) from account;
+-----+---+
| count(1) |
+-----+---+
| 10000000 |
+-----+---+
1 row selected (1.97 seconds)

```

11.2.3 TiSpark

11.2.3.1 PySpark TiSpark

TiSpark on PySpark TiSpark Python Python PySpark python

PySpark TiSpark

```
./bin/pyspark --jars /PATH/tispark-${name_with_version}.jar
# Query as you are in spark-shell
spark.sql("show databases").show()
spark.sql("use tpch_test")
spark.sql("show tables").show()
spark.sql("select count(*) from customer").show()
# Result
+-----+
|count(1)|
+-----+
|     150|
+-----+
```

pip pytispark TiSpark

pip pytispark

```
pip install pytispark
```

python test.py

```
import pytispark.pytispark as pti
from pyspark.sql import SparkSession
spark = SparkSession.builder.getOrCreate()
ti = pti.TiContext(spark)
ti.tidbMapDatabase("tpch_test")
spark.sql("select count(*) from customer").show()
```

spark-submit

```
./bin/spark-submit --jars /PATH/tispark-${name_with_version}.jar test.py
# Result:
+-----+
|count(1)|
+-----+
|     150|
+-----+
```

11.2.3.2 TiSparkR TiSpark

TiSparkR TiSpark R R PySpark TiSparkR library

PySpark TiSpark

```
./bin/sparkR --jars /PATH/tispark-${name_with_version}.jar
sql("use tpch_test")
count <- sql("select count(*) from customer")
head(count)
# Result
+-----+
|count(1)|
+-----+
|     150|
+-----+
```

SparkR TiSpark R test.R

```
library(SparkR)
sparkR.session()
sql("use tpch_test")
count <- sql("select count(*) from customer")
head(count)
```

spark-submit

```
./bin/spark-submit --jars /PATH/tispark-${name_with_version}.jar test.R
# Result:
+-----+
|count(1)|
+-----+
|      150|
+-----+
```

11.2.3.3 TiSpark TiFlash

TiSpark TiFlash

11.3 TiSpark on TiFlash

TiFlash TiSpark TiKV TiSpark TiFlash TiSpark TiFlash TiKV TiFlash Region TiSpark TiKV
TiFlash TiSpark TiFlash TiDB TiFlash TiSpark

11.3.1 TiSpark TiFlash

TiSpark TiSpark JAR TiExtension PD TiFlash TiSpark

TiSpark TiDB engine TiFlash:

```
spark.tispark.use.tiflash true ( false)
```

:

1. spark-defaults.conf

```
spark.tispark.use.tiflash true
```

1. spark-shell ThriftServer

```
--conf spark.tispark.use.tiflash = true
```

1. Spark SThell :

```
spark.conf.set("spark.tispark.use.tiflash", true)
```

1. ThriftServer beeline :

```
set spark.tispark.use.tiflash = true;
```

true TiFlash false TiKV true TiFlash TiFlash

11.3.2

1. Region TiSpark TiFlash Region TiKV Region Spark Region TiSpark Region Region
48M 2 TiSpark Split Region

```
spark.tispark.partition_per_split 2
```

2. TiFlash TiSpark

```
spark.locality.wait 0s
```

11.4 TiSpark

TiDB

11.4.1 Hive

Hadoop Hive Hive TiDB Hadoop TiDB

1. SQL Hive/TiDB +
2. Hive
3. hive-site.xml Spark hive-site.xml SPARK_HOME/conf hive-site.xml Hive Metastore Spark Hive

11.4.2 beeline + Livy + Spark + Tispark

1.

- Livy 0.6 Spark
- beeline 3.1.1 Livy
- Spark 2.4.0 yarn Spark
- Tispark Jar 2.1.8 TiKV
- YARN

2.

```
function runMixSQLOnLivy(){
export HIVE_HOME=/usr/local/share/apache-Hive-3.1.1-bin
/usr/local/share/apache-Hive-3.1.1-bin/bin/beeline -n hdfs_user_name -p hdfs_user_pwd --verbose=false --color=false \
-u "jdbc:Hive2://bj0000,bj0001,bj0002:2222;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=mix-livy" \
--Hiveconf livy.session.conf.spark.sql.extensions=org.apache.spark.sql.TiExtensions \
--Hiveconf livy.session.conf.spark.tispark.pd.addresses=10.10.10.10:2379,10.10.10.10:2379,10.10.10.10:2379 \
--Hiveconf livy.session.conf.spark.jars=hdfs://com-hdfs/user/spark/tispark-core-2.1.4-spark_2.4-jar-with-dependencies.jar \
--Hiveconf livy.session.name=session_name_${RANDOM}_$2 \
--Hiveconf livy.session.queue=your_yarn_queue_name \
-e "$1"
}
```

3.

- shell 2 \$1 SQL \$2 \$RANDOM session
- beeline hdfs
- -u JDBC URL livy
- TiDB a) spark.sql.extensions=org.apache.spark.sql.TiExtensions
- b) spark.tispark.pd.addresses=10.10.10.11:2379,10.10.10.12:2379,10.10.13.10:2379
- c) spark.jars=hdfs://com-hdfs/user/spark/tispark-core-2.1.4-spark_2.4-jar-with-dependencies.jar
- livy.session.name \$2
- livy.session.queue YARN
- -e "\$1"

4.

(1)

```
0: jdbc:Hive2://bj0000,bj0001,bj0002:2222> show databases;
+-----+
|      databaseName      |
+-----+
| sales_db               |
| db_em                  |
| db_test                |
| db_shr                 |
+-----+
```

sales_db Hive db_em TiDB

(2) TiDB

```
0: jdbc:Hive2://bj0000,bj0001,bj0002:2222> select count(1) from db_em.app_war_room_fpyr_rt;
RSC client is executing SQL query: select count(1) from db_em.app_war_room_fpyr_rt, statementId = 91583fc3-0837-
4b26-b579-b438a53f151e, session = SessionHandle [b6933d96-8ec0-47b9-a767-ff2da5c5b2b2]
[Stage 0:>                                              (0 + 1) / 1]
+-----+
| count(1)   |
+-----+
| 224        |
+-----+
1 row selected (6.505 seconds)
```

(3) Hive

```
0: jdbc:Hive2://bj0000,bj0001,bj0002:2222> select count(1) from sales_db.mdms_tsqa_syyt;
RSC client is executing SQL query: select count(1) from sales_db.mdms_tsqa_syyt, statementId = 324d7d1d-03bc-4d9
b-849f-18967305a454, session = SessionHandle [b6933d96-8ec0-47b9-a767-ff2da5c5b2b2]
[Stage 2:>                                              (0 + 1) / 151]
...
[Stage 3:>                                              (0 + 0) / 1]
+-----+
| count(1)   |
+-----+
| 142380109  |
+-----+
1 row selected (25.169 seconds)
```

(4) Hive :

```
0: jdbc:Hive2://bj0000,bj0001,bj0002:2222> insert into dc_tmp.test_for_mix select count(1) as cnt from sales_d
b.mdms_tsqa_syyt union select count(1) as cnt from db_em.app_war_room_fpyr_rt;
RSC client is executing SQL query: insert into dc_tmp.test_for_mix select count(1) as cnt from sales_db.mdms_ts
qa_syyt union select count(1) as cnt from db_em.app_war_room_fpyr_rt, statementId = 7e87c512-d5c5-4ba1-bd62-d013
ff16a4e7, session = SessionHandle [b6933d96-8ec0-47b9-a767-ff2da5c5b2b2]
[Stage 5:>                                              (0 + 0) / 151]
.....
[Stage 10:>                                              (0 + 0) / 1]
+-----+
| Result   |
+-----+
+-----+
No rows selected (31.221 seconds)
```

(5)

```
0: jdbc:Hive2://bj0000,bj0001,bj0002:2222> select cnt from dc_tmp.test_for_mix;
RSC client is executing SQL query: select cnt from dc_tmp.test_for_mix, statementId = 18247b15-d9fb-4b97-87a1-6f
a9cc3afad8, session = SessionHandle [b6933d96-8ec0-47b9-a767-ff2da5c5b2b2]
[Stage 11:>                                              (0 + 1) / 1]
.....
[Stage 11:>                                              (0 + 1) / 1]
+-----+
|   cnt   |
+-----+
| 142380109 |
| 224       |
+-----+
2 rows selected (2.977 seconds)
```

HiveTiDB

11.4.3

1. TiDB TiDB 4.0 TiSpark TiDB (1) Spark JDBC TiDB MySQL (2) TiSpark
TiSpark TiDB 4.0 TiSpark TiDB TiDB TiDB

[Github](#) ↗

2.
TiDB Hive TiSpark TiDB TiSpark TiDB TiDB TiSpark tidb_ TiDB

```
spark.tispark.db_prefix "tidb_"
```

11.4.4 TiSpark

TiSpark Apache Spark Spark Apache Spark

1.1.1 TiUP

RPM Linux Yum Anaconda python TiDB Prometheus

TiDB 4.0 TiUP TiDB TiDBPDTiKV TiDB TiUP

<https://tiup.io/>

1.

TiUP

```
curl --proto '=https' --tlsv1.2 -sSf https://tiup-mirrors.pingcap.com/install.sh | sh
```

TiUP \$HOME/.tiup \$HOME/.tiup/bin Shell Profile TiUP TiUP

```
tiup --version
```

TiUP v0.0.3 TiUP

2.

TiUP TiUP tiup help

```

-$ tiup help
TiUP is a command-line component management tool that can help to download and install
TiDB platform components to the local system. You can run a specific version of a component via
"tiup <component>[:version]". If no version number is specified, the latest version installed
locally will be used. If the specified component does not have any version installed locally,
the latest stable version will be downloaded from the repository.

Usage:
  tiup [flags] <command> [args...]
  tiup [flags] <component> [args...]

Available Commands:
  install      Install a specific version of a component
  list         List the available TiDB components or versions
  uninstall    Uninstall components or versions of a component
  update       Update tiup components to the latest version
  status        List the status of instantiated components
  clean        Clean the data of instantiated components
  help         Help about any command or component

Available Components:

Flags:
  -B, --binary <component>[:version]  Print binary path of a specific version of a component <component>[:version]
                                         and the latest version installed will be selected if no version specified
  --binpath string                    Specify the binary path of component instance
  -h, --help                          help for tiup
  --skip-version-check               Skip the strict version check, by default a version must be a valid SemVer str
  ing
  -T, --tag string                   Specify a tag for component instance

Component instances with the same "tag" will share a data directory ($TIUP_HOME/data/$tag):
$ tiup --tag mycluster playground

Examples:
$ tiup playground                      # Quick start
$ tiup playground nightly                # Start a playground with the latest nightly version
$ tiup install <component>[:version]   # Install a component of specific version
$ tiup update --all                     # Update all installed components to the latest version
$ tiup update --nightly                 # Update all installed components to the nightly version
$ tiup update --self                   # Update the "tiup" to the latest version
$ tiup list --refresh                  # Fetch the latest supported components list
$ tiup status                         # Display all running/terminated instances
$ tiup clean <name>                   # Clean the data of running/terminated instance (Kill process if it's running)
$ tiup clean --all                     # Clean the data of all running/terminated instances

Use "tiup [command] --help" for more information about a command.

```

TiUP

```

tiup [flags] <command> [args...]
tiup [flags] <component> [args...]

```

TiUP

1. tiup: TiUP
2. flags:
3. command component:
4. args:

- list:
- install:
- update:
- uninstall:

- status:
- clean:
- help:

flags

- --binary :
- --binpath :
- --tag : tag ID tag

3

```
tiup help <command|component>    tiup <command|component> --help    tiup <command|component> -h
install    tiup help install    tiup install --help    tiup install -h
```

(1) tiup list

TiUP list

```
~$ tiup help list
List the available TiDB components if you don't specify any component name,
or list the available versions of a specific component. Display a list of
local caches by default. You must use --refresh to force TiUP to fetch
the latest list from the mirror server. Use the --installed flag to hide
components or versions which have not been installed.

# Refresh and list all available components
tiup list --refresh

# List all installed components
tiup list --installed

# List all installed versions of TiDB
tiup list tidb --installed

Usage:
  tiup list [component] [flags]

Flags:
  -h, --help      help for list
  --installed    List installed components only.
  --refresh     Refresh local components/version list cache.

Global Flags:
  --skip-version-check  Skip the strict version check, by default a version must be a valid SemVer string
```

- tiup list :
- tiup list <component> :

flag

- --installed :
- --refresh :

```
tiup list --installed
```

TiKV

```
tiup list tikv --refresh
```

(2) tiup install

tiup install

```
$ tiup help install
Install a specific version of a component. The component can be specified
by <component> or <component>:<version>. The latest stable version will
be installed if there is no version specified.

You can install multiple components at once, or install multiple versions
of the same component:

tiup install tidb:v3.0.5 tikv pd
tiup install tidb:v3.0.5 tidb:v3.0.8 tikv:v3.0.9

Usage:
  tiup install <component1>[:version] [component2...N] [flags]

Flags:
  -h, --help    help for install

Global Flags:
  --skip-version-check  Skip the strict version check, by default a version must be a valid SemVer string
```

- `tiup install <component> :`
- `tiup install <component>:[version] :`

TiUP TiDB

```
tiup install tidb
```

TiUP nightly TiDB

```
tiup install tidb:nightly
```

TiUP v3.0.6 TiKV

```
tiup install tikv:v3.0.6
```

(3) tiup update

TiUP

```
$ tiup help update
Update some components to the latest version. Use --nightly
to update to the latest nightly version. Use --all to update all components
installed locally. Use <component>:<version> to update to the specified
version. Components will be ignored if the latest version has already been
installed locally, but you can use --force explicitly to overwrite an
existing installation. Use --self which is used to update TiUP to the
latest version. All other flags will be ignored if the flag --self is given.

$ tiup update --all           # Update all components to the latest stable version
$ tiup update --nightly --all   # Update all components to the latest nightly version
$ tiup update playground:v0.0.3 --force # Overwrite an existing local installation
$ tiup update --self            # Update TiUP to the latest version

Usage:
tiup update [component1][:version] [component2..N] [flags]

Flags:
--all      Update all components
--force    Force update a component to the latest version
-h, --help   help for update
--nightly  Update the components to nightly version
--self     Update tiup to the latest version

Global Flags:
--skip-version-check  Skip the strict version check, by default a version must be a valid SemVer string
```

install flag

- --all :
- --nightly : nightly
- --self : TiUP
- --force :

```
tiup update --all
```

nightly

```
tiup update --all --nightly
```

TiUP

```
tiup update --self
```

(4) tiup <component>

TiUP

```
tiup [flags] <component>[:version] [args...]
```

TiUP tag tag tag

```
--tag tag
```

v3.0.8 TiDB

```
tiup tidb:v3.0.8
```

tag TiKV

```
tiup --tag=experiment tikv
```

(5) tiup status

```
tiup status
```

```
~$ tiup help status
List the status of instantiated components

Usage:
  tiup status [flags]

Flags:
  -h, --help    help for status

Global Flags:
  --skip-version-check  Skip the strict version check, by default a version must be a valid SemVer string
```

(6) tiup clean

```
tiup clean kill
```

```
~$ tiup help clean
Clean the data of instantiated components

Usage:
  tiup clean [flags]

Flags:
  --all      Clean all data of instantiated components
  -h, --help    help for clean

Global Flags:
  --skip-version-check  Skip the strict version check, by default a version must be a valid SemVer string
```

tag experiment

```
tiup clean experiment
```

```
tiup clean --all
```

(4) tiup uninstall

TiUP TiUP

```
-$ tiup help uninstall
If you specify a version number, uninstall the specified version of
the component. You must use --all explicitly if you want to remove all
components or versions which are installed. You can uninstall multiple
components or multiple versions of a component at once. The --self flag
which is used to uninstall tiup.

# Uninstall tiup
tiup uninstall --self

# Uninstall the specific version a component
tiup uninstall tidb:v3.0.10

# Uninstall all version of specific component
tiup uninstall tidb --all

# Uninstall all installed components
tiup uninstall --all

Usage:
  tiup uninstall <component>[:version] [flags]

Flags:
  --all      Remove all components or versions.
  -h, --help   help for uninstall
  --self     Uninstall tiup and clean all local data

Global Flags:
  --skip-version-check  Skip the strict version check, by default a version must be a valid SemVer string
```

v3.0.8 TiDB

```
tiup uninstall tidb:v3.0.8
```

TiKV

```
tiup uninstall tikv --all
```

```
tiup uninstall --all
```

1.1.2 TiUP

TiDB TiDB 3 PD 3 TiKV 2 TiDB TiDB TiUP TiUP playground client TiDB

1. playground

playground TiUP TiDB/PD/TiKV

list playground

```
tiup list playground
```

install playground

```
tiup install playground
```

```
tiup playground playground
```

```
~$ tiup help playground
Bootstrap a TiDB cluster in your local host, the latest release version will be chosen
if you don't specified a version.
```

Examples:

```
$ tiup playground nightly          # Start a TiDB nightly version local cluster
$ tiup playground v3.0.10 --db 3 --pd 3 --kv 3    # Start a local cluster with 10 nodes
$ tiup playground nightly --monitor      # Start a local cluster with monitor system
$ tiup playground --pd.config ~/config/pd.toml   # Start a local cluster with specified configuration file,
$ tiup playground --db.binpath /xx/tidb-server     # Start a local cluster with component binary path
```

Usage:

```
tiup playground [version] [flags]
```

Flags:

--db int	TiDB instance number (default 1)
--db.binpath string	TiDB instance binary path
--db.config string	TiDB instance configuration file
-h, --help	help for tiup
--host string	Playground cluster host (default "127.0.0.1")
--kv int	TiKV instance number (default 1)
--kv.binpath string	TiKV instance binary path
--kv.config string	TiKV instance configuration file
--monitor	Start prometheus component
--pd int	PD instance number (default 1)
--pd.binpath string	PD instance binary path
--pd.config string	PD instance configuration file

playground

-
-
- --host host IP
- --monitor Prometheus

```
tiup playground
```

- TiUP playground v0.0.6 tiup playground:v0.0.6
- playground v0.0.6 TiUP

- playground TiDB/PD/TiKV release v4.0.0 tiup playground:v0.0.6 v4.0.0
 - playground 1 TiDB1 TiKV 1 PD
 - playground TiUP TiDB/PD/TiKV tiup tidb:v4.0.0 TiDB
 - playground MySQL dashboard
- tag TiUP tag tag tag

```
tiup --tag=my-cluster playground
```

TiUP v3.0.9

```
tiup playground v3.0.9
```

TiUP nightly

```
tiup playground nightly
```

TiKV 3 Prometheus

```
tiup playground --kv=3 --monitor
```

IP x.x.x.x

```
tiup playground --host x.x.x.x
```

2. playground

TiDB 2 TiDB 3 TiKV 3 PD playground

```
tiup playground --db=2 --kv=3 --pd=3
```

playground --monitor

```
tiup playground --db=2 --kv=3 --pd=3 --monitor
```

playground MySQL

```
CLUSTER START SUCCESSFULLY, Enjoy it ^-^
To connect TiDB: mysql --host 127.0.0.1 --port 4000 -u root
```

playground TiUP client

```
tiup client
```

client DOS

tag

```
tiup client <tag>
```


1.1.4 TiOps

1.TiOps

1.1

TiDB TiDB Ansible TiDBTiDB Ansible TiDB TiOps

1.2

TiOps TiDB TiDB

2.

2.1

TiOps TiDB (SSH)

2.2

TiDB //

3.

- CentOS 7.3 Linux el7
- epel-release EPEL
- `$HOME` > 10GB
-

(1)

```
yum install -y epel-release
```

(2) TiOps

rpm tiup nightly

v0.2.0

```
wget https://download.pingcap.org/tiops-v0.2.0-1.el7.x86_64.rpm
yum localinstall -y tiops-v0.2.0-1.dev.el7.x86_64.rpm
```

(3) TiOps

```
tiops -h TiOps
```

```

tiops -h
usage: tiops [-h] [-v]

{bootstrap-local,bootstrap-ssh,bootstrap-host,deploy,start,stop,restart,upgrade,display,destroy,reload,edit-config,sc
ale-out,scale-in,exec,quickdeploy,version}
...

positional arguments:
{bootstrap-local,bootstrap-ssh,bootstrap-host,deploy,start,stop,restart,upgrade,display,destroy,reload,edit-config,
scale-out,scale-in,exec,quickdeploy,version}
    tidb cluster commands
  bootstrap-local      init localhost
  bootstrap-ssh        init ssh for cluster hosts
  bootstrap-host       init environment for cluster hosts
  deploy               create tidb cluster
  start                start tidb cluster
  stop                 stop tidb cluster
  restart              restart tidb cluster
  upgrade              upgrade tidb cluster
  display              display tidb cluster
  destroy              destroy tidb cluster
  reload               reload tidb cluster
  edit-config          edit tidb cluster configuration
  scale-out            scale out tidb cluster
  scale-in             scale in tidb cluster
  exec                run shell command on host in the tidb cluster
  quickdeploy          deploy a tidb cluster in demo mode
  version              show TiOps version and exit

optional arguments:
-h, --help            show this help message and exit
-v, --verbose         Print verbose output.

```

(4)

ssh key `~/.ssh/id_rsa` `~/.ssh/id_rsa.pub` ssh key ssh

```
tiops bootstrap-local
```

(5)

IP TiDB sudo

```
tiops bootstrap-ssh -H 10.9.1.1,10.9.1.2,10.9.1.3 -u root -d tidb
```

```

-H |--host hosts IP IP `,` IP IP /usr/share/tiops/hosts.yaml.example

-u | --user root
-p | --password
-d |--deploy-user TiDB
-s | --ssh-port SSH 22
-f | --forks 5

```

- ssh TiOps

- root root
-

(6)

NTP Swapirqbalance

```
tiops bootstrap-host -H 10.9.1.1,10.9.1.2,10.9.1.3 -d tidb

-H |--host hosts IP  IP  `,`  IP  IP  /usr/share/tiops/hosts.yaml.example

-d |--deploy-user  TiDB
-s | --ssh-port  SSH  22
--enable-check-ntp  NTP          disable
--ntp-server  NTP  NTP example: 10.0.0.10 or 10.0.0.10,10.0.0.11
--timezone  Asia/Shanghai
--enable-swap  swap          disable
--disable-irqbalance  irqbalance      enable
--enable-checks          disable
-f | --forks 5
```

- TiDB

(7)

TiOps

```
cat /usr/share/tiops/topology.yaml.example
```

- pd_servertidb_serverstikv_serversmonitoring_servergrafana_server
- IP data_dir deploy_dir port /usr/share/tiops/topology.yaml.example.full

4.TiOps

4.1

```
/usr/share/tiops/topology.yaml.example  /usr/share/tiops/topology.yaml.example.full
```

```
cp /usr/share/tiops/topology.yaml.example topology.yaml
```

```
vim topology.yaml
```

```
pd_servertidb_serverstikv_serversmonitoring_servergrafana_server IP
```

```

---
pd_servers:
  - ip: 10.9.1.1
  - ip: 10.9.1.2
  - ip: 10.9.1.3

tidb_servers:
  - ip: 10.9.1.1
  - ip: 10.9.1.2

tikv_servers:
  - ip: 10.9.1.1
  - ip: 10.9.1.2
  - ip: 10.9.1.3

monitoring_server:
  - ip: 10.9.1.1

grafana_server:
  - ip: 10.9.1.1

```

4.2

(1)

-T TiOps mai start

```
tiops deploy -c mai -T topology.yaml -d tidb
```

```

-c | --cluster_name name
-T | --topology
-t | --tidb-version TiDB 3.0.5
-d | --deploy-user TiDB
--enable-check-cpu CPU vcores           disable
--enable-check-mem Memory Size         disable
--enable-check-disk Disk Available Space      disable
--enable-check-iops IOPS latency          disable
--enable-check-all CPUMemoryDiskIOPS       disable
--enable-firewall                         disable
--enable-check-config                     disable
--local-pkg
-f | --forks 5

```

5.TiOps

5.1 //

:

```
tiops start -c mai
```

```
Finished start. mysql client TiDB
```

```
mysql -h 10.9.1.1 -P4000 -u root
```

```
Server version: 5.7.25-TiDB-v3.0.9 MySQL Community Server (Apache License 2.0) TiDB
```

:

```
-c | --cluster_name name
-T | --topology
-t | --tidb-version TiDB 3.0.5
-d | --deploy-user TiDB
--enable-check-cpu CPU vcores           disable
--enable-check-mem Memory Size         disable
--enable-check-disk Disk Available Space disable
--enable-check-iops IOPS latency       disable
--enable-check-all CPUMemoryDiskIOPS    disable
--enable-firewall                      disable
--enable-check-config                  disable
--local-pkg
-f | --forks 5
```

:

- ID
-

5.2 //

:

```
tiops restart -c mai
```

TiOps grafana, blackbox_exporter, node_exporter, prometheus, tidb, tikv, pd

:

```
-c | --cluster_name cluster_name
-r | --role role TiDB                 "pd", "tikv", "pump", "tidb", "drainer", "monitoring", "monitored", "grafana", "alertmanager"
-n | --node-id node_id ID ID display
-f | --forks 5
```

5.3 //

:

```
tiops stop -c mai
```

TiOps grafana, blackbox_exporter, node_exporter, prometheus, tidb, tikv, pd

:

```
-c | --cluster_name cluster_name
-r | --role role TiDB                 "pd", "tikv", "pump", "tidb", "drainer", "monitoring", "monitored", "grafana", "alertmanager"
-n | --node-id node_id ID ID display
-f | --forks 5
```

5.4

:

```
tiops destroy -c mai
```

```
-c | --cluster_name cluster_name
-f | --forks 5
```

•

5.5

```
tiops upgrade -c mai -d tidb
```

```
-c | --cluster_name cluster_name
-r | --role role TiDB          "pd", "tikv", "pump", "tidb", "drainer", "monitoring", "monitored", "grafana", "alertmanager"
-n | --node-id node_id ID    ID  display
--force
--local-pkg
--enable-check-config        disable
-f | --forks 5
```

5.6

```
tiops scale-out -c mai -T topology.yaml
```

```
-c | --cluster_name name
-T | --topology
--check-cpu CPU vcores           disable
--enable-check-mem Memory Size   disable
--enable-check-disk Disk Available Space disable
--enable-check-iops IOPS latency  disable
--enable-check-all CPUMemoryDiskIOPS disable
--local-pkg
-f | --forks 5
```

5.7

```
tiops scale-in -c mai
```

```
-c | --cluster_name name
-n | --node_id nodes ID ID display
-f | --forks 5
```

5.8

```
tiops reload -c mai
```

```
-c | --cluster_name cluster_name
-r | --role role TiDB          "pd", "tikv", "pump", "tidb", "drainer", "monitoring", "monitored", "grafana", "alertmanager"
-n | --node_id node_id ID ID display
-f | --forks 5
```

5.9

```
tiops edit-config -c mai
```

```
-c | --cluster_name cluster_name
```

- reload

5.10

//

```
tiops display -c mai
```

```
-c | --cluster_name cluster_name
-r | --role role TiDB          "pd", "tikv", "pump", "tidb", "drainer", "monitoring", "monitored", "grafana", "alertmanager"
-n | --node_id node_id ID ID display
--status
```

```
--status TiDB
```

```
TiDB cluster mai, version 3.0.9
Node list:
ID      Role      Host       Ports      Status     Data Dir           Deploy Dir
f5961fad Monitoring 10.9.1.1  9090/9091 -          /home/tidb/data/mai/monitoring-f5961fad /home/tidb/deploy/mai
i/monitoring-f5961fad
eb57cedb PD        10.9.1.1  2379/2380 Health    /home/tidb/data/mai/pd-eb57cedb   /home/tidb/deploy/mai
i/pd-eb57cedb
hb6289b2 PD        10.9.1.2  2379/2380 Health|L  /home/tidb/data/mai/pd-hb6289b2   /home/tidb/deploy/mai
/pd-hb6289b2
dd25d798 PD        10.9.1.3  2379/2380 Health    /home/tidb/data/mai/pd-dd25d798   /home/tidb/deploy/mai
/pd-dd25d798
dded5656 TiDB     10.9.1.1  4000/10080 Up       -          /home/tidb/deploy/mai
/tidb-dded5656
dedfd1cb TiDB     10.9.1.2  4000/10080 Up       -          /home/tidb/deploy/mai
/tidb-dedfd1cb
jbab8302 Monitored 10.9.1.1  9100/9115 -          -          /home/tidb/deploy/mai
/monitored-jbab8302
db46f98f Monitored 10.9.1.2  9100/9115 -          -          /home/tidb/deploy/mai
i/monitored-db46f98f
h1f9fce0 Monitored 10.9.1.3  9100/9115 -          -          /home/tidb/deploy/mai
/monitored-h1f9fce0
g0418cd7 TiKV     10.9.1.1  20160/20180 Up       /home/tidb/data/mai/tikv-g0418cd7 /home/tidb/deploy/mai
i/tikv-g0418cd7
c7e55608 TiKV     10.9.1.2  20160/20180 Up       /home/tidb/data/mai/tikv-c7e55608 /home/tidb/deploy/mai
/tikv-c7e55608
hf2c6bb4 TiKV     10.9.1.3  20160/20180 Up       /home/tidb/data/mai/tikv-hf2c6bb4 /home/tidb/deploy/mai
/tikv-hf2c6bb4
i701ff8d Grafana  10.9.1.1  3000      -          -          /home/tidb/deploy/mai
i/grafana-i701ff8d
```

```
TiDB Up/Down
PD Health/Unhealth, PD Down; PD Leader L
TiKV Store Up/Offline/Tombstone TiKV Down
Pump Node Pump Down
Drainer Node Drainer Down
```

Status -

5.11 TiOps

```
tiops version
```

TiOps

5.12 quickdeploy

```
tiops quickdeploy -c mai -d tidb -H 10.9.1.1,10.9.1.2,10.9.1.3 -T topology.yaml
```

```
-c | --cluster_name name
-T | --topology
-H | --host hosts IP IP `` IP IP /usr/share/tiops/hosts.yaml.example

-u | --user root
-d | --deploy-user TiDB
-t | --tidb-version TiDB 3.0.5
-p | --password
-s | --ssh-port SSH 22
--enable-check-cpu CPU vcores          disable
--enable-check-mem Memory Size        disable
--enable-check-disk Disk Available Space      disable
--enable-check-iops IOPS latency        disable
--enable-check-all CPUMemoryDiskIOPS      disable
--enable-firewall                      disable
--enable-check-config                  disable
--local-pkg
-f | --forks 5
```

6.

TiOps TiSpark TiFlash

1.1.5 TiOps

- CentOS 7.3 Linux

	CPU					
PD	>= 4 Core	>= 8 GB	SSD	>= 300 GB	>= 1	3
TiDB	>= 16 Core	>= 32 GB	SAS/SSD	>= 300 GB	>= 1	2
TiKV	>= 16 Core	>= 32 GB	SSD	<= 2 TB	>= 1	3
Prometheus	>= 8 Core	>= 16 GB	SAS/SSD	>= 300 GB	>= 1	1

- TiDB
 - >= 2 tidb-server
 - SAS/SSD >= 300 GB
- TiKV
 - >= 3 tikv-server
 - label label
 - SSD PCI-E SSD <= 2 TB, SSD <= 1.5 TB
- PD
 - >= 3
 - SSD >= 300 GB
- Prometheus
 - 1
- Other
 - TiDBPDPrometheus
 -
 -

TiDB TiDB

TiDB	4000	DBA
TiDB	10080	TiDB
TiKV	20160	TiKV
PD	2379	TiDB PD
PD	2380	PD
Pump	8250	Pump
Drainer	8249	Drainer
Prometheus	9090	Prometheus
Pushgateway	9091	TiDBTiKVPD
Node_exporter	9100	TiDB
Blackbox_exporter	9115	Blackbox_exporter TiDB
Grafana	3000	Web ()
Grafana	8686	grafana_collector Dashboard PDF
Kafka_exporter	9308	Kafka_exporter binlog kafka

topology

```

---
pd_servers:
  - ip: 10.9.1.1
  - ip: 10.9.1.2
  - ip: 10.9.1.3

tidb_servers:
  - ip: 10.9.1.2
  - ip: 10.9.1.3

tikv_servers:
  - ip: 10.9.1.4
    ## The value of label can be customized, for example: 'zone=z1,rack=r1,host=h1' or 'a=a1,b=b1,c=c1', etc
    ## Can only set in tikv_servers
    # label: host=h1
  - ip: 10.9.1.5
  - ip: 10.9.1.6

monitoring_server:
  - ip: 10.9.1.7

grafana_server:
  - ip: 10.9.1.7

```

label 'zone=z1,rack=r1,host=h1' 'a=a1,b=b1,c=c1' label tikv_servers

"1.1.4 " TiOps TiDB

1.2.1 TiDB Operator

1.2.1.1

2020

“ TiDB ”

“ 2020 TiDB TiDB ”

“ deiK8s ”

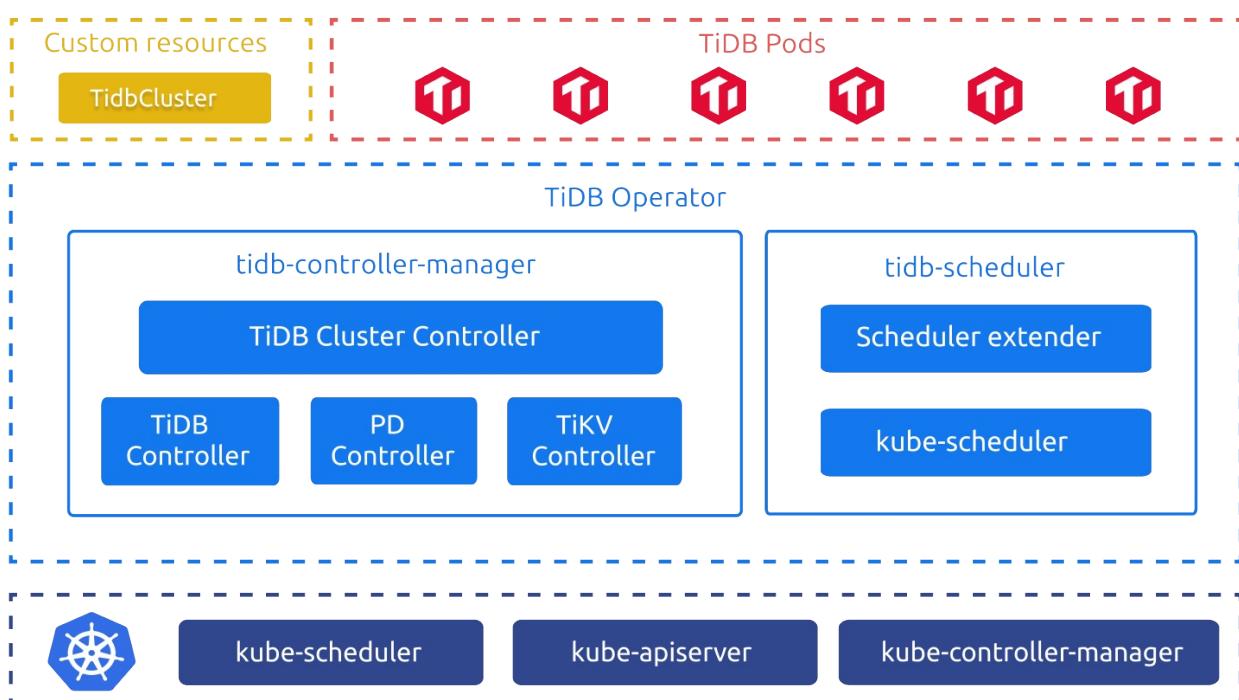
1.2.1.2

TiDB Operator Kubernetes TiDB TiDB TiDB OperatorTiDB Kubernetes

—— PingCAP

TiDB Operator “” TiDB “”“” Operator “What to do“ Operator “How to do” TiDB Operator “What to do“
auto-scaler

1.2.1.3 TiDB Operator



1.2.1.4 TiDB Operator

- TiDB Cluster CRD `CustomResourceDefinition` `TidbCluster` Kubernetes TiDB Cluster
`Deployment` `StatefulSet` Kubernetes TiDB Operator v1.1.0 CRD
`TidbCluster` `Backup` `Restore` `BackupSchedule` `TidbMonitor` `TidbInitializer` `TidbClusterAutoScaler`
- `tidb-controller-manager`
- `tidb-scheduler` Kubernetes Kubernetes TiDB Node TiDB TiKV

1.2.1.5

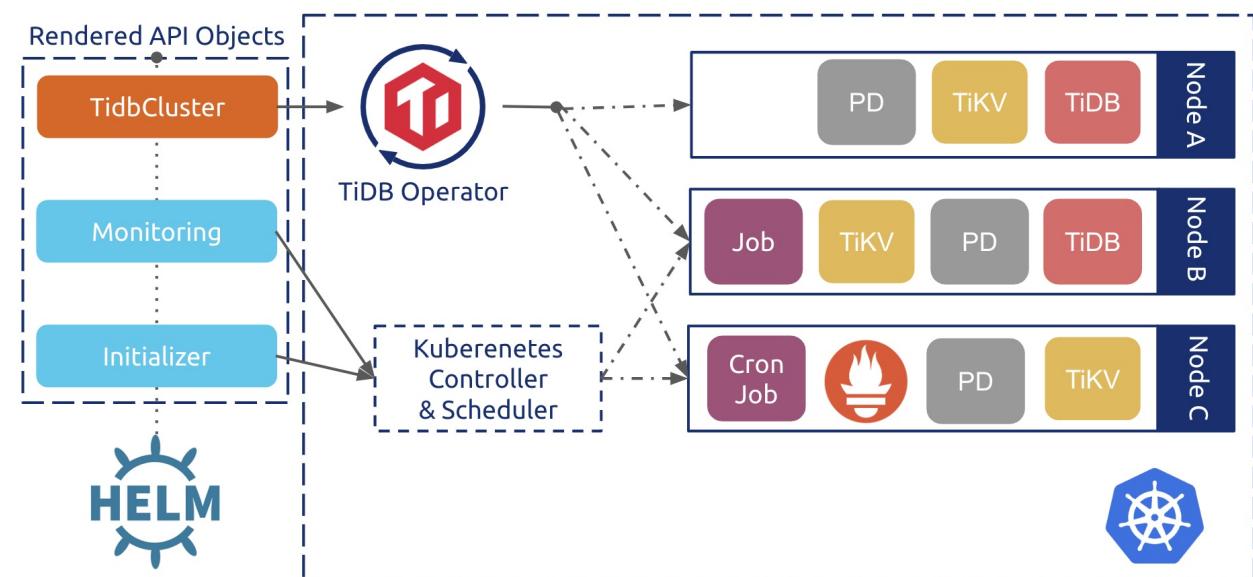
- TiDB Cluster CR CustomResource TiDB Cluster TidbCluster

1.2.1.6 Kubernetes

- kube-apiserver Kubernetes API Server
- kube-controller-manager TidbCluster CR StatefulSet Deployment CronJob K8s
- kube-scheduler TiDB Cluster Pod filtering kube-scheduler tidb-scheduler kube-scheduler scoring Pod

Deployment StatefulSet CRD K8s CR

1.2.1.7



TiDB Operator Helm Chart TiDB

1. Helm TidbCluster Kubernetes CronJob
 2. TiDB Operator Kubernetes API Server watch TidbCluster PDTiKVTiDB StatefulSet Service
 3. Kubernetes StatefulSet Deployment CronJob Pod
 4. PDTiKVTiDB Pod tidb-scheduler tidb-scheduler Pod TiDB
- TiDB Operator TidbCluster “”

1.2.2 TiDB-Operator

1.2.2.1

Linux MacOS kind Kubernetes TiDB Operator TiDB

1. kind K8s
2. K8s TiDB Operator
3. K8s TiDB

- 4GB+CPU 2 cores
- Docker 17.03+
- Go 1.10+
- net.ipv4.ip_forward 1

1.2.2.2 kind K8s

1.

```
# cd /root & git clone --depth=1 https://github.com/pingcap/tidb-operator && cd tidb-operator
```

1.

```
# cd /root/tidb-operator && hack/kind-cluster-build.sh
```

:

```
##### success create cluster:[kind] #####
To start using your cluster, run:
  kubectl config use-context kind-kind
```

1. K8s PATH

```
# export PATH=$PATH:/root/tidb-operator/output/bin/
```

1. K8s

```
# kubectl cluster-info
Kubernetes master is running at https://127.0.0.1:32771
KubeDNS is running at https://127.0.0.1:32771/api/v1/namespaces/kube-system/services/kube-dns:dns/proxy
```

K8s

```
# helm version
Client: &version.Version{SemVer:"v2.9.1", GitCommit:"20adb27c7c5868466912eebdf6664e7390ebe710", GitTreeState:"clean"}
Server: &version.Version{SemVer:"v2.9.1", GitCommit:"20adb27c7c5868466912eebdf6664e7390ebe710", GitTreeState:"clean"}
```

Helm

1.2.2.3 K8s TiDB Operator

1. helm TiDB Operator

TiDB CRD

```
# kubectl apply -f https://raw.githubusercontent.com/pingcap/tidb-operator/master/manifests/crd.yaml && kubectl get crd tidbclusters.pingcap.com
customresourcedefinition.apirextensions.k8s.io/tidbclusters.pingcap.com unchanged
customresourcedefinition.apirextensions.k8s.io/backups.pingcap.com unchanged
customresourcedefinition.apirextensions.k8s.io/restores.pingcap.com unchanged
customresourcedefinition.apirextensions.k8s.io/backupschedules.pingcap.com unchanged
customresourcedefinition.apirextensions.k8s.io/tidbmonitors.pingcap.com unchanged
customresourcedefinition.apirextensions.k8s.io/tidbinitializers.pingcap.com unchanged
customresourcedefinition.apirextensions.k8s.io/tidbclusterautoscalers.pingcap.com unchanged
NAME                CREATED AT
tidbclusters.pingcap.com  2020-03-06T13:38:32Z
```

TiDB Operator Helm chart

```
# mkdir -p /root/chart/
https://github.com/pingcap/tidb-operator/releases/tidb-operator-chart-v1.0.6.tgz /root/chart/
# cd /root/chart/ && tar xvf tidb-operator-chart-v1.0.6.tgz
```

/root/tidb-operator/charts/tidb-operator/values.yaml scheduler.kubeSchedulerImageName registry.cn-hangzhou.aliyuncs.com/google_containers/kube-scheduler

1. TiDB Operator

```
# helm install --namespace=tidb-admin --name=tidb-operator /root/tidb-operator/charts/tidb-operator -f /root/tidb-operator/charts/tidb-operator/values.yaml
NAME: tidb-operator
LAST DEPLOYED: Fri Mar 6 14:24:09 2020
NAMESPACE: tidb-admin
STATUS: DEPLOYED
...
```

1. Operator

```
# kubectl get pods -n tidb-admin
NAME                               READY   STATUS    RESTARTS   AGE
tidb-controller-manager-85d8d498bf-2n8km  1/1    Running   0          19s
tidb-scheduler-7c67d6c77b-qd54r        2/2    Running   0          19s
```

Operator

1.2.2.4 K8s TiDB

1. TiDB Cluster helm chart

```
# mkdir -p /root/chart/
https://github.com/pingcap/tidb-operator/releases/tidb-cluster-chart-v1.0.6.tgz /root/chart/
```

1. TiDB Cluster

```
# cd /root/chart/ && tar xvf tidb-cluster-chart-v1.0.6.tgz
# helm install --namespace dba-test --name=test /root/tidb-operator/charts/tidb-cluster -f /root/tidb-operator/charts/tidb-cluster/values.yaml
NAME: test
LAST DEPLOYED: Fri Mar 6 14:50:25 2020
NAMESPACE: dba-test
STATUS: DEPLOYED
```

TiDB Cluster

1. TiDB POD

```
# kubectl get pods -n dba-test
NAME                      READY   STATUS    RESTARTS   AGE
test-discovery-668b48577c-lqqbz  1/1    Running   0          7m37s
test-monitor-5b586d8cb-227qx   3/3    Running   0          7m37s
test-pd-0                   1/1    Running   0          7m37s
test-pd-1                   1/1    Running   0          7m37s
test-pd-2                   1/1    Running   1          7m37s
test-tidb-0                  2/2    Running   0          6m18s
test-tidb-1                  2/2    Running   0          6m18s
test-tikv-0                  1/1    Running   0          6m58s
test-tikv-1                  1/1    Running   0          6m58s
test-tikv-2                  1/1    Running   0          6m58s
```

TiDB Cluster Pod

1. TiDB

```
# nohup kubectl port-forward svc/test-tidb 4000:4000 --namespace=dba-test &
# yum install -y mysql
# mysql -h 127.0.0.1 -uroot -P4000
mysql -h 127.0.0.1 -P 4000 -uroot
Welcome to the MariaDB monitor. Commands end with ; or \g.
Your MySQL connection id is 1
Server version: 5.7.25-TiDB-v3.0.5 MySQL Community Server (Apache License 2.0)
Copyright (c) 2000, 2018, Oracle, MariaDB Corporation Ab and others.
Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.
MySQL [(none)]>
```

TiDB

1.2.3 TiDB-Operator

TiDB-Operator TiDB

TiDB

- [AWS EKS TiDB](#)
- [GCP GKE TiDB](#)
- [TiDB](#)
- [TiDB](#)

TiDB TiDB PV TiDB Operator TiDB TiDB

1.2.3.1 TiDB

TiDB

- [AWS EKS TiDB](#)
- [GCP GKE TiDB](#)
- [TiDB](#)
- [TiDB](#)

1.2.3.1.1 AWS EKS TiDB

Linux macOS AWS EKS (Elastic Kubernetes Service) TiDB

1.

- `awscli` >= 1.16.73 AWS

AWS `awscli aws configure :`

```
aws configure
```

AWS Access Key ID AWS Secret Access Key

```
AWS Access Key ID [None]: IOSF0DNN7EXAMPLE
AWS Secret Access Key [None]: wJaXUtnFEMI/K7MDENG/bPxRfiCYEXAMPLEKEY
Default region name [None]: us-west-2
Default output format [None]: json
```

Access key VPC EBS EC2 Role

- `terraform` >= 0.12

- `kubectl` >= 1.11
- `helm` >= 2.11.0 < 3.0.0
- `jq`
- `aws-iam-authenticator` AWS PATH

`aws-iam-authenticator`

Linux

```
curl -o aws-iam-authenticator https://amazon-eks.s3-us-west-2.amazonaws.com/1.12.7/2019-03-27/bin/linux/amd64/aws-iam-authenticator
```

macOS

```
curl -o aws-iam-authenticator https://amazon-eks.s3-us-west-2.amazonaws.com/1.12.7/2019-03-27/bin/darwin/amd64/aws-iam-authenticator
```

```
chmod +x ./aws-iam-authenticator && \
sudo mv ./aws-iam-authenticator /usr/local/bin/aws-iam-authenticator
```

2.

VPC t2.micro ec2 EKS

- 3 m5.xlarge PD

- 3 c5d.4xlarge TiKV
- 2 c5.4xlarge TiDB
- 1 c5.2xlarge

Github

```
git clone --depth=1 https://github.com/pingcap/tidb-operator && \
cd tidb-operator/deploy/aws
```

terraform

terraform init

terraform apply

terraform apply "yes"

10 **terraform apply**

```
Apply complete! Resources: 67 added 0 changed 0 destroyed.
```

Outputs:

```
bastion_ip = [
  "34.219.204.217",
]
default-cluster_monitor-dns = a82db513ba84511e9af170283460e413-1838961480.us-west-2.elb.amazonaws.com
default-cluster_tidb-dns = a82df6d13a84511e9af170283460e413-d3ce3b9335901d8c.elb.us-west-2.amazonaws.com
eks_endpoint = https://9A9A5ABB8303DDD35C0C2835A1801723.y14.us-west-2.eks.amazonaws.com
eks_version = 1.12
kubeconfig_filename = credentials/kubeconfig_my-cluster
region = us-west-21
```

terraform output

1.14 EKS NLB TiDB

AWS NLB

3.

terraform apply **ssh** MySQL client TiDB

<>)

```
ssh -i credentials/<eks_name>.pem centos@<bastion_ip>
```

```
mysql -h <default-cluster_tidb-dns> -P 4000 -u root
```

eks_name my-cluster DNS

kubectl helm kubeconfig credentials/kubeconfig_<eks_name> EKS

- **--kubeconfig**

```
kubectl --kubeconfig credentials/kubeconfig_<eks_name> get po -n <default_cluster_name>
```

```
helm --kubeconfig credentials/kubeconfig_<eks_name> ls
```

- KUBECONFIG

```
export KUBECONFIG=$PWD/credentials/kubeconfig_<eks_name>
```

```
kubectl get po -n <default_cluster_name>
```

```
helm ls
```

4. Grafana

```
<monitor-dns>:3000 Grafana
```

Grafana

- admin
- admin

5. TiDB

```
TiDB      terraform.tfvars  default_cluster_version      terraform apply
```

```
TiDB 4.0.1      default_cluster_version  v4.0.1
```

```
default_cluster_version= "v4.0.1"
```

```
kubectl --kubeconfig credentials/kubeconfig_<eks_name> get po -n <default_cluster_name> --watch
```

6. TiDB

```
TiDB      terraform.tfvars  default_cluster_tikv_count  default_cluster_tidb_count      terraform apply
```

```
default_cluster_tidb_count  2 4 TiDB
```

```
default_cluster_tidb_count = 4
```

- TiDB
- ```
kubectl --kubeconfig credentials/kubeconfig_<eks_name> get po -n <default_cluster_name> --watch
```

## 7.

```
terraform.tfvars
```

## AWS

```
TiDB Internal Elastic Load Balancer Amazon EC2 TiDB MySQL Sysbench SSH ELB TiDB VPC EC2
 create_bastion false

TiDB terraform.tfvars
```

## TiDB

```
Terraform EKS TiDB clusters.tf override_values TiDB values.yaml
 ./default-cluster.yaml values.yaml ""

EKS values.yaml NodeSelector Tolerations NodeSelector Tolerations Terraform TiDB
cluster.tf tidb-cluster module
```

```
values.yaml tidb-cluster module
```

```
pd:
 storageClassName: ebs-gp2
tikv:
 storageClassName: local-storage
tidb:
 service:
 type: LoadBalancer
 annotations:
 service.beta.kubernetes.io/aws-load-balancer-internal: '0.0.0.0/0'
 service.beta.kubernetes.io/aws-load-balancer-type: nlb
 service.beta.kubernetes.io/aws-load-balancer-cross-zone-load-balancing-enabled: >'true'
 separateSlowLog: true
monitor:
 storage: 100Gi
 storageClassName: ebs-gp2
 persistent: true
grafana:
 config:
 GF_AUTH_ANONYMOUS_ENABLED: "true"
 service:
 type: LoadBalancer
```

## TiDB Operator

```
terraform.tfvars operator_values values.yaml TiDB Operator
operator_values = "./operator_values.yaml"
}
```

## 8. TiDB

```
tidb-cluster TiDB clusters.tf tidb-cluster TiDB
```

```

module example-cluster {
 source = "../modules/aws/tidb-cluster"

 # The target EKS, required
 eks = local.eks
 # The subnets of node pools of this TiDB cluster, required
 subnets = local.subnets
 # TiDB cluster name, required
 cluster_name = "example-cluster"

 # Helm values file
 override_values = file("example-cluster.yaml")
 # TiDB cluster version
 cluster_version = "v3.0.0"
 # SSH key of cluster nodes
 ssh_key_name = module.key-pair.key_name
 # PD replica number
 pd_count = 3
 # TiKV instance type
 pd_instance_type = "t2.xlarge"
 # TiKV replica number
 tikv_count = 3
 # TiKV instance type
 tikv_instance_type = "t2.xlarge"
 # The storage class used by TiKV, if the TiKV instance type do not have local SSD, you should change it to storage
 class
 # TiDB replica number
 tidb_count = 2
 # TiDB instance type
 tidb_instance_type = "t2.xlarge"
 # Monitor instance type
 monitor_instance_type = "t2.xlarge"
 # The version of tidb-cluster helm chart
 tidb_cluster_chart_version = "v1.0.0"
 # Decides whether or not to create the tidb-cluster helm release.
 # If this variable is set to false, you have to
 # install the helm release manually
 create_tidb_cluster_release = true
}

```

cluster\_name

kubectl TiDB Terraform outputs.tf

```

output "example-cluster_tidb_hostname" {
 value = module.example-cluster.tidb_hostname
}

output "example-cluster_monitor_hostname" {
 value = module.example-cluster.monitor_hostname
}

```

terraform init terraform apply

tidb-cluster TiDB EC2

## 9.

Terraform Kubernetes TiDB Operator

- clusters.tf TiDB create\_tidb\_cluster\_release

```
module "default-cluster" {
 ...
 create_tidb_cluster_release = false
}
```

`create_tidb_cluster_release` `false` Terraform TiDB TiDB Helm

`create_tidb_cluster_release` `false` TiDB TiDB

## 10.

`terraform destroy`

- EKS EKS TiDB
- `terraform destroy` AWS EBS

## 11. Kubernetes

Kubernetes EKS TiDB

Terraform Terraform

- `tidb-operator` EKS EKS [TiDB Operator](#)
- `tidb-cluster` TiDB TiDB
- EKS TiDB `vpc` `key-pair` bastion

Kubernetes Kubernetes Terraform Terraform

```
mkdir -p deploy/aws-staging
vim deploy/aws-staging/main.tf
```

`deploy/aws-staging/main.tf`

```
provider "aws" {
 region = "us-west-1"
}

ssh key Kubernetes
module "key-pair" {
 source = "../modules/aws/key-pair"

 name = "another-eks-cluster"
 path = "${path.cwd}/credentials/"
}

VPC
module "vpc" {
 source = "../modules/aws/vpc"

 vpc_name = "another-eks-cluster"
}

VPC EKS tidb-operator
module "tidb-operator" {
 source = "../modules/aws/tidb-operator"
```

```

eks_name = "another-eks-cluster"
config_output_path = "credentials/"
subnets = module.vpc.private_subnets
vpc_id = module.vpc.vpc_id
ssh_key_name = module.key-pair.key_name
}

helm EKS
resource "local_file" "kubeconfig" {
 depends_on = [module.tidb-operator.eks]
 sensitive_content = module.tidb-operator.eks.kubeconfig
 filename = module.tidb-operator.eks.kubeconfig_filename
}
provider "helm" {
 alias = "eks"
 insecure = true
 install_tiller = false
 kubernetes {
 config_path = local_file.kubeconfig.filename
 }
}

EKS TiDB
module "tidb-cluster-a" {
 source = "../modules/aws/tidb-cluster"
 providers = {
 helm = "helm.eks"
 }

 cluster_name = "tidb-cluster-a"
 eks = module.tidb-operator.eks
 ssh_key_name = module.key-pair.key_name
 subnets = module.vpc.private_subnets
}

EKS TiDB
module "tidb-cluster-b" {
 source = "../modules/aws/tidb-cluster"
 providers = {
 helm = "helm.eks"
 }

 cluster_name = "tidb-cluster-b"
 eks = module.tidb-operator.eks
 ssh_key_name = module.key-pair.key_name
 subnets = module.vpc.private_subnets
}

#
module "bastion" {
 source = "../modules/aws/bastion"

 bastion_name = "another-eks-cluster-bastion"
 key_name = module.key-pair.key_name
 public_subnets = module.vpc.public_subnets
 vpc_id = module.vpc.vpc_id
 target_security_group_id = module.tidb-operator.eks.worker_security_group_id
 enable_ssh_to_workers = true
}

tidb-cluster-a TiDB
output "cluster-a_tidb-dns" {
 description = "tidb service endpoints"
 value = module.tidb-cluster-a.tidb_hostname
}

tidb-cluster-b
output "cluster-b_monitor-dns" {
 description = "tidb service endpoint"
 value = module.tidb-cluster-b.monitor_hostname
}

```

```
IP
output "bastion_ip" {
 description = "Bastion IP address"
 value = module.bastion.bastion_ip
}
```

bastion Terraform Terraform

Terraform variables.tf

Terraform Terraform Terraform

- Terraform [hashicorp/terraform#2430](#) Terraform helm provider
- Terraform source
- tidb-operator modules

Terraform deploy/aws Kubernetes terraform apply Terraform

### 1.2.3.1.2 GCP GKE TiDB

Linux macOS GCP GKE TiDB

## 1.

- [Git](#)
- [Google Cloud SDK](#)
- [Terraform](#) >= 0.12
- [kubectl](#) >= 1.14
- [Helm](#) >= 2.11.0 < 3.0.0
- [jq](#)

## 2.

Google Cloud SDK API Terraform

```
git clone --depth=1 https://github.com/pingcap/tidb-operator && \
cd tidb-operator/deploy/gcp
```

## Google Cloud SDK

Google Cloud SDK `gcloud init`

## API

GCP API

```
gcloud services enable clouddresourcemanagement.googleapis.com \
cloudbilling.googleapis.com iam.googleapis.com \
compute.googleapis.com container.googleapis.com
```

## Terraform

Terraform 3 Terraform `.tfvars`

- `GCP_CREDENTIALS_PATH` GCP
  - Terraform `./create-service-account.sh`
  - `deploy/gcp` `JSON` `JSON`
- `GCP_REGION` `us-west1`
- `GCP_PROJECT` GCP

3 Terraform

(1) `GCP_REGION` GCP Region

```
```bash
echo GCP_REGION=\"us-west1\" >> terraform.tfvars
```

```

## (2) GCP\_PROJECT GCP GCP

```
```bash
echo "GCP_PROJECT=$(gcloud config get-value project)" >> terraform.tfvars
```

```

## (3) Terraform

```
```bash
terraform init
```

```

## (4) Terraform

```
```bash
./create-service-account.sh
```

```

Terraform terraform.tfvars \*.auto.tfvars      Terraform GCP\_REGION      GCP\_PROJECT      terraform.tfvars  
 GCP\_CREDENTIALS\_PATH      credentials.auto.tfvars

### 3. TiDB

## TiDB

## (1)

```
- TiDB

```bash
cat small.tfvars >> terraform.tfvars
```

- benchmark

```bash
cat prod.tfvars >> terraform.tfvars
```

`prod.tfvars` VPC f1-micro GKE

* 3 n1-standard-4 PD
* 3 n1-highmem-8 TIKV
* 3 n1-standard-16 TiDB
* 3 n1-standard-2

91 CPU GCP [](https://cloud.google.com/compute/quotas) CPU

> ****
>
> Region Region 3 `us-central1` 4 [Regions and Zones](https://cloud.google.com/compute/docs/regions-zones/) []
#)
```

## (2) TiDB

```

```bash
terraform apply
```

> ****
>
> 3 `terraform apply` 3 [Terraform](#-terraform)

10 `terraform apply` :

```
Apply complete! Resources: 23 added, 0 changed, 0 destroyed.

Outputs:

how_to_connect_to_default_cluster_tidb_from_bastion = mysql -h 172.31.252.20 -P 4000 -u root
how_to_ssh_to_bastion = gcloud compute ssh tidb-cluster-bastion --zone us-west1-b
how_to_set_reclaim_policy_of_pv_for_default_tidb_cluster_to_delete = kubectl --kubeconfig ./credentials/kubeconfig
tidb-cluster get pvc -n tidb-cluster -o jsonpath='{.items[*].spec.volumeName}'|fmt -1 | xargs -I {} kubectl --kubeco
nfig ./credentials/kubeconfig_tidb-cluster patch pv {} -p '{"spec":{"persistentVolumeReclaimPolicy":"Delete"}}'
kubeconfig_file = ./credentials/kubeconfig_tidb-cluster
monitor_lb_ip = 35.227.134.146
monitor_port = 3000
region = us-west1
tidb_version = v3.0.1
```

```

## 4. TiDB

terraform apply TiDB TiDB <>

(1) ssh

```

```bash
gcloud compute ssh <gke-cluster-name>-bastion --zone <zone>
```

```

(2) MySQL TiDB

```

```bash
mysql -h <tidb_ilb_ip> -P 4000 -u root
```

> ****
>
> MySQL TiDB MySQL

```

## 5. GKE

kubectl helm kubeconfig credentials/kubeconfig\_<gke\_cluster\_name> GKE

gke\_cluster\_name tidb-cluster variables.tf gke\_name

- kubeconfig

```
kubectl --kubeconfig credentials/kubeconfig_<gke_cluster_name> get po -n <tidb_cluster_name>
```

```
helm --kubeconfig credentials/kubeconfig_<gke_cluster_name> ls
```

- KUBECONFIG

```
export KUBECONFIG=$PWD/credentials/kubeconfig_<gke_cluster_name>
```

```
kubectl get po -n <tidb_cluster_name>
```

```
helm ls
```

## 6. TiDB

TiDB

```
(1) terraform.tfvars (2) terraform apply
```

```
TiDB 4.0.0-rc.2 tidb_version v4.0.0-rc.2
```

```
tidb_version = "v4.0.0-rc.2"
```

```
kubectl --kubeconfig credentials/kubeconfig_<gke_cluster_name> get po -n <tidb_cluster_name> --watch
```

```
tidb_version() TiDB
```

```
select tidb_version();
```

```
***** 1. row *****
tidb_version(): Release Version: v3.0.0-rc.2
Git Commit Hash: 06f3f63d5a87e7f0436c0618cf524fea7172eb93
Git Branch: HEAD
UTC Build Time: 2019-05-28 12:48:52
GoVersion: go version go1.12 linux/amd64
Race Enabled: false
TiKV Min Version: 2.1.0-alpha.1-ff3dd160846b7d1aed9079c389fc188f7f5ea13e
Check Table Before Drop: false
1 row in set (0.001 sec)
```

## 7. TiDB

```
tidb-cluster GKE TiDB TiDB
```

```
(1) tidbclusters.tf tidb-cluster
```

```
```hcl
module "example-tidb-cluster" {
providers = {
    helm = "helm.gke"
}
source                  = "../modules/gcp/tidb-cluster"
cluster_id              = module.tidb-operator.cluster_id
tidb_operator_id        = module.tidb-operator.tidb_operator_id
gcp_project              = var.GCP_PROJECT
gke_cluster_location    = local.location
gke_cluster_name         = <gke-cluster-name>
cluster_name             = <example-tidb-cluster>
cluster_version          = "v3.0.1"
kubeconfig_path          = local.kubeconfig
tidb_cluster_chart_version = "v1.0.0"
pd_instance_type         = "n1-standard-1"
tikv_instance_type       = "n1-standard-4"
tidb_instance_type       = "n1-standard-2"
monitor_instance_type    = "n1-standard-1"
pd_node_count            = 1
tikv_node_count          = 2
tidb_node_count          = 1
monitor_node_count       = 1
}
```
> ****
>
> - `cluster_name`

> - Region

`kubectl` TiDB Terraform `outputs.tf` `output`


```hcl
output "how_to_connect_to_example_tidb_cluster_from_bastion" {
value = module.example-tidb-cluster.how_to_connect_to_tidb_from_bastion
}
```
```
TiDB

```

(2)

```
```bash
terraform init
```
  

```bash
terraform apply
```

8.

TiDB

(1) `terraform.tfvars`    `tikv_count` `tidb_count` (2)    `terraform apply`

tikv count

```
kubectl --kubeconfig credentials/kubeconfig_<gke_cluster_name> get po -n <tidb_cluster_name> --watch
```

```
tidb_count 1 2 TiDB
```

```
tidb_count = 2
```

## 9.

```
terraform.tfvars
```

### GCP

GCP n1-standard-1 SSD

### TiDB

Terraform GKE TiDB      tidbclusters.tf      TiDB      override\_values      override\_values\_file      override\_values

```
override_values = <<EOF
discovery:
 image: pingcap/tidb-operator:v1.0.1
 imagePullPolicy: IfNotPresent
resources:
 limits:
 cpu: 250m
 memory: 150Mi
 requests:
 cpu: 30m
 memory: 30Mi
EOF
```

```
override_values_file = "./test-cluster.yaml"
```

```
deploy/modules/gcp/tidb-cluster values/default.yaml
```

GKE      values.yaml      NodeSelector      Tolerations      NodeSelector      Tolerations      Terraform      TiDB

```
tidbclusters.tf tidb-cluster module
```

```
values.yaml tidb-cluster module
```

```

pd:
 storageClassName: pd-ssd
tikv:
 storageClassName: local-storage
tidb:
 service:
 type: LoadBalancer
 annotations:
 cloud.google.com/load-balancer-type: "Internal"
 separateSlowLog: true
monitor:
 storageClassName: pd-ssd
 persistent: true
grafana:
 config:
 GF_AUTH_ANONYMOUS_ENABLED: "true"
 service:
 type: LoadBalancer

```

## TiDB Operator

TiDB Operator    operator\_helm\_values    operator\_helm\_values\_file    operator\_helm\_values    tidb-operator

```

operator_helm_values = <<EOF
controllerManager:
 resources:
 limits:
 cpu: 250m
 memory: 150Mi
 requests:
 cpu: 30m
 memory: 30Mi
EOF

```

```
operator_helm_values_file = "./test-operator.yaml"
```

GKE    Fluentd StackdriverFluentd CPU RAM

Fluent Bit Fluentd Fluent Bit

GKE    Fluent Bit

(regional) (zonal) GKE Grafana

gcloud

GKE    gcloud compute instances delete GKE

(1)

```
```bash
gcloud compute instance-groups managed list | grep monitor
```
```
gke-tidb-monitor-pool-08578e18-grp us-west1-b zone gke-tidb-monitor-pool-08578e18 0 0 gke-tidb-mo
nitor-pool-08578e18 no
gke-tidb-monitor-pool-7e31100f-grp us-west1-c zone gke-tidb-monitor-pool-7e31100f 1 1 gke-tidb-mo
nitor-pool-7e31100f no
gke-tidb-monitor-pool-78a961e5-grp us-west1-a zone gke-tidb-monitor-pool-78a961e5 1 1 gke-tidb-mo
nitor-pool-78a961e5 no
```
```
```

(2)

```
```bash
gcloud compute instance-groups managed list-instances <the-name-of-the-managed-instance-group> --zone <zone>
```

```bash
gcloud compute instance-groups managed list-instances gke-tidb-monitor-pool-08578e18-grp --zone us-west1-b
```

```
NAME ZONE STATUS ACTION INSTANCE_TEMPLATE VERSION
_NAME LAST_ERROR
gke-tidb-monitor-pool-08578e18-c7vd us-west1-b RUNNING NONE gke-tidb-monitor-pool-08578e18

```

(3)

```
```bash
gcloud compute instance-groups managed delete-instances gke-tidb-monitor-pool-08578e18-grp --instances=gke-tidb-monit
or-pool-08578e18-c7vd --zone us-west1-b
```

```

10. TiDB

TiDB

terraform destroy

```
 terraform destroy Error reading Container Cluster "tidb": Cluster "tidb" has status "RECONCILING" with
message"" GCP Kubernetes master GCP terraform destroy
```

- Google Cloud Console gcloud

- terraform destroy Kubernetes PV (Persistent Volume) Delete terraform destroy kubectl
 

```
kubectl --kubeconfig /path/to/kubeconfig/file get pvc -n namespace-of-tidb-cluster -o jsonpath='{.items[*].spec.volumeName}'|fmt -1 | xargs -I {} kubectl --kubeconfig /path/to/kubeconfig/file patch pv {} -p '{"spec":{"persistentVolumeReclaimPolicy":"Delete"}}'
```

TiDB PVC (Persistent Volume Claim) PV Delete terraform destroy PVC  
 change-pv-reclaimpolicy.sh deploy/gcp
 

```
./change-pv-reclaimpolicy.sh /path/to/kubeconfig/file <tidb-cluster-namespace>
```

## 11. Kubernetes

Kubernetes Kubernetes TiDB

TiDB Terraform

- tidb-operator TiDB [Kubernetes Control Plane](#) TiDB Operator
- tidb-cluster Kubernetes TiDB
- vpc bastion project-credentials GKE TiDB

Kubernetes

(1) Kubernetes (2) Terraform

Terraform Kubernetes

```
mkdir -p deploy/gcp-staging && \
vim deploy/gcp-staging/main.tf
```

deploy/gcp-staging/main.tf

```
provider "google" {
 credentials = file(var.GCP_CREDENTIALS_PATH)
 region = var.GCP_REGION
 project = var.GCP_PROJECT
}

// required for taints on node pools
provider "google-beta" {
 credentials = file(var.GCP_CREDENTIALS_PATH)
 region = var.GCP_REGION
 project = var.GCP_PROJECT
}

locals {
 gke_name = "another-gke-name"
 credential_path = "${path.cwd}/credentials"
 kubeconfig = "${local.credential_path}/kubeconfig_${var.gke_name}"
}

module "project-credentials" {
 source = "../modules/gcp/project-credentials"

 path = local.credential_path
}

module "vpc" {
 source = "../modules/gcp/vpc"
 create_vpc = true
 gcp_project = var.GCP_PROJECT
 gcp_region = var.GCP_REGION
```

```

vpc_name = "${locals.gke_name}-vpc-network"
private_subnet_name = "${locals.gke_name}-private-subnet"
public_subnet_name = "${locals.gke_name}-public-subnet"
}

module "tidb-operator" {
 source = "../modules/gcp/tidb-operator"
 gke_name = locals.gke_name
 vpc_name = module.vpc.vpc_name
 subnetwork_name = module.vpc.private_subnetwork_name
 gcp_project = var.GCP_PROJECT
 gcp_region = var.GCP_REGION
 kubeconfig_path = local.kubeconfig
 tidb_operator_version = "v1.0.0"
}

module "bastion" {
 source = "../modules/gcp/bastion"
 vpc_name = module.vpc.vpc_name
 public_subnet_name = module.vpc.public_subnetwork_name
 gcp_project = var.GCP_PROJECT
 bastion_name = "${locals.gke_name}-tidb-bastion"
}

HACK: Helm GKE
data "local_file" "kubeconfig" {
 depends_on = [module.tidb-operator.cluster_id]
 filename = module.tidb-operator.kubeconfig_path
}
resource "local_file" "kubeconfig" {
 depends_on = [module.tidb-operator.cluster_id]
 content = data.local_file.kubeconfig.content
 filename = module.tidb-operator.kubeconfig_path
}

provider "helm" {
 alias = "gke"
 insecure = true
 install_tiller = false
 kubernetes {
 config_path = local_file.kubeconfig.filename
 }
}
module "tidb-cluster-a" {
 providers = {
 helm = "helm.gke"
 }
 source = "../modules/gcp/tidb-cluster"
 gcp_project = var.GCP_PROJECT
 gke_cluster_location = var.GCP_REGION
 gke_cluster_name = locals.gke_name
 cluster_name = "tidb-cluster-a"
 cluster_version = "v3.0.1"
 kubeconfig_path = module.tidb-operator.kubeconfig_path
 tidb_cluster_chart_version = "v1.0.0"
 pd_instance_type = "n1-standard-1"
 tikv_instance_type = "n1-standard-4"
 tidb_instance_type = "n1-standard-2"
 monitor_instance_type = "n1-standard-1"
}

module "tidb-cluster-b" {
 providers = {
 helm = "helm.gke"
 }
 source = "../modules/gcp/tidb-cluster"
 gcp_project = var.GCP_PROJECT
 gke_cluster_location = var.GCP_REGION
 gke_cluster_name = locals.gke_name
 cluster_name = "tidb-cluster-b"
 cluster_version = "v3.0.1"
 kubeconfig_path = module.tidb-operator.kubeconfig_path
}

```

```

tidb_cluster_chart_version = "v1.0.0"
pd_instance_type = "n1-standard-1"
tikv_instance_type = "n1-standard-4"
tidb_instance_type = "n1-standard-2"
monitor_instance_type = "n1-standard-1"
}

output "how_to_ssh_to_bastion" {
 value= module.bastion.how_to_ssh_to_bastion
}

output "connect_to_tidb_cluster_a_from_bastion" {
 value = module.tidb-cluster-a.how_to_connect_to_default_cluster_tidb_from_bastion
}

output "connect_to_tidb_cluster_b_from_bastion" {
 value = module.tidb-cluster-b.how_to_connect_to_default_cluster_tidb_from_bastion
}

```

- `*.tf` `module`
- `variables.tf` `terraform.tfvars`

- Terraform `source`
- `tidb-operator` `modules`
- Terraform [hashicorp/terraform#2430](#) **HACK: Helm GKE** Helm provider `tf`

Terraform `deploy/gcp` Kubernetes `terraform apply`

### 1.2.3.1.3 TiDB

Linux macOS TiDB

#### 1.

- `aliyun-cli` >= 3.0.15 `aliyun-cli`

Access Key

- `kubectl` >= 1.12
- `helm` >= 2.11.0
- `jq` >= 1.6
- `terraform` 0.12.\*

- AliyunECSFullAccess
- AliyunESSFullAccess
- AliyunVPCFullAccess
- AliyunSLBFullAccess
- AliyunCSFullAccess
- AliyunEIPFullAccess
- AliyunECIFullAccess
- AliyunVPNGatewayFullAccess
- AliyunNATGatewayFullAccess

#### 2.

- VPC
- ECS
- ACK Kubernetes worker
  - 2 ECS 2 2 GB Kubernetes CoreDNS
  - 3 `ecs.g5.large` PD
  - 3 `ecs.i2.2xlarge` TiKV
  - 2 `ecs.c5.4xlarge` TiDB
  - 1 `ecs.c5.xlarge`
  - 100 GB

(Auto-scaling Group)

#### 3.

- (1) Region      `terraform`

```
```shell
export TF_VAR_ALICLOUD_REGION=<YOUR_REGION> && \
export TF_VAR_ALICLOUD_ACCESS_KEY=<YOUR_ACCESS_KEY> && \
export TF_VAR_ALICLOUD_SECRET_KEY=<YOUR_SECRET_KEY>
```

`variables.tf` ` -var`
```

## (2) Terraform

```
```shell
git clone --depth=1 https://github.com/pingcap/tidb-operator && \
cd tidb-operator/deploy/aliyun
```

```shell
terraform init
```

`apply` `yes`

```shell
terraform apply
```

`terraform apply` `terraform apply`

5 10 `terraform output`

```
Apply complete! Resources: 3 added, 0 changed, 1 destroyed.

Outputs:

bastion_ip = 47.96.174.214
cluster_id = c2d9b20854a194f158ef2bc8ea946f20e
kubeconfig_file = /tidb-operator/deploy/aliyun/credentials/kubeconfig
monitor_endpoint = 121.199.195.236:3000
region = cn-hangzhou
ssh_key_file = /tidb-operator/deploy/aliyun/credentials/my-cluster-keyZ.pem
tidb_endpoint = 172.21.5.171:4000
tidb_version = v3.0.0
vpc_id = vpc-bp1v8i5rwsctyh8dwyp5
```

```

## (3) kubectl helm

```
```shell
export KUBECONFIG=$PWD/credentials/kubeconfig
```

```shell
kubectl version
```

```shell
helm ls
```

```

**4.**

TiDB

```
ssh -i credentials/<cluster_name>-key.pem root@<bastion_ip>
```

```
mysql -h <tidb_slb_ip> -P 4000 -u root
```

## 5.

```
<monitor_endpoint> Grafana
```

- admin
- admin

```
VPN VPC deploy/modules/aliyun/tidb-cluster/values/default.yaml monitor.grafana.service.annotations
service.beta.kubernetes.io/alicloud-loadbalancer-address-type intranet
```

## 6. TiDB

```
terraform.tfvars tidb_version terraform apply
```

```
kubectl get pods --namespace <tidb_cluster_name> -o wide --watch
```

## 7. TiDB

```
terraform.tfvars tikv_count tidb_count terraform apply TiDB
```

## 8.

```
terraform destroy
```

```
Kubernetes destroy Kubernetes
```

```
terraform state list
```

```
terraform state rm module.ack.alicloud_cs_managed_kubernetes.k8s
```

## 9.

### TiDB Operator

```
variables.tf TiDB Operator variable operator_helm_values TiDB Operator values.yaml
```

- `terraform.tfvars` `operator_helm_values`

```
operator_helm_values = "./my-operator-values.yaml"
```

- main.tf operator\_helm\_values

```
operator_helm_values = file("./my-operator-values.yaml")
```

Terraform VPC VPC variable.tf vpc\_id VPC vswitch Kubernetes

## TiDB

TiDB ./my-cluster.yaml values.yaml TiDB Kubernetes TiDB

## 10. TiDB

Kubernetes TiDB ./main.tf tidb-cluster

```
module "tidb-cluster-dev" {
 source = "../modules/aliyun/tidb-cluster"
 providers = {
 helm = helm.default
 }

 cluster_name = "dev-cluster"
 ack = module.tidb-operator

 pd_count = 1
 tikv_count = 1
 tidb_count = 1
 override_values = file("dev-cluster.yaml")
}

module "tidb-cluster-staging" {
 source = "../modules/aliyun/tidb-cluster"
 providers = {
 helm = helm.default
 }

 cluster_name = "staging-cluster"
 ack = module.tidb-operator

 pd_count = 3
 tikv_count = 3
 tidb_count = 2
 override_values = file("staging-cluster.yaml")
}
```

TiDB cluster\_name tidb-cluster

|                            |                         |                   |
|----------------------------|-------------------------|-------------------|
|                            |                         |                   |
| ack                        | Kubernetes              | nil               |
| cluster_name               | TiDB                    | nil               |
| tidb_version               | TiDB                    | v3.0.1            |
| tidb_cluster_chart_version | tidb-cluster helm chart | v1.0.1            |
| pd_count                   | PD                      | 3                 |
| pd_instance_type           | PD                      | ecs.g5.large      |
| tikv_count                 | TiKV                    | 3                 |
| tikv_instance_type         | TiKV                    | ecs.i2.2xlarge    |
| tidb_count                 | TiDB                    | 2                 |
| tidb_instance_type         | TiDB                    | ecs.c5.4xlarge    |
| monitor_instance_type      |                         | ecs.c5.xlarge     |
| override_values            | TiDB values.yaml file() | nil               |
| local_exec_interpreter     |                         | ["/bin/sh", "-c"] |

## 11. Kubernetes

Kubernetes Terraform Terraform Module

.tf

deploy/aliyun deploy/modules Terraform tidb-operator

(1)

```
```shell
mkdir -p deploy/aliyun-staging
```

```

(2) deploy/aliyun main.tf

```

```hcl
provider "alicloud" {
  region      = <YOUR_REGION>
  access_key  = <YOUR_ACCESS_KEY>
  secret_key  = <YOUR_SECRET_KEY>
}

module "tidb-operator" {
  source      = "../modules/aliyun/tidb-operator"

  region      = <YOUR_REGION>
  access_key  = <YOUR_ACCESS_KEY>
  secret_key  = <YOUR_SECRET_KEY>
  cluster_name = "example-cluster"
  key_file    = "ssh-key.pem"
  kubeconfig_file = "kubeconfig"
}

provider "helm" {
  alias     = "default"
  insecure = true
  install_tiller = false
  kubernetes {
    config_path = module.tidb-operator.kubeconfig_filename
  }
}

module "tidb-cluster" {
  source = "../modules/aliyun/tidb-cluster"
  providers = {
    helm = helm.default
  }

  cluster_name = "example-cluster"
  ack         = module.tidb-operator
}

module "bastion" {
  source = "../modules/aliyun/bastion"

  bastion_name      = "example-bastion"
  key_name          = module.tidb-operator.key_name
  vpc_id            = module.tidb-operator.vpc_id
  vswitch_id        = module.tidb-operator.vswitch_ids[0]
  enable_ssh_to_worker = true
  worker_security_group_id = module.tidb-operator.security_group_id
}
```

```

module "bastion"

```

deploy/aliyun terraform apply clone

```

## 12.

```
pod cidr service cidr
```

### 1.2.3.1.4 TiDB

## 1. Kubernetes

<https://docs.jdcloud.com/cn/jcs-for-kubernetes/create-to-cluster>

## 2.

<https://docs.jdcloud.com/cn/jcs-for-kubernetes/connect-to-cluster>

## 3. Helm

- (1) <https://github.com/helm/helm/releases> Helm TiDB Operator Helm < 3.0

```
wget https://get.helm.sh/helm-v2.16.1-linux-amd64.tar.gz
```

- (2)

```
tar -zxvf helm-v2.16.1-linux-amd64.tar.gz
```

- (3)

```
mv linux-amd64/helm /usr/local/bin/helm
chmod +x /usr/local/bin/helm
```

- (4)

```
helm help
```

- (5) Tiller [Role-based Access Control rbac-config.yaml](#)

```
apiVersion: v1
kind: ServiceAccount
metadata:
 name: tiller
 namespace: kube-system

apiVersion: rbac.authorization.k8s.io/v1beta1
kind: ClusterRoleBinding
metadata:
 name: tiller
 roleRef:
 apiGroup: rbac.authorization.k8s.io
 kind: ClusterRole
 name: cluster-admin
 subjects:
 - kind: ServiceAccount
 name: tiller
 namespace: kube-system
```

- (6) Helm Tiller

```
helm init --upgrade --service-account tiller
```

```
--tiller-image
```

(7)

```
helm version
```

```
Client: &version.Version{SemVer:"v2.16.1", GitCommit:"bbdfe5e7803a12bbdf97e94cd847859890cf4050", GitTreeState:"clean"}
Server: &version.Version{SemVer:"v2.16.1", GitCommit:"bbdfe5e7803a12bbdf97e94cd847859890cf4050", GitTreeState:"clean"}
```

(8) PingCAP chart

```
helm repo add pingcap https://charts.pingcap.org/
```

## 4. TiDB Operator

TiDB Operator CRD (Custom Resource Definition) Kubernetes TiDB Operator TidbCluster Kubernetes

```
kubectl apply -f https://raw.githubusercontent.com/pingcap/tidb-operator/master/manifests/crd.yaml && kubectl get crd tidbclusters.pingcap.com
```

TidbCluster Kubernetes TiDB Operator

(1) tidb-operator chart values.yaml

```
mkdir -p /home/tidb/tidb-operator && \
helm inspect values pingcap/tidb-operator --version=<chart-version> > /home/tidb/tidb-operator/values-tidb-operator.yaml
```

<chart-version> chart v1.0.0 helm search -l tidb-operator

(2) TiDB Operator

TiDB Operator k8s.gcr.io/kube-scheduler /home/tidb/tidb-operator/values-tidb-operator.yaml  
scheduler.kubeSchedulerImageName

(3) TiDB Operator

```
helm install pingcap/tidb-operator --name=tidb-operator --namespace=tidb-admin --version=<chart-version> -f /home/tidb/tidb-operator/values-tidb-operator.yaml && \
kubectl get po -n tidb-admin -l app.kubernetes.io/name=tidb-operator
```

## 5. TiDB Operator

/home/tidb/tidb-operator/values-tidb-operator.yaml TiDB Operator values.yaml /home/tidb/tidb-operator/values-tidb-operator.yaml

TiDB Operator

- tidb-controller-manager
- tidb-scheduler

Deployment    values.yaml    limit    request    replicas

values.yaml

```
helm upgrade tidb-operator pingcap/tidb-operator --version=<chart-version> -f /home/tidb/tidb-operator/values-tidb-operator.yaml
```

## 6. ulimit

TiDB Docker    ulimit    1048576

- Kubernetes Node

云主机 | [REDACTED] ● 运行中

资源信息    弹性网卡    监控    基盘    安全组

**基本信息**

ID : [REDACTED] 名称 : [REDACTED] 描述 : kubernetes cluster node  
计费类型 : 按配置 创建时间 : 2020-03-07 01:46:31 地域 : 华北-北京  
可用区 : 可用区A 高可用组 : k8s-vm-ag-c9ma8cqo9 故障域 : 4  
标签 : --

**配置信息**

镜像 : k8s-node-cloud-1123 规格 : g.n2.xlarge (4核16GB 通用 标准型) 系统盘 : 通用型SSD云盘 (100GB)  
密钥 : --

**网络信息**

所属网络 : drds\_vpc 子网 : k8s-2aa08twq0y-node-subnet-3bf 内网IPv4 : 192.168.160.6  
公网IPv4 : -- IPv6 : --

操作▼

远程连接  
停止  
重启  
制作镜像  
重置系统  
**重置密码**  
相同配置创建  
调整配置  
续费  
删除  
编辑标签

- Node

- ulimit    ulimit

```
sudo vim /etc/security/limits.conf
```

root soft hard nofile 1048576

- Docker    ulimit

```
sudo vim /etc/systemd/system/docker.service
```

LimitNOFILE 1048576

- Docker

```
systemctl daemon-reload
systemctl restart docker
```

LimitNOFILE 1048576 infinity systemd bug infinity systemd 65536

## 7. TiDB

### StorageClass

Kubernetes [kubernetes.io/jdcloud-ebs](https://kubernetes.io/jdcloud-ebs) provisioner Kubernetes Kubernetes StorageClass:

```
kubectl get storageclass
NAME PROVISIONER AGE
default (default) kubernetes.io/jdcloud-ebs 39d
jdcloud-hdd kubernetes.io/jdcloud-ebs 39d
jdcloud-ssd kubernetes.io/jdcloud-ebs 39d
```

### StorageClass

```
kind: StorageClass
apiVersion: storage.k8s.io/v1
metadata:
 name: mystorageclass-hdd1
provisioner: kubernetes.io/jdcloud-ebs
parameters:
 zones: cn-north-1a, cn-north-1b
 fstype: ext4
reclaimPolicy: Retain
```

(1) provisioner kubernetes.io/jdcloud-ebs Provisioner

(2) reclaimPolicy storage class Persistent Volume reclaimPolicy Delete Retain storageClass reclaimPolicy Delete

(3) parameters

type ssd(gp1) ssd(io1) hdd(std1) SSD SSD HDD

| StorageClass type |     |               |       |
|-------------------|-----|---------------|-------|
| hdd.std1          | hdd | [20-16000]GiB | 10GiB |
| ssd(gp1)          | ssd | [20-16000]GiB | 10GiB |
| ssd.io1           | ssd | [20-16000]GiB | 10GiB |

fstype xfs fstype ext4 fstype ext4 fstype=ext4

<https://docs.jdcloud.com/cn/jcs-for-kubernetes/deploy-storageclass>

## Values

tidb-cluster chart values.yaml

```
mkdir -p /home/tidb/<release-name> && \
helm inspect values pingcap/tidb-cluster --version=<chart-version> > /home/tidb/<release-name>/values-<release-name>.yaml
```

- /home/tidb
- release-name Kubernetes PodService      helm ls -q release-name
- chart-version tidb-cluster chart      helm search -l tidb-cluster
- values.yaml      /home/tidb/<release-name>/values-<release-name>.yaml

3 PD Pod3 TiKV Pod2 TiDB Pod 1 PodTiDB Operator Kubernetes 3 Kubernetes 3 PD Pod Pending  
TiKV TiDB Pod

Kubernetes 3 TiDB PD TiKV Pod 1      values.yaml    schedulerName Kubernetes default-scheduler

default-scheduler    default-scheduler    TiDB    [TiDB Pod StableScheduling](#)

TiDB

## 8. TiDB

[20-16000]GiB 10G values.yaml PD TiKV Monitor Drainer PV

Kubernetes values.yaml tidb.service.annotations service.beta.kubernetes.io/jdcloud-load-balancer-spec TiDB Service Port listeners

```
tidb:
 service:
 type: LoadBalancer
 exposeStatus: true
 annotations:
 service.beta.kubernetes.io/jdcloud-load-balancer-spec: |
 version: "v1" # "v1"
 loadBalancerType: nlb # JD LB
 internal: true # true LB IPfalse IP IP

 listeners:
 - protocol: "tcp" # port LB listener , ports
 # listener , alb: Tcp, Http,Https,Tls;nlb: Tcp;dnlb: Tcp

 connectionIdleTimeSeconds: 1800 # alb/nlb
 backend: # JD LB backend
 connectionDrainingSeconds: 300 # nlb target 300s [0-3600] (Optional)

 sessionStickyTimeout: 300 # nlb sessionStickiness 300s, [1-3600] (Optional)

 algorithm: "IpHash" # , [IpHash, RoundRobin, LeastConn] IP Hash RoundRobin (Optional)
 nlbdnlnlb:
 - protocol: "tcp"
 connectionIdleTimeSeconds: 1800
 backend:
 connectionDrainingSeconds: 300
 sessionStickyTimeout: 300
 algorithm: "IpHash"
```

LoadBalance <https://docs.jdcloud.com/cn/jcs-for-kubernetes/deploy-service-new>

Secret

```
kubectl create secret generic <tidb-secretname> --from-literal=root=<password> --namespace=<namespace>
```

values.yaml tidb passwordSecretName <tidb-secretname> TiDB

values.yaml TiDB

```
helm install pingcap/tidb-cluster --name=<release-name> --namespace=<namespace> --version=<chart-version> -f /home/tidb/<release-name>/values-<release-name>.yaml
```

namespace release-name

Pod

```
kubectl get po -n <namespace> -l app.kubernetes.io/instance=<release-name>
```

Kubernetes TiDB Operator TiDB release-name namespace namespace

TiDB TiDB Service ClusterIP

```
$kubectl -n jddb get svc -l app.kubernetes.io/instance=jddb
```

| NAME                  | TYPE         | CLUSTER-IP      | EXTERNAL-IP    |                                    |
|-----------------------|--------------|-----------------|----------------|------------------------------------|
| jddb-discovery        | ClusterIP    | 192.168.189.43  | <none>         | 10261/TCP 65m                      |
| jddb-grafana          | NodePort     | 192.168.190.132 | <none>         | 3000:32445/TCP 65m                 |
| jddb-monitor-reloader | NodePort     | 192.168.188.141 | <none>         | 9089:30732/TCP 65m                 |
| jddb-pd               | ClusterIP    | 192.168.188.210 | <none>         | 2379/TCP 65m                       |
| jddb-pd-peer          | ClusterIP    | None            | <none>         | 2380/TCP 65m                       |
| jddb-prometheus       | NodePort     | 192.168.186.63  | <none>         | 9090:30415/TCP 65m                 |
| jddb-tidb             | LoadBalancer | 192.168.188.201 | 192.168.176.10 | 4000:30487/TCP,10080:32045/TCP 65m |
| jddb-tidb-peer        | ClusterIP    | None            | <none>         | 10080/TCP 62m                      |
| jddb-tikv-peer        | ClusterIP    | None            | <none>         | 20160/TCP 63m                      |

jddb-tidb TiDB Service Pod CLUSTER-IP TiDB VPC EXTERNAL-IP TiDB



### 1.2.3.2.1

Kubernetes TiDB

#### 1.

|            |                       |
|------------|-----------------------|
|            |                       |
| Docker     | Docker CE 18.09.6     |
| Kubernetes | v1.12.5+              |
| CentOS     | CentOS 7.6 3.10.0-957 |

#### 2.

|                                   |         |
|-----------------------------------|---------|
|                                   |         |
| net.core.somaxconn                | 32768   |
| vm.swappiness                     | 0       |
| net.ipv4.tcp_syncookies           | 0       |
| net.ipv4.ip_forward               | 1       |
| fs.file-max                       | 1000000 |
| fs.inotify.max_user_watches       | 1048576 |
| fs.inotify.max_user_instances     | 1024    |
| net.ipv4.conf.all.rp_filter       | 1       |
| net.ipv4.neigh.default.gc_thresh1 | 80000   |
| net.ipv4.neigh.default.gc_thresh2 | 90000   |
| net.ipv4.neigh.default.gc_thresh3 | 100000  |

Kubernetes swap

```
swapoff -a
```

swap

```
free -m
```

swap 0 swap

swap /etc/fstab swap

SMP IRQ Affinity CPU CPU TiDB CPU

```
cat /proc/interrupts|grep <iface-name>|awk '{print $1,$NF}'
```

CPU

```
cat /proc/irq/<ir_num>/smp_affinity
```

CPU SMP IRQ Affinity

```
cat /proc/irq/<ir_num>/smp_affinity_list
```

CPU

CPU SMP IRQ Affinity CPU

- irqbalance centos7 :

```
systemctl start irqbalance
```

- irqbalance CPU set\_irq\_affinity.sh

RPS/RFS (RSS) irqbalance RPSRFS

### 3.

binary TiDB Intel x86-64 64 TiDB TiDB

diskmemoryCPU Kubernetes master etcd worker master worker master kubelet Kubernetes

3 Kubernetes master + 3 etcd + worker Kubernetes master

## 4. Kubernetes

- SAS 1T Docker kubelet Docker kubelet emptyDir
- Kubernetes , Prometheus SAS SAS SAS

RAID5 RAID5

- etcd Kubernetes master master etcd etcd SSD

## 5. TiDB

TiDB PDTiKVTiDB TiDB TiDB 3 PD + 3 TiKV + 2 TiDB

- PD PD 2C 4GB

PD master 5 TiDB 3 master 5 PD 5 PD SSD GB bind mount SSD 5  
filesystem by multiple filesystem PVs

TiDB master PD master worker PD PD PD TiDB

SSD master PD 7 7 TiDB 3 PD PD

- TiKV TiKV I/O TiKV NVMe 8C 32GB TiKV buffer
- TiDB TiDB CPU 8C 32 GB

## 6. TiDB

5 3 PD + 3 TiKV + 2 TiDB PD 2C 4GBTiDB 8C 32GBTiKV 8C 32GBKubernetes 7 3 master  
worker 4 worker

- master
  - 1 etcd (2C 4GB) + 2 PD (2 \* 2C 2 \* 4GB) + 3 TiKV (3 \* 8C 3 \* 32GB) + 1 TiDB (8C 32GB) 38C 140GB
  - SSD etcd 2 PD
  - RAID5 SAS Docker kubelet
  - NVMe TiKV
- worker
  - 3 PD (3 \* 2C 3 \* 4GB) + 2 TiKV (2 \* 8C 2 \* 32GB) + 2 TiDB (2 \* 8C 2 \* 32GB) 38C 140GB
  - SSD PD
  - RAID5 SAS Docker kubelet
  - NVMe TiKV

5 TiDB 7 3 master worker 4 worker

- master worker 48C 192GB2 SSD RAID5 SAS NVMe
- worker 48C 192GB1 SSD RAID5 SAS NVMe

master TiDB TiDB master worker master

### 1.2.3.2.2 Kubernetes

#### 1.

TiDB PDTiKV TiDB Binlog Kubernetes

[PersistentVolume \(PV\)](#)Kubernetes

•

•

IDC

PV volume provisioner PV Pod

[PersistentVolumeClaim \(PVC\)](#)

PV PV PVC PV volume provisioner PVC

PV PVC PV

PV volume provisioner

#### 2. TiDB

TiKV Raft PD TiKV SSD

PD Raft IO SAS SSD AWS gp2 EBS GCP SSD

TiDB Binlog TiDB Binlog pump drainer IO AWS io1 EBS GCP SSD

TiDB Operator TiDB values.yaml storageClassName local-storage

#### 3. PV

Kubernetes 1.11 [PV](#) StorageClass

```
kubectl patch storageclass <storage-class-name> -p '{"allowVolumeExpansion": true}'
```

PV

(1) PVC

```
PVC 10 Gi 100 Gi
```
kubectl patch pvc -n <namespace> <pvc-name> -p '{"spec": {"resources": {"requests": {"storage": "100Gi"}}}}'
```
```

(2) PV

```
`kubectl get pvc -n <namespace> <pvc-name>` PV
```
kubectl get pv | grep <pvc-name>
```
`
```

## 4. PV

Kubernetes local-static-provisioner local-volume-provisioner

(1) Kubernetes

(2) local-volume-provisioner

```
```shell
kubectl apply -f https://raw.githubusercontent.com/pingcap/tidb-operator/master/manifests/local-dind/local-volume-provisioner.yaml
```

Pod PV

```shell
kubectl get po -n kube-system -l app=local-volume-provisioner && \
kubectl get pv | grep local-storage
```

`local-volume-provisioner` (discovery directory) PV GKE 375 GiB
```

Kubernetes local-static-provisioner

- Local PV UUID
- IO
- 

local-static-provisioner

TiDB Binlog SAS StorageClass

- bind mount /mnt/disks local-storage StorageClass

TiDB 1 1 PV TiDB 1 PV

- TiDB Binlog bind mount /mnt/backup backup-storage StorageClass

TiDB Pump 1 1 PV Pump 1 PV drainer 1 PV Ad-hoc 1 PV 1 PV

- PD bind mount /mnt/sharedssd shared-ssd-storage StorageClass

TiDB PD 1 1 PV PD PV

- TiKV /mnt/ssd ssd-storage StorageClass

local-volume-provisioner.yaml StorageClass.yaml

```
apiVersion: storage.k8s.io/v1
kind: StorageClass
metadata:
 name: "local-storage"
provisioner: "kubernetes.io/no-provisioner"
volumeBindingMode: "WaitForFirstConsumer"

```

```

apiVersion: storage.k8s.io/v1
kind: StorageClass
metadata:
 name: "ssd-storage"
provisioner: "kubernetes.io/no-provisioner"
volumeBindingMode: "WaitForFirstConsumer"

apiVersion: storage.k8s.io/v1
kind: StorageClass
metadata:
 name: "shared-ssd-storage"
provisioner: "kubernetes.io/no-provisioner"
volumeBindingMode: "WaitForFirstConsumer"

apiVersion: storage.k8s.io/v1
kind: StorageClass
metadata:
 name: "backup-storage"
provisioner: "kubernetes.io/no-provisioner"
volumeBindingMode: "WaitForFirstConsumer"

apiVersion: v1
kind: ConfigMap
metadata:
 name: local-provisioner-config
 namespace: kube-system
data:
 nodeLabelsForPV: |
 - kubernetes.io/hostname
 storageClassMap: |
 shared-ssd-storage:
 hostDir: /mnt/sharedssd
 mountDir: /mnt/sharedssd
 ssd-storage:
 hostDir: /mnt/ssd
 mountDir: /mnt/ssd
 local-storage:
 hostDir: /mnt/disks
 mountDir: /mnt/disks
 backup-storage:
 hostDir: /mnt/backup
 mountDir: /mnt/backup

 volumeMounts:

 - mountPath: /mnt/ssd
 name: local-ssd
 mountPropagation: "HostToContainer"
 - mountPath: /mnt/sharedssd
 name: local-sharedssd
 mountPropagation: "HostToContainer"
 - mountPath: /mnt/disks
 name: local-disks
 mountPropagation: "HostToContainer"
 - mountPath: /mnt/backup
 name: local-backup
 mountPropagation: "HostToContainer"
 volumes:

 - name: local-ssd
 hostPath:
 path: /mnt/ssd
 - name: local-sharedssd
 hostPath:
 path: /mnt/sharedssd

```

```

- name: local-disks
 hostPath:
 path: /mnt/disks
- name: local-backup
 hostPath:
 path: /mnt/backup
.....

```

```
kubectl apply -f local-volume-provisioner
```

```
kubectl apply -f https://raw.githubusercontent.com/pingcap/tidb-operator/master/manifests/local-dind/local-volume-provisioner.yaml
```

TiDB StorageClass

## 5.

PVC PV provisioner StorageClass (reclaim policy) Retain PV Retain Retain PV

- 

|              |                  |              |        |              |        |             |        |
|--------------|------------------|--------------|--------|--------------|--------|-------------|--------|
| StorageClass | provisioner      | StorageClass | GKE pd | StorageClass | Delete | pd-standard | Retain |
| TiDB         | storageClassName | pd-standard  |        |              |        |             |        |

```

apiVersion: storage.k8s.io/v1
kind: StorageClass
metadata:
 name: pd-standard
parameters:
 type: pd-standard
provisioner: kubernetes.io/gce-pd
reclaimPolicy: Retain
volumeBindingMode: WaitForFirstConsumer

```

- PV

```
kubectl patch pv <pv-name> -p '{"spec":{"persistentVolumeReclaimPolicy":"Retain"}}'
```

TiDB Operator PD TiKV PV Retain

PV Retain PV PV

(1) PV PVC

```

```shell
kubectl delete pvc <pvc-name> --namespace=<namespace>
```

```

(2) PV Delete PV

```

```shell
kubectl patch pv <pv-name> -p '{"spec":{"persistentVolumeReclaimPolicy":"Delete"}}'
```

```

PV PV



### 1.2.3.2.3 Kubernetes TiDB Operator

Kubernetes TiDB Operator

#### 1.

TiDB Operator

- Kubernetes v1.12
- Helm >= v2.11.0 && < v3.0.0

- TiDB Operator TiDB TiDB
- Kubernetes v1.12

#### Helm

helm tiller RBAC tiller

```
kubectl apply -f https://raw.githubusercontent.com/pingcap/tidb-operator/master/manifeststiller-rbac.yaml && \
helm init --service-account=tiller --upgrade
```

```
tiller Pod running
```

```
kubectl get po -n kube-system -l name=tiller
```

Kubernetes RBAC tiller

```
helm init --upgrade
```

Kubernetes Helm chart PingCAP helm chart <https://charts.pingcap.org/>

```
helm repo add pingcap https://charts.pingcap.org/
```

```
helm search PingCAP chart
```

```
helm search pingcap -l
```

| NAME                  | CHART VERSION | APP VERSION | DESCRIPTION                             |
|-----------------------|---------------|-------------|-----------------------------------------|
| pingcap/tidb-backup   | v1.0.0        |             | A Helm chart for TiDB Backup or Restore |
| pingcap/tidb-cluster  | v1.0.0        |             | A Helm chart for TiDB Cluster           |
| pingcap/tidb-operator | v1.0.0        |             | tidb-operator Helm chart for Kubernetes |

chart helm repo update

```
helm repo update
```

## 2. TiDB Operator

TiDB Operator CRD (Custom Resource Definition) Kubernetes TiDB Operator TiDBCluster Kubernetes

```
kubectl apply -f https://raw.githubusercontent.com/pingcap/tidb-operator/master/manifests/crd.yaml && \
kubectl get crd tidbclusters.pingcap.com
```

TidbCluster Kubernetes TiDB Operator

(1) tidb-operator chart values.yaml

```
```shell
mkdir -p /home/tidb/tidb-operator && \
helm inspect values pingcap/tidb-operator --version=<chart-version> > /home/tidb/tidb-operator/values-tidb-operator.yaml
```

> ****
>
> `<chart-version>` chart `v1.0.0` `helm search -l tidb-operator`
```

(2) TiDB Operator

```
TiDB Operator `k8s.gcr.io/kube-scheduler` `/home/tidb/tidb-operator/values-tidb-operator.yaml` `scheduler.kubeSchedulerImageName` `registry.cn-hangzhou.aliyuncs.com/google_containers/kube-scheduler`
```

(3) TiDB Operator

```
```shell
helm install pingcap/tidb-operator --name=tidb-operator --namespace=tidb-admin --version=<chart-version> -f /home/tidb/tidb-operator/values-tidb-operator.yaml && \
kubectl get po -n tidb-admin -l app.kubernetes.io/name=tidb-operator
```

```

## 3. TiDB Operator

```
/home/tidb/tidb-operator/values-tidb-operator.yaml TiDB Operator values.yaml /home/tidb/tidb-operator/values-tidb-operator.yaml
```

TiDB Operator

- tidb-controller-manager
- tidb-scheduler

```
Deployment values.yaml limit request replicas
```

```
values.yaml
```

```
helm upgrade tidb-operator pingcap/tidb-operator --version=<chart-version> -f /home/tidb/tidb-operator/values-tidb-operator.yaml
```

## 4. chart

```
PingCAP chart wget http://charts.pingcap.org/tidb-cluster-<chart-version>.tgz chart
```

```
chart /home/tidb/ v1.0.6 :
```

```
cd /home/tidb/ && tar -xvf tidb-operator-chart-v1.0.6.tgz
```

```
/home/tidb/tidb-operator/values.yaml TiDB Operator
```

```
helm install /home/tidb/tidb-operator --namespace=tidb-admin --name=tidb-operator -f /home/tidb/tidb-operator/values.yaml
```

#### TiDB Operator

```
kubectl get pods -n tidb-admin
```

| NAME                                     | READY | STATUS  | RESTARTS | AGE |
|------------------------------------------|-------|---------|----------|-----|
| tidb-controller-manager-85d8d498bf-2n8km | 1/1   | Running | 0        | 19s |
| tidb-scheduler-7c67d6c77b-qd54r          | 2/2   | Running | 0        | 19s |

#### TiDB Operator

### 1.2.3.2.4 TiDB

#### (1) TiDB Cluster helm chart

```
mkdir -p /root/charts/
https://github.com/pingcap/tidb-operator/releases/tidb-cluster-chart-v1.0.6.tgz /root/charts/
```

#### (2) TiDB Cluster

```
cd /root/charts/ && tar xvf tidb-cluster-chart-v1.0.6.tgz
helm install --namespace dba-test --name=test /root/charts/tidb-cluster -f /root/charts/tidb-cluster/values.yaml
1
NAME: test
LAST DEPLOYED: Sat Mar 7 05:27:57 2020
NAMESPACE: dba-test
STATUS: DEPLOYED
...
...
```

TiDB Cluster

#### (3) TiDB Cluster Pod

```
kubectl get pods -n dba-test
NAME READY STATUS RESTARTS AGE
test-discovery-854fb5b46c-hbg4q 1/1 Running 0 4m41s
test-monitor-66589f9748-q28lp 3/3 Running 0 4m41s
test-pd-0 1/1 Running 1 4m40s
test-pd-1 1/1 Running 0 4m40s
test-pd-2 1/1 Running 0 4m40s
test-tidb-0 2/2 Running 0 2m13s
test-tidb-1 2/2 Running 0 2m13s
test-tikv-0 1/1 Running 0 2m45s
test-tikv-1 1/1 Running 0 2m45s
test-tikv-2 1/1 Running 0 2m45s
```

TiDB Cluster Pod

#### (4) TiDB Cluster

```
kubectl get svc -n dba-test
NAME TYPE CLUSTER-IP EXTERNAL-IP PORT(S) AGE
test-discovery ClusterIP 10.102.244.238 <none> 10261/TCP 3m58s
test-grafana NodePort 10.104.130.193 <none> 3000:32326/TCP 3m58s
test-monitor-reloader NodePort 10.106.105.144 <none> 9089:30818/TCP 3m58s
test-pd ClusterIP 10.96.183.196 <none> 2379/TCP 3m58s
test-pd-peer ClusterIP None <none> 2380/TCP 3m58s
test-prometheus NodePort 10.107.17.45 <none> 9090:31800/TCP 3m58s
test-tidb NodePort 10.104.37.71 <none> 4000:30169/TCP,10080:30286/TCP 3m58s
test-tidb-peer ClusterIP None <none> 10080/TCP 90s
test-tikv-peer ClusterIP None <none> 20160/TCP 2m2s
```

test-tidb Service CLUSTER-IP TiDB Cluster

```
mysql -h 10.104.37.71 -uroot -P4000
Welcome to the MySQL monitor. Commands end with ; or \g.
Your MySQL connection id is 1
Server version: 5.7.25-TiDB-v3.0.5 MySQL Community Server (Apache License 2.0)

Copyright (c) 2000, 2019, Oracle and/or its affiliates. All rights reserved.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql>
```

TiDB Cluster

### 1.2.3.2.5 TiDB Kubernetes

## 1. TiDB PD

4 kind 5 5 node

PD TiDB

```
kubectl get pod --all-namespaces -o wide | grep worker5|grep dba-test
dba-test test-pd-2 1/1 Running 0 5m18s 10.244.5.4 kind-worker5
ker5 <none> <none>
```

kind-worker5 PD      kubectl cordon Pod

```
kubectl cordon kind-worker5
node/kind-worker5 cordoned
```

kubectl drain

```
kubectl drain kind-worker5 --ignore-daemonsets --delete-local-data
node/kind-worker5 already cordoned
WARNING: ...
pod/test-pd-2 evicted
```

kind-worker5 Pod

```
kubectl get pods -n dba-test|grep kind-worker5
```

kind-worker5 Pod TiDB Pod

```
kubectl get pods -n dba-test
NAME READY STATUS RESTARTS AGE
test-discovery-854fb5b46c-c8lbg 1/1 Running 0 21m
test-monitor-59468bcd58-btbpcc 3/3 Running 0 10m
test-pd-0 1/1 Running 2 21m
test-pd-1 1/1 Running 0 21m
test-pd-2 1/1 Running 0 4m50s
test-tidb-0 2/2 Running 0 19m
test-tidb-1 2/2 Running 0 19m
test-tikv-0 1/1 Running 0 20m
test-tikv-1 1/1 Running 0 20m
test-tikv-2 1/1 Running 0 20m
```

Pod

1.

2.

```
kubectl uncordon kind-worker5
```

K8s K8s

```
kubectl delete node kind-worker5
node "kind-worker5" deleted
```

## 2. TiKV

TiKV node

4 TiKV 3

kubectl cordon Pod

```
kubectl cordon kind-worker4
node/kind-worker4 cordoned
```

TiKV

```
kubectl get pods -n dba-test -owide|grep tikv|grep kind-worker4
test-tikv-2 1/1 Running 0 5m28s 10.244.1.7 kind-worker4 <none> <
```

TiKV store-id

```
kubectl get tc test -ojson -n dba-test| jq '.status.tikv.stores | .[] | select (.podName == "test-tikv-2") | .id'
"115"
```

PD

```
nohup kubectl port-forward svc/test-pd 2379:2379 -n dba-test &
[1] 8968
```

pd-ctl TiKV

```
pd-ctl -d store delete 115
Success!
```

TiKV store state\_name Tombstone

```
pd-ctl -d store 115|grep state_name
"state_name": "Tombstone"
```

TiKV

Pod PVC

```
kubectl get -n dba-test pod test-tikv-2 -ojson | jq '.spec.volumes | .[] | select (.name == "tikv") | .persistentVolumeClaim.claimName'
"tikv-test-tikv-2"
```

**PVC**

```
kubectl delete pvc/tikv-test-tikv-2 -n dba-test
persistentvolumeclaim "tikv-test-tikv-2" deleted
```

**TiKV**

```
kubectl delete pod/test-tikv-2 -n dba-test
pod "test-tikv-2" deleted
```

**TiKV**

```
kubectl get pods -n dba-test -owide|grep tikv
test-tikv-0 1/1 Running 0 29m 10.244.2.6 kind-worker2 <none> <no
ne>
test-tikv-1 1/1 Running 0 29m 10.244.4.6 kind-worker3 <none> <no
ne>
test-tikv-2 1/1 Running 0 34s 10.244.3.6 kind-worker <none> <no
ne>
test-tikv-3 1/1 Running 0 25m 10.244.3.5 kind-worker <none> <no
ne>
```

test-tikv-2 kind-worker4 kind-worker

:

1.

2.

```
kubectl uncordon kind-worker4
```

**K8s K8s**

```
kubectl delete node kind-worker4
node "kind-worker4" deleted
```

### 1.2.3.2.6 TiDB

“test” TiDB

## 1. helm TiDB Cluster

```
helm list
NAME REVISION UPDATED STATUS CHART APP VERSION NAMESPACE
test 1 Sat Mar 7 22:30:16 2020 DEPLOYED tidb-cluster-v1.0.6
tidb-operator 1 Sat Mar 7 05:02:15 2020 DEPLOYED tidb-operator-v1.0.6
n
helm delete test --purge
release "test" deleted
```

TiDB

## 2. PVC

```
kubectl delete pvc -n dba-test -l app.kubernetes.io/instance=test,app.kubernetes.io/managed-by=tidb-operator
persistentvolumeclaim "pd-test-pd-0" deleted
persistentvolumeclaim "pd-test-pd-1" deleted
persistentvolumeclaim "pd-test-pd-2" deleted
persistentvolumeclaim "pd-test-pd-3" deleted
persistentvolumeclaim "tikv-test-tikv-0" deleted
persistentvolumeclaim "tikv-test-tikv-1" deleted
persistentvolumeclaim "tikv-test-tikv-2" deleted
```

TiDB PVC

## 3. PV

```
kubectl get pv -l app.kubernetes.io/namespace=dba-test,app.kubernetes.io/managed-by=tidb-operator,app.kubernetes.io
/instance=test -o name|xargs -I {} kubectl patch {} -p '{"spec":{"persistentVolumeReclaimPolicy":"Delete"}}'
persistentvolume/local-pv-2c956bbd patched
persistentvolume/local-pv-3a4dae53 patched
persistentvolume/local-pv-3c7e9ebb patched
persistentvolume/local-pv-5ebe9899 patched
persistentvolume/local-pv-682d37c9 patched
persistentvolume/local-pv-af00e20c patched
persistentvolume/local-pv-d4cf548e patched
```

PV

```
kubectl get pv
NAME CAPACITY ACCESS MODES RECLAIM POLICY STATUS CLAIM STORAGECLASS REASON AGE
local-pv-2c956bbd 1468Mi RWO Delete Available local-storage 64s
local-pv-3a4dae53 1468Mi RWO Delete Available local-storage 72s
local-pv-3c7e9ebb 1468Mi RWO Delete Available local-storage 54s
local-pv-5cb252d7 1974Mi RWO Delete Available local-storage 24h
local-pv-5ebe9899 1468Mi RWO Delete Available local-storage 58s
local-pv-682d37c9 1468Mi RWO Delete Available local-storage 72s
local-pv-af00e20c 1974Mi RWO Delete Available local-storage 24h
local-pv-d4cf548e 1468Mi RWO Delete Available local-storage 58s
local-pv-eb0e3c9f 1974Mi RWO Delete Available local-storage 24h
```

PV “test” TiDB

## 1.2.4 Kubernetes TiDB

### 1.2.4.1 Kubernetes TiDB

Kubernetes TiDB TiDB service      <release-name>-tidb.<namespace> TiDB      tidb-cluster Helm chart  
values.yaml    tidb.service

```
tidb:
 service:
 type: NodePort
 # externalTrafficPolicy: Cluster
 # annotations:
 # cloud.google.com/load-balancer-type: Internal
```

#### 1. NodePort

LoadBalancer NodePort NodePort

- externalTrafficPolicy=Cluster TiDB TCP

Cluster IP TiDB TiDB Pod TiDB Pod

TiDB IP IP IP IP

- externalTrafficPolicy=Local TiDB TCP TiDB

Local tidb-scheduler StableScheduling tidb-scheduler TiDB TiDB

- NodePort IP/PORT

Service Node Port TiDB Service

```
kubectl -n <namespace> get svc <release-name>-tidb -ojsonpath=".spec.ports[?(@.name=='mysql-client')].nodePort{`\n'"}
```

IP TiDB

- externalTrafficPolicy Cluster IP
- externalTrafficPolicy Local TiDB

```
kubectl -n <namespace> get pods -l "app.kubernetes.io/component=tidb,app.kubernetes.io/instance=<release-name>" -ojsonpath="{range .items[*]}{.spec.nodeName}{`\n'{end}"
```

- LoadBalancer

LoadBalancer GCP/AWS LoadBalancer

[Kubernetes Service](#) Service Load Balancer

### 1.2.4.2 Kubernetes TiDB

Kubernetes TiDB TiDB Kubernetes TiDB Operator

#### 1. TiDB

TiDB Prometheus Grafana TiDB TiDB Operator TiDB TiDB TiDB Namespace Prometheus Grafana

values.yaml monitor.persistent true storageClass

1.

Grafana   NodePort   Kubernetes      values.yaml      monitor.grafana.service.type      LoadBalancer      helm upgrade

Grafana      values.yaml      monitor.grafana.create      false

1.

Prometheus   NodePort   Kubernetes      values.yaml      monitor.prometheus.service.type      LoadBalancer      helm  
upgrade

1. Kubernetes

TiDB TiDB Kubernetes TiDB Operator Kubernetes

1.

Kubernetes Kubernetes

Kubernetes Kubernetes

- [CollectD](#)
- [Nagios](#)
- [Prometheus & node\\_exporter](#)
- [Zabbix](#)

[Prometheus Operator](#) Kubernetes [Node Exporter](#) Prometheus Kubernetes

1. Kubernetes

Kubernetes Kubernetes

Kubernetes Kubernetes

TiDB Operator Kubernetes Kubernetes TiDB Operator

[Prometheus Operator](#) [Node Exporter](#) Prometheus

## 1.2.4.3

1. TiDB

TiDB Prometheus Prometheus Alerts

charts

Prometheus Prometheus      [AlertManager](#)

AlertManager      values.yaml      monitor.prometheus.alertmanagerURL      Prometheus AlertManager

1. Kubernetes

Prometheus Operator Kubernetes AlertManager      [kube-prometheus](#)

Kubernetes

## 1.2.5 BR TiDB

TiDB AWS S3 AWS S3 “Ad-hoc” BR AWS

TiDB Operator v1.1 Custom Resource Definition(CRD)

### AWS

:

(1) AWS AccessKey SecretKey :

AWS AWS\_ACCESS\_KEY\_ID AWS\_SECRET\_ACCESS\_KEY

(2) IAM Pod : IAM Pod Pod kube2iam

- kube2iam Kubernetes kube2iam TiDB Operator TiDB
- hostNetwork spec.tikv.hostNetwork false

(3) IAM ServiceAccount :

IAM Kubeneters serviceAccount ServiceAccount Pod EKS Pod Identity Webhook

- AWS EKS TiDB Operator TiDB

### Ad-hoc

Ad-hoc Backup Custom Resource (CR) TiDB Operator Backup

Ad-hoc Kubernetes test1 namespace TiDB demo1

### Ad-hoc

#### AccessKey SecretKey

(1) backup-rbac.yaml test1 namespace RBAC

```
kubectl apply -f backup-rbac.yaml -n test1
```

(2) s3-secret secret secret S3

```
kubectl create secret generic s3-secret --from-literal=access_key=xxx --from-literal=secret_key=yyy --namespace=test1
```

(3) backup-demo1-tidb-secret secret secret TiDB

```
kubectl create secret generic backup-demo1-tidb-secret --from-literal=password=<password> --namespace=test1
```

### IAM Pod

(1) backup-rbac.yaml test1 namespace RBAC

```
kubectl apply -f backup-rbac.yaml -n test1
```

(2) backup-demo1-tidb-secret secret secret TiDB

```
kubectl create secret generic backup-demo1-tidb-secret --from-literal=password=<password> --namespace=test1
```

(3) IAM

[AWS IAM](#) [AWS IAM](#) [Backup](#) [AWS S3 IAM](#) [AmazonS3FullAccess](#)

(4) IAM TiKV Pod

BR TiKV Pod BR Pod S3 TiKV Pod annotation IAM

```
kubectl edit tc demo1 -n test1
```

```
spec.tikv.annotations annotation arn:aws:iam::123456789012:role/user TiKV Pod Pod annotation
```

arn:aws:iam::123456789012:role/user (3) IAM

## IAM ServiceAccount

(1) backup-rbac.yaml test1 namespace RBAC

```
kubectl apply -f backup-rbac.yaml -n test2
```

(2) backup-demo1-tidb-secret secret secret TiDB root

```
kubectl create secret generic backup-demo1-tidb-secret --from-literal=password=<password> --namespace=test1
```

(3) IAM

[AWS EKS IAM](#)

(4) IAM

[AWS IAM](#) [AmazonS3FullAccess](#) [Trust relationships](#)

(5) IAM ServiceAccount

```
kubectl annotate sa tidb-backup-manager -n eks.amazonaws.com/role-arn=arn:aws:iam::123456789012:role/user --namespace =test1
```

arn:aws:iam::123456789012:role/user (4) IAM

(6) ServiceAccount TiKV Pod

```
kubectl edit tc demo1 -n test1
```

```
spec.tikv.serviceAccount tidb-backup-manager TiKV Pod Pod serviceAccountName
```

## BR Amazon S3

- [Backup](#) CR accessKey secretKey :

```
kubectl apply -f backup-aws-s3.yaml
```

```
backup-aws-s3.yaml
```

```

apiVersion: pingcap.com/v1alpha1
kind: Backup
metadata:
 name: demo1-backup-s3
 namespace: test1
spec:
 backupType: full
 br:
 cluster: demo1
 clusterNamespace: test1
 # enableTLSClient: false
 # logLevel: info
 # statusAddr: <status-addr>
 # concurrency: 4
 # rateLimit: 0
 # timeAgo: <time>
 # checksum: true
 # sendCredToTikv: true
 from:
 host: <tidb-host-ip>
 port: <tidb-port>
 user: <tidb-user>
 secretName: backup-demo1-tidb-secret
 s3:
 provider: aws
 secretName: s3-secret
 region: us-west-1
 bucket: my-bucket
 prefix: my-folder
```

- Backup CR IAM Pod:

```
kubectl apply -f backup-aws-s3.yaml
```

```
backup-aws-s3.yaml
```

```

apiVersion: pingcap.com/v1alpha1
kind: Backup
metadata:
 name: demo1-backup-s3
 namespace: test1
 annotations:
 iam.amazonaws.com/role: arn:aws:iam::123456789012:role/user
spec:
 backupType: full
 br:
 cluster: demo1
 sendCredToTikv: false
 clusterNamespace: test1
 # enableTLSClient: false
 # logLevel: info
 # statusAddr: <status-addr>
 # concurrency: 4
 # rateLimit: 0
 # timeAgo: <time>
 # checksum: true
 from:
 host: <tidb-host-ip>
 port: <tidb-port>
 user: <tidb-user>
 secretName: backup-demo1-tidb-secret
 s3:
 provider: aws
 region: us-west-1
 bucket: my-bucket
 prefix: my-folder
```

- Backup CR IAM ServiceAccount :

```
kubectl apply -f backup-aws-s3.yaml
```

```
backup-aws-s3.yaml
```

```

apiVersion: pingcap.com/v1alpha1
kind: Backup
metadata:
 name: demo1-backup-s3
 namespace: test1
spec:
 backupType: full
 serviceAccount: tidb-backup-manager
 br:
 cluster: demo1
 sendCredToTikv: false
 clusterNamespace: test1
 # enableTLSClient: false
 # logLevel: info
 # statusAddr: <status-addr>
 # concurrency: 4
 # rateLimit: 0
 # timeAgo: <time>
 # checksum: true
 from:
 host: <tidb-host-ip>
 port: <tidb-port>
 user: <tidb-user>
 secretName: backup-demo1-tidb-secret
 s3:
 provider: aws
 region: us-west-1
 bucket: my-bucket
 prefix: my-folder

```

Amazon S3 Amazon S3      acl endpoint storageClass

Amazon S3 access-control list (ACL)

- private
- public-read
- public-read-write
- authenticated-read
- bucket-owner-read
- bucket-owner-full-control

ACL      private      AWS

Amazon S3      storageClass

- STANDARD
- REDUCED\_REDUNDANCY
- STANDARD\_IA
- ONEZONE\_IA
- GLACIER
- DEEP\_ARCHIVE

storageClass      STANDARD\_IA      AWS

Backup CR

```
kubectl get bk -n test1 -o wide
```

Backup CR :

- .spec.metadata.namespace      Backup CR namespace
- .spec.from.host      TiDB
- .spec.from.port      TiDB

- `.spec.from.user` TiDB
- `.spec.from.tidbSecretName` TiDB    `.spec.from.user` secret

### S3 provider

- `alibaba` Alibaba Cloud Object Storage System (OSS) formerly Aliyun
- `digitalocean` Digital Ocean Spaces
- `dreamhost` Dreamhost DreamObjects
- `ibmcos` IBM COS S3
- `minio` Minio Object Storage
- `netease` Netease Object Storage (NOS)
- `wasabi` Wasabi Object Storage
- `other` Any other S3 compatible provider

TiDB    `BackupSchedule` CR Ad-hoc

[Ad-hoc](#)

## BR Amazon S3

- `BackupSchedule` CR TiDB `accessKey` `secretKey`

```
kubectl apply -f backup-scheduler-aws-s3.yaml
```

`backup-scheduler-aws-s3.yaml`

```

apiVersion: pingcap.com/v1alpha1
kind: BackupSchedule
metadata:
 name: demo1-backup-schedule-s3
 namespace: test1
spec:
 #maxBackups: 5
 #pause: true
 maxReservedTime: "3h"
 schedule: "*/*/*/*/*"
 backupTemplate:
 backupType: full
 br:
 cluster: demo1
 clusterNamespace: test1
 # enableTLSClient: false
 # logLevel: info
 # statusAddr: <status-addr>
 # concurrency: 4
 # rateLimit: 0
 # timeAgo: <time>
 # checksum: true
 # sendCredToTikv: true
 from:
 host: <tidb-host-ip>
 port: <tidb-port>
 user: <tidb-user>
 secretName: backup-demo1-tidb-secret
 s3:
 provider: aws
 secretName: s3-secret
 region: us-west-1
 bucket: my-bucket
 prefix: my-folder

```

- **BackupSchedule CR TiDB IAM Pod**

```
kubectl apply -f backup-scheduler-aws-s3.yaml
```

```
backup-scheduler-aws-s3.yaml
```

```

apiVersion: pingcap.com/v1alpha1
kind: BackupSchedule
metadata:
 name: demo1-backup-schedule-s3
 namespace: test1
 annotations:
 iam.amazonaws.com/role: arn:aws:iam::123456789012:role/user
spec:
 #maxBackups: 5
 #pause: true
 maxReservedTime: "3h"
 schedule: "*/*/*/*/*"
 backupTemplate:
 backupType: full
 br:
 cluster: demo1
 sendCredToTikv: false
 clusterNamespace: test1
 # enableTLSClient: false
 # logLevel: info
 # statusAddr: <status-addr>
 # concurrency: 4
 # rateLimit: 0
 # timeAgo: <time>
 # checksum: true
 from:
 host: <tidb-host-ip>
 port: <tidb-port>
 user: <tidb-user>
 secretName: backup-demo1-tidb-secret
 s3:
 provider: aws
 region: us-west-1
 bucket: my-bucket
 prefix: my-folder

```

- `BackupSchedule CR TiDB IAM ServiceAccount`

```
kubectl apply -f backup-scheduler-aws-s3.yaml
```

```
backup-scheduler-aws-s3.yaml
```

```

apiVersion: pingcap.com/v1alpha1
kind: BackupSchedule
metadata:
 name: demo1-backup-schedule-s3
 namespace: test1
spec:
 #maxBackups: 5
 #pause: true
 maxReservedTime: "3h"
 schedule: "*/*/*/*/*"
 serviceAccount: tidb-backup-manager
 backupTemplate:
 backupType: full
 br:
 cluster: demo1
 sendCredToTikv: false
 clusterNamespace: test1
 # enableTLSClient: false
 # logLevel: info
 # statusAddr: <status-addr>
 # concurrency: 4
 # rateLimit: 0
 # timeAgo: <time>
 # checksum: true
 from:
 host: <tidb-host-ip>
 port: <tidb-port>
 user: <tidb-user>
 secretName: backup-demo1-tidb-secret
 s3:
 provider: aws
 region: us-west-1
 bucket: my-bucket
 prefix: my-folder

```

```
kubectl get bks -n test1 -o wide
```

```
kubectl get bk -l tidb.pingcap.com/backup-schedule=demo1-backup-schedule-s3 -n test1
```

|                         | backupSchedule | backupSchedule | backupTemplate          | backupTemplate | S3   | Ad-hoc | S3 | BR | Amazon S3 |
|-------------------------|----------------|----------------|-------------------------|----------------|------|--------|----|----|-----------|
| <b>backupSchedule</b>   |                |                |                         |                |      |        |    |    |           |
| ● .spec.maxBackups      |                | 0              |                         |                |      |        |    |    |           |
| ● .spec.maxReservedTime |                | 24h 24         | func ParseDuration      |                |      |        |    |    |           |
| ● .spec.schedule        | Cron           | Cron           |                         |                |      |        |    |    |           |
| ● .spec.pause           | false          | true           | Garbage Collection (GC) |                | true | false  |    |    |           |

## BR AWS

Amazon S3 TiDB

**AccessKey SecretKey**

(1) `backup-rbac.yaml` test2 namespace RBAC

```
kubectl apply -f backup-rbac.yaml -n test2
```

(2) `s3-secret` secret secret S3

```
kubectl create secret generic s3-secret --from-literal=access_key=xxx --from-literal=secret_key=yyy --namespace=test2
```

(3) `restore-demo2-tidb-secret` secret secret TiDB root

```
kubectl create secret generic restore-demo2-tidb-secret --from-literal=password=<password> --namespace=test2
```

## IAM Pod

(1) `backup-rbac.yaml` test2 namespace RBAC

```
kubectl apply -f backup-rbac.yaml -n test2
```

(2) `restore-demo2-tidb-secret` secret secret TiDB root

```
kubectl create secret generic restore-demo2-tidb-secret --from-literal=password=<password> --namespace=test2
```

(3) IAM

[AWS IAM](#)    [AWS IAM](#)    [Restore](#)    [AWS S3 IAM](#)    [AmazonS3FullAccess](#)

(4) IAM TiKV Pod:

BR TiKV Pod BR Pod S3 TiKV Pod annotation IAM

```
kubectl edit tc demo2 -n test2
```

```
spec.tikv.annotations , annotation arn:aws:iam::123456789012:role/user , , TiKV Pod Pod annotation
```

arn:aws:iam::123456789012:role/user (3) IAM

## IAM ServiceAccount

(1) `backup-rbac.yaml` test2 namespace RBAC

```
kubectl apply -f backup-rbac.yaml -n test2
```

(2) `restore-demo2-tidb-secret` secret secret TiDB root:

```
kubectl create secret generic restore-demo2-tidb-secret --from-literal=password=<password> --namespace=test2
```

(3) IAM

[AWS EKS IAM](#)

(4) IAM :

[AWS IAM](#)    [AmazonS3FullAccess](#)    [Trust relationships](#)

(5) IAM ServiceAccount

```
kubectl annotate sa tidb-backup-manager -n eks.amazonaws.com/role-arn=arn:aws:iam::123456789012:role/user --namespace
=test2
```

## (6) ServiceAccount TiKV Pod

```
kubectl edit tc demo2 -n test2
```

```
spec.tikv.serviceAccount tidb-backup-manager , TiKV Pod Pod serviceAccountName
```

```
arn:aws:iam::123456789012:role/user (4) IAM
```

**TiDB**

- `Restore` CR accessKey secretKey

```
kubectl apply -f resotre-aws-s3.yaml
```

```
restore-aws-s3.yaml
```

```

apiVersion: pingcap.com/v1alpha1
kind: Restore
metadata:
 name: demo2-restore-s3
 namespace: test2
spec:
 br:
 cluster: demo2
 clusterNamespace: test2
 # enableTLSClient: false
 # logLevel: info
 # statusAddr: <status-addr>
 # concurrency: 4
 # rateLimit: 0
 # timeAgo: <time>
 # checksum: true
 # sendCredToTikv: true
 to:
 host: <tidb-host-ip>
 port: <tidb-port>
 user: <tidb-user>
 secretName: restore-demo2-tidb-secret
 s3:
 provider: aws
 secretName: s3-secret
 region: us-west-1
 bucket: my-bucket
 prefix: my-folder
```

- `Restore` CR IAM Pod

```
kubectl apply -f restore-aws-s3.yaml
```

```
restore-aws-s3.yaml
```

```

apiVersion: pingcap.com/v1alpha1
kind: Restore
metadata:
 name: demo2-restore-s3
 namespace: test2
 annotations:
 iam.amazonaws.com/role: arn:aws:iam::123456789012:role/user
spec:
 br:
 cluster: demo2
 sendCredToTikv: false
 clusterNamespace: test2
 # enableTLSClient: false
 # logLevel: info
 # statusAddr: <status-addr>
 # concurrency: 4
 # rateLimit: 0
 # timeAgo: <time>
 # checksum: true
 to:
 host: <tidb-host-ip>
 port: <tidb-port>
 user: <tidb-user>
 secretName: restore-demo2-tidb-secret
 s3:
 provider: aws
 region: us-west-1
 bucket: my-bucket
 prefix: my-folder

```

- `Restore CR IAM ServiceAccount`

```
kubectl apply -f restore-aws-s3.yaml
```

```
restore-aws-s3.yaml
```

```

apiVersion: pingcap.com/v1alpha1
kind: Restore
metadata:
 name: demo2-restore-s3
 namespace: test2
spec:
 serviceAccount: tidb-backup-manager
 br:
 cluster: demo2
 sendCredToTikv: false
 clusterNamespace: test2
 # enableTLSClient: false
 # logLevel: info
 # statusAddr: <status-addr>
 # concurrency: 4
 # rateLimit: 0
 # timeAgo: <time>
 # checksum: true
 to:
 host: <tidb-host-ip>
 port: <tidb-port>
 user: <tidb-user>
 secretName: restore-demo2-tidb-secret
 s3:
 provider: aws
 region: us-west-1
 bucket: my-bucket
 prefix: my-folder

```

Restore CR

```
kubectl get rt -n test2 -o wide
```

Restore CR

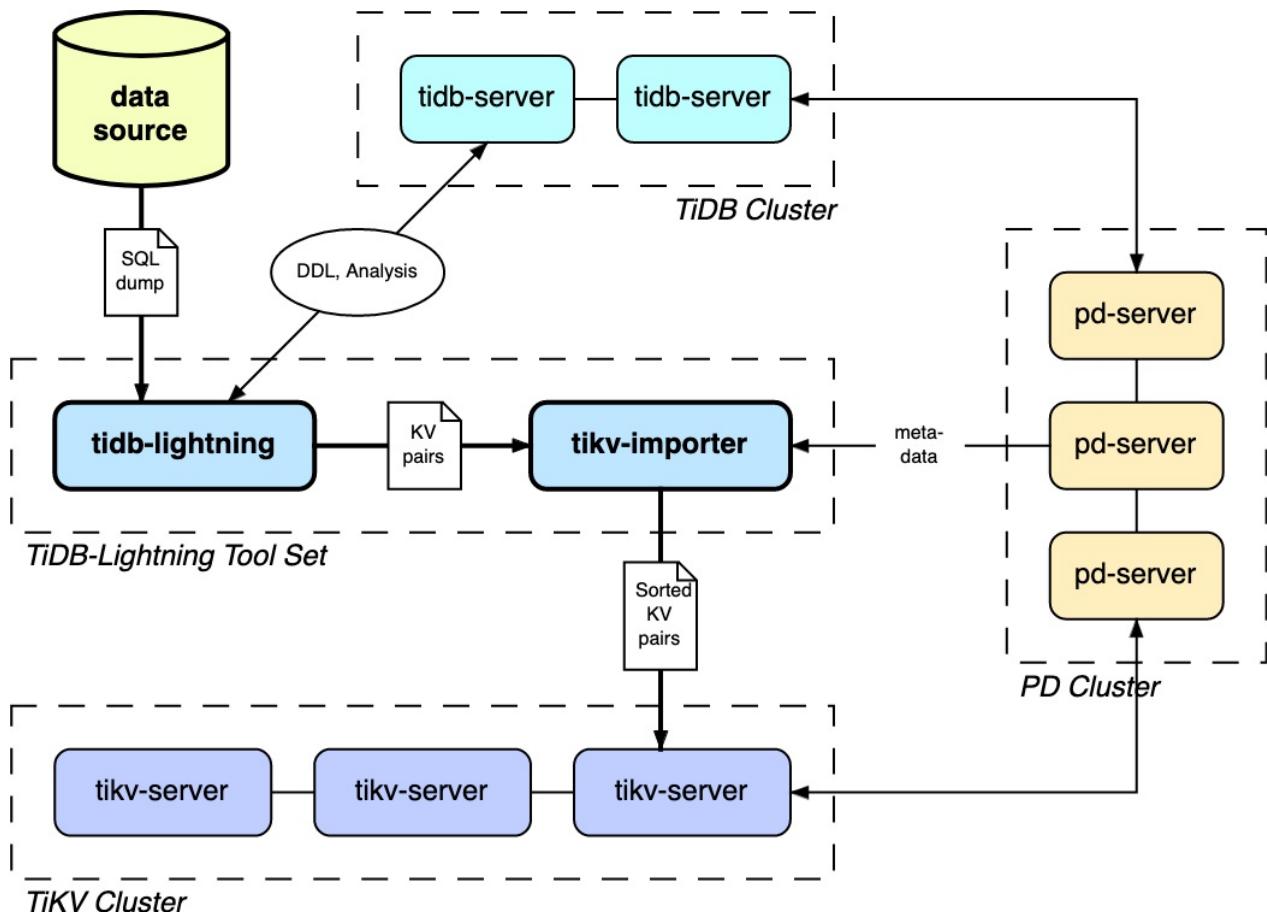
- .spec.metadata.namespace Restore CR namespace
- .spec.to.host TiDB
- .spec.to.port TiDB
- .spec.to.user TiDB
- .spec.to.tidbSecretName TiDB .spec.to.user secret

## 1.2.6 Kubernetes Lightning

### 1.

Mydumper + Loader TiDB SQL TiDB —— SQL KV TiKV

TiDB Lightning



TiDB Lightning

- **tidb-lightning** "" TiDB KV                              **tikv-importer**
- **tikv-importer** "" TiKV                              **tidb-lightning** TiKV

Kubernetes tikv-importer TiDB Helm chart 1 (replicas=1) StatefulSet tidb-lightning Helm chart Job

TiDB Lightning tikv-importer tidb-lightning

## 2. tikv-importer

tikv-importer TiDB TiDB

- TiDB tikv-importer

```
(1) tidb-cluster values.yaml importer.create true
```

```
(2)
```

```
helm install pingcap/tidb-cluster --name=<tidb-cluster-release-name> --namespace=<namespace> -f values.yaml --version=<chart-version>
```

- TiDB tikv-importer

(1) TiDB      `values.yaml`      `importer.create`      `true`

(2) TiDB

```
helm upgrade <tidb-cluster-release-name> pingcap/tidb-cluster -f values.yaml --version=<chart-version>
```

### 3. tidb-lightning

(1) TiDB Lightning

TiDB Lightning

```
helm inspect values pingcap/tidb-lightning --version=<chart-version> > tidb-lightning-values.yaml
```

tidb-lightning Helm chart



Mydumper Kubernetes      `dataSource.local.nodeName`      `dataSource.local.hostPath`      Mydumper

- PVC

PVC Mydumper TiDB namespace PVC      `dataSource.adhoc.pvcName`      Mydumper PVC



`rclone` Mydumper tarball PV `rclone`      Google Cloud Storage (GCS) AWS S3 Ceph Object Storage

- `values.yaml`      `dataSource.local.nodeName`      `dataSource.local.hostPath`

- `rclone`      Secret `rclone`      `rclone`

```

```yaml
apiVersion: v1
kind: Secret
metadata:
  name: cloud-storage-secret
type: Opaque
stringData:
  rclone.conf: |
    [s3]
    type = s3
    provider = AWS
    env_auth = false
    access_key_id = <my-access-key>
    secret_access_key = <my-secret-key>
    region = us-east-1
    [ceph]
    type = s3
    provider = Ceph
    env_auth = false
    access_key_id = <my-access-key>
    secret_access_key = <my-secret-key>
    endpoint = <ceph-object-store-endpoint>
    region = :default-placement
    [gcs]
    type = google cloud storage
    # Storage Object Viewer
    # `cat <service-account-file.json> | jq -c .` 
    service_account_credentials = <service-account-json-file-content>
```

```

secret.yaml    kubectl apply -f secret.yaml -n <namespace>    Secret

- dataSource.remote.storageClassName    Kubernetes

## (2) TiDB Lightning

```
helm install pingcap/tidb-lightning --name=<tidb-lightning-release-name> --namespace=<namespace> --set failFast=true
-f tidb-lightning-values.yaml --version=<chart-version>
```

## 4. Demo

### Mydumper

- 

namespace test-cluster cluster-1cluster-2

cluster-1 Mysql

```
'select * from cloud.test_tbl;'
```

| id | title      | author | date       |
|----|------------|--------|------------|
| 1  | K8s        | shonge | 2020-03-07 |
| 2  | operator   | shonge | 2020-03-07 |
| 3  | kubernetes | shonge | 2020-03-07 |

- secret

```
kubectl create secret generic backup-secret --namespace=test-backup --from-literal=user=root --from-literal=password=<root_password>
```

```
helm install pingcap/tidb-backup --version=v1.1.0-beta.2 --name backup-cluster-1 --namespace test-backup --set-string clusterName=cluster-1,storage.size=500Gi
```

- ```
kubectl -n test-backup get job -l app.kubernetes.io/instance=backup-cluster-1
NAME              COMPLETIONS  DURATION   AGE
basic-fullbackup-202003080800  1/1        3s         3m32s
```

PV

```
kubectl -n test-cluster get pvc -l app.kubernetes.io/instance=backup-cluster-1
NAME      STATUS  VOLUME      CAPACITY  ACCESS MODES  STORAGECLASS  AGE
fullbackup-202003080800  Bound   local-pv-2a2853fb  77Gi     RWO  local-storage  62m

kubectl describe pv local-pv-2a2853fb
Name:          local-pv-2a2853fb
Labels:        kubernetes.io/hostname=tidb-operator-worker2
Annotations:   pv.kubernetes.io/bound-by-controller: yes
               pv.kubernetes.io/provisioned-by: local-volume-provisioner-tidb-operator-worker2-9d6bdbba-89ff-4180-9917-35b4dda3a3db
Finalizers:    [kubernetes.io/pv-protection]
StorageClass:  local-storage
Status:        Bound
Claim:         test-cluster/fullbackup-202003080800
Reclaim Policy: Delete
Access Modes:  RWO
VolumeMode:   Filesystem
Capacity:    500Gi
Node Affinity:
  Required Terms:
    Term 0:  kubernetes.io/hostname in [tidb-operator-worker2]
Message:
Source:
  Type: LocalVolume (a persistent volume backed by local storage on a node)
  Path: /mnt/disks/20
Events: <none>
```

kind worker node

```
docker exec -ti tidb-operator-worker2 ls /mnt/disks/20/fullbackup-202003080800
cloud-schema-create.sql           mysql.opt_rule_blacklist-schema.sql
cloud.test_tbl-schema.sql        mysql.role_edges-schema.sql
cloud.test_tbl.sql               mysql.stats_buckets-schema.sql
metadata                         mysql.stats_feedback-schema.sql
mysql-schema-create.sql          mysql.stats_histograms-schema.sql
mysql.GLOBAL_VARIABLES-schema.sql mysql.stats_histograms.sql
mysql.GLOBAL_VARIABLES.sql       mysql.stats_meta-schema.sql
mysql.bind_info-schema.sql      mysql.stats_meta.sql
mysql.columns_priv-schema.sql   mysql.stats_top_n-schema.sql
mysql.db-schema.sql             mysql.tables_priv-schema.sql
mysql.default_roles-schema.sql   mysql.tidb-schema.sql
mysql.expr_pushdown_blacklist-schema.sql mysql.tidb.sql
mysql.gc_delete_range-schema.sql mysql.user-schema.sql
mysql.gc_delete_range_done-schema.sql mysql.user.sql
mysql.global_priv-schema.sql    test-schema-create.sql
mysql.help_topic-schema.sql
```

Lightning

- cluster-2 importer

```
helm upgrade cluster-2 --set-string importer.create=true pingcap/tidb-cluster
```

- lightning

```
helm install pingcap/tidb-lightning --version=v1.1.0-beta.2 --name restore-cluster-1 --namespace test-cluster --set-string dataSource.adhoc.pvcName='fullbackup-202003080800',targetTidbCluster.name='cluster-2'
```

-

```
kubectl -n test-cluster get job -l app.kubernetes.io/name='restore-cluster-1-tidb-lightning'
```

NAME	COMPLETIONS	DURATION	AGE
restore-cluster-1-tidb-lightning	1/1	3s	9m3s

- cluster-2 TiDB

```
MySQL [(none)]> select * from cloud.test_tbl;
+----+-----+-----+-----+
| id | title | author | date      |
+----+-----+-----+-----+
| 1 | K8s   | shonge | 2020-03-07 |
| 2 | operator | shonge | 2020-03-07 |
| 3 | kubernetes | shonge | 2020-03-07 |
+----+-----+-----+-----+
3 rows in set (0.01 sec)
```

1.2.7 Kubernetes TiDB

Kubernetes TiDB Kubernetes TiDB Kubernetes TiDB

1.2.7.1 Kubernetes PD Control

[PD Control](#) PD

PD Control Kubernetes TiDB PD :

```
kubectl exec -n <namespace> -it <pd-pod-name> sh
```

```
127.0.0.1:2379 PD pd-ctl
```

```
./pd-ctl config show
```

1.2.7.2 Kubernetes TiKV Control

[TiKV Control](#) TiKV TiKV Control Kubernetes TiDB TiKV Control

- [tikv-ctl](#) TiKV PD [kubectl port-forward](#) PD TiKV

```
kubectl port-forward -n <namespace> svc/<cluster-name>-pd 2379:2379 &>/tmp/portforward-pd.log &
```

```
kubectl port-forward -n <namespace> <tikv-pod-name> 20160:20160 &>/tmp/portforward-tikv.log &
```

PD TiKV

```
tikv-ctl --host 127.0.0.1:20160 <subcommand>
```

```
tikv-ctl --pd 127.0.0.1:2379 compact-cluster
```

- [TiKV](#) TiKV [TiKV](#) [TiKV](#) [tkctl debug](#) TiKV Pod [tikv-ctl](#)

1.

```
kubectl annotate pod <tikv-pod-name> -n <namespace> runmode=debug
```

2. TiKV

```
kubectl exec <tikv-pod-name> -n <namespace> -c tikv -- kill -s TERM 1
```

3. debug

```
tkctl debug <tikv-pod-name> -c tikv
```

4. [tikv-ctl](#) [tikv](#) [/proc/1/root](#)

```
tikv-ctl --db /path/to/tikv/db size -r 2
```

```
Kubernetes TiKV debug db          /proc/1/root/var/lib/tikv/db size -r 2
```

1.2.7.3 Kubernetes TiDB Control

[TiDB Control](#) TiDB TiDB Control TiDB PD `kubectl port-forward TiDB PD`

```
kubectl port-forward -n <namespace> svc/<cluster-name>-pd 2379:2379 &>/tmp/portforward-pd.log &
```

```
kubectl port-forward -n <namespace> <tidb-pod-name> 10080:10080 &>/tmp/portforward-tidb.log &
```

```
tidb-ctl
```

```
tidb-ctl schema in mysql
```

1.2.7.4 Helm

[Helm](#) Kubernetes 2.9.0 ≤ Helm < 3.0.0 [Helm](#)

1. Helm

```
curl https://raw.githubusercontent.com/kubernetes/helm/release-2.16/scripts/get | bash
```

macOS Homebrew Helm

```
brew install helm@2
brew link --force helm@2
```

2. Helm

helm tiller RBAC tiller

```
kubectl apply -f https://raw.githubusercontent.com/pingcap/tidb-operator/master/manifests/tiller-rbac.yaml && \
helm init --service-account=tiller --upgrade
```

tiller Pod running

```
kubectl get po -n kube-system -l name=tiller
```

Kubernetes RBAC tiller

```
helm init --upgrade
```

Kubernetes helm chartPingCAP Kubernetes TiDB Helm chart

- `tidb-operator` TiDB Operator
- `tidb-cluster` TiDB
- `tidb-backup` TiDB

chart PingCAP helm chart <https://charts.pingcap.org/>

```
helm repo add pingcap https://charts.pingcap.org/
```

```
helm search PingCAP chart
```

```
helm search pingcap -l
```

NAME	CHART VERSION	APP VERSION	DESCRIPTION
pingcap/tidb-backup	v1.0.0		A Helm chart for TiDB Backup or Restore
pingcap/tidb-cluster	v1.0.0		A Helm chart for TiDB Cluster
pingcap/tidb-operator	v1.0.0		tidb-operator Helm chart for Kubernetes

chart helm repo update

```
helm repo update
```

Helm helm install helm upgrade helm del helm ls Helm chart YAML Helm chart YAML
values.yaml

```
helm ls
```

```
helm ls
```

chart chart-name release-name values.yaml chart chart --version chart-version (GA

-

```
helm install <chart-name> --name=<release-name> --namespace=<namespace> --version=<chart-version> -f <values-file>
```

- chart-version chart values.yaml

```
helm upgrade <release-name> <chart-name> --version=<chart-version> -f <values-file>
```

helm

```
helm del --purge <release-name>
```

Helm3 GA Helm3 helm [Helm](#)

1.2.7.5 Terraform

[Terraform](#) Infrastructure as Code [Kubernetes](#) [TiDB](#) [Terraform](#) [TiDB](#)

[Terraform](#) [Terraform](#)

1.2.8 TiDB Operator

TiDB Operator TiDB Operator TiDB Operator

values.yaml helm upgrade

```
helm upgrade tidb-operator pingcap/tidb-operator --version=<chart-version> -f /home/tidb/tidb-operator/values-tidb-operator.yaml
```

tidb-operator values.yaml operatorImage tidb-operator chart values.yaml values.yaml values.yaml

TiDB Operator TiDB TiDB TiDB Operator TiDB TiDB

TiDB Operator

1.2.8.1 Kubernetes

Kubernetes kubeSchedulerImageTag Helm

helm upgrade

1.3.1 TiOps

TiOps Ansible TiOps

1.3.1.1

TiOps SSH PD join PD

```
tidb-test TiKV PD
1. scale.yaml TiKV PD IP
```

```
---
pd_servers:
  - ip: 10.9.20.15
tikv_servers:
  - ip: 10.9.20.15
```

1. TiOps scale.yaml

```
$ tiops scale-out -c tidb-test -T scale.yaml
```

```
-c | --cluster-name
-T | --topology
--local-pkg
-f | --forks 5
--enable-check-cpu CPU vcores disable
--enable-check-mem Memory Size disable
--enable-check-disk Disk Available Space disable
--enable-check-iops TiKV IOPS latency disable
--enable-check-all CPUMemoryDiskIOPS disable
```

```
tiops display -c tidb-test --status
```

1.3.1.2

TiKV Binlog API TiKV Binglog

- TiKV Binlog
 - TiOps API
 - TiKV Binlog
 -
 -
 -
- - PD API delete PD
 -
- ID display ID ID d294683d PD

```
$ tiops display -c tidb-test
TiDB cluster tidb-test, version 4.0.0-beta.1
Node list:
ID      Role      Host      Ports      Data Dir          Deploy Dir
e5f0e2d2 Monitoring 10.9.26.126 9090/9091 /home/tidb/data/tidb-test/monitoring-e5f0e2d2 /home/tidb/deploy/tid
b-test/monitoring-e5f0e2d2
f8cd24a7 PD        10.9.26.126 2379/2380 /home/tidb/data/tidb-test/pd-f8cd24a7 /home/tidb/deploy/tid
b-test/pd-f8cd24a7
g993a624 PD        10.9.170.207 2379/2380 /home/tidb/data/tidb-test/pd-g993a624 /home/tidb/deploy/tid
b-test/pd-g993a624
d294683d PD        10.9.20.15   2379/2380 /home/tidb/data/tidb-test/pd-d294683d /home/tidb/deploy/tid
b-test/pd-d294683d
a8e4c35a TiDB     10.9.26.126 4000/10080 -           /home/tidb/deploy/tid
b-test/tidb-a8e4c35a
abb1e0ca TiDB     10.9.170.207 4000/10080 -           /home/tidb/deploy/tid
b-test/tidb-abb1e0ca
db591398 Monitored 10.9.165.44 9100/9115 -           /home/tidb/deploy/tid
b-test/monitored-db591398
d444b608 Monitored 10.9.170.207 9100/9115 -           /home/tidb/deploy/tid
b-test/monitored-d444b608
hcbc7435 Monitored 10.9.26.126 9100/9115 -           /home/tidb/deploy/tid
b-test/monitored-hcbc7435
de205f0f TiKV     10.9.26.126 20160/20180 /home/tidb/data/tidb-test/tikv-de205f0f /home/tidb/deploy/tid
b-test/tikv-de205f0f
jd4610dc TiKV     10.9.170.207 20160/20180 /home/tidb/data/tidb-test/tikv-jd4610dc /home/tidb/deploy/tid
b-test/tikv-jd4610dc
bcb3e186 TiKV     10.9.165.44 20160/20180 /home/tidb/data/tidb-test/tikv-bcb3e186 /home/tidb/deploy/tid
b-test/tikv-bcb3e186
h592456d TiKV     10.9.20.15   20160/20180 /home/tidb/data/tidb-test/tikv-h592456d /home/tidb/deploy/tid
b-test/tikv-h592456d
h5f8ca0d Grafana   10.9.26.126 3001      -           /home/tidb/deploy/tid
b-test/grafana-h5f8ca0d
```

1.

```
$ tiops scale-in -c tidb-test -n d294683d
```

```
-c | --cluster-name
-n | --node-id ID
-f | --forks 5
```

```
tiops display -c tidb-test --status
```

1.3.2 TiDB Operator

TiDB Operator TiDB

1.

- TiDB Operator TiDB
- release test dba-test chart /home/charts/tidb-cluster TiDB
 - 3 PD
 - 2 TiDB
 - 3 TiKV

2.

TiDB TiDB replicas PDTiKVTiDB

(1) value.yaml pd.replicas tidb.replicas tikv.replicas 334

(2) helm upgrade

```
```shell
helm upgrade test /home/charts/tidb-cluster -f /home/charts/tidb-cluster/values.yaml
Release "test" has been upgraded. Happy Helming!
LAST DEPLOYED: Sun Mar 8 15:39:57 2020
NAMESPACE: dba-test
STATUS: DEPLOYED
```

```

(3)

```
```shell
watch kubectl get po -n dba-test
```

Pod 3 PD 3 TiDB 4 TiKV `Running`


```shell
NAME READY STATUS RESTARTS AGE
test-discovery-668b48577c-zw4jh 1/1 Running 0 116m
test-monitor-86797cd996-9ggfh 3/3 Running 0 116m
test-pd-0 1/1 Running 0 116m
test-pd-1 1/1 Running 0 116m
test-pd-2 1/1 Running 1 116m
test-tidb-0 2/2 Running 0 112m
test-tidb-1 2/2 Running 0 112m
test-tidb-2 2/2 Running 0 2m52s
test-tikv-0 1/1 Running 0 114m
test-tikv-1 1/1 Running 0 114m
test-tikv-2 1/1 Running 0 114m
test-tikv-3 1/1 Running 0 2m52s
```

```

- PDTiKV
- TiKV PD TiKV Region TiKV TiKV Pod Running Pod TiKV `kubectl get tidbcluster -n <namespace> <release-name> -o json | jq '.status.tikv.stores'` TiKV Offline
- PDTiKV PVC PV Reclaim Policy Retain PVC
- TiKV [TiKV Store Tombstone](#)

3.

(1) `values.yaml` `tidb.resources` `tikv.resources` `pd.resources`

(2) `helm upgrade`

```
```shell
helm upgrade test /home/charts/tidb-cluster -f /home/charts/tidb-cluster/values.yaml
Release "test" has been upgraded. Happy Helming!
LAST DEPLOYED: Sun Mar 8 15:57:03 2020
NAMESPACE: dba-test
STATUS: DEPLOYED
```

```

(3)

```
```shell
watch kubectl -n <namespace> get pod -o wide
```

Pod `Running`
```

- `requests` PDTiKV Local PV Pod Pending
- TiDB TiDB

1.4.1 TiOps

TiDB

1.4.1.1

- PD
 - Leader
 - Leader Leader
 - PD Leader
 - Leader Leader
 -
- TiKV
 - PD TiKV region leader Leader
 - Leader TiKV
 - TiKV Leader
-
-

1.4.1.2

4.0.0-beta.1

```
$ tiops upgrade -c tidb-test -t 4.0.0-beta.1
```

```
-c | --cluster-name
-t | --tidb-version
--enable-check-configdisable
--local-pkg
-f | --forks 5
-r | --role role TiDB "pd", "tikv", "pump", "tidb", "drainer", "monitoring", "monitored", "grafana", "alertmanager"
-n | --node-id ID
--force
```



1.4.2 TiDB-Operator

TiDB PDTiKV TiDB Pod Pod Pod Pod
 PDTiKV Leader TiDB DDL Owner PD * 3TiKV * 3TiDB * 2 TiKVPD
 TiDB TiDB TiDB
 TiDB

1.4.2.1 TiDB

1. values.yaml tidb.image tikv.image pd.image
2. helm upgrade

```
helm upgrade <release-name> pingcap/tidb-cluster -f values.yaml --version=<chart-version>
```

- 3.

```
watch kubectl -n <namespace> get pod -o wide
```

| | |
|-----|---------|
| Pod | Running |
|-----|---------|

1.4.2.2 TiDB

TiDB

1. values.yaml enableConfigMapRollout true
2. values.yaml
3. helm upgrade

```
helm upgrade <release-name> pingcap/tidb-cluster -f values.yaml --version=<chart-version>
```

- 4.

```
watch kubectl -n <namespace> get pod -o wide
```

| | |
|-----|---------|
| Pod | Running |
|-----|---------|

- enableConfigMapRollout PDTiKV TiDB

1.4.2.3 TiDB

PD PD PD tag NodeAffinity TiDB TiDB TiDB
 force-upgrade TiDB Operator > v1.0.0-beta.3 annotation
 kubectl annotate --overwrite tc <release-name> -n <namespace> tidb.pingcap.com/force-upgrade=true

```
helm upgrade
```

```
helm upgrade <release-name> pingcap/tidb-cluster -f values.yaml --version=<chart-version>
```

1.5

1.5.1

TiDB v4.0

-
-

1.5.2

TiDB v4.0 pd-ctl component TiDBTiKVPD

(1)

```
4.0 TiDBTiKVPD      enable-dynamic-config = false
```

(2)

```
/ # ./pd-ctl component --help
manipulate components' configs

Usage:
pd-ctl component [command]

Available Commands:
delete      delete component config with a given component ID (e.g. 127.0.0.1:20160)
ids        get all component IDs with a given component (e.g. tikv)
set        set the component config (set option with value)
show       show component config with a given component ID (e.g. 127.0.0.1:20160)

Global Flags:
-h, --help      Help message.
-u, --pd string  Address of pd. (default "http://127.0.0.1:2379")
```

(3)

```
component set <component> <key> <value>      component      tidb  tikv  pd    key    value
```

```
>> component set tikv gc.batch-keys 1024
```

TiKV GC batch-keys 1024

(4)

```
component set <address> <key> <value>      address  IP 127.0.0.1:20160
```

```
>> component set 127.0.0.1:20160 gc.batch-keys 1024
```

127.0.0.1:20160 TiKV GC batch-keys 1024

(5)

```
component show <address>
```

```
>> component show 127.0.0.1:20160
```

1.5.3

(1) tidbcluster

```
apiVersion: pingcap.com/v1alpha1
kind: TidbCluster
metadata:
  name: basic
spec:
  version: nightly
  timezone: UTC
  pvReclaimPolicy: Delete
  pd:
    baseImage: pingcap/pd
    replicas: 1
    requests:
      storage: "1Gi"
    config: {}
  tikv:
    baseImage: pingcap/tikv
    replicas: 3
    requests:
      storage: "1Gi"
    config: {}
  tidb:
    baseImage: pingcap/tidb
    replicas: 1
    service:
      type: NodePort
    config: {}
```

(2) PD Pod

- ```
/ # ./pd-ctl component show basic-tikv-0.basic-tikv-peer.test-cluster.svc:20160 | grep batch-keys
batch-keys = 512
```

- ```
/ # ./pd-ctl component set tikv gc.batch-keys 1024
Success!
```

- ```
Success! 30s
```

- ```
/ # ./pd-ctl component show basic-tikv-0.basic-tikv-peer.test-cluster.svc:20160 | grep batch-keys
batch-keys = 512
```

pd-ctl

2.1 4.0 CDC

TiCDC (TiDB Change Data Capture) TiDB 4.0 TiKV kv change log TiDB TiCDC TiCDC TiCDC TiDB
TiDB MySQL

2.1.1 CDC

TiCDC TiDB-Tools TiDB-Binlog TiDB-Binlog TiDB binlog TiCDC TiKV kv change log

1.

- TiCDC
 - TiCDC
 - TiCDC SQL admin UI

2.

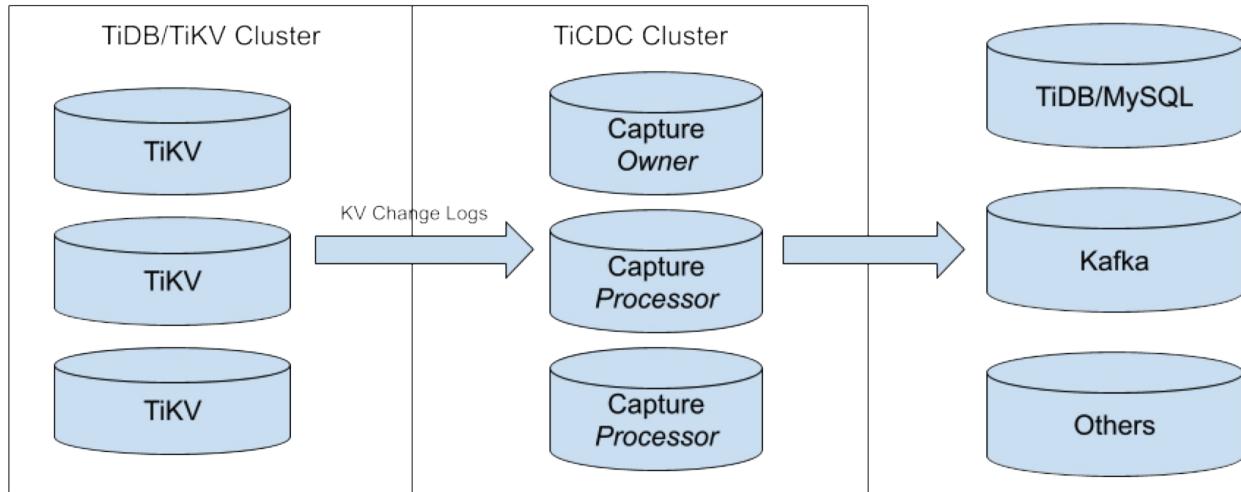
- TiCDC
 - TiCDC TiDB TiDB-Binlog TiCDCPump Drainer TiDB
 - TiDB-Binlog (Commit binlog) TiKV 10minTiCDC ms
 - TiDB-Binlog TiDB transaction
 - big transaction
 - PD commit ts
 - 1 pc

3.

- TiCDC
 - TiCDC watch KV TiCDC PD etcd
 - TiDB-Binlog

2.1.2 CDC

TiCDC TiDB Change Data Capture TiDB/TiKV TiDB TiDB Binlog TiCDC TiDB



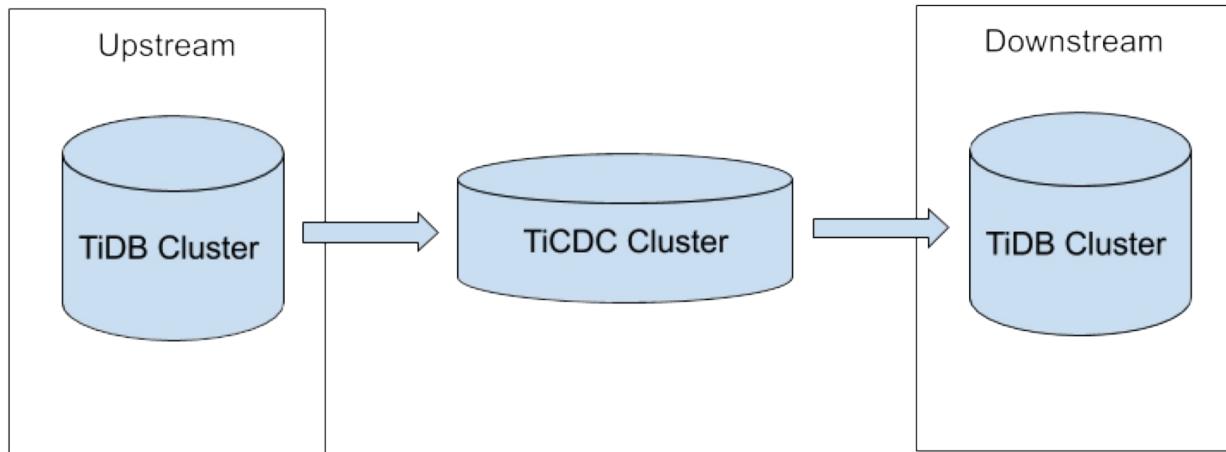
TiCDC

- **TiKV** KV change logs KV TiKV row changed events KV TiKV KV TiCDC
- **capture** TiCDC TiCDC capture KV TiCDC capture owner processor
 - owner owner capture owner
 - processor capture capture processor processor processor capture TiCDC
 - capture PD etcd capture owner processor
 - "" changefeed TiDB TSO KV TiCDC task capture
- **Sink** TiCDC TiCDC
 - TiDB MySQL NOT NULL
 - TiCDC open protocol Kafka Kafka
 -

2.1.3 CDC

TiCDC TiDB

1.



- TiDB PD 10.1.1.10:2379
- TiDB SQL 10.3.1.30:4000
- TiCDC 3 capture
 - 10.2.1.20:8300
 - 10.2.1.21:8300
 - 10.2.1.22:8300

2.

(1)

- CentOS 7.3 Linux x86_64 (amd64)
- TiCDC Go >= 1.13
-

(2) [Github](#) bin

```
$ make
```

(3) TiCDC

```
$ cdc server --pd=http://10.1.1.10:2379 --status-addr=127.0.0.1:8300
```

- pd : TiDB PD
- status-addr : capture

TiCDC TiKV capture

```
$ cdc cli capture list --pd=http://10.1.1.10:2379
[
  {
    "id": "5d1fd3bd-efc9-4cdf-9e8a-6d955f65b3b0",
    "is-owner": true
  },
  {
    "id": "629ec61e-16a3-466c-8fd4-2b2b457dabf7",
    "is-owner": false
  },
  {
    "id": "c5cd08b4-f601-456f-995c-62c97044444b",
    "is-owner": false
  }
]
```

capture owner processor

3.

```
$ cdc cli changefeed create --pd=http://10.1.1.10:2379 --sink-uri="mysql://user:password@10.3.1.30:4000/" --config=~/cdc-config.toml --start-ts=0
```

- pd TiDB PD
- sink-url TiDB DSN
- config TSO
- start-ts TSO 0 TSO

```
test mysql information_schema sns.user sns.following
```

```
ignore-txn-commit-ts = []
filter-case-sensitive = false

[filter-rules]
ignore-dbs = ["test", "mysql""information_schema", ]

[[filter-rules.do-tables]]
db-name = "sns"
tbl-name = "user"

[[filter-rules.do-tables]]
db-name = "sns"
tbl-name = "following"
```

4.

```
$ cdc cli changefeed list --pd=http://10.1.1.10:2379
[
  {
    "id": "004a0ea8-2ef1-45b4-8ce1-b3281e7dc24d"
  }
]
```

ID

```
$ cdc cli changefeed query --pd=http://10.1.1.10:2379 --changefeed-id=004a0ea8-2ef1-45b4-8ce1-b3281e7dc24d
{
  "info": {
    "sink-uri": "mysql://root:123456@127.0.0.1:3306/",
    "opts": {},
    "create-time": "2020-03-13T16:17:33.965778+08:00",
    "start-ts": 415259021527482369,
    "target-ts": 0,
    "admin-job-type": 0,
    "config": {
      "filter-case-sensitive": false,
      "filter-rules": null,
      "ignore-txn-commit-ts": null
    }
  },
  "status": {
    "resolved-ts": 415259037347348481,
    "checkpoint-ts": 415259036823060481,
    "admin-job-type": 0
  }
}
```

```
$ cdc cli processor query --pd=http://10.1.1.10:2379 --changefeed-id=004a0ea8-2ef1-45b4-8ce1-b3281e7dc24d --capture-id=5d1fd3bd-efc9-4cdf-9e8a-6d955f65b3b0
{
  "status": {
    "table-infos": [
      {
        "id": 45,
        "start-ts": 415259021527482369
      }
    ],
    "table-p-lock": null,
    "table-c-lock": null,
    "admin-job-type": 0
  },
  "position": {
    "checkpoint-ts": 415259059393658881,
    "resolved-ts": 415259059917946881
  }
}
```

5. HTTP

TiCDC HTTP

capture

```
$ curl http://10.2.1.20:8300/status
{
  "version": "0.0.1",
  "git_hash": "",
  "id": "4a54c85b-fc1d-4897-9934-1be3b9aa6a45",
  "pid": 31652
}
```

`id` TiCDC capture ID `pid` ID

owner TiCDC owner owner

```
$ curl -X POST http://10.2.1.20:8300/capture/owner/resign
```

owner processor

```
$ curl -X POST -d "admin-job=X&cf-id=136a3bee-621c-42d5-80ec-4c1aaf6ddb53" http://10.2.1.20:8300/capture/owner/admin
```

admin-job

- admin-job=1 processor
- admin-job=2
- admin-job=3 processor

owner

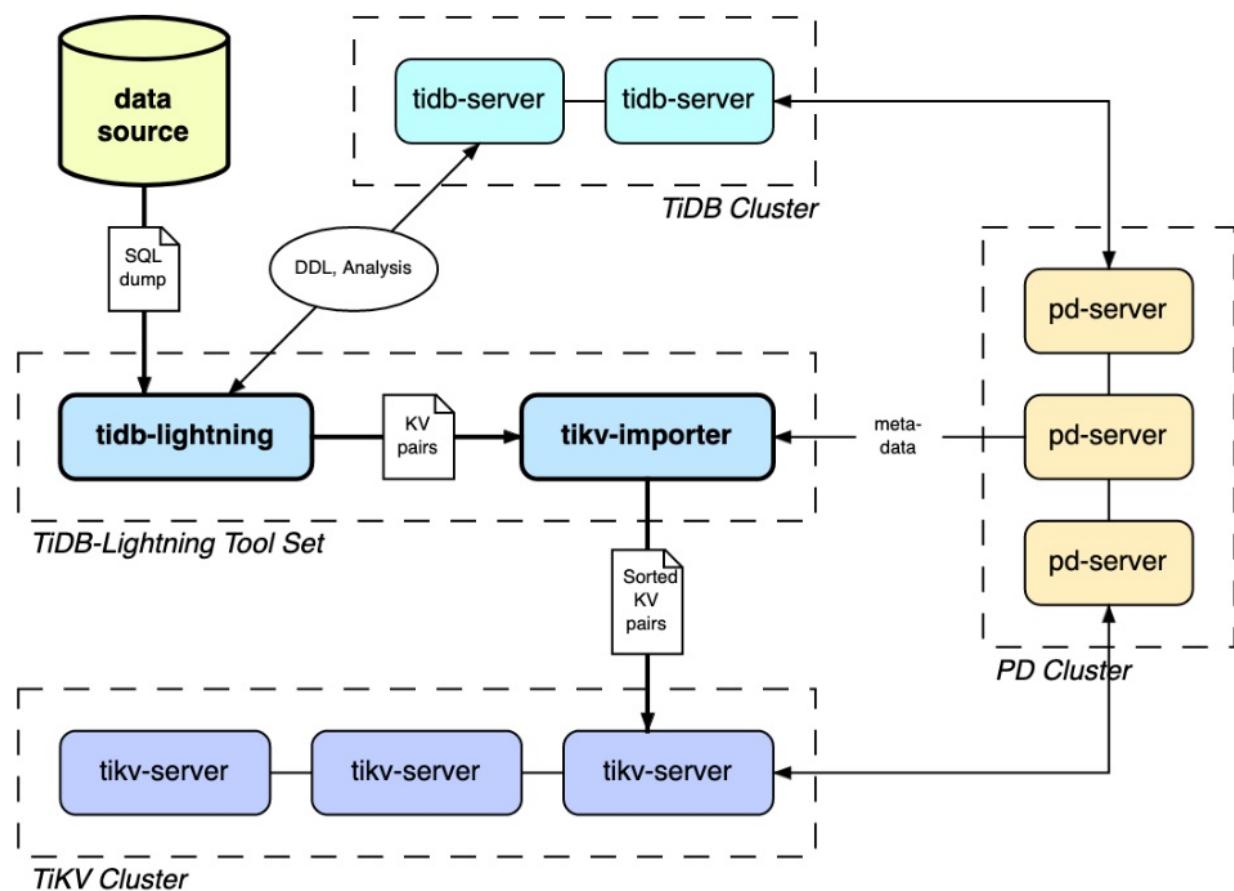
owner processors etcd

```
$ curl http://10.2.1.20:8300/debug/info
```


2.2.1 TiDB Lightning

TiDB Lightning TiDB SQL 3 300 GB Lightning Mydumper CSV

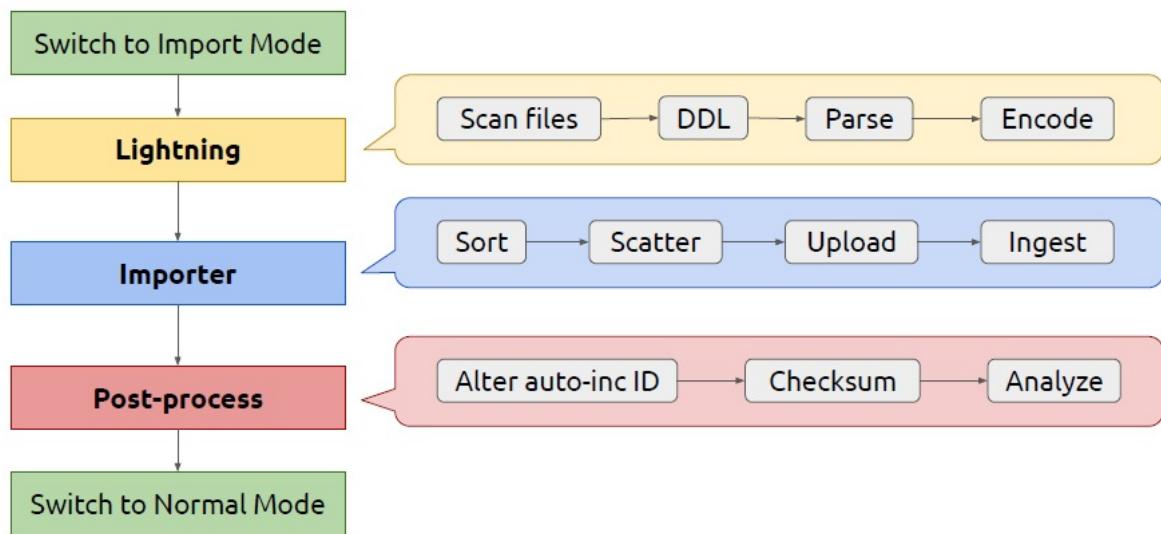
1.



TiDB Lightning

- **TiDB Lightning**“” TiDB tikv-importer
- **tikv-importer**“” TiKV TiDB Lightning TiKV

Import Process



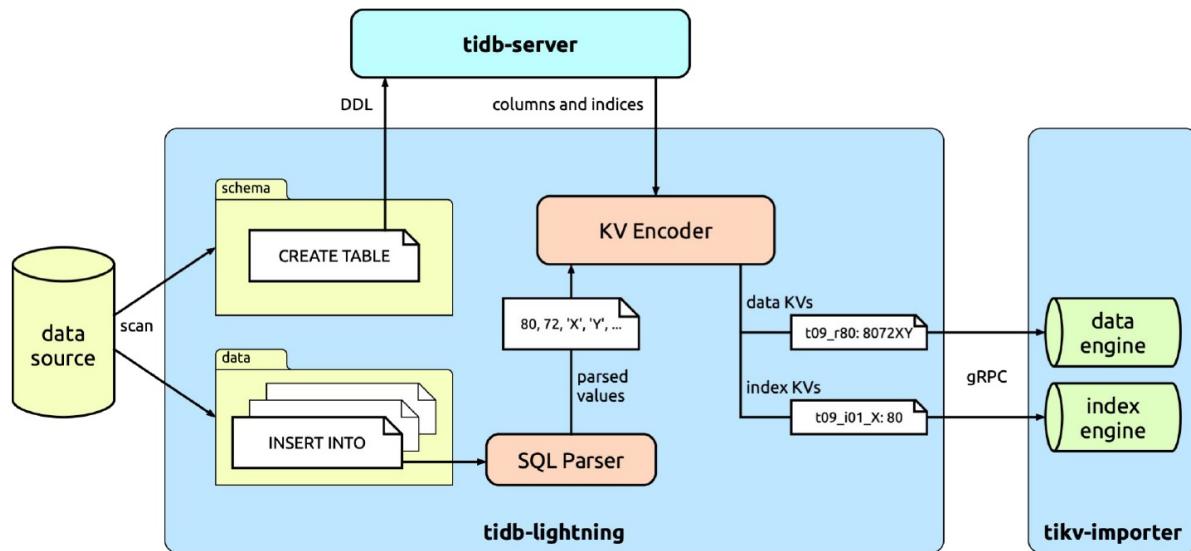
1. TiDB Lightning TiKV “”import mode
2. TiDB Lightning
3. 200 GB+
4. TiDB Lightning gRPC tikv-importer “engine file”TiDB Lightning SQL dump TiDB tikv-importer
5. tikv-importer TiKV TiKV
6. TiDB Lightning checksum TiDB ANALYZETiDB Lightning AUTO_INCREMENT ID ID
TiDB ID
7. TiDB Lightning TiKV “”normal mode TiDB

Lightning “”

- TiKV SST
- write stall triggers

Lightning “”

2. Lightning



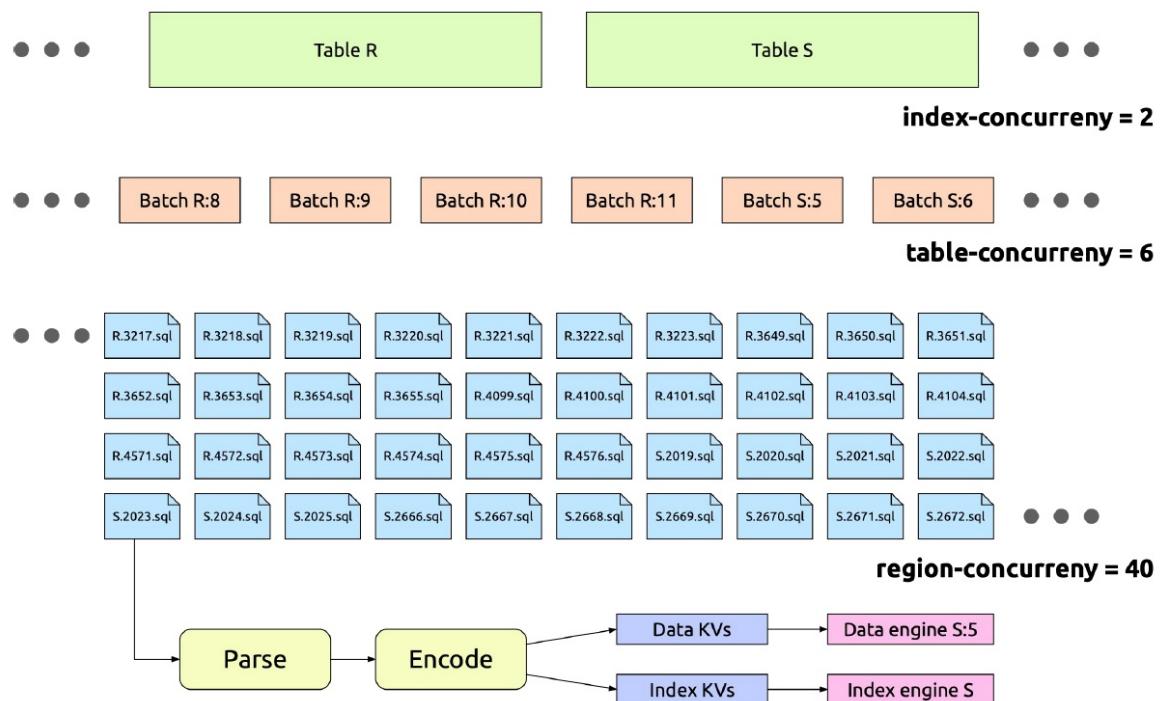
Lightning SQL CREATE TABLE INSERT TiDB

Lightning INSERT

```
INSERT INTO `tbl` VALUES (1, 2, 3), (4, 5, 6), (7, 8, 9);
INSERT INTO `tbl` VALUES (10, 11, 12), (13, 14, 15), (16, 17, 18);
INSERT INTO `tbl` VALUES (19, 20, 21), (22, 23, 24), (25, 26, 27);
```

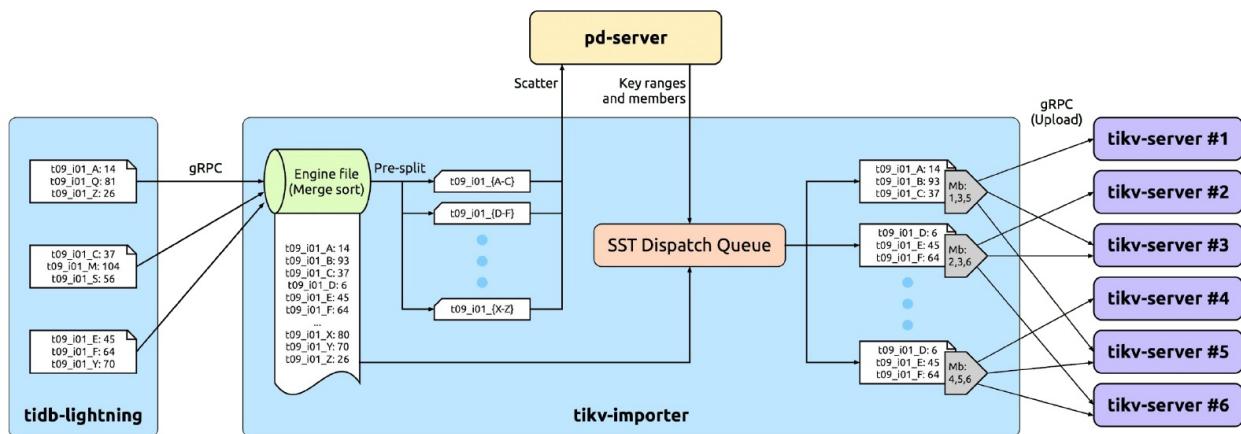
Lightning Lightning TiDB SQL “”KV encoder TiDB KV Lightning INSERT Lightning Importer

Lightning



- batch-size 5 TB+ Importer Lightning batch-size 100 GiB batch-size region balance leader balance region TiKV
 - table-concurrency batch-size
 - index-concurrency table-concurrency + index-concurrency Importer max-open-engines
 - io-concurrency IO cache miss IO
 - block-size Lightning block-size 64 KiB
 - region-concurrency → → Importer
- I/O io-concurrency
- CPU region-concurrency 50 ms 20 block-size 64 KiB 1.28 MB region-concurrency = 60 75 MB/s

3. Importer

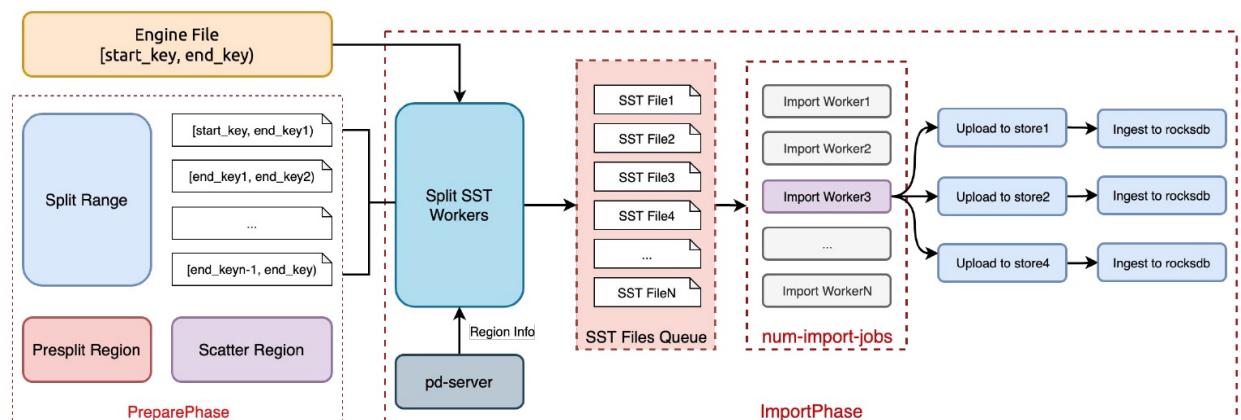


Importer Importer

RocksDB RocksDB RocksDB SST

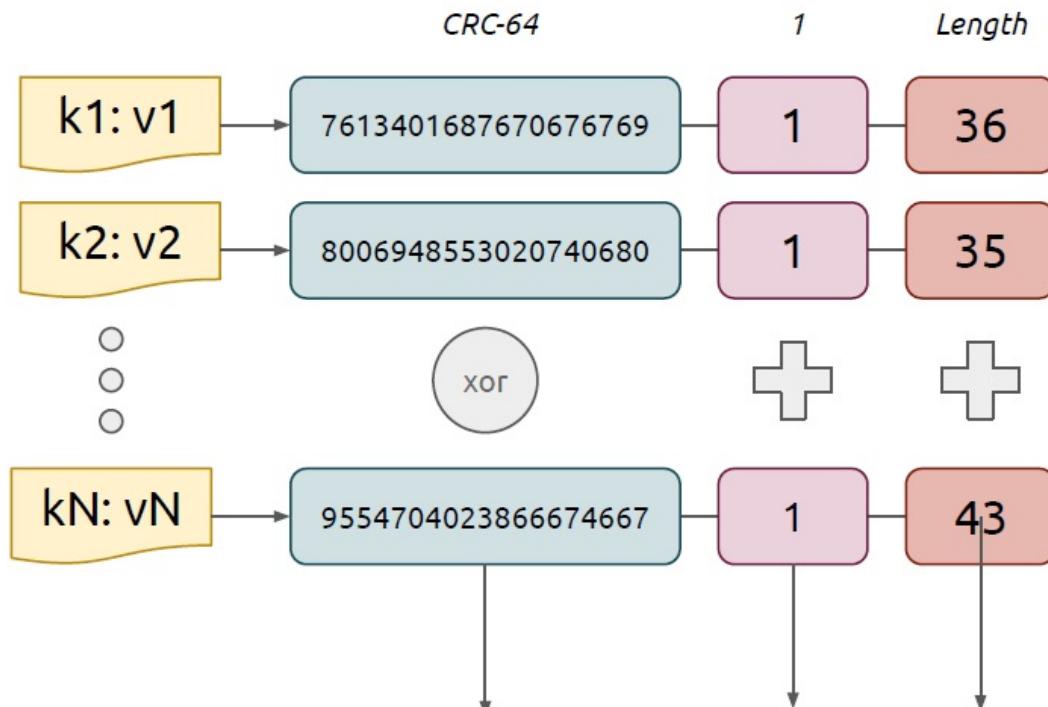
SST TiKV Regions 96 MiB Importer Region PD Region TiKV

Importer SST Region Leader Ingest SST Raft group Region



- max-open-engines Lightning Importer Lightning Lightning index-concurrency + table-concurrency
- Lightning index-concurrency + table-concurrency Lightning batch-size 7 Importer
- num-import-jobs : Lightning batch-size Import TiKV
- region-split-size : 100 GiB TiKV SST SST 96 MiBSST Ingest

4.



```
mysql> admin checksum table `datetime`;
+-----+-----+-----+-----+-----+
| Db_name | Table_name | Checksum_crc64_xor | Total_kvs | Total_bytes |
+-----+-----+-----+-----+-----+
| vt     | datetime   | 4772761481773448246 | 210      | 8336       |
+-----+-----+-----+-----+-----+
1 row in set (0.13 sec)
```

Lightning Region Importer Checksum TiKV Checksum

Checksum Hash TiKV Checksum Lightning Checksum Checksum SST SHA-256

CRC64 XOR 64 Checksum 3

- Lightning encode
- TiDBSQL
 - ADMIN CHECKSUM TABLE xxxx ;

5.

Lightning

```
ANALYZE TABLE `xxxx`;
ALTER TABLE `xxxx` AUTO_INCREMENT=123456;
```

6.

Lightning

-
-

Lightning Lightning Lightning

7.

- Lightning Importer Importer TiKV ~300 MB/s ~30 MB/s
- Lightning
- Importer 20 5 TB Importer TiKV Importer 5 4 TB ~2 TB
- TiKV ($\times 4$) 5 TB SQL 20 TB TiKV Lightning

8.

- io-concurrency cache miss
- Lightning → Importer
 - 10 Gb/s 300 MB/s Importer Lightning

```
max speed = bandwidth (1.2 GB/s) / replicas (3)
```

- Lightning Importer

```
ram usage = (max-open-engines (8) × write-buffer-size (1 GB) × 2)
            + (num-import-jobs (24) × region-split-size (512 MB) × 2)
```

- Importer $\approx N$, $N = \max(index\text{-}concurrency, table\text{-}concurrency)$

2.2.2 Lightning

Lightning

2.2.2.1 TiDB Lightning

- TiDB Lightning TiDB
- `tidb-lightning` “” CPU `tidb-lightning-ctl` “”

```
.../tidb-ansible/resource/bin/tidb-lightning-ctl -switch-mode=normal
```

TiDB Lightning TiDB

| | SELECT | INSERT | UPDATE | DELETE | CREATE | DROP | ALTER |
|--|---------------|---------------|---------------|---------------|-------------------|-------------------|--------------|
| | Tables | Tables | Tables | Tables | Databases, tables | Databases, tables | Tables |

```
checksum = true TiDB Lightning TiDB admin
```

```
tidb-lightning tikv-importer
```

```
tidb-lightning
```

- 32+ CPU
- SSD
- 300 MB/s
- CPU `tidb-server` `region-concurrency` `tidb-lightning` CPU

```
tikv-importer
```

- 32+ CPU
- 40 GB+
- 1 TB+ SSD IOPS ≥8000
- N N = max(index-concurrency, table-concurrency)
- 300 MB/s
- CPUI/O

```
tidb-lightning + tikv-importer
```

TiDB-Ansible TiDB Lightning

TiDB-Ansible sudo

- (1) inventory.ini IP `tidb-lightning` `tikv-importer`

```

...
[importer_server]
# import_dir
IS1 ansible_host=172.16.4.1 deploy_dir=/data/deploy tikv_importer_port=8287 import_dir=/data/import
[lightning_server]
# data_source_dir
LS1 ansible_host=172.16.4.2 deploy_dir=/data/deploy tidb_lightning_pprof_port=8289 data_source_dir=/data/wanted
...

```

(2) data_source_dir mydumper sql csv csv conf/tidb-lightning.yml

```

...
[mydumper]
no-schema: true
[mydumper.csv]
# ASCII
separator = ','
# ASCII
delimiter = """
# CSV
# true
header = true
# CSV NULL
# trueCSV NULL
not-null = false
# `not-null` false CSV NULL
# NULL
null = '\N'
#
backslash-escape = true
#
trim-last-separator = false
...

```

(3) Lightning Importer

```
$ ansible-playbook bootstrap.yml -l IS1LS
```

(4) Lightning Importer

```
$ ansible-playbook deploy.yml -l IS1LS
$ ansible-playbook deploy.yml --tags=lightning
```

(5) Importer Lightning

Importer Lightning

1. Importer Lightning
- 2.
3. Importer *scripts/start_importer.sh* Importer
4. Lightning *scripts/start_lightning.sh*

(6)

- grafana
-

lightning ["the whole procedure completed"] ["tidb lightning exit"]

(7) Importer

Importer *scripts/stop_importer.sh*

2.2.2.2 TiDB Lightning

TiDB Lightning bin tikv-importer.tomltidb-lightning.tomlTiDB Lightning ansible

2.2.2.3 TiDB Lightning TiDB-Backend

Importer-backend TiDB-backend

TiDB Lightning tidb-lightning TiDB Lightning Importer-backend TiDB-backend

- Importer-backendtidb-lightning SQL CSV tikv-importer Ingest TiKV
- TiDB-backendtidb-lightning INSERT TiDB SQL

| | Importer-backend | TiDB-backend |
|------|-------------------------|---------------------|
| | | |
| | (~300 GB/) | (~50 GB/) |
| | | |
| ACID | | |
| | | |

TiDB Lightning TiDB-Backend TiDB

TiDB-backend TiDB Lightning TiDB SQL

- 16 CPU
- SSD
-

TiDB-backend

TiDB-backend tikv-importer TiDB-backend

- tikv-importer
- TiDB-backend

(1) inventory.ini[importer_server]

```
...
[importer_server]
# keep empty
[lightning_server]
# data_source_dir
LS1 ansible_host=172.16.4.2 deploy_dir=/data/deploy tidb_lightning_pprof_port=8289 data_source_dir=/data/wante
...
```

(2) data_source_dir conf/tidb-lightning.yml backend

```
tikv_importer:
backend: "tidb" # <-- "tidb"
```

```
mydumper sql csv csv conf/tidb-lightning.yml
```

```
...
[mydumper]
no-schema: true
[mydumper.csv]
# ASCII
separator = ','
# ASCII
delimiter = """
# CSV
# true
header = true
# CSV NULL
# trueCSV NULL
not-null = false
# `not-null` false CSV NULL
# NULL
null = '\N'
#
backslash-escape = true
#
trim-last-separator = false
...
```

(3) Lightning

```
$ ansible-playbook bootstrap.yml -l LS
```

(4) Lightning

```
$ ansible-playbook deploy.yml -l LS
$ ansible-playbook deploy.yml --tags=lightning
```

(5) Lightning

1. Lightning
- 2.
3. Lightning *scripts/start_lightning.sh*

(6)

gafana

2.2.2.4 TiDB Lightning Web

TiDB Lightning TiDB Lightning Web

TiDB Lightning web

- `tidb-lightning --server-mode`

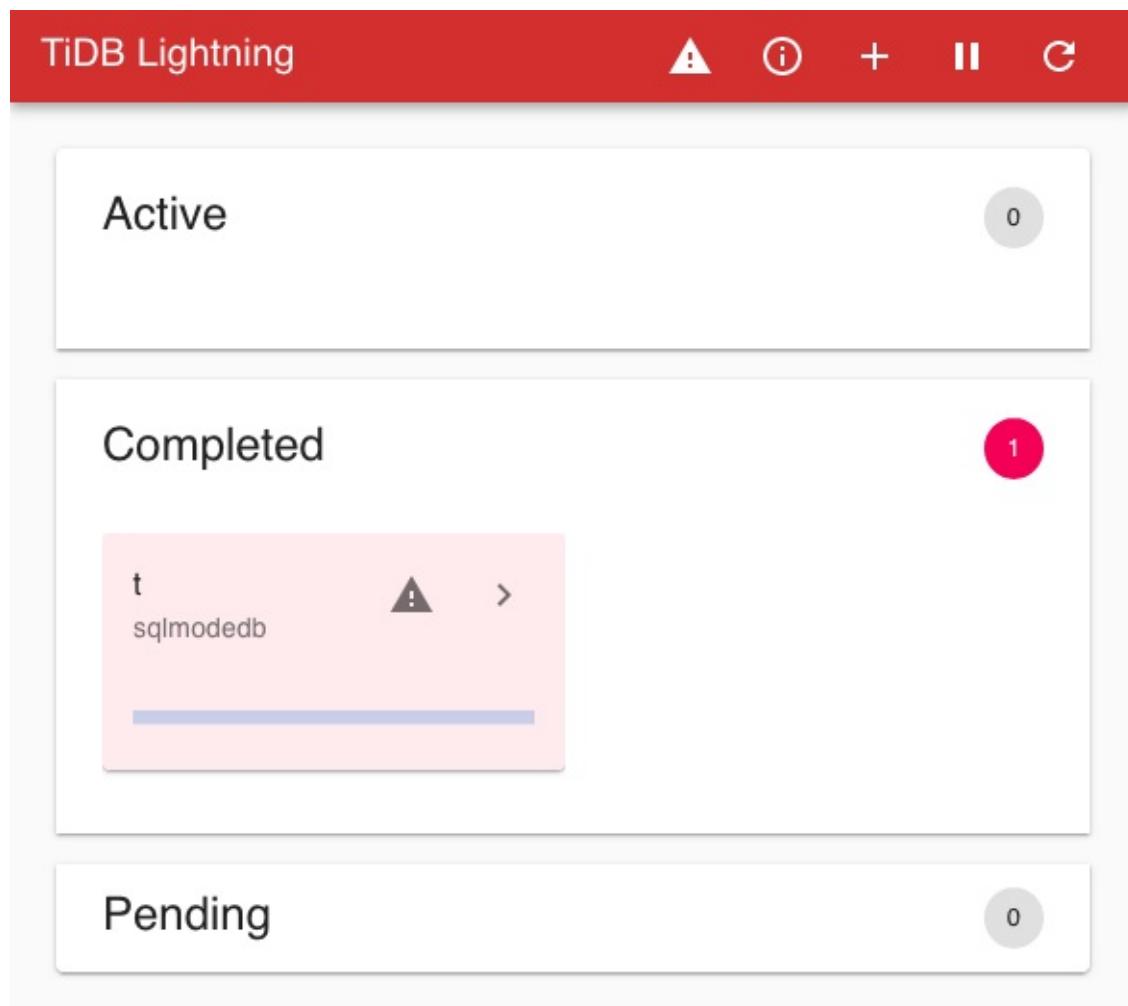
```
./tidb-lightning --server-mode --status-addr :8289
```

- `tidb-lightning.toml lightning.server-mode`

```
[lightning]
server-mode = true
status-addr = ':8289'
```

TiDB Lightning <http://127.0.0.1:8289> URL status-addr

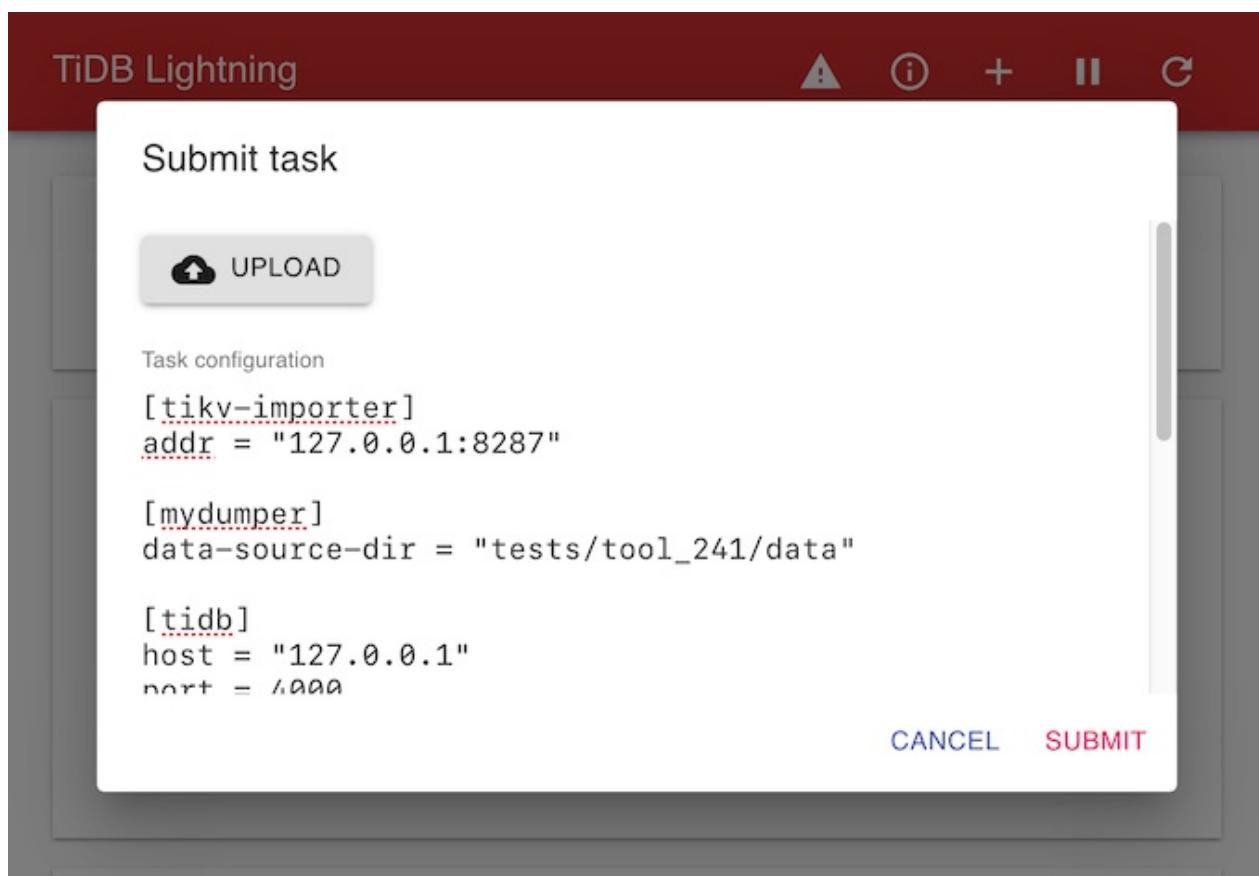
TiDB Lightning Web



| | |
|------------------|---|
| | |
| “TiDB Lightning” | |
| △ | |
| + | |
| / ► | / |
| ⟳ | |

- Active
- Completed
- Pending

+



(task) TOML [TiDB Lightning](#) UPLOAD TOML

SUBMIT

>

TiDB Lightning

`mocker_test`.`tbl_multi_index` •••••••• writing

| Engines | | |
|-----------|---------------|-------|
| Engine ID | Status | Files |
| :1 | ••••• writing | 0 |
| :0 | ••••• writing | 1 |

| Files | | | |
|---|--------|----------------------------------|-----------------------------------|
| Chunk | Engine | Progress | |
| lightning/mydump/examples/mocker_test.tbl_multi_index.sql:0 | :0 | <div style="width: 10%;">—</div> | <div style="width: 100%;">—</div> |

TiDB Lightning

TiDB Lightning (1) + II C

| Current | ... | ... |
|----------------------|-----|--|
| 8/12/2019 2:57:53 AM | ... | {
"id": 1565549873900573000,
"lightning": {
"table-concurrency": 1,
"index-concurrency": 2,
"region-concurrency": 4,
"io-concurrency": 5,
"check-requirements": false
},
"tidb": {
"host": "127.0.0.1",
"port": 4000,
"user": "root",
"status-port": 10080,
"pd-addr": "127.0.0.1:2379",
"sql-mode": "ONLY_FULL_GROUP_BY STRICT_TRANS_TABLES_N
} |
| Queue | ... | |
| 8/12/2019 2:58:28 AM | ... | |

JSON :

2.2.2.5

TiDB Lightning

```
[black-white-list]
do-dbs = ["pattern1", "pattern2", "pattern3"]
ignore-dbs = ["pattern4", "pattern5"]
```

- [black-white-list] do-dbs do-dbs
- do-dbs ignore-dbs
- do-dbs ignore-dbs

```
[[black-white-list.do-tables]]
db-name = "db-pattern-1"
tbl-name = "table-pattern-1"
#
[[black-white-list.do-tables]]
db-name = "db-pattern-2"
tbl-name = "table-pattern-2"
[[black-white-list.ignore-tables]]
db-name = "db-pattern-3"
tbl-name = "table-pattern-3"
[[black-white-list.ignore-tables]]
db-name = "db-pattern-4"
tbl-name = "table-pattern-4"
```

- do-tables do-tables
- do-tables ignore-tables
- do-tables ignore-tables

| | |
|-------------------|---|
| | |
| logs | messages_2016messages_2017messages_2018 |
| forum | messages |
| forum_backup_2016 | messages |
| forum_backup_2017 | messages |
| forum_backup_2018 | messages |
| admin | secrets |

```
[black-white-list]
do-dbs = [
    "forum_backup_2018",   # A
    "~^(logs|forum)$",     # B ~ Go
]
ignore-dbs = [
    "~^forum_backup_",
]                                     # C
[[black-white-list.do-tables]]          # D
db-name = "logs"
tbl-name = "~_2018$"

[[black-white-list.ignore-tables]]      # E
db-name = "~.*"
tbl-name = "~^messages.*"

[[black-white-list.do-tables]]          # F
db-name = "~^forum.*"
tbl-name = "messages"
```

| | | |----:|----:| logs | B | | forum | B | | forum_backup_2016 | C | | forum_backup_2017 | C |
| forum_backup_2018 | A C | | admin | do-dbs |

| | |
|---|-----------|
| logs . messages_2016 logs . messages_2017 | E |
| logs . messages_2018 | D E |
| forum . users | do-tables |
| forum . messages | F E |
| forum_backup_2016 . messages forum_backup_2017 . messages | |
| forum_backup_2018 .`message | F E |
| admin . secrets | |

2.2.2.6

Lightning

```
[checkpoint]
#
# Lightning
enable = true

#
# - file v2.1.1
# - mysql MySQL
driver = "file"

#
# driver = "mysql"
# schema = "tidb_lightning_checkpoint"

#
#/tmp/CHECKPOINT_SCHEMA.pb`
# driver = "mysql"dsn":@tcp(:)/"
dsn = "/tmp/tidb_lightning_checkpoint.pb"
#
# keep-after-success = false
```

TiDB Lightning MySQL

- driver = "file" dsn RAM disk
- driver = "mysql" MySQL 5.7 MariaDB TiDB() MySQL tidb-lightning

- tidb-lightning tidb-lightning-ctl

```
# `schema`.`table` DROP""
# `schema`.`table`
$ tidb-lightning-ctl --checkpoint-error-destroy='`schema`.`table`'
```

- “all”

```
$ tidb-lightning-ctl --checkpoint-error-destroy=all
```

- schema . table “all”

```
$ tidb-lightning-ctl --checkpoint-error-ignore='`schema`.`table`' &&
$ tidb-lightning-ctl --checkpoint-error-ignore=all
```

(CHECKSUM)

- --checkpoint-remove

```
$ tidb-lightning-ctl --checkpoint-remove='`schema`.`table`' &&
$ tidb-lightning-ctl --checkpoint-remove=all
```

- --checkpoint-dump

```
# driver = "mysql"
$ tidb-lightning-ctl --checkpoint-dump=output/directory
```

2.2.2.7 CSV

TiDB Lightning CSV TSV

CSV db_name.table_name.csv db_name table_name

CSV CSV db_name.table_name.003.csv

*.csv

CSV TiDB

- DDL CREATE TABLE db_name.table_name-schema.sql
- TiDB tidb-lightning.toml [mydumper] no-schema = true

CSV tidb-lightning.toml [mydumper.csv] MySQL [LOAD DATA](#)

```
[mydumper.csv]
# ASCII
separator = ','
# ASCII
delimiter = ""
# CSV
# true
header = true
# CSV NULL
# trueCSV NULL
not-null = false
# `not-null` false CSV NULL
# NULL
null = '\N'
#
backslash-escape = true
#
trim-last-separator = false
```

separator

-
- ASCII
- - CSV ','
 - TSV "\t"
- LOAD DATA FIELDS TERMINATED BY

delimiter

-
- delimiter
- - RFC 4180
 - "
- LOAD DATA FIELDS ENCLOSED BY

header

- CSV
- true false

not-null null

- not-null
- not-null false null SQL NULL

CSV

```
A,B,C
\n,"\"\\N",
```

- not-null = false; null = '\N' A and B TiDB NULL C " NULL

backslash-escape

- backslash-escape true


```
| | | |:----|----| |\0 | (U+0000) || \b | (U+0008) || \n | (U+000A) || \r | (U+000D) || \t | (U+0009) || \Z | Windows
EOF (U+001A) |
```
- \"""


```
• LOAD DATA FIELDS ESCAPED BY \"
```

trim-last-separator

separator

CSV

- A,,B,,
- trim-last-separator = false 5 ('A', ', 'B', ', ',)
- trim-last-separator = true 3 ('A', ', 'B')

TiDB Lightning LOAD DATA

- CR\rLF\n CRLF\r\n LINES TERMINATED BY
- LINES STARTING BY
- IGNORE n LINES
- ASCII

CSV

RFC 4180

```
[mydumper.csv]
separator = ','
delimiter = ""
header = true
not-null = false
null = '\N'
backslash-escape = true
trim-last-separator = false
```

```
ID,Region,Count
1,"East",32
2,"South",\N
3,"West",10
4,"Nor
```

TSV

```
[mydumper.csv]
separator = "\t"
delimiter = ''
header = true
not-null = false
null = 'NULL'
backslash-escape = false
trim-last-separator = false
```

```
ID Region Count
1 East 32
2 South NULL
3 West 10
4 North 39
```

TPC-H DBGEN

```
[mydumper.csv]
separator = '|'
delimiter = ''
header = false
not-null = true
backslash-escape = false
trim-last-separator = true
```

```
1|East|32|
2|South|0|
3|West|10|
4|North|39|
```

2.2.2.8 TiDB Lightning

TiDB Lightning TiDB Lightning TiKV Importer

TiDB Lightning “”“” TiDB Lightning

TiDB Lightning

```
### tidb-lightning

[lightning]
# web Prometheus HTTP 0
status-addr = ':8289'

# web
# "TiDB Lightning web"
server-mode = false

#
level = "info"
file = "tidb-lightning.log"
max-size = 128 # MB
max-days = 28
max-backups = 14
```

TiDB Lightning

```
### tidb-lightning

[lightning]
#
# check-requirements = true

#
# """
#
# tikv-importer
# tikv-importer max-open-engines
index-concurrency = 2
table-concurrency = 6

# CPU
# CPU 75% CPU
# region-concurrency =

# I/O I/O
# I/O
#
io-concurrency = 5

[checkpoint]
#
# TiDB Lightning
# Lightning
enable = true
#
schema = "tidb_lightning_checkpoint"
#
# - file
# - mysql MySQL
driver = "file"

# dsn (data source name)
# driver = "file" dsn
#"/tmp/CHECKPOINT_SCHEMA.pb"
# driver = "mysql" dsn ":@tcp(:)/*" URL
# URL [tidb] TiDB
# TiDB MySQL
# dsn = "/tmp/tidb_lightning_checkpoint.pb"
```

```

# false
#
# keep-after-success = false

[tikv-importer]
# "importer" "tidb"
# backend = "importer"
# "importer" tikv-importer
addr = "172.16.31.10:8287"
# "tidb"
# - replace
# - ignore
# - error
# on-duplicate = "replace"

[mydumper]
#
read-block-size = 65536 # Byte ( 64 KB)

#
# Lightning
batch-size = 107_374_182_400 # Byte ( 100 GB)

#
# Lightning
#
# """ 1 GB
# /
# "" 0
# 0 <= batch-import-ratio < 1
batch-import-ratio = 0.75

# mydumper
data-source-dir = "/data/my_database"
# no-schema = false TiDB Lightning TiDB
# `CREATE TABLE`
no-schema = false
# `CREATE TABLE`
# - utf8mb4 UTF-8 Lightning
# - gb18030 GB-18030 Lightning
# - auto UTF-8 GB-18030
# - binary
# **** binary
character-set = "auto"

# CSV
[mydumper.csv]
# ASCII
separator = ','
# ASCII
delimiter = """
# CSV
# header = true
# CSV NULL
# not-null = trueCSV NULL
not-null = false
# not-null = false CSV NULL
# NULL
null = '\N'
# \""
backslash-escape = true
#
trim-last-separator = false

[tidb]
# tidb-server
host = "172.16.31.1"
port = 4000
user = "root"
password = ""
# TiDB "status-port"

```

```
status-port = 10080
# pd-server
pd-addr = "172.16.31.4:2379"
# tidb-lightning TiDB
# TiDB
log-level = "error"

# TiDB Checksum Analyze
# "Analyze"
build-stats-concurrency = 20
distsql-scan-concurrency = 100
index-serial-scan-concurrency = 20
checksum-table-concurrency = 16

# SQL SQL
sql-mode = "STRICT_TRANS_TABLES,NO_ENGINE_SUBSTITUTION"
# `max-allowed-packet`
# `max_allowed_packet` 0
# global `max_allowed_packet`
max-allowed-packet = 67_108_864

# tidb-lightning ChecksumCompact Analyze
# true
# Checksum -> Compact -> Analyze
[post-restore]
# true `ADMIN CHECKSUM TABLE <table>`
#
checksum = true
# true level-1 Compact
# false
level-1-compact = false
# true TiKV full Compact
# false
compact = false
# true `ANALYZE TABLE <table>`
analyze = true

#
# hms
[cron]
# Lightning TiKV
switch-mode = "5m"
#
log-progress = "5m"

# "TiDB Lightning "
# [black-white-list]
# ...
```

TiKV Importer

```

# TiKV Importer

#
log-file = "tikv-importer.log"
# trace, debug, info, warn, error off
log-level = "info"

[server]
# tikv-importer tidb-lightning
addr = "192.168.20.10:8287"
# gRPC
grpc-concurrency = 16

[metric]
# Prometheus job
job = "tikv-importer"
# Prometheus
interval = "15s"
# Prometheus Pushgateway
address = ""

[rocksdb]
# background job
max-background-jobs = 32

[rocksdb.defaultcf]
#
write-buffer-size = "1GB"
#
max-write-buffer-number = 8

#
# 0 KV
# 6 SST
# 1 5
compression-per-level = ["lz4", "no", "no", "no", "no", "no", "no", "lz4"]

[rocksdb.writecf]
#
compression-per-level = ["lz4", "no", "no", "no", "no", "no", "no", "lz4"]

[import]
#
import-dir = "/mnt/ssd/data.import/"
# RPC
num-threads = 16
# job
num-import-jobs = 24
# Region
# max-prepare-duration = "5m"
# Region
#region-split-size = "512MB"
# stream-channel-window
# channel stream
# stream-channel-window = 128
#
max-open-engines = 8
# Importer TiKV /
# upload-speed-limit = "512MB"
# store_available_space/store_capacity
# Importer SST
# PD Regions
min-available-ratio = 0.05

```

tidb-lightning

tidb-lightning

| | | |
|--------------------------------------|--------------------------------------|--------------------------|
| | | |
| <code>-config file</code> | <i>file</i> | |
| <code>-V</code> | | |
| <code>-d directory</code> | | mydumper.data-source-dir |
| <code>-L level</code> | debuginfo warner error fatal (info) | lightning.log-level |
| <code>-backend backend</code> | importer tidb | tikv-importer.backend |
| <code>-log-file file</code> | | lightning.log-file |
| <code>-status-addr ip:port</code> | TiDB Lightning | lightning.status-port |
| <code>-importer host:port</code> | TiKV Importer | tikv-importer.addr |
| <code>-pd-urls host:port</code> | PD endpoint | tidb.pd-addr |
| <code>-tidb-host host</code> | TiDB Server host | tidb.host |
| <code>-tidb-port port</code> | TiDB Server 4000 | tidb.port |
| <code>-tidb-status port</code> | TiDB Server 10080 | tidb.status-port |
| <code>-tidb-user user</code> | TiDB | tidb.user |
| <code>-tidb-password password</code> | TiDB | tidb.password |

```
cfg.toml ./tidb-lightning -L debug --config cfg.toml "debug"
```

tidb-lightning-ctl**tidb-lightning-ctl**

| | |
|--|--------------------------|
| | |
| <code>-compact</code> | full compact |
| <code>-switch-mode mode</code> | TiKV Store normal import |
| <code>-import-engine uuid</code> | TiKV Importer TiKV |
| <code>-cleanup-engine uuid</code> | TiKV Importer |
| <code>-checkpoint-dump folder</code> | CSV |
| <code>-checkpoint-error-destroy tablename</code> | |
| <code>-checkpoint-error-ignore tablename</code> | |
| <code>-checkpoint-remove tablename</code> | |

tablename db . tb1 all

```
tidb-lightning tidb-lightning-ctl
```

tikv-importer**tikv-importer**

| | | |
|---------------------------|-----------------------------|-------------------|
| | | |
| -C, --config <i>file</i> | <i>file</i> | |
| -V, --version | | |
| -A, --addr <i>ip:port</i> | TiKV Importer | server.addr |
| -import-dir <i>dir</i> | | import.import-dir |
| -log-level <i>level</i> | tracedebuginfowarnerror off | log-level |
| -log-file <i>file</i> | | log-file |

2.2.2.9 TiDB Lightning

tidb-lightning tikv-importer [Prometheus](#) (metrics) TiDB Lightning

- TiDB Ansible Lightning inventory.ini [monitored_servers]
- Lightning

tikv-importer

tikv-importer v2.1 [Pushgateway](#) tikv-importer.toml Pushgateway

```
[metric]

# Prometheus
job = "tikv-importer"

# Prometheus
interval = "15s"

# Prometheus Pushgateway
address = ""
```

Prometheus tidb-lightning tidb-lightning.toml

```
[lightning]
# Prometheus HTTP 0
pprof-port = 8289
...
```

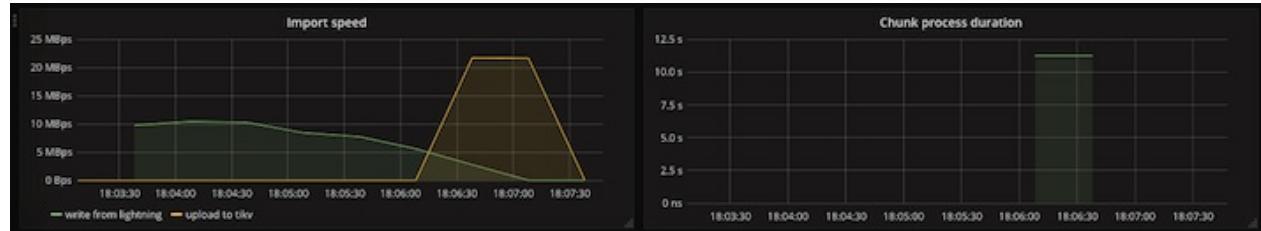
Prometheus Lightning

```
...
scrape_configs:
  - job_name: 'tidb-lightning'
    static_configs:
      - targets: ['192.168.20.10:8289']
```

Grafana

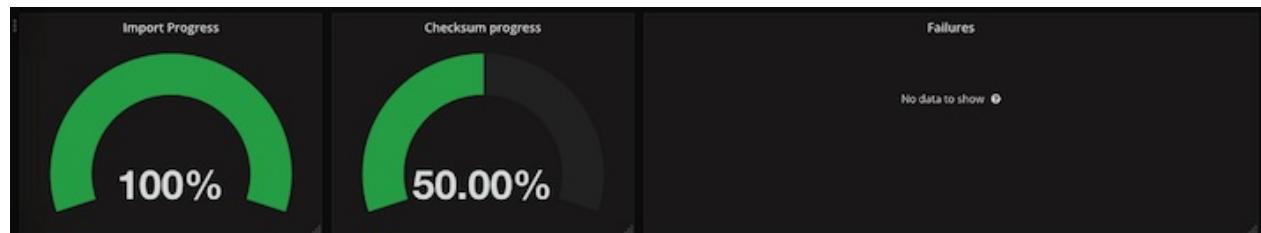
[Grafana](#) Prometheus

TiDB Ansible TiDB Grafana + Prometheus

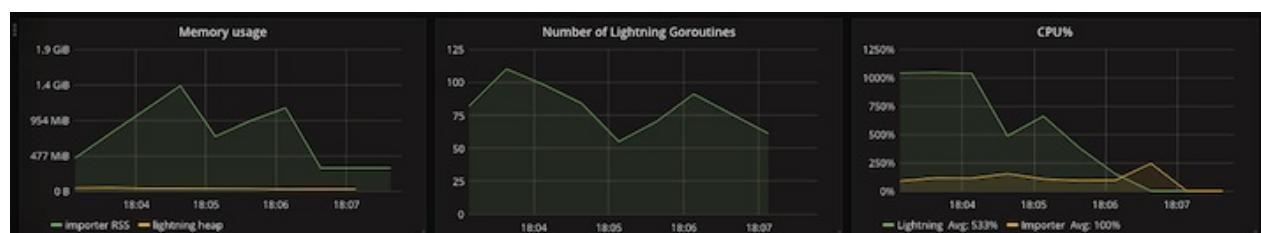
TiDB Lightning [JSON](#)

| | | |
|------------------------|----------------------|------------------------------|
| Import speed | write from lightning | TiDB Lightning TiKV Importer |
| Import speed | upload to tikv | TiKV Importer SST TiKV |
| Chunk process duration | | |

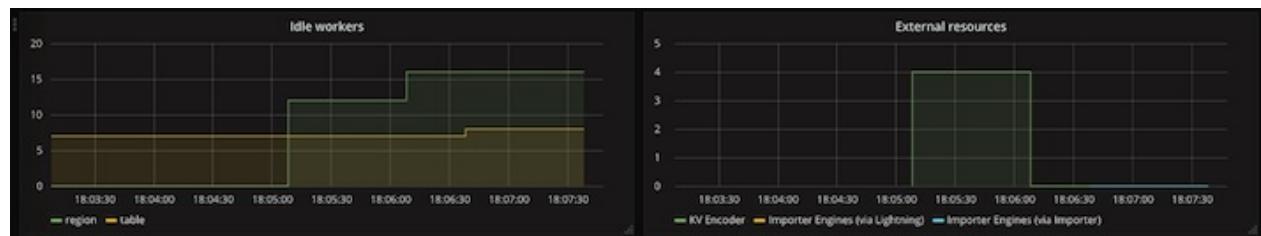
0



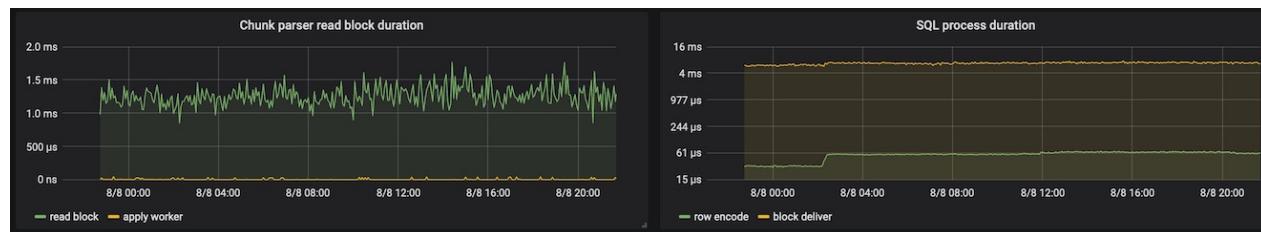
| | |
|-------------------|--|
| Import progress | |
| Checksum progress | |
| Failures | |



| | |
|--------------------------------|---------------------------|
| Memory usage | |
| Number of Lightning Goroutines | TiDB Lightning goroutines |
| CPU% | CPU |



| Idle workers | io | io-concurrency 50 |
|--------------------|------------------|--|
| Idle workers | closed-engine | index-concurrency table-concurrency 80 TiDB Lightning TiKV Importer TiDB Lightning |
| Idle workers | table | table-concurrency 0 |
| Idle workers | index | index-concurrency 0 |
| Idle workers | region | region-concurrency 0 |
| External resources | KV Encoder | KV encoder region-concurrency |
| External resources | Importer Engines | max-open-engines |



| Chunk parser read block duration | read block | |
|----------------------------------|---------------|----------------|
| Chunk parser read block duration | apply worker | io-concurrency |
| SQL process duration | row encode | |
| SQL process duration | block deliver | TiKV Importer |

TiDB Lightning I/O



| | | |
|------------------|--------------------|----------------|
| | | |
| SQL process rate | data deliver rate | TiKV Importer |
| SQL process rate | index deliver rate | TiKV Importer |
| SQL process rate | total deliver rate | |
| Total bytes | parser read size | TiDB Lightning |
| Total bytes | data deliver size | TiKV Importer |
| Total bytes | index deliver size | TiKV Importer |
| Total bytes | storage_size/3 | TiKV 1/33 |



| | | |
|----------------------|----------------|------------|
| | | |
| Delivery duration | Range delivery | range TiKV |
| Delivery duration | SST delivery | SST TiKV |
| SST process duration | Split SST | SST |
| SST process duration | SST upload | SST |
| SST process duration | SST ingest | ingest SST |
| SST process duration | SST size | SST |

tikv-importer tidb-lightning

tikv-importer

tikv-importer tikv *import**

tikv_import_rpc_duration

- RPC
 - request RPC
 - switch_mode — TiKV import/normal
 - open_engine —
 - write_engine —
 - close_engine —
 - import_engine — TiKV
 - cleanup_engine —
 - compact_cluster — TiKV
 - upload — SST
 - ingest — Ingest SST
 - compact — TiKV

- resultRPC
 - ok
 - error

tikv_import_write_chunk_bytes

- Lightning

tikv_import_write_chunk_duration

- tidb-lightning

tikv_import_upload_chunk_bytes

- TiKV SST

tikv_import_range_delivery_duration

- range dispatch-job

tikv_import_split_sst_duration

- range SST

tikv_import_sst_delivery_duration

- SST dispatch-job ImportSSTJob

tikv_import_sst_recv_duration

- ImportSSTJob dispatch-job SST

tikv_import_sst_upload_duration

- ImportSSTJob SST TiKV

tikv_import_sst_chunk_bytes

- TiKV SST

tikv_import_sst_ingest_duration

- SST TiKV

tikv_import_each_phase

- 1 0
 - phaseprepare / import

tikv_import_wait_store_available_count

- TiKV SST
 - store_id TiKV ID

tikv_import_upload_chunk_duration

- TiKV

tidb-lightning

tidb-lightning lightning_*

lightning_importer_engine

- - type:
 - open

- closed

lightning_idle_workers

- worker
 - name
 - table — table-concurrency 0
 - index — index-concurrency 0
 - region — region-concurrency 0
 - io — io-concurrency 5 0
 - closed-engine — index-concurrency table-concurrency 8 0 TiDB Lightning TiKV Importer TiDB Lightning

lightning_kv_encoder

- KV KV TiDB SQL INSERT
 - type:
 - open
 - closed

lightning_tables

- - state
 - pending —
 - written —
 - closed —
 - imported —
 - altered_auto_inc — ID
 - checksum —
 - analyzed —
 - completed —
 - result
 - success —
 - failure —

lightning_engines

- - state
 - pending —
 - written —
 - closed —
 - imported —
 - completed —
 - result
 - success —
 - failure —

lightning_chunks

- Chunks
 - state: Chunk Chunk
 - estimated — Chunk
 - pending —
 - running —
 - finished — Chunk
 - failed —

lightning_import_seconds

-

lightning_row_read_bytes

- SQL

lightning_row_encode_seconds

- SQL

lightning_row_kv_deliver_seconds

- SQL

lightning_block_deliver_seconds

- tikv-importer

lightning_block_deliver_bytes

- Importer

lightning_chunk_parser_read_block_seconds

- SQL

lightning_checksum_seconds

- Checksum

lightning_apply_worker_seconds

- worker (lightning_idle_workers)

- name
 - table
 - index
 - region
 - io
 - closed-engine

2.3 4.0 BR

BR(<https://github.com/pingcap/br>) BR TiKV SQL dump BR

2.3.1 BR

2.3.1.1

BR TiKV

```
* BR TiKV
* TiKV region leader iterator
* TiKV SST IO
* TiKV SST https3
* TiKV BR
```

BR

```
* TiKV
* TiKV Streaming
*
*
* PD leader
* schema
```

TiKV TiDB/TiKV percolator 3 CF defaultlockwrite TiKV SST TiKV write default

TiKV ts backup_ts write default key-value

TiKV

(backup_ts, current_ts]

SI write CF

- Put
- Delete

- Lockselect for update
- Rollback

1. region leader
2. SST SST rocksdb /
3. S3HTTP
4. SI point-in-time

2.3.1.1

1. KV Scan
2. s3 TiKV SST IO
3. TiDB

- RegionError region split/mergenot leader
- KeyLocked
- Server is busy TiKV

1. GC GC GCBR BR GC GC 10

2.3.1.2

- database table
- table SST Split & Scatter Region
- SST TiKV
- table ID SST Key Rewrite
- SST Ingest TiKV

Key RewriteSplit & Scatter Region, Ingest SST

1. Key Rewrite TiDB TiKV

| | |
|---|---|
| Key: tablePrefix{tableID}_recordPrefixSep{rowID} Value: [col1, col2, col3, col4] | Key: tablePrefix{tableID}_indexPrefixSep{indexID}_indexedColumnsValue Value: rowID |
|---|---|

Key tableID SST TiKV tableID

SST key tableID tableID indexID

1. Split & Scatter TiKV Region 96MBSplit Region Region Region Scatter

SST Region SST Region Region

1. Ingest SST Ingest SST sst_importer SST Raft Ingest

| |
|------------------------------------|
| * region leader leader region TiKV |
| * Key Rewrite IO |

BR

2.3.2 BR

1.

BR Ansible

1. ansible v4.0.0-beta.1
2. pip install -r ./requirements.txt
3.
 - i. 6IP 101-106
 - ii. 101-106 TiKV 101-103 PD 104-106 TiDB

:

br tidb-toolkit-v4.0.0-beta.1.tar.gz ansible tidb-ansible downloads

2.

TiDB

```
MySQL [(none)]> create database br_test;
Query OK, 0 rows affected (0.11 sec)

MySQL [(none)]> use br_test;
Database changed

MySQL [br_test]> create table br_table(id int primary key,c varchar(128),ctime timestamp);
Query OK, 0 rows affected (0.12 sec)
```

python

```
import mysql.connector
import time
mydb = mysql.connector.connect(
    host="xxxx.104", # tidb-server ip
    user='root',
    port=4000,
    database='br_test'
)
mycursor = mydb.cursor()

for i in range (100000):
    mycursor.execute('insert into br_table values(%s,%s,now())',(i,str(i)+'xxxx'))
    if i%100==0:
        mycursor.execute('commit')

mycursor.execute('commit')
mycursor.close()
mydb.close()
```

TiDB

```
MySQL [br_test]> select count(1) from br_table;
+-----+
| count(1) |
+-----+
| 100000 |
+-----+
1 row in set (0.04 sec)
```

6 TiKV

3.

1. tikv_gc_life_time

```
# gc
#
SELECT VARIABLE_VALUE FROM mysql.tidb WHERE VARIABLE_NAME = 'tikv_gc_life_time';
10m0s
# 720h
UPDATE mysql.tidb SET VARIABLE_VALUE = '720h' WHERE VARIABLE_NAME = 'tikv_gc_life_time';
#
SELECT * FROM mysql.tidb WHERE VARIABLE_NAME = 'tikv_gc_life_time';
720h
```

1.

SMB/NFS

| S3GCS

TiKV, BR TiDBPD

NFS

```
mount -t nfs //nfs_address/:/data /data_nfs1
```

4.

BR

gRPC TiKV BR PD

5.

```
--ca--cert--key TLS
--concurrency 4
--log-file,--log-level
-u, --pd PD 127.0.0.1:2379
--ratelimit MB
-s, --storage "local:///data_nfs1"
```

6.

```
bin/br backup full --pd "192.168.122.101:2379" --storage "local:///data_nfs1/backup"
```

TiKV /data_nfs1/backup

```
* PD TiKV
* infoSchema
* {"cluster_id":6801677637806839235,"start_key":"dIAAAAAAAAAAvX3IAAAAAAAAAAA=","end_key":"dIAAAAAAAAAAvX3L//////////wA",
=","end_version":415142848617512967,"concurrency":4,"storage_backend":{"Backend":{"Local":{"path":"/data_nfs1/backup"}}
}}"
* TiKV BR
* checksum [table=`br_test`.`br_table`] [Crc64Xor=12896770389982935753] [TotalKvs=100000] [TotalBytes=4788890]
*
*
```

protobuf

5_2_23_80992061af3e5194c3f28a5b79d486c5e9db2feda1afb3f84b4ca229ddce9932_write.sst

7.

```
MySQL [br_test]> drop table br_table;
```

TiKV

```
bin/br restore full --pd "192.168.122.101:2379" --storage "local:///data_nfs1/backup"
```

1. PD
2. DDLCREATE DATABASE / !32312 IF NOT EXISTS/ CREATE TABLE IF NOT EXISTS
3. alter auto incrementID , id
4. sst Region
5. ["Full restore summary: total restore tables: 1, total success: 1, total failed: 0, total take(s): 0.25, total size(MB): 2.28, avg speed(MB/s): 9.08, total kv: 50001"] ["restore files"=1] ["restore ranges"=1] ["split region"=6.373065871s] ["restore checksum"=45.843202ms]

```
MySQL [br_test]> select count(1) from br_table;
+-----+
| count(1) |
+-----+
| 100000|
+-----+
1 row in set (0.12 sec)
```

8.

```
bin/br backup db --db "br_test" --pd "192.168.122.101:2379" --storage "local:///data_nfs1/backup"
```

```
bin/br restore db --db "br_test" --pd "192.168.122.101:2379" --storage "local:///data_nfs1/backup"
```

9.

```
bin/br backup table --db "br_test" --table "br_table" --pd "192.168.122.101:2379" --storage "local:///data_nfs1/backup"
```

```
bin/br restore table --db "br_test" --table "br_table" --pd "192.168.122.101:2379" --storage "local:///data_nfs1/backup"
```

BR BR (<https://github.com/pingcap/br>),

2.4.1 Dumpling

1.

Dumpling Dumpling PingCAP Mydumper fork TiDB TiDB Mydumper Mydumper

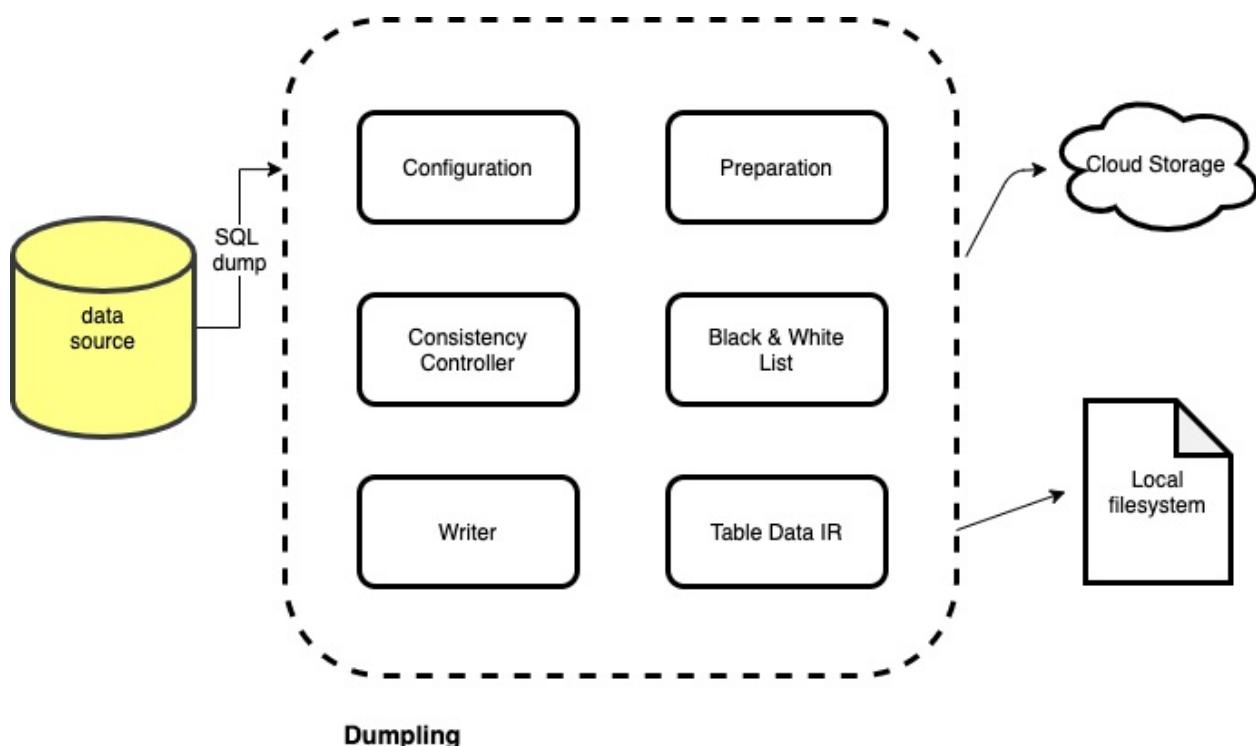
- TiDB Lighting
- Mydumper TiDB fork
- Mydumper C GLib Go TiDB DM
- Mydumper TiDB

Dumpling Dumpling

- Golang TiDB
- Mydumper MySQL
- SQLCSV
- S3

2.

Dumpling



Dumpling Black & White

-
-
-
-
-
-
- Black & White

-
-
-
- - API

2.4.2 Dumpling

1.

- SELECT
- RELOAD
- LOCK TABLES
- REPLICATION CLIENT

TiDB : <https://pingcap.com/docs-cn/stable/reference/security/privilege-system/#%E6%9D%83%E9%99%90%E7%AE%A1%E7%90%86>

2.

```
dumpling -B tidb -F 2048 -H 127.0.0.1 -u root -P 4000 --loglevel debug
```

```
Release version:  
Git commit hash: a35708fb6a9ca19294b92598b88d7894cf130ca6  
Git branch: master  
Build timestamp: 2020-03-08 03:15:54Z  
Go version: go version go1.13.1 darwin/amd64  
  
[2020/03/08 18:09:43.780 +08:00] [DEBUG] [config.go:72] ["parse server info"] ["server info string"=5.7.25-TiDB-v4.0.0-beta-313-g2d5d2fde27]  
[2020/03/08 18:09:43.780 +08:00] [INFO] [config.go:85] ["detect server type"] [type=TiDB]  
[2020/03/08 18:09:43.780 +08:00] [INFO] [config.go:103] ["detect server version"] [version=4.0.0-beta-313-g2d5d2fde27]  
[2020/03/08 18:09:43.781 +08:00] [DEBUG] [prepare.go:27] ["list all the tables"]  
[2020/03/08 18:09:43.788 +08:00] [DEBUG] [black_white_list.go:78] ["filter tables"]  
[2020/03/08 18:09:43.788 +08:00] [WARN] [black_white_list.go:70] ["unsupported dump schema in TiDB now"] [schema=mysql]
```

3.

Dumpling :

| | |
|-------------------------|--|
| | |
| --consistency level | : auto, none, flush, lock, snapshot, auto |
| -B, --database database | |
| -F, --filesize size | , Bytes |
| -H, --host hostname | , 127.0.0.1 |
| --loglevel level | : debug, info, warn, error, dpanic, panic, fatal, info |
| -W, --no-views | , true |
| -o, --output dir | , ./export-2020-03-08T11:37:05+08:00 |
| -p, --password password | |
| -P, --port port | , 4000 |
| --snapshot position | , snapshot |
| -t, --threads num | , 4 |
| -u, --user user | , root |

1 SQL

SQL SQL TiDB SQL Join SQL Hint SQL SQL Hint SQL Plan Management TiDB
SQL Batch SQL

SQL

1.1

[SQL](#) [SQL](#)

1.1.1 EXPLAIN

TiDB [EXPLAIN](#) [SQL](#)

TiDB EXPLAIN 5 id estRows task access object operator info 5 EXPLAIN

| | |
|---------------|---|
| | |
| id | ID TiDB 2.1 ID |
| estRows | 4.0 count |
| task | task task root task tidb-server cop task TiKV TiFlash task root task cop taskroot task
cop task cop task TiDB TiKV TiFlash cop task TiKV TiFlash |
| access object | table partition index |
| operator info | operator info |

1.1.2 EXPLAIN ANALYZE

[EXPLAIN](#) [EXPLAIN ANALYZE](#) [SQL](#) [EXPLAIN](#) [EXPLAIN ANALYZE](#) actRows , execution info , memory , disk

| | |
|----------------|------------------------|
| | |
| actRows | |
| execution info | time (loops)loops rows |
| memory | |
| disk | |

```
mysql> explain analyze select * from t where a < 10;
+-----+-----+-----+-----+-----+
| id | estRows | actRows | task | access object | execution info |
| disk | | | | operator info | memory |
+-----+-----+-----+-----+-----+
-----+
| IndexLookUp_10 | 9.00 | 9 | root | | time:641.245µs, loops:2,
| rpc num: 1, rpc time:242.648µs, proc keys:0 | 9.23046875 KB |
| N/A |
| |-IndexRangeScan_8(Build) | 9.00 | 9 | cop[tikv] | table:t, index:idx_a(a) | time:142.94µs, loops:10,
| | range:[-inf,10), keep order:false | N/A |
| N/A |
| \-TableRowIDScan_9(Probe) | 9.00 | 9 | cop[tikv] | table:t | time:141.128µs, loops:10
| | keep order:false | N/A |
| N/A |
+-----+-----+-----+-----+-----+
-----+
3 rows in set (0.00 sec)
```

estRows actRows IndexLookUp_10 9 KB SQL

1.1.3

TiDB SQL CPU

Build Probe Build Probe

ID Build Probe TiDB Build Probe

```
TiDB(root@127.0.0.1:test) > explain select * from t use index(idx_a) where a = 1;
+-----+-----+-----+-----+
| id | estRows | task | access object | operator info |
| | | | operator info | |
+-----+-----+-----+-----+
-----+
| IndexLookUp_7 | 10.00 | root | | |
| |-IndexRangeScan_5(Build) | 10.00 | cop[tikv] | table:t, index:idx_a(a) | range:[1,1], keep order:false, stat
| | s:pseudo |
| \-TableRowIDScan_6(Probe) | 10.00 | cop[tikv] | table:t | keep order:false, stats:pseudo |
| |
+-----+-----+-----+-----+
-----+
3 rows in set (0.00 sec)
```

IndexLookUp_7 IndexRangeScan_5(Build) TableRowIDScan_6(Probe) IndexRangeScan_5(Build) RowID
 TableRowIDScan_6(Probe) RowID

```
TiDB(root@127.0.0.1:test) > explain select * from t t1 use index(idx_a) join t t2 use index() where t1.a = t2.a;
+-----+-----+-----+-----+
| id | estRows | task | access object | operator info |
+-----+-----+-----+-----+
| HashJoin_22 | 12487.50 | root | | inner join, inner:TableReader_
26, equal:[eq(test.t.a, test.t.a)] |
| └─TableReader_26(Build) | 9990.00 | root | | data:Selection_25
|   |
|   └─Selection_25 | 9990.00 | cop[tikv] | | not(isnull(test.t.a))
|   |
|   └─TableFullScan_24 | 10000.00 | cop[tikv] | table:t2 | keep order:false, stats:pseudo
|   |
|   └─IndexLookUp_29(Probe) | 9990.00 | root | | keep order:false, stats:pseudo
|   |
|   └─IndexFullScan_27(Build) | 9990.00 | cop[tikv] | table:t1, index:idx_a(a) | keep order:false, stats:pseudo
|   |
|   └─TableRowIDScan_28(Probe) | 9990.00 | cop[tikv] | table:t1 | keep order:false, stats:pseudo
|   |
+-----+-----+-----+-----+
7 rows in set (0.00 sec)
```

HashJoin_22 TableReader_26(Build) IndexLookUp_29(Probe) IndexLookUp_29(Probe) IndexFullScan_27(Build)
 TableRowIDScan_28(Probe) TableRowIDScan_28(Probe)

1.1.4

TiKV Block Cache

- **TableFullScan** “”
- **TableRangeScan**
- **TableRowIDScan** RowID
- **IndexFullScan** “”
- **IndexRangeScan**

TiDB TiKV/TiFlash “”

- **TableReader** TiKV TableFullScan TableRangeScan
- **IndexReader** TiKV IndexFullScan IndexRangeScan
- **IndexLookUp** Build TiKV RowID Probe RowID TiKV Build IndexFullScan IndexRangeScan Probe
 TableRowIDScan
- **IndexMerge** IndexLookupReader Build Probe Build TiKV RowID Probe RowID TiKV Build
 IndexFullScan IndexRangeScan Probe TableRowIDScan

IndexLookUp

```
mysql> explain select * from t use index(idx_a);
+-----+-----+-----+-----+
| id | estRows | task | access object | operator info |
+-----+-----+-----+-----+
IndexLookUp_6	10000.00	root		keep order:false, stats:pseudo
└─IndexFullScan_4(Build)	10000.00	cop[tikv]	table:t, index:idx_a(a)	keep order:false, stats:pseudo
└─TableRowIDScan_5(Probe)	10000.00	cop[tikv]	table:t	keep order:false, stats:pseudo
+-----+-----+-----+-----+
3 rows in set (0.00 sec)
```

IndexLookUp_6 IndexFullScan_4(Build) TableRowIDScan_5(Probe) IndexFullScan_4(Build) a RowID
 TableRowIDScan_5(Probe) RowID TableReader IndexLookUp

TableReader

```
mysql> explain select * from t where a > 1 or b >100;
+-----+-----+-----+-----+
| id | estRows | task | access object | operator info |
+-----+-----+-----+-----+
TableReader_7	8000.00	root		data:Selection_6
└Selection_6	8000.00	cop[tikv]		or(gt(test.t.a, 1), gt(test.t.b, 100))
└TableFullScan_5	10000.00	cop[tikv]	table:t	keep order:false, stats:pseudo
+-----+-----+-----+-----+
3 rows in set (0.00 sec)
```

TableReader_7 Selection_6 Cop Task TiKV Cop Task TableFullScan_5 Selection SQL SQL
WHERE / HAVING / ON TableFullScan_5 IndexMerge

IndexMerge

TIDB Index Merge 4.0 RC 4.0 Index Merge or and Index Merge session global set
@@tidb_enable_index_merge = 1;

```
mysql> set @@tidb_enable_index_merge = 1;
mysql> explain select * from t use index(idx_a, idx_b) where a > 1 or b > 1;
+-----+-----+-----+-----+
| id | estRows | task | access object | operator info |
+-----+-----+-----+-----+
IndexMerge_16	6666.67	root		
└─IndexRangeScan_13(Build)	3333.33	cop[tikv]	table:t, index:idx_a(a)	range:(1,+inf], keep order:false, stats:pseudo
└─IndexRangeScan_14(Build)	3333.33	cop[tikv]	table:t, index:idx_b(b)	range:(1,+inf], keep order:false, stats:pseudo
└─TableRowIDScan_15(Probe)	6666.67	cop[tikv]	table:t	keep order:false, stats:pseudo
+-----+-----+-----+-----+
4 rows in set (0.00 sec)
```

IndexMerge IndexMerge_16 IndexRangeScan_13 IndexRangeScan_14 RowID TableRowIDScan_15 RowID

1.1.5

Hash Aggregate

TiDB Hash Aggregation Hash Aggregate

```
TiDB(root@127.0.0.1:test) > explain select /*+ HASH_AGG() */ count(*) from t;
+-----+-----+-----+-----+
| id | estRows | task | access object | operator info |
+-----+-----+-----+-----+
HashAgg_11	1.00	root		funcs:count(Column#7)->Column#4
└TableReader_12	1.00	root		data:HashAgg_5
└HashAgg_5	1.00	cop[tikv]		funcs:count(1)->Column#7
└TableFullScan_8	10000.00	cop[tikv]	table:t	keep order:false, stats:pseudo
+-----+-----+-----+-----+
4 rows in set (0.00 sec)
```

TiDB Hash Aggregate TiKV/TiFlash Coprocessor TiDB coprocessor Task

Stream Aggregate

| TiDB | Stream Aggregation | Hash Aggregate | Hash Aggregate | Stream Aggregate | Stream Aggregate |
|---|---|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| TiDB(root@127.0.0.1:test) > explain select /*+ STREAM_AGG() */ count(*) from t; | | | | | |
| +-----+-----+-----+-----+-----+ | id estRows task access object operator info | +-----+-----+-----+-----+-----+ | +-----+-----+-----+-----+-----+ | +-----+-----+-----+-----+-----+ | +-----+-----+-----+-----+-----+ |
| StreamAgg_16 | 1.00 | root | | funcs:count(Column#7)->Column#4 | |
| └TableReader_17 | 1.00 | root | | data:StreamAgg_8 | |
| └StreamAgg_8 | 1.00 | cop[tikv] | | funcs:count(1)->Column#7 | |
| └TableFullScan_13 | 10000.00 | cop[tikv] | table:t | keep order:false, stats:pseudo | |
| +-----+-----+-----+-----+-----+ | | | | | |
| 4 rows in set (0.00 sec) | | | | | |

| Hash Aggregate | TiDB | Stream Aggregate | TiKV/TiFlash | Coprocessor | TiDB | coprocessor | Task |
|----------------|------|------------------|--------------|-------------|------|-------------|------|
|----------------|------|------------------|--------------|-------------|------|-------------|------|

1.1.6 Join

TiDB Join

- Hash Join
- Merge Join
- Index Hash Join
- Index Merge Join
- Apply

Join

Hash Join

| TiDB | Hash Join | Hash Join |
|------|-----------|-----------|
|------|-----------|-----------|

| | | | | | |
|--|---|---------------------------------|---------------------------------|---|---------------------------------|
| mysql> explain select /*+ HASH_JOIN(t1, t2) */ * from t t1 join t2 on t1.a = t2.a; | | | | | |
| +-----+-----+-----+-----+-----+ | id estRows task access object operator info | +-----+-----+-----+-----+-----+ | +-----+-----+-----+-----+-----+ | +-----+-----+-----+-----+-----+ | +-----+-----+-----+-----+-----+ |
| HashJoin_33 | 10000.00 | root | | inner join, inner:TableReader_43, equal:[eq(test.t.a, test.t2.a)] | |
| └TableReader_43(Build) | 10000.00 | root | | data:Selection_42 | |
| └Selection_42 | 10000.00 | cop[tikv] | | not(isnull(test.t2.a)) | |
| └TableFullScan_41 | 10000.00 | cop[tikv] | table:t2 | keep order:false | |
| └TableReader_37(Probe) | 10000.00 | root | | data:Selection_36 | |
| └Selection_36 | 10000.00 | cop[tikv] | | not(isnull(test.t.a)) | |
| └TableFullScan_35 | 10000.00 | cop[tikv] | table:t1 | keep order:false | |
| +-----+-----+-----+-----+-----+ | | | | | |
| 7 rows in set (0.00 sec) | | | | | |

| Hash Join | Build | Hash Table | Probe | ProbeBuild | Hash Table |
|-----------|-------|------------|-------|------------|------------|
|-----------|-------|------------|-------|------------|------------|

Merge Join

| TiDB | Merge Join | Hash Join | Merge Join |
|------|------------|-----------|------------|
|------|------------|-----------|------------|

```
mysql> explain select /*+ SM_JOIN(t1) */ * from t t1 join t t2 on t1.a = t2.a;
+-----+-----+-----+-----+
| id | estRows | task | access object | operator info
+-----+-----+-----+-----+
MergeJoin_6	10000.00	root		inner join, left key:test.t.a, right key:test.t.a			
	-IndexLookUp_13(Build)	10000.00	root				
		-IndexFullScan_11(Build)	10000.00	cop[tikv]	table:t2, index:idx_a(a)	keep order:true	
		TableRowIDScan_12(Probe)	10000.00	cop[tikv]	table:t2	keep order:false	
		IndexLookUp_10(Probe)	10000.00	root			
			-IndexFullScan_8(Build)	10000.00	cop[tikv]	table:t1, index:idx_a(a)	keep order:true
			TableRowIDScan_9(Probe)	10000.00	cop[tikv]	table:t1	keep order:false
+-----+-----+-----+-----+
7 rows in set (0.00 sec)
```

Merge Join Build Join Group Probe Build Join Group “” Join Group Join Key

Index Hash Join

INL_HASH_JOIN(t1_name [, tl_name]) Index Nested Loop Hash Join Index Nested Loop Join

```
mysql> explain select /*+ INL_HASH_JOIN(t1) */ * from t t1 join t t2 on t1.a = t2.a;
+-----+-----+-----+-----+
| id | estRows | task | access object | operator info
+-----+-----+-----+-----+
IndexHashJoin_32	10000.00	root		inner join, inner:IndexLookUp_23, outer key:test.t.a, inner key:test.t.a						
	-TableReader_35(Build)	10000.00	root		data:Selection_34					
		Selection_34	10000.00	cop[tikv]		not(isnull(test.t.a))				
			TableFullScan_33	10000.00	cop[tikv]	table:t2	keep order:false			
			IndexLookUp_23(Probe)	1.00	root					
				Selection_22(Build)	1.00	cop[tikv]		not(isnull(test.t.a))		
					IndexRangeScan_20	1.00	cop[tikv]	table:t1, index:idx_a(a)	range: decided by [eq(test.t.a , test.t.a)], keep order:false	
						TableRowIDScan_21(Probe)	1.00	cop[tikv]	table:t1	keep order:false
+-----+-----+-----+-----+
8 rows in set (0.00 sec)
```

Index Merge Join

INL_MERGE_JOIN(t1_name [, tl_name]) Index Nested Loop Merge Join INL_JOIN INL_JOIN join keys index index join keys

```
mysql> explain select /*+ INL_MERGE_JOIN(t1) */ * from t t1 where t1.a in ( select t2.a from t2 where t2.b < t1.b);
+-----+-----+-----+-----+
| id      | estRows | task      | access object | operator info
|         |          |          |
+-----+-----+-----+-----+
| HashJoin_26        | 8000.00 | root      |          | semi join, inner:TableReader_49, equal:[eq(test.t.a, test.t2.a)], other cond:lt(test.t2.b, test.t.b)
|   | TableReader_49(Build) | 10000.00 | root      |          | data:Selection_48
|   |   | Selection_48       | 10000.00 | cop[tikv] |          | not(isnull(test.t2.a)), not(isnull(test.t2.b))
|   |   | TableFullScan_47    | 10000.00 | cop[tikv] | table:t2  | keep order:false
|   |   |   |                   |
|   |   | TableReader_38(Probe)| 10000.00 | root      |          | data:Selection_37
|   |   |   |                   |
|   |   | Selection_37        | 10000.00 | cop[tikv] |          | not(isnull(test.t.a)), not(isnull(test.t.b))
|   |   | TableFullScan_36    | 10000.00 | cop[tikv] | table:t1  | keep order:false
|   |   |   |                   |
+-----+-----+-----+-----+
7 rows in set, 1 warning (0.01 sec)
```

Apply

```
mysql> explain select /*+ INL_MERGE_JOIN(t1) */ * from t t1 where t1.a in ( select avg(t2.a) from t2 where t2.b < t1.b);
+-----+-----+-----+-----+
| id      | estRows | task      | access object | operator info
|         |          |          |
+-----+-----+-----+-----+
| Projection_10        | 10000.00 | root      |          | test.t.id, test.t.a, test.t.b
|   |   |           |
|   |   | Apply_12      | 10000.00 | root      |          | semi join, inner:StreamAgg_30, equal:[eq(Column#8, Column#7)]
|   |   |   |           |
|   |   |   | Projection_13(Build) | 10000.00 | root      |          | test.t.id, test.t.a, test.t.b, cast(test.t.a, decimal(20,0) BINARY)->Column#8
|   |   |   |   | TableReader_15      | 10000.00 | root      |          | data:TableFullScan_14
|   |   |   |   |   |           |
|   |   |   |   |   | TableFullScan_14    | 10000.00 | cop[tikv] | table:t1  | keep order:false
|   |   |   |   |   |           |
|   |   |   |   | StreamAgg_30(Probe) | 1.00     | root      |          | funcs:avg(Column#12, Column#13)->Column#7
|   |   |   |   |   |           |
|   |   |   |   | TableReader_31      | 1.00     | root      |          | data:StreamAgg_19
|   |   |   |   |   |           |
|   |   |   |   | StreamAgg_19        | 1.00     | cop[tikv] |          | funcs:count(test.t2.a)->Column#12, funcs:sum(test.t2.a)->Column#13
|   |   |   |   | Selection_29       | 8000.00  | cop[tikv] |          | lt(test.t2.b, test.t.b)
|   |   |   |   |   |           |
|   |   |   |   | TableFullScan_28    | 10000.00 | cop[tikv] | table:t2  | keep order:false
|   |   |   |   |   |           |
+-----+-----+-----+-----+
10 rows in set, 1 warning (0.00 sec)
```

1.1.7 EXPLAIN FOR CONNECTION

`EXPLAIN FOR CONNECTION` `EXPLAIN TiDB MySQL`

- MySQL TiDB
- MySQL MySQL `PROCESS TiDB` `SUPER`

1.2

TiDB System R SQL Join

1.2.1

1

TiDB

- DataSource `select * from t t`
- Selection `select * from t where a = 5 where a = 5`
- Projection `select c, a + b from t c a + b`
- Join `select t1.b, t2.c from t1 join t2 on t1.a = t2.a t1 join t2 on t1.a = t2.a t1 t2 Join`

SelectionProjectionJoin SPJ 3

2

| 1 | 2 |
|------------|----------------|
| 1 | 6 |
| 2 | 7 |
| 3 | 8 |
| 4Max / Min | 9 |
| 5 | 10TopN / Limit |

1

- 1LogicalJoin LogicalJoin outer plan
- 2
 - 2.1LogicalJoin join key inner plan
 - 2.2LogicalJoin

1 2 2.1 2.2

1 2.1

```
select t1.a from t1 left join t2 on t1.b = t2.b;
```

```
select t1.a from t1;
```

2 Max / Min

```
Max / Min     Max / Min
```

```
select min(id) from t;
```

```
select id from t order by id desc limit 1;
```

TableScan Aggregation TableScan + Sort + Limit id Sort TableScan/IndexLookUp + Limit

“”

1.2.2

-
-
-
-

1.2.3

TiDB NULL TOPN

1

```
ANALYZE      tidb_enable_fast_analyze 0 1 person fast analyze
```

```
set @@tidb_enable_fast_analyze = 1;
analyze table person;
```

```
show analyze status      where
```

```
mysql> show analyze status where job_info = 'analyze columns';
+-----+-----+-----+-----+
| Table_schema | Table_name | Job_info          | Start_time       | State   |
+-----+-----+-----+-----+
test	person	analyze columns	2020-03-07 06:22:34	finished
test	customer	analyze columns	2020-03-07 06:32:19	finished
test	person	analyze columns	2020-03-07 06:35:27	finished
+-----+-----+-----+-----+
3 rows in set (0.01 sec)
```

2

DML TiDB 1 20 * stats-lease

```
stats-lease 3s 0
```

3

meta

```
mysql> show stats_meta where table_name = 'person';
+-----+-----+-----+-----+-----+
| Db_name | Table_name | Partition_name | Update_time | Modify_count | Row_count |
+-----+-----+-----+-----+-----+
| test | person | | 2020-03-07 07:20:54 | 0 | 4 |
+-----+-----+-----+-----+
1 row in set (0.01 sec)
```

```
mysql> show stats_healthy where table_name = 'person';
+-----+-----+-----+
| Db_name | Table_name | Partition_name | Healthy |
+-----+-----+-----+
| test | person | | 100 |
+-----+-----+-----+
1 row in set (0.00 sec)
```

SHOW STATS_HISTOGRAMS NULL:

```
mysql> show stats_histograms where table_name = 'person';
+-----+-----+-----+-----+-----+-----+
| Db_name | Table_name | Partition_name | Column_name | Is_index | Update_time | Distinct_count | Null_count
| Avg_col_size | Correlation |
+-----+-----+-----+-----+-----+-----+
| test | person | | name | 0 | 2020-03-07 07:20:54 | 4 | 0
| 6.25 | -0.2 |
+-----+-----+-----+-----+-----+-----+
+-----+
1 row in set (0.00 sec)
```

SHOW STATS_BUCKETS

```
mysql> show stats_buckets;
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Db_name | Table_name | Partition_name | Column_name | Is_index | Bucket_id | Count | Repeats | Lower_Bound | Upper_Bound |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| test | person | | name | 0 | 0 | 1 | 1 | jack | jack
| |
| test | person | | name | 0 | 1 | 2 | 1 | peter | peter
| |
| test | person | | name | 0 | 2 | 3 | 1 | smith | smith
| |
| test | person | | name | 0 | 3 | 4 | 1 | tom | tom
| |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
+-----+
4 rows in set (0.01 sec)
```

4

DROP STATS

```
mysql> DROP STATS person;
```

5

```
 ${db_name}  ${table_name}  json  
  
http://${tidb-server-ip}:${tidb-server-status-port}/stats/dump/${db_name}/${table_name}
```

test person

```
curl -G "http://127.0.0.1:10080/stats/dump/test/person" > person.json
```

json

```
mysql> LOAD STATS 'file_name';
```

file_name

1.3 SQL Plan Management

1.3.1

SQL SQL TiDB SQL Plan Management SQL

1.3.2 SQL Bind

SQL Bind SQL Plan Management TiDB 3.0 GA SQL SQL Bind

SQL

```
CREATE [GLOBAL | SESSION] BINDING FOR SelectStmt USING SelectStmt;
```

GLOBAL SESSION SQL SESSION SQL SQL SQL Hint

“” SQL “?”

```
TiDB(root@127.0.0.1:test) > create binding for select * from t where a = 1 using select * from t use index(idx_a) where a = 1;
Query OK, 0 rows affected (0.00 sec)
```

binding original_sql SQL

```
TiDB(root@127.0.0.1:test) > show bindings;
+-----+-----+-----+-----+
| Original_sql | Bind_sql | Default_db | Status | Create_time
| Update_time   | Charset | Collation |
+-----+-----+-----+-----+
| select * from t where a = ? | select * from t use index(idx_a) where a = 1 | test      | using   | 2020-03-08 14:00:28.819 | 2020-03-08 14:00:28.819 | utf8     | utf8_general_ci |
+-----+-----+-----+-----+
1 row in set (0.00 sec)
```

binding

```
TiDB(root@127.0.0.1:test) > drop binding for select * from t where a = 1;
Query OK, 0 rows affected (0.00 sec)
```

```
TiDB(root@127.0.0.1:test) > show bindings;
Empty set (0.00 sec)
```

1.3.3 Baseline Evolution

Binding 4.0 TiDB Binding tidb_capture_plan_baselines on SQL TiDB SQL SQL TiDB 4.0
Statements Summary SQL

```
set tidb_enable_stmt_summary = 1;          -- statement summary
set tidb_capture_plan_baselines = 1;        --
```

```
TiDB(root@127.0.0.1:test) > select * from t;  
Empty set (0.01 sec)
```

```
TiDB(root@127.0.0.1:test) > select * from t;  
Empty set (0.00 sec)
```

global bindings binding

```
TiDB(root@127.0.0.1:test) > show global bindings;  
+-----+-----+-----+-----+  
| Original_sql | Bind_sql | Default_db | Status | Create_time  
| Update_time | Charset | Collation |  
+-----+-----+-----+-----+  
| select * from t | SELECT /*+ USE_INDEX(@`sel_1` `test`.`t` ) */ * FROM `t` | test | using | 2020-03-08 14:09:  
30.129 | 2020-03-08 14:09:30.129 | | |  
+-----+-----+-----+-----+  
1 row in set (0.00 sec)
```

4.0 set global tidb_evolve_plan_baselines = 1 SQL SQL

SQL TiDB

```
tidb_evolve_plan_task_max_time      tidb_evolve_plan_task_start_time  
tidb_evolve_plan_task_end_time
```

1.4

1.4.1

TiDB

```
TiDB(root@127.0.0.1:test) > show variables like "%tidb%opt%";
+-----+-----+
| Variable_name | Value |
+-----+-----+
tidb_opt_agg_push_down	0
tidb_opt_concurrency_factor	3
tidb_opt_copcpu_factor	3
tidb_opt_correlation_exp_factor	1
tidb_opt_correlation_threshold	0.9
tidb_opt_cpu_factor	3
tidb_opt_desc_factor	3
tidb_opt_disk_factor	1.5
tidb_opt_insubq_to_join_and_agg	1
tidb_opt_join_reorder_threshold	0
tidb_opt_memory_factor	0.001
tidb_opt_network_factor	1
tidb_opt_scan_factor	1.5
tidb_opt_seek_factor	20
tidb_opt_write_row_id	0
tidb_optimizer_selectivity_level	0
+-----+-----+
16 rows in set (0.01 sec)
```

1.4.2

10

```
TiDB(root@127.0.0.1:test) > show variables like "%tidb%factor%";
+-----+-----+
| Variable_name | Value |
+-----+-----+
tidb_opt_concurrency_factor	3
tidb_opt_copcpu_factor	3
tidb_opt_correlation_exp_factor	1
tidb_opt_cpu_factor	3
tidb_opt_desc_factor	3
tidb_opt_disk_factor	1.5
tidb_opt_memory_factor	0.001
tidb_opt_network_factor	1
tidb_opt_scan_factor	1.5
tidb_opt_seek_factor	20
+-----+-----+
10 rows in set (0.01 sec)
```

TiKV `tidb_opt_desc_factor`

:

```
TiDB(root@127.0.0.1:test) > desc select * from t order by a desc;
+-----+-----+-----+-----+
| id | estRows | task | access object | operator info |
+-----+-----+-----+-----+
| Projection_13 | 10000.00 | root | | test.t.a, test.t.b
|   |
|   IndexLookUp_12 | 10000.00 | root | | |
|   |
|   IndexFullScan_10(Build) | 10000.00 | cop[tikv] | table:t, index:idx_a(a) | keep order:true, desc, stats:ps
eudo |
|   TableRowIDScan_11(Probe) | 10000.00 | cop[tikv] | table:t | keep order:false, stats:pseudo
|   |
+-----+-----+-----+-----+
-----+
4 rows in set (0.00 sec)
```

TiKV

```
TiDB(root@127.0.0.1:test) > set @@tidb_opt_desc_factor = 10;
Query OK, 0 rows affected (0.00 sec)

TiDB(root@127.0.0.1:test) > desc select * from t order by a desc;
+-----+-----+-----+-----+
| id | estRows | task | access object | operator info |
+-----+-----+-----+-----+
| Sort_4 | 10000.00 | root | | test.t.a:desc
| TableReader_8 | 10000.00 | root | | data:TableFullScan_7
|   TableFullScan_7 | 10000.00 | cop[tikv] | table:t | keep order:false, stats:pseudo
+-----+-----+-----+-----+
3 rows in set (0.00 sec)
```

1.4.3

TiDB Join Join SQL

Join

```
TiDB(root@127.0.0.1:test) > desc select count(*) from t t1 join t t2 on t1.a = t2.a group by t1.a;
+-----+-----+-----+-----+
| id | estRows | task | access object | operator info |
+-----+-----+-----+-----+
| HashAgg_10 | 7992.00 | root | | group by:test.t.a, funcs:count(1)
->Column#7 |
| MergeJoin_13 | 12487.50 | root | | inner join, left key:test.t.a, right key:test.t.a |
|   IndexReader_41(Build) | 9990.00 | root | | index:IndexFullScan_40
|   |
|   IndexFullScan_40 | 9990.00 | cop[tikv] | table:t2, index:idx_a(a) | keep order:true, stats:pseudo
|   |
|   IndexReader_39(Probe) | 9990.00 | root | | index:IndexFullScan_38
|   |
|   IndexFullScan_38 | 9990.00 | cop[tikv] | table:t1, index:idx_a(a) | keep order:true, stats:pseudo
|   |
+-----+-----+-----+-----+
-----+
6 rows in set (0.00 sec)
```

`tidb_opt_agg_push_down` Join

```
TiDB(root@127.0.0.1:test) > set tidb_opt_agg_push_down = 1;
Query OK, 0 rows affected (0.00 sec)

TiDB(root@127.0.0.1:test) > desc select count(*) from t t1 join t t2 on t1.a = t2.a group by t1.a;
+-----+-----+-----+-----+
| id | estRows | task | access object | operator info |
+-----+-----+-----+-----+
| HashAgg_11 | 7992.00 | root | | group by:test.t.a, funcs:count(Column#8->Column#7)
| HashJoin_24 | 9990.00 | root | | inner join, inner:HashAgg_37, equal:[eq(test.t.a, test.t.a)]
| HashAgg_37(Build) | 7992.00 | root | | group by:test.t.a, funcs:count(1)->Column#8, funcs:firstrow(test.t.a)->test.t.a |
| IndexReader_44 | 9990.00 | root | | index:IndexFullScan_43
| IndexFullScan_43 | 9990.00 | cop[tikv] | table:t2, index:idx_a(a) | keep order:false, stats:pseudo
| IndexReader_48(Probe) | 9990.00 | root | | index:IndexFullScan_47
| IndexFullScan_47 | 9990.00 | cop[tikv] | table:t1, index:idx_a(a) | keep order:false, stats:pseudo
+-----+-----+-----+-----+
7 rows in set (0.00 sec)
```

TiDB mysql.opt_rule_blacklist

```
TiDB(root@127.0.0.1:test) > desc select * from t where a > 10;
+-----+-----+-----+-----+
| id | estRows | task | access object | operator info |
+-----+-----+-----+-----+
| TableReader_7 | 3333.33 | root | | data:Selection_6
| Selection_6 | 3333.33 | cop[tikv] | | gt(test.t.a, 10)
| TableFullScan_5 | 10000.00 | cop[tikv] | table:t | keep order:false, stats:pseudo
+-----+-----+-----+-----+
3 rows in set (0.00 sec)
```

predicate_pushdown

```
TiDB(root@127.0.0.1:test) > insert into mysql.opt_rule_blacklist values("predicate_push_down");
Query OK, 1 row affected (0.00 sec)
```

session reload

```
TiDB(root@127.0.0.1:test) > admin reload opt_rule_blacklist;
Query OK, 0 rows affected (0.00 sec)
```

TiKV

```
TiDB(root@127.0.0.1:test) > desc select * from t where a > 10;
+-----+-----+-----+-----+
| id | estRows | task | access object | operator info |
+-----+-----+-----+-----+
| Selection_5 | 8000.00 | root | | gt(test.t.a, 10)
| TableReader_7 | 10000.00 | root | | data:TableFullScan_6
| TablefullScan_6 | 10000.00 | cop[tikv] | table:t | keep order:false, stats:pseudo
+-----+-----+-----+-----+
3 rows in set (0.00 sec)
```

TiDB TiKV TiDB TiKV TiDB TiDB

```
mysql.expr_pushdown_blacklist
```

- `mysql.expr_pushdown_blacklist`
- `admin reload expr_pushdown_blacklist`

- `mysql.expr_pushdown_blacklist`
- `admin reload expr_pushdown_blacklist`

```
TiDB(root@127.0.0.1:test) > explain select * from t where a < 2;
+-----+-----+-----+-----+
| id      | estRows | task      | access object | operator info      |
+-----+-----+-----+-----+
TableReader_7	3323.33	root		data:Selection_6
└Selection_6	3323.33	cop[tikv]		lt(test.t.a, 2)
└TableFullScan_5	10000.00	cop[tikv]	table:t	keep order:false, stats:pseudo
+-----+-----+-----+-----+
3 rows in set (0.00 sec)
```

<

```
TiDB(root@127.0.0.1:test) > insert into mysql.expr_pushdown_blacklist values('<');
Query OK, 1 row affected (0.00 sec)
```

```
TiDB(root@127.0.0.1:test) > admin reload expr_pushdown_blacklist;
Query OK, 0 rows affected (0.00 sec)
```

< TiKV

```
TiDB(root@127.0.0.1:test) > explain select * from t where a < 2;
+-----+-----+-----+-----+
| id      | estRows | task      | access object | operator info      |
+-----+-----+-----+-----+
Selection_5	8000.00	root		lt(test.t.a, 2)
└TableReader_7	10000.00	root		data:TableFullScan_6
└TableFullScan_6	10000.00	cop[tikv]	table:t	keep order:false, stats:pseudo
+-----+-----+-----+-----+
3 rows in set (0.00 sec)
```

1.5 SQL

SQLSQLSQLLOOM

1.5.1 SQL

SQL

- 1.
2. oom-action
 - o "log"
 - o TiDB SQL mem-quota-query
 - o ["log", "cancel"]
 - o "log" SQL TiDB log LOG SQL OOM LOG SQL
 - o "cancel" SQL TiDB SQL errorerror SQL
3. mem-quota-query
 - o 3435973836832GB
 - o SQL mem-quota-query
 - o oom-action
4. oom-use-tmp-storage
 - o true
 - o SQL mem-quota-query
5. tmp-storage-path
 - o </>/tidb/tmp-storage
 - o SQL mem-quota-query
 - o oom-use-tmp-storage true
6. session
7. tidb_mem_quota_query
 - o 32 GB
 - o TiDB OOMAction
8. tidb_mem_quota_hashjoin
 - o 32 GB
 - o HashJoin HashJoin TiDB OOMAction
9. tidb_mem_quota_mergejoin
 - o 32 GB
 - o MergeJoin MergeJoin TiDB OOMAction
10. tidb_mem_quota_sort
 - o 32 GB
 - o Sort Sort TiDB OOMAction
11. tidb_mem_quota_topn
 - o 32 GB
 - o TopN TopN TiDB OOMAction
12. tidb_mem_quota_indexlookupreader
 - o 32 GB
 - o IndexLookupReader IndexLookupReader TiDB OOMAction
13. tidb_mem_quota_indexlookupjoin
 - o 32 GB
 - o IndexLookupJoin IndexLookupJoin TiDB OOMAction
14. tidb_mem_quota_nestedloopapply
 - o 32 GB

- o NestedLoopApply NestedLoopApply TiDB OOMAction

SQL 8GB

```
set @@tidb_mem_quota_query = 8 << 30;
```

SQL 8MB

```
set @@tidb_mem_quota_query = 8 << 20;
```

1. Optimizer Hints
2. `memory_quota`
 - o Hint MB GB
 - o log

SQL1024 MB

```
select /*+ MEMORY_QUOTA(1024 MB) */ * from t;
```

1.5.2 SQL

1. session/global
2. `max_execution_time`
 - o ms
 - o statement SELECT 100ms

10

```
set @@global.MAX_EXECUTION_TIME=10000
```

1. Optimizer Hints
2. `max_execution_time`
 - o ms
 - o N

SQL1000 1

```
select /*+ MAX_EXECUTION_TIME(1000) */ * from t1 inner join t2 where t1.id = t2.id;
```

1.5.3

SQL/ `tidb.log` Expensive query

```
[2020/02/05 15:32:25.096 +08:00] [WARN] [expensivequery.go:167] [expensive_query] [cost_time=60.008338935s] [wait_time=0s] [request_count=1] [total_keys=70] [process_keys=65] [num_cop_tasks=1] [process_avg_time=0s] [process_p90_time=0s] [process_max_time=0s] [process_max_addr=10.0.1.9:20160] [wait_avg_time=0.002s] [wait_p90_time=0.002s] [wait_max_time=0.002s] [wait_max_addr=10.0.1.9:20160] [stats=t:pseudo] [conn_id=60026] [user=root] [database=test] [table_ids="[1 22]"] [txn_start_ts=414420273735139329] [mem_max="1035 Bytes (1.0107421875 KB)"] [sql="insert into t select sleep(1) from t"]
```

- `cost_time`
- `stats` pesudo analyze
- `table_ids` ID
- `txn_start_ts` ID TiDB
- `sql` SQL

- `mem_max` Bytes MBGB

- `user`
- `conn_id` ID con:60026 TiDB
- `database` database

TiKV Coprocessor Task

- `wait_time` TiKV TiKV Coprocessor Coprocessor
- `request_count` Coprocessor
- `total_keys` Coprocessor key
- `processed_keys` Coprocessor key `total_keys` `processed_keys` MVCC `processed_keys` `total_keys`
- `num_cop_tasks` Coprocessor
- `process_avg_time` Coprocessor task
- `process_p90_time` Coprocessor task P90
- `process_max_time` Coprocessor task
- `process_max_addr` task Coprocessor
- `wait_avg_time` Coprocessor task
- `wait_p90_time` Coprocessor task P90
- `wait_max_time` Coprocessor task
- `wait_max_addr` task Coprocessor

2.1

TiDB Dashboard Key Visualizer (KeyVis) TiDB DBA CT

1.

- Bucket

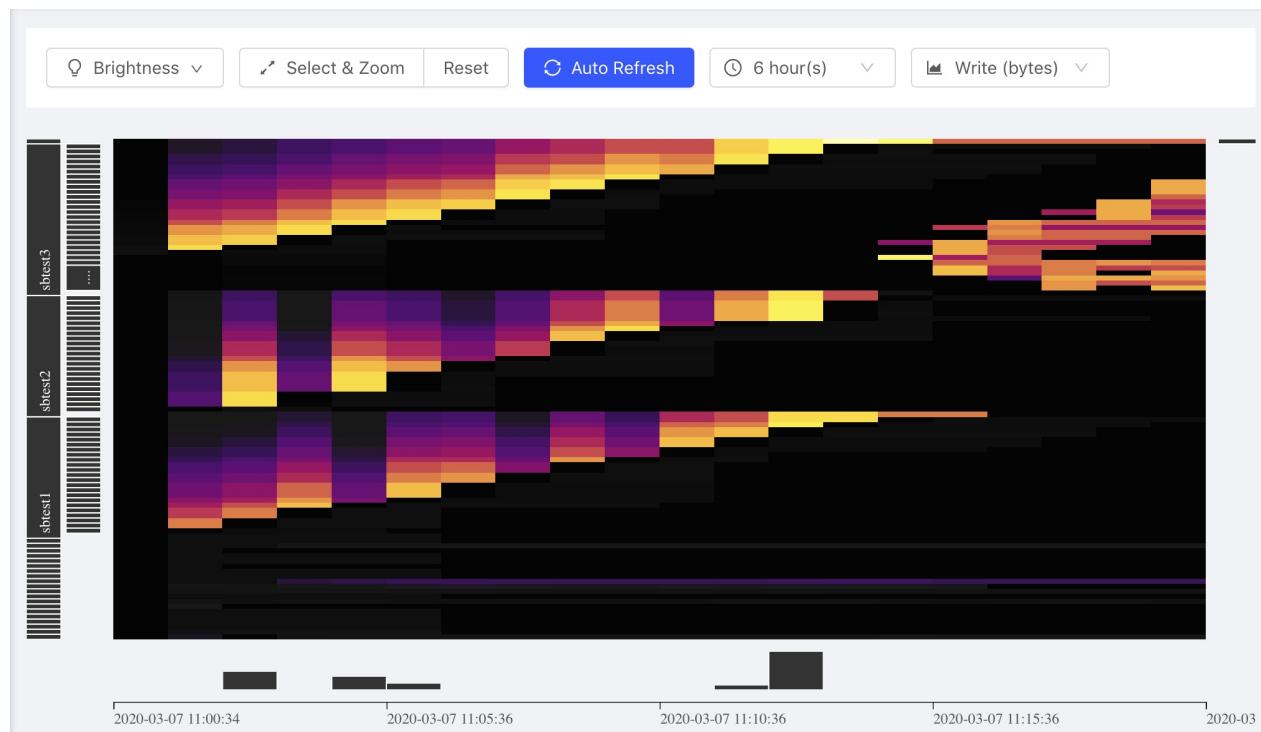
TiDB Region Region Region 1500 Bucket Region Region Bucket Region Bucket

-

X Y Region TiDB cold Region hot

2.

-
-
-



3. Region

6 Bucket Region

-
- Select & Zoom
- Reset Region
-

Region

4.

Bucket cold Region hot Brightness

5.

Auto Refresh Region

6.

- Read (bytes)
- Write (bytes)
- Read (keys)
- Write (keys)
- All

7.

Region TiKV KeyVis Region DBA

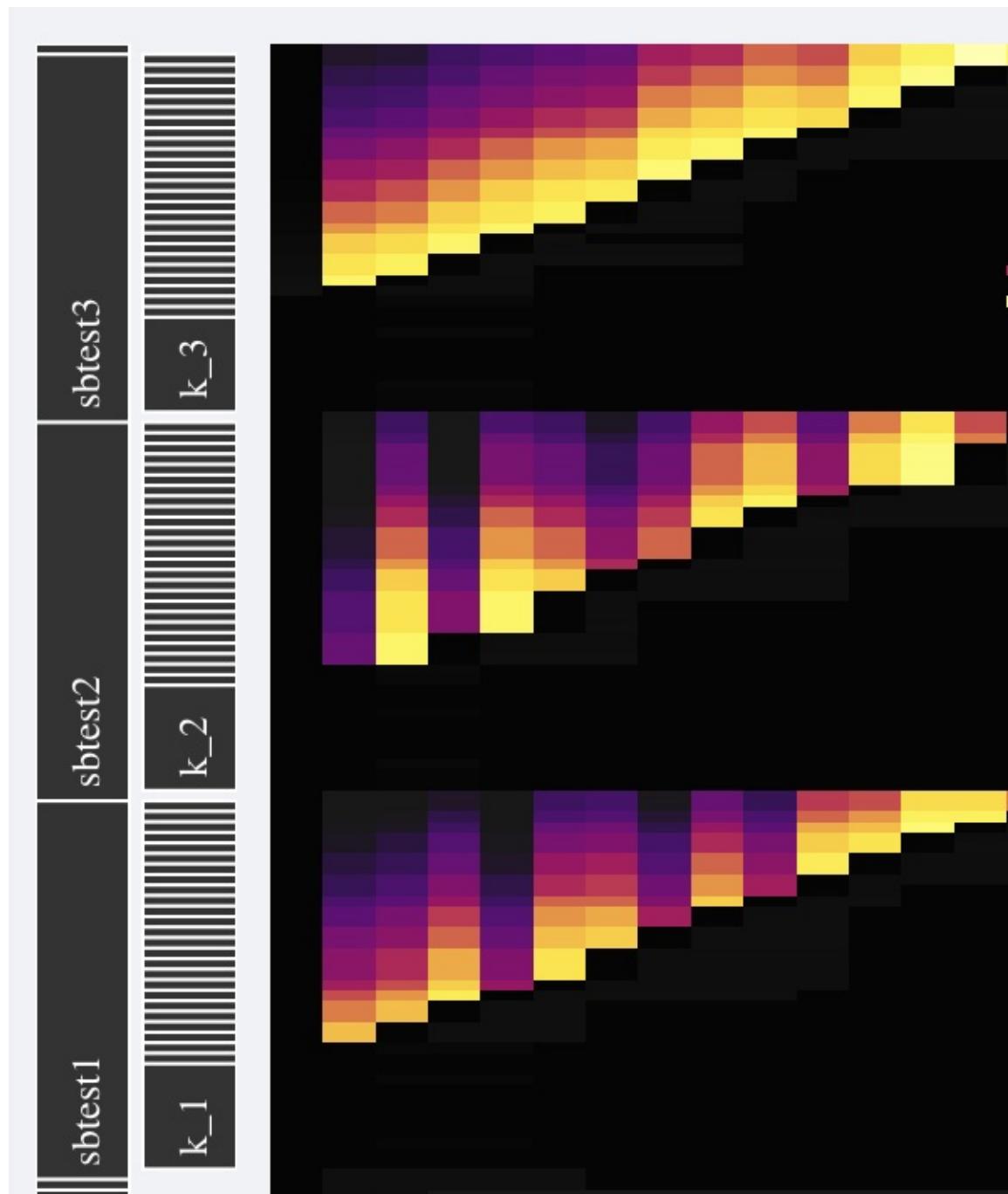
-
-
-
-
-
-
- scale 10
- ()

CRUD TiDB Dashboard DBA

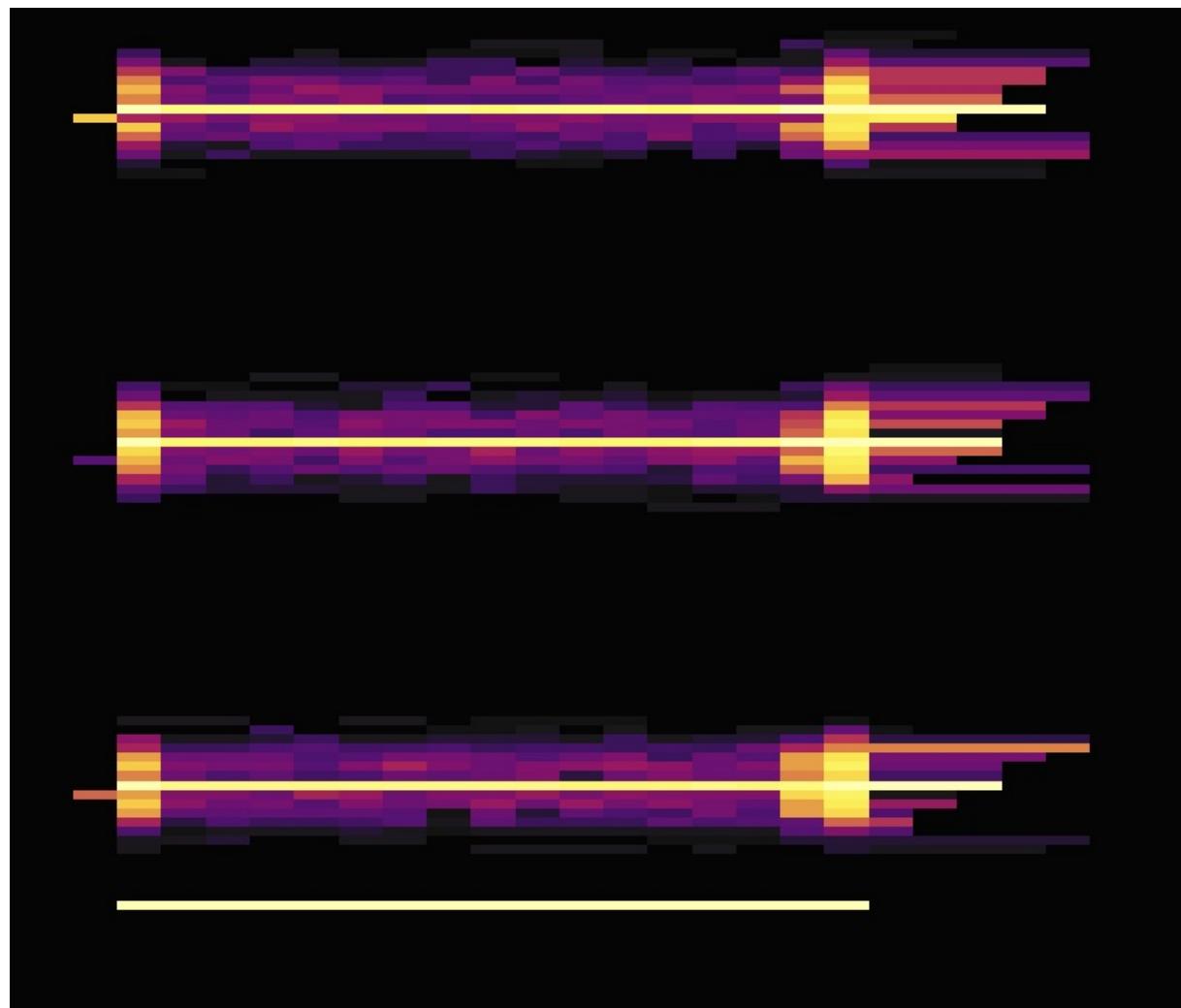
8.

-

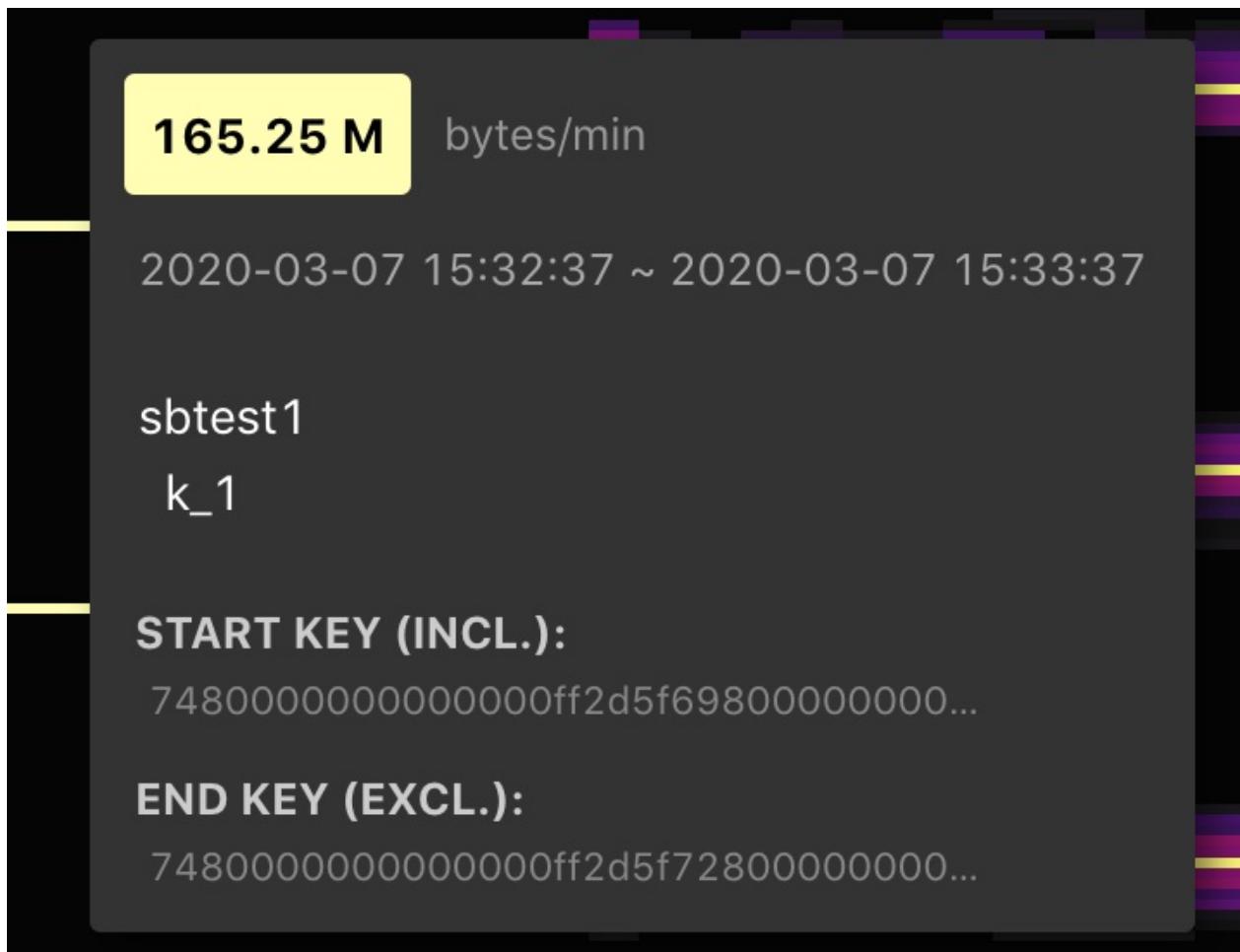
(sysbench prepare) sbtest1, sbtest2, sbtest3,

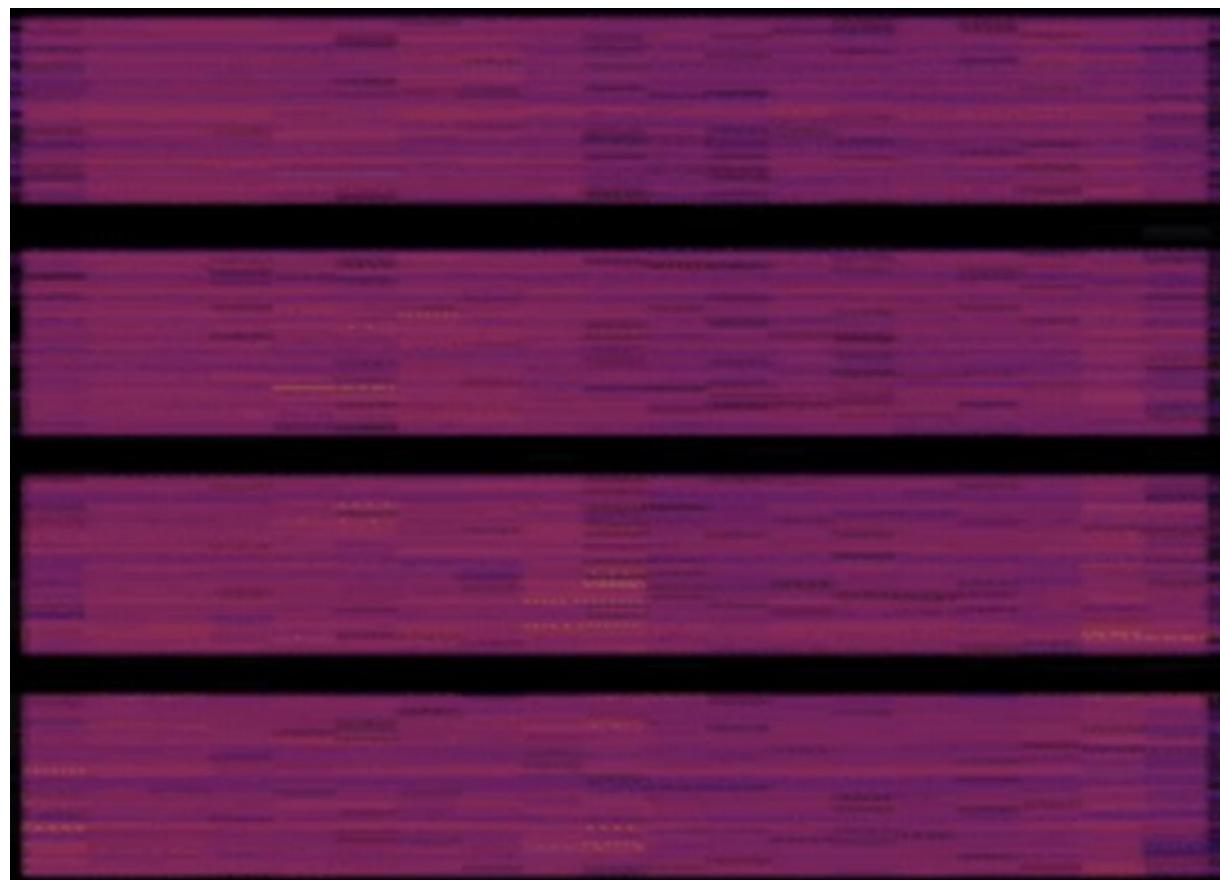


• KeyVis

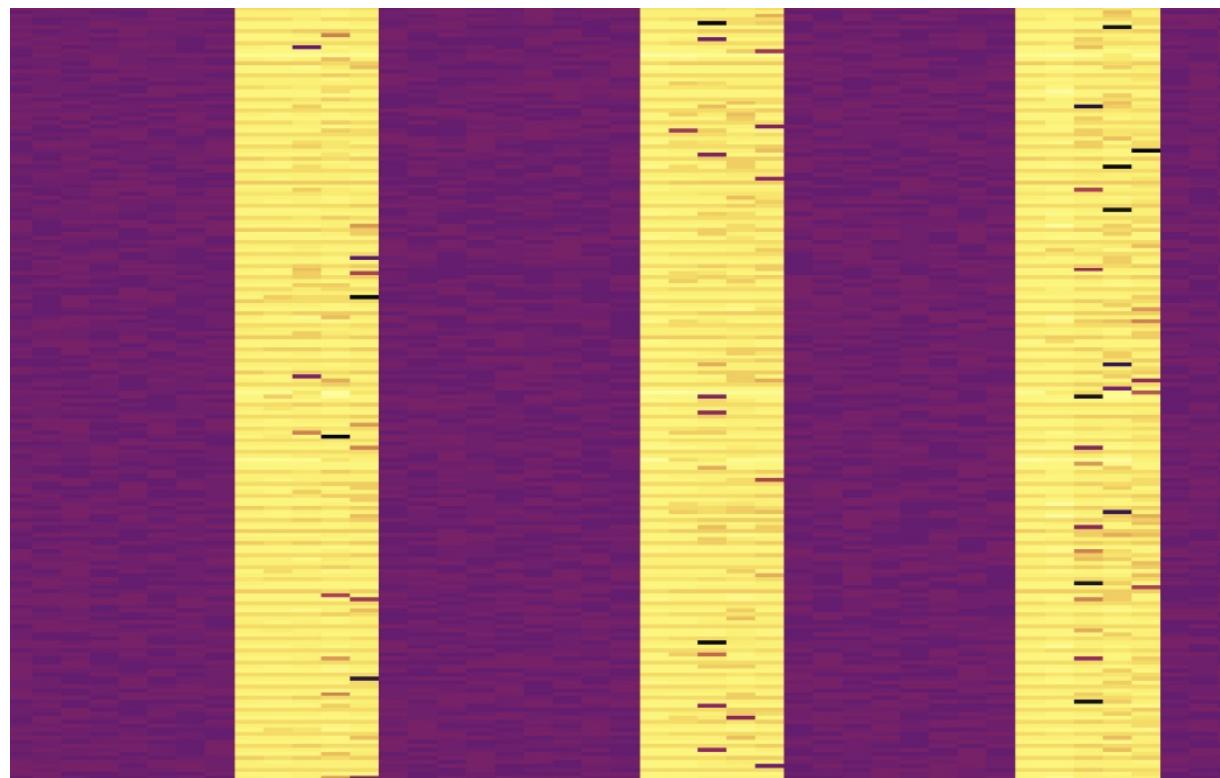


165.25 , sbtest1 k_1, key start key end key





•



2.2 SQL

KeyVis TiDB

TiDB Dashboard Statements SQL SQL

1. Statements

Statement SQL

SQL TiDB Dashboard Statements SQL

SQL

2. Statements

TiDB

1. Grafana SQL
2. Slow log SQL
3. General log
4. Explain analyze
5. Profile

Statements SQL

3. Statements

SQL

SQL

- 1.
- 2.
- 3.

SQL SQL SQL SQL

```
SELECT * FROM employee WHERE id IN (1, 2, 3);
select * from EMPLOYEE where ID in (4, 5);
```

```
select * from employee where id in (...);
```

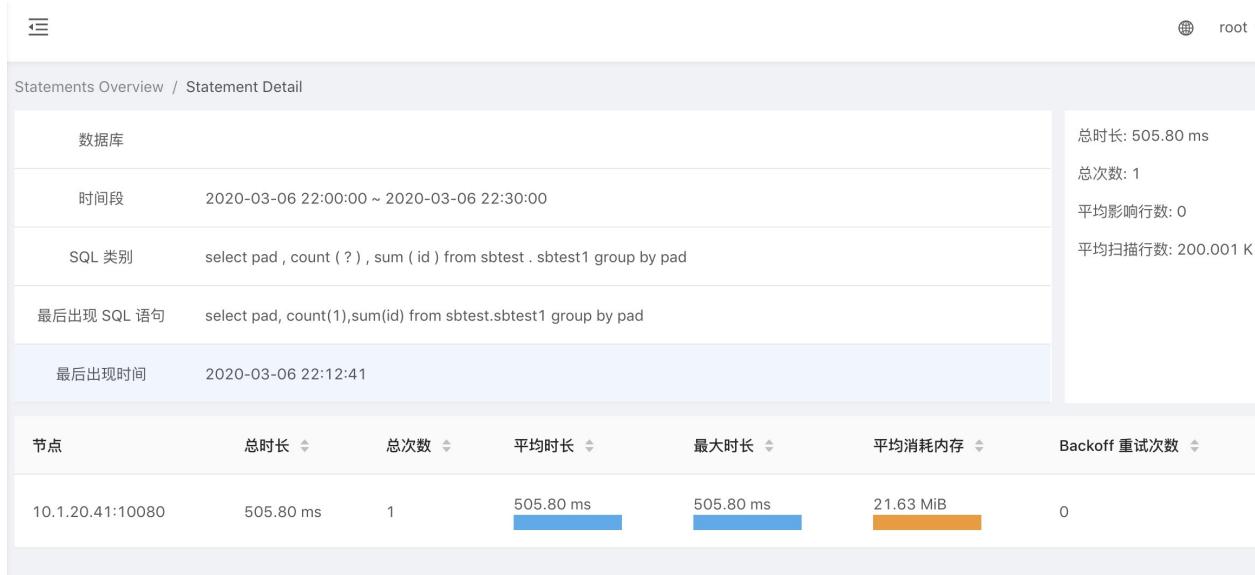


4. Statements

SQL SQL SQL SQL

Statements

1. SQL
- 2.
- 3.
- 4.



5. Statements

- `tidb_enable_stmt_summary`

Statements

```
set global tidb_enable_stmt_summary = true;
```

- `tidb_stmt_summary_refresh_interval`

```
performance_schema.events_statements_summary_by_digest (s) 1800
```

```
set global tidb_stmt_summary_refresh_interval = 1800;
```

- `tidb_stmt_summary_history_size`

```
performance_schema.events_statements_summary_by_digest_history SQL 24
```

```
set global tidb_stmt_summary_history_size = 24;
```

Statements SQL SQL config.toml [stmt-summary]

- `max-stmt-count SQL 200 SQL max-stmt-count SQL`
- `max-sql-length DIGEST_TEXT QUERY_SAMPLE_TEXT 4096`

```
tidb_stmt_summary_history_size max-stmt-count max-sql-length
```

Statements SQL KeyVis

2.3

+

TiDB DBA

html

2.3.1

1.

PD pd-server.metric-storage prometheus

```
curl -X POST -d '{"metric-storage":"http://[PROMETHEUS_ADDRESS]"}' http://[PD_ADDRESS]/pd/api/v1/config
```

```
curl -X POST -d '{"metric-storage":"http://127.0.0.1:9090"}' http://127.0.0.1:2379/pd/api/v1/config
```

2.

2 min ~ 60 min 1

3.



报告状态

时间段 2020-03-05 16:30:00 ~ 2020-03-05 16:50:00

进度 100%

完整报告 </dashboard/api/diagnose/reports/3>

返回

4.

2 start end 10

Step 1: 选择第一个时间范围

2020-03-05 16:30:00 ~ 2020-03-05 16:40:00

生成诊断报告 取消对比

对比

Step 2: 选择第二个时间范围

2020-03-05 17:30:00 ~ 2020-03-05 17:40:00

生成对比诊断报告

Step 3: 点击生成对比报告

5.

- i i
- expand expand instance, label

| Load | the disk write latency in each node | instance | Avg | Max | Min |
|-----------------------------|-------------------------------------|-------------------|--------|--------|----------|
| METRIC_NAME | | | | | |
| node_disk_write_latency | fold | | 0.0002 | 0.0003 | 0.000007 |
| l-- node_disk_write_latency | | 172.16.5.40:19110 | 0.0002 | 0.0003 | 0.000007 |
| node_disk_read_latency | expand | | | | |
| node_cpu_usage | expand | | 80.56% | 81.55% | 79.4% |
| node_mem_usage | expand | | 27.9% | 28.71% | 26.89% |

2.3.2

TiDB 4.0 INSPECTION_RESULT

2.3.3

1. Header

(1) Report Time Range

| START_TIME | END_TIME |
|---------------------|---------------------|
| 2020-03-05 16:30:00 | 2020-03-05 16:50:00 |

(2) cluster hardware

| HOST | INSTANCE | CPU_CORES | MEMORY (GB) | DISK (GB) | UPTIME (DAY) |
|-------------|-----------------------|-----------|-------------|--|--------------------|
| 172.16.5.40 | tidb*1
tikv*4 pd*1 | 20/40 | 122.696 | sda3: 1465.187500
nvme0n1: 343.715603
sda1: 0.185237 | 253.14520396237572 |

- HOST
- INSTANCE tidb*1 tikv*4 pd*1 1 TiDB 4 TiKV 1 PD
- CPU_CORES CPU /
- MEMORY GB
- DISK GB
- UPTIME: Day

(3) cluster info

TiDB information_schema.cluster_info

| TYPE | INSTANCE | STATUS_ADDRESS | VERSION | GIT_HASH |
|------|-------------------|-------------------|--|--|
| tidb | 172.16.5.40:4009 | 172.16.5.40:10089 | 5.7.25-TiDB-v4.0.0-beta-311-gd93c06149 | d93c061491d8094751a53c510bd8de886722952a |
| pd | 172.16.5.40:24799 | 172.16.5.40:24799 | 4.1.0-alpha | 61d9f9cc35d3f191eb5e7ea1eb4f8e29eb73eda0 |
| tikv | 172.16.5.40:21150 | 172.16.5.40:21151 | 4.1.0-alpha | 8fa0e059e14c3a1433fcb581452f9ea0a14a72ce |
| tikv | 172.16.5.40:22150 | 172.16.5.40:22151 | 4.1.0-alpha | 8fa0e059e14c3a1433fcb581452f9ea0a14a72ce |
| tikv | 172.16.5.40:23150 | 172.16.5.40:23151 | 4.1.0-alpha | 8fa0e059e14c3a1433fcb581452f9ea0a14a72ce |
| tikv | 172.16.5.40:20150 | 172.16.5.40:20151 | 4.1.0-alpha | 8fa0e059e14c3a1433fcb581452f9ea0a14a72ce |

2. Load info

(1) node load

/ CPUMemory AVG MAX MIN

| METRIC_NAME | instance | AVG | MAX | MIN |
|-------------------------|----------|--------|--------|----------|
| node_disk_write_latency | | 0.0002 | 0.0003 | 0.000007 |
| node_disk_read_latency | | | | |
| node_cpu_usage | | 80.56% | 81.55% | 79.4% |
| node_mem_usage | | 27.9% | 28.71% | 26.89% |

(2) process cpu usage

TiDB / PD / TiKV CPU AVG MAX MIN

| instance | job | AVG | MAX | MIN |
|-------------------|------|-------|-------|-------|
| 172.16.5.40:10089 | tidb | 18.38 | 19.16 | 17.76 |
| 172.16.5.40:20151 | tikv | 9.68 | 10.03 | 9.42 |
| 172.16.5.40:24799 | pd | 0.6 | 0.63 | 0.56 |
| 172.16.5.40:22151 | tikv | 0.36 | 1.08 | 0.04 |
| 172.16.5.40:23151 | tikv | 0.3 | 1.07 | 0.04 |
| 172.16.5.40:21151 | tikv | 0.27 | 0.71 | 0.04 |

(3) TiKV Thread CPU Usage

TiKV CPU

| METRIC_NAME | instance | AVG | MAX | MIN |
|--|--------------|-------------------|---------|--------|
| raftstore fold >) | | 0.01 | 0.06 | 0 |
| \ | -- raftstore | 172.16.5.40:20151 | 0.06 | 0.06 |
| \ | -- raftstore | 172.16.5.40:22151 | 0.00003 | 0.0002 |
| \ | -- raftstore | 172.16.5.40:23151 | 0 | 0 |
| \ | -- raftstore | 172.16.5.40:21151 | 0 | 0 |
| apply expand >) | | 0 | 0 | 0 |
| sched_worker expand >) | | 0 | 0 | 0 |
| snap expand >) | | 0.000002 | 0.0002 | 0 |
| cop expand >) | | 0.01 | 0.18 | 0 |
| grpc expand >) | | 0.16 | 0.85 | 0 |

(4) goroutines count

TiDB / PD goroutines AVG MAX MIN

| instance | job | AVG | MAX | MIN |
|-------------------|------|-----|-----|-----|
| 172.16.5.40:10089 | tidb | 899 | 996 | 651 |
| 172.16.5.40:24799 | pd | 114 | 115 | 113 |

3. Overview

(1) Time Consume

| METRIC_NAME | LABEL | TIME_RATIO | TOTAL_TIME | TOTAL_COUNT | P999 | P99 |
|---|-------|------------|------------|-------------|--------|----------|
| tidb_query expand) | | 1 | 184598.05 | 9467997 | 0.4 | 0.12 |
| tidb_get_token expand) | | 0.00007 | 12.1 | 9467600 | 0.0002 | 0.000001 |
| tidb_parse expand) | | 0.000007 | 1.23 | 3593 | 0.04 | 0.008 |
| tidb_compile expand) | | 0.00001 | 1.82 | 3593 | 0.04 | 0.03 |
| tidb_execute expand) | | 0.000007 | 1.26 | 3593 | 0.2 | 0.17 |
| tidb_distsql_execution expand) | | 0.0003 | 56.72 | 1953 | 1.41 | 0.5 |
| tidb_cop expand) | | 0.001 | 206.21 | 2756 | 2.01 | 1.66 |
| tidb_transaction expand) | | 0.09 | 15920.8 | 9456815 | 0.49 | 0.06 |

- METRIC_NAME
- LABEL LABEL expand LABEL
- TIME_RATIO TIME_RATIO 1

```
tidb_cop  tidb_query  0.001
```

- TOTAL_TIME
- TOTAL_COUNT
- P999 P999
- P99 P99
- P90 P90
- P80 P80

(2) Error

error

| METRIC_NAME | LABEL | TOTAL_COUNT | |
|--|---|-----------------|-----|
| tidb_binlog_error_total_count | | 0 | |
| tidb_handshake_error_total_count | | 0 | |
| tidb_transaction_retry_error_total_count | | | |
| tidb_kv_region_error_total_count <code>fold></code> | | 739 | |
| \ | -- tidb_kv_region_error_total_count | not_leader | 481 |
| \ | -- tidb_kv_region_error_total_count | stale_command | 256 |
| \ | -- tidb_kv_region_error_total_count | epoch_not_match | 1 |
| tidb_schema_lease_error_total_count | | | |
| tikv_grpc_error_total_count | | | |
| tikv_critical_error_total_count | | | |
| tikv_scheduler_is_busy_total_count | | | |
| tikv_channel_full_total_count | | | |
| tikv_coprocessor_request_error_total_count <code>fold></code> | | 8 | |
| \ | -- tikv_coprocessor_request_error_total_count | meet_lock | 7 |
| \ | -- tikv_coprocessor_request_error_total_count | not_leader | 1 |
| tikv_engine_write_stall | | 0 | |
| tikv_server_report_failures_total_count | | 0 | |
| tikv_storage_async_request_error <code>fold></code> | | 793 | |
| \ | -- tikv_storage_async_request_error | snapshot | 793 |
| tikv_lock_manager_detect_error_total_count | | 0 | |
| tikv_backup_errors_total_count | | | |
| node_network_in_errors_total_count | | 0 | |
| node_network_out_errors_total_count | | 0 | |

4. TiDB

TiDB

(1) Time Consume

TiDB Overview label

(2) Transaction

TiDB

| METRIC_NAME | LABEL | TOTAL_VALUE | TOTAL_COUNT | P999 | P99 | P90 |
|--|-------|-------------|-------------|---------|--------|------|
| tidb_transaction_retry_num | | 0 | 0 | 0 | 0 | 0 |
| tidb_transaction_statement_num
expand>) | | 17524099 | 17500676 | 511 | 501 | 398 |
| tidb_txn_region_num expand>) | | 900385 | 900095 | 4 | 1 | 1 |
| tidb_txn_kv_write_num
expand>) | | 658293 | 1379 | 1023 | 1011 | 896 |
| tidb_txn_kv_write_size
expand>) | | 216307980 | 1379 | 1043333 | 996147 | 2487 |
| tidb_load_safepoint_total_num
expand>) | | 67 | | | | |
| tidb_lock_resolver_total_num
expand>) | | 40 | | | | |

- TOTAL_VALUE
- TOTAL_COUNT
- P999 P999
- P99 P99
- P90 P90
- P80 P80

tidb_txn_kv_write_size 1379 kv kv 216307980 P999P99P90P80 1043333996147248705235266 byte

(3) DDL-owner

TiDB DDL owner

| MIN_TIME | DDL OWNER |
|---------------------|-------------------|
| 2020-03-05 16:30:00 | 172.16.5.40:10089 |

owner owner TiDB ddl_worker DDL owner ddl_worker DDL job owner

5. PD / TiKV

PD TiKV i

6. Config

(1) Scheduler Config

PD Scheduler change history

| MIN_TIME | CONFIG_ITEM | VALUE | CHANGE_COUNT |
|----------------------------|---------------------------------|--------|--------------|
| 2020-03-05 16:30:00.000000 | enable-makeup-replica | 1 | 1 |
| 2020-03-05 16:30:00.000000 | enable-remove-down-replica | 1 | 1 |
| 2020-03-05 16:30:00.000000 | enable-remove-extra-replica | 1 | 1 |
| 2020-03-05 16:30:00.000000 | enable-replace-offline-replica | 1 | 1 |
| 2020-03-05 16:30:00.000000 | high-space-ratio | 0.6 | 1 |
| 2020-03-05 16:30:00.000000 | hot-region-cache-hits-threshold | 3 | 1 |
| 2020-03-05 16:30:00.000000 | hot-region-schedule-limit | 4 | 1 |
| 2020-03-05 16:30:00.000000 | leader-schedule-limit | 4 | 2 |
| 2020-03-05 16:46:00.000000 | leader-schedule-limit | 8 | 2 |
| 2020-03-05 16:30:00.000000 | low-space-ratio | 0.8 | 1 |
| 2020-03-05 16:30:00.000000 | max-merge-region-keys | 200000 | 1 |
| 2020-03-05 16:30:00.000000 | max-merge-region-size | 20 | 1 |
| 2020-03-05 16:30:00.000000 | max-pending-peer-count | 16 | 1 |
| 2020-03-05 16:30:00.000000 | max-replicas | 3 | 1 |
| 2020-03-05 16:30:00.000000 | max-snapshot-count | 3 | 1 |
| 2020-03-05 16:30:00.000000 | merge-schedule-limit | 8 | 1 |
| 2020-03-05 16:30:00.000000 | region-schedule-limit | 2048 | 1 |
| 2020-03-05 16:30:00.000000 | replica-schedule-limit | 64 | 1 |
| 2020-03-05 16:30:00.000000 | store-balance-rate | 15 | 1 |
| 2020-03-05 16:30:00.000000 | tolerant-size-ratio | 0 | 1 |

- MIN_TIME
- CONFIG_ITEM
- VALUE
- CHANGE_COUNT

- leader-schedule-limit
 - 2020-03-05 16:30:00 leader-schedule-limit 4
 - 2020-03-05 16:46:00 leader-schedule-limit 8 2020-03-05 16:46:00

leader-schedule-limit 2020-03-05 16:46:00 prometheus

(2) TiDB GC Config

TiDB GC change history

| MIN_TIME | CONFIG_ITEM | VALUE | CHANGE_COUNT |
|----------------------------|----------------------|-------|--------------|
| 2020-03-03 17:00:00.000000 | tikv_gc_life_time | 6000 | 1 |
| 2020-03-03 17:00:00.000000 | tikv_gc_run_interval | 60 | 1 |

(3) TiDB / PD / TiKV Current Config

TiDB information_schema.cluster_config

current Config

2.4

TiDB Dashboard TiDB

1.

- :
-
-
-
- regex

The screenshot shows the 'Log Searching' section of the TiDB Dashboard. It includes search filters for time range (2020-03-06 00:00:00 ~ 2020-03-08 11:59:59), log level (WARN), and components (10.1.20.41:4000, 10.1.20.41:20160). The search results table lists three log entries related to 'server' for TiKV components. To the right, a '搜索进度' (Search Progress) panel shows successful search runs for both TiDB and TiKV components.

| 时间 | 日志等级 | 组件 | 日志 |
|---------------------------|------|------|--|
| 2020-03-07T20:04:39+08:00 | WARN | TiDB | [base_client.go:170] "[pd] failed to get cluster id" [url=http://10.... |
| 2020-03-07T20:04:48+08:00 | WARN | TiKV | [server.rs:790] ["check: kernel"] [err="kernel parameters net.core.... |
| 2020-03-07T20:04:48+08:00 | WARN | TiKV | [server.rs:790] ["check: kernel"] [err="kernel parameters net.ipv4.t...] |

2.

-
- Dashboard
-
- ()
- Cancel
- Retry

3.

4

-
- WARNERROR
- TiDBTiKVPD
- 512

[base_client.go:170] ["[pd] failed to
get cluster id"]
[url=http://10.1.20.41:2379]
[error="error:rpc error: code =
Unknown desc = server not started
target:10.1.20.41:2379
status:READY"]

日志

[base_client.go:170] ["[pd] failed to get cluster id"] [url=http://10....

4.

搜索进度

3 成功

下载选中日志

Cancel

Retry

- ▼ TiDB 1 成功
 - 10.1.20.41:4000
- ▼ TiKV 1 成功
 - 10.1.20.41:20160
- ▼ PD 1 成功
 - 10.1.20.41:2379

- zip log
- tar zip zip log

2.5 CPU

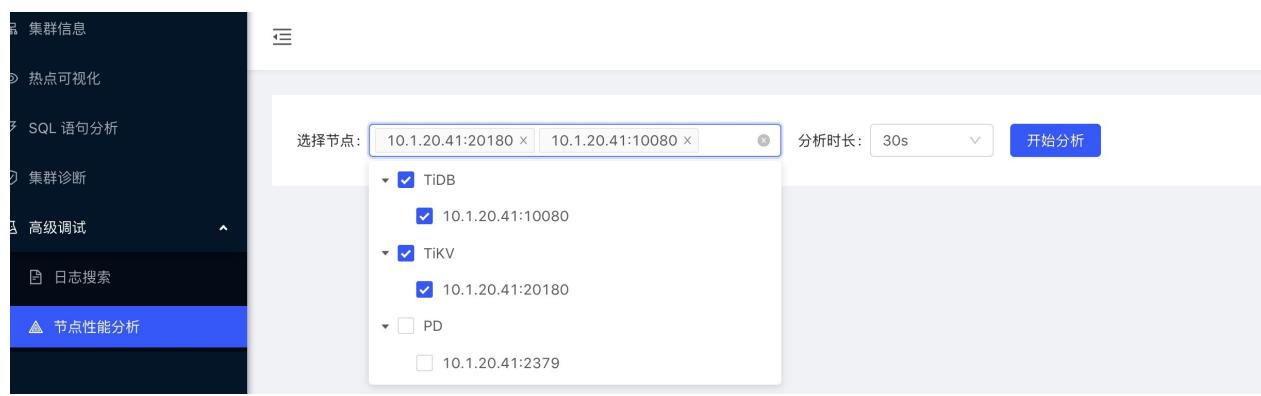
TiDBTiKVPD

go profile CPU

1.

Dashboard →

(30 120)



2.

1

下载性能分析结果

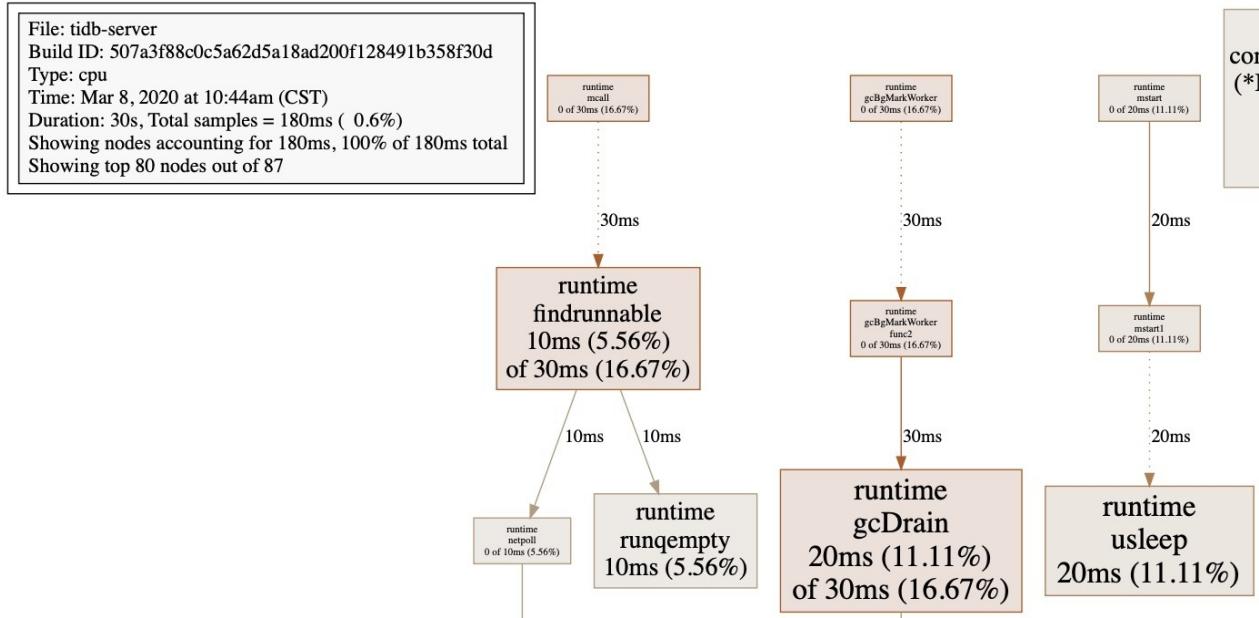
| Node | Kind | Status |
|------------------|------|------------------------------------|
| 10.1.20.41:20180 | tikv | <div style="width: 13%;">13%</div> |
| 10.1.20.41:10080 | tidb | <div style="width: 13%;">13%</div> |

Dashboard PD gogo pprofGraphViz TiDB PD TiKV

3.

TiDBTiKV

- TiDB Profile



- TiKV CPU

__read

procinfo::parsers::read_to_end::hbf1978b53ba4b293

```

procinfo::pid::stat::stat_task::hbfffc6a4d54a436f2/procinfo::pid::stat::stat_file::h62c32fb8a991c85b
tikv::server::load_statistics::linux::ThreadLoadStatistics::record::hc3968d280f130402
<futures::stream::for_each::ForEach<S,F,U> as futures::future::Future>::poll::h6a25e5cae30875c5/tikv::serv
tokio_threadpool::worker::Worker::run_task::hf9246b3e69e970c3/tokio_threadpool::worker::Worker::run_ta
tokio_threadpool::worker::Worker::run::ha03a92852493b565/tokio_threadpool::worker::Worker::try_run_ta
std::sys_common::backtrace::__rust_begin_short_backtrace::h870a8fbb202512f0/tokio_threadpool::pool::Po
core::ops::function::FnOnce::call_once::{vtable.shim}::hc7882ad627a4de8e/std::thread::Builder::spawn_u
<alloc::boxed::Box<F> as core::ops::function::FnOnce<A>>::call_once::hbabfae4f70ab9609
std::sys::unix::thread::Thread::new::thread_start::h3f42d1616b825de5/std::sys_common::thread::start_threa
start_thread
__clone
Unknown
transport-stats
all

```

Function: procinfo::parsers::read_to_end::hbf1978b53ba4b293 (106 samples, 32.02%)

3

TiDB 4.0 ——— SQL SQL TiDB 4.0 SQL

3.1 SQL TiDB 4.0 Statements SQL

3.2 3.4 SQL TiDB 4.0 metrics_schema SQL SQL TiDB SQL

3.3 SQL SQL SQL SQL select * from inspection_result

3.5 3.6 processlist slow_query SQL cluster_processlist cluster_slow_query cluster_processlist TiDB
cluster_slow_query

3.7 Statement Summary Statement Summary SQL SQL

3.1

3.1.1

TiDB TiDB 4.0 SQL

TiDB 4.0

3.1.2

TiDB 4.0

- `information_schema.cluster_info` Git Hash
- `information_schema.cluster_config` TiDB 4.0 HTTP API
- `information_schema.cluster_hardware`
- `information_schema.cluster_load`
- `information_schema.cluster_systeminfo` sysctl
- `information_schema.cluster_log` grep

TiDB 4.0 TiDB 4.0 TiDB

3.1.3

1.

`information_schema.cluster_info`

```
mysql> select type, instance, status_address, uptime from cluster_info;
+-----+-----+-----+-----+
| type | instance      | status_address | uptime      |
+-----+-----+-----+-----+
tidb	127.0.0.1:4000	127.0.0.1:10080	11m6.204302s
tidb	127.0.0.1:4001	127.0.0.1:10081	11m6.204306s
tidb	127.0.0.1:4002	127.0.0.1:10082	11m6.204307s
pd	127.0.0.1:2380	127.0.0.1:2380	11m16.204308s
pd	127.0.0.1:2381	127.0.0.1:2381	11m16.20431s
pd	127.0.0.1:2382	127.0.0.1:2382	11m16.204311s
tikv	127.0.0.1:20161	127.0.0.1:20181	11m11.204312s
tikv	127.0.0.1:20162	127.0.0.1:20182	11m11.204313s
tikv	127.0.0.1:20160	127.0.0.1:20180	11m11.204314s
tikv	127.0.0.1:20163	127.0.0.1:20183	11m11.204315s
+-----+-----+-----+-----+
10 rows in set (0.01 sec)
```

:

- TYPE pd/tikv/tidb
- INSTANCE IP:PORT
- STATUS_ADDRESSHTTP API tikv-ctl / pd-ctl / tidb-ctl HTTP API HTTP API
- VERSIONTiDB MySQL \${mysql-version}-\${tidb-version}
- GIT_HASH git commit hash
- START_TIME
- UPTIME

2.

`information_schema.cluster_config` TiDB 4.0 HTTP API TiKV Coprocessor

```
mysql> select * from cluster_config where type='tikv' and `key`='coprocessor.batch-split-limit';
+-----+-----+-----+
| TYPE | INSTANCE | KEY | VALUE |
+-----+-----+-----+
tikv	127.0.0.1:20163	coprocessor.batch-split-limit	10
tikv	127.0.0.1:20161	coprocessor.batch-split-limit	10
tikv	127.0.0.1:20162	coprocessor.batch-split-limit	10
tikv	127.0.0.1:20160	coprocessor.batch-split-limit	10
+-----+-----+-----+
4 rows in set (2.98 sec)
```

TiDB SQL log.file.filename / port / status.status-port

```
mysql> select `key`,count(distinct value) as c from cluster_config where type='tidb' group by `key` having c > 1;
+-----+---+
| key | c |
+-----+---+
log.file.filename	3
port	3
status.status-port	3
+-----+---+
3 rows in set (0.01 sec)
```

- TYPE information_schema.cluster_info TYPE tidb/pd/tikv
- INSTANCE information_schema.cluster_info STATUS_ADDRESS
- KEY
- VALUE

3.

information_schema.cluster_hardware CPU

```
mysql> select type, instance, name, value from cluster_hardware where name='cpu-logical-cores';
+-----+-----+-----+
| type | instance | name | value |
+-----+-----+-----+
tidb	127.0.0.1:10080	cpu-logical-cores	8
tidb	127.0.0.1:10081	cpu-logical-cores	8
tidb	127.0.0.1:10082	cpu-logical-cores	8
pd	127.0.0.1:2380	cpu-logical-cores	8
pd	127.0.0.1:2381	cpu-logical-cores	8
pd	127.0.0.1:2382	cpu-logical-cores	8
tikv	127.0.0.1:20160	cpu-logical-cores	8
tikv	127.0.0.1:20163	cpu-logical-cores	8
tikv	127.0.0.1:20161	cpu-logical-cores	8
tikv	127.0.0.1:20162	cpu-logical-cores	8
+-----+-----+-----+
10 rows in set (0.78 sec)
```

- TYPE information_schema.cluster_info TYPE tidb/pd/tikv
- INSTANCE information_schema.cluster_info STATUS_ADDRESS
- DEVICE_TYPE cpu/memory/disk/net
- DEVICE_NAME DEVICE_TYPE
 - cpu
 - disk
 - netNIC
 - memory
- NAME cpu cpu-logical-cores / cpu-physical-cores select name from cluster_hardware where device_type='cpu'

- group by name NAME
- VALUECPU

4.

information_schema.cluster_load SQL CPU

```
mysql> select type, instance, name, value from cluster_load where device_type='cpu' and device_name='cpu' and name='load1';
+-----+-----+-----+
| type | instance | name   | value |
+-----+-----+-----+
tidb	127.0.0.1:10080	load1	3.10
tidb	127.0.0.1:10081	load1	3.10
tidb	127.0.0.1:10082	load1	3.10
pd	127.0.0.1:2380	load1	3.10
pd	127.0.0.1:2381	load1	3.10
pd	127.0.0.1:2382	load1	3.10
tikv	127.0.0.1:20163	load1	3.1015625
tikv	127.0.0.1:20161	load1	3.1015625
tikv	127.0.0.1:20162	load1	3.1015625
tikv	127.0.0.1:20160	load1	3.1015625
+-----+-----+-----+
10 rows in set (0.55 sec)
```

- TYPE information_schema.cluster_info TYPE tidb/pd/tikv
- INSTANCE information_schema.cluster_info STATUS_ADDRESS
- DEVICE_TYPE cpu/memory/disk/net
- DEVICE_NAME DEVICE_TYPE
 - cpu cpu
 - disk
 - netNIC
 - memory memory
- NAME cpu load1/load5/load15 CPU 1min/5min/15min select name from cluster_load where device_type='cpu' group by name NAME
- VALUE CPU 1min/5min/15min

5.

information_schema.cluster_systeminfo sysctl TiDB fd

```
mysql> select type, instance, name, value from cluster_systeminfo where type='tidb' and name like '%fd%';
+-----+-----+-----+
| type | instance | name           | value |
+-----+-----+-----+
tidb	127.0.0.1:10080	net.inet6.ip6.maxifdefrouters	16
tidb	127.0.0.1:10080	net.necp.client_fd_count	89
tidb	127.0.0.1:10080	net.necp.observer_fd_count	0
tidb	127.0.0.1:10081	net.inet6.ip6.maxifdefrouters	16
tidb	127.0.0.1:10081	net.necp.client_fd_count	89
tidb	127.0.0.1:10081	net.necp.observer_fd_count	0
tidb	127.0.0.1:10082	net.inet6.ip6.maxifdefrouters	16
tidb	127.0.0.1:10082	net.necp.client_fd_count	89
tidb	127.0.0.1:10082	net.necp.observer_fd_count	0
+-----+-----+-----+
9 rows in set (0.04 sec)
```

- TYPE information_schema.cluster_info TYPE tidb/pd/tikv

- INSTANCE information_schema.cluster_info STATUS_ADDRESS
- SYSTEM_TYPE system
- SYSTEM_NAME SYSTEM_NAME sysctl
- NAMEsysctl
- VALUEsysctl

6.

information_schema.cluster_log TiDB grep warning

```
mysql> select * from cluster_log where level='warn'\G
***** 1. row *****
TIME: 2020/03/08 12:17:41.329
TYPE: pd
INSTANCE: 127.0.0.1:2382
LEVEL: WARN
MESSAGE: [grpclog.go:60] ["transport: http2Server.HandleStreams failed to read frame: read tcp 127.0.0.1:2382->127.0.0.1:57030: use of closed network connection"]
***** 2. row *****
TIME: 2020/03/08 12:17:41.338
TYPE: pd
INSTANCE: 127.0.0.1:2382
LEVEL: WARN
MESSAGE: [grpclog.go:60] ["transport: http2Server.HandleStreams failed to read frame: read tcp 127.0.0.1:2382->127.0.0.1:57029: use of closed network connection"]
***** 3. row *****
TIME: 2020/03/08 12:17:41.361
TYPE: pd
INSTANCE: 127.0.0.1:2382
LEVEL: WARN
MESSAGE: [grpclog.go:60] ["transport: http2Server.HandleStreams failed to read frame: read tcp 127.0.0.1:2382->127.0.0.1:57031: use of closed network connection"]
3 rows in set (0.01 sec)
```

- TIME
- TYPE information_schema.cluster_info TYPE tidb/pd/tikv
- INSTANCE information_schema.cluster_info INSTANCE
- LEVEL
- MESSAGE

```
select      from cluter_log where instance='tikv-1' tikv-1 message like regexp pattern regexp message
grep pipeline select      from cluster_log where message like 'coprocessor%' and message regexp '.slow.' grep
'coprocessor' xxx.log | grep -E '.slow.'
```

TiDB 4.0 TiDB 4.0 region id region txn id key

3.2

TiDB 4.0

TiDB 4.0 metrics_schema SQL SQL

3.2.1

tidb_query_duration COMMENT TiDB query P999P99P90

```
metrics_schema> show create table tidb_query_duration;
+-----+-----+
| Table          | Create Table
|               |
+-----+-----+
| tidb_query_duration | CREATE TABLE `tidb_query_duration` (
|               |
|               |   `time` datetime unsigned DEFAULT NULL,
|               |
|               |   `instance` varchar(512) DEFAULT NULL,
|               |
|               |   `sql_type` varchar(512) DEFAULT NULL,
|               |
|               |   `quantile` double unsigned DEFAULT NULL,
|               |
|               |   `value` double unsigned DEFAULT NULL
|               |
|               | ) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_bin COMMENT='The quantile of TiDB que
ry durations(second)'
+-----+-----+
-----+
```

P90 TiDB Query Select Query P90 0.0384 internal P90 0.00327instance TiDB

```
metrics_schema> select * from tidb_query_duration where value is not null and time=now() and quantile=0.90;
+-----+-----+-----+-----+-----+
| time          | instance      | sql_type | quantile | value      |
+-----+-----+-----+-----+-----+
| 2020-03-08 13:34:40 | 172.16.5.40:10089 | Select    | 0.9      | 0.0384     |
| 2020-03-08 13:34:40 | 172.16.5.40:10089 | internal  | 0.9      | 0.00327692307692 |
+-----+-----+-----+-----+-----+
```

3.2.2

information_schema.metrics_tables SQL select * from information_schema.metrics_tables

```
mysql> desc metrics_tables;
+-----+-----+-----+-----+-----+
| Field       | Type        | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+
TABLE_NAME	varchar(64)	YES		NULL	
PROMQL	varchar(64)	YES		NULL	
LABELS	varchar(64)	YES		NULL	
QUANTILE	double unsigned	YES		NULL	
COMMENT	varchar(256)	YES		NULL	
+-----+-----+-----+-----+-----+
5 rows in set (0.00 sec)
```

:

- TABLE_NAME metrics_schema
- PROMQL SQL PromQL prometheus SQL PromQL
- LABELS label label SQL PromQL
- QUANTILE 0
- COMMENT

3.2.3

PROMQL TiDB SQL PromQL prometheus

SQL MemTableScan PromQL start_time end_time step 1 prometheus range query HTTP API

```
metrics_schema> desc select * from tidb_query_duration where value is not null and time=now() and quantile=0.90;
+-----+-----+-----+
| id | estRows | task | operator info
+-----+-----+-----+
| Selection_5 | 8000.00 | root | not(isnull(Column#5))
|   |
|   +---+---+---+---+---+---+
|   |   +---+---+---+---+---+---+
|   |   |   +---+---+---+---+---+---+
|   |   |   |   +---+---+---+---+---+---+
|   |   |   |   |   +---+---+---+---+---+---+
|   |   |   |   |   |   +---+---+---+---+---+---+
|   |   |   |   |   |   |   +---+---+---+---+---+---+
|   |   |   |   |   |   |   |   +---+---+---+---+---+---+
|   |   |   |   |   |   |   |   |   +---+---+---+---+---+---+
|   |   |   |   |   |   |   |   |   |   +---+---+---+---+---+---+
|   |   |   |   |   |   |   |   |   |   |   +---+---+---+---+---+---+
```

SQL Where time 10

3.2.4 session

2 session session

- tidb_metric_query_step prometheus query_range startendstep step
- tidb_metric_query_range_duration PromQL PromQL \$RANGE_DURATION 60

range query prometheus query

- query=: PromQL
- start=:
- end=:
- step=: query resolution
- timeout=:

prometheus PromQL

- [start, end] start step
- start=10end=20step=2 ts=10ts=12ts=14ts=16ts=18ts=206 6 6

60

```
set @@session.tidb_metric_query_step=60
```

```
TiDB PD      prometheus      prometheus PD      prometheus      instance      sql_type      ['2020-03-08 13:23:00',
'2020-03-08 13:33:00')  P99 avg, max, min
```

```
metrics_schema> select instance,sql_type, avg(value),max(value),min(value) from tidb_query_duration where time >= '2020-03-08 13:23:00' and time < '2020-03-08 13:33:00' and value is not null and quantile=0.99 group by instance,sql_type;
+-----+-----+-----+-----+
| instance | sql_type | avg(value) | max(value) | min(value) |
+-----+-----+-----+-----+
172.16.5.40:10089	Select	0.00800917072138	0.00824108821892	0.00790462559176
172.16.5.40:10089	internal	0.012384	0.01554	0.0062
172.16.5.40:10089	Insert	0.00687276884265	0.0069763539823	0.00670463917526
172.16.5.40:10089	general	0.00092395833333	0.00133333333333	0.0006666666666667
+-----+-----+-----+-----+
```

3.3

1.

3.1 3.2 TiDB 4.0 SQL information_schema.cluster_config information_schema.cluster_info

SQL TiDB 4.0

- information_schema.inspection_result
- information_schema.inspection_summary

SQL select * from information_schema.inspection_result

 warning / critical details

2.

information_schema.inspection_result

```
mysql> desc inspection_result;
+-----+-----+-----+-----+-----+
| Field | Type      | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+
RULE	varchar(64)	YES		NULL	
ITEM	varchar(64)	YES		NULL	
TYPE	varchar(64)	YES		NULL	
INSTANCE	varchar(64)	YES		NULL	
VALUE	varchar(64)	YES		NULL	
REFERENCE	varchar(64)	YES		NULL	
SEVERITY	varchar(64)	YES		NULL	
DETAILS	varchar(256)	YES		NULL	
+-----+-----+-----+-----+-----+
8 rows in set (0.00 sec)
```

- RULE select * from inspection_rules where type='inspection'
 - config warning
 - version githash critical
 - current-load warning
 - critical-error warning
 - threshold-check
- ITEM
- TYPE tidb/tikv/pd
- INSTANCE
- VALUE
- REFERENCE VALUE
- SEVERITY warning/critical
- DETAILS SQL

```
mysql> select * from inspection_rules where type='inspection';
+-----+-----+-----+
| NAME      | TYPE      | COMMENT |
+-----+-----+-----+
config	inspection	
version	inspection	
current-load	inspection	
critical-error	inspection	
threshold-check	inspection	
+-----+-----+-----+
5 rows in set (0.00 sec)
```

3.

Coprocessor 8 Coprocessor CPU 750% workload

information_schema.inspection_summary

information_schema.inspection_summary

```
mysql> desc inspection_summary;
+-----+-----+-----+-----+-----+-----+
| Field     | Type      | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+-----+
RULE	varchar(64)	YES		NULL	
INSTANCE	varchar(64)	YES		NULL	
METRICS_NAME	varchar(64)	YES		NULL	
LABEL	varchar(64)	YES		NULL	
QUANTILE	double unsigned	YES		NULL	
AVG_VALUE	double(22,6) unsigned	YES		NULL	
MIN_VALUE	double(22,6) unsigned	YES		NULL	
MAX_VALUE	double(22,6) unsigned	YES		NULL	
+-----+-----+-----+-----+-----+-----+
8 rows in set (0.00 sec)
```

- RULE `select * from inspection_rules where type='summary'`
- INSTANCE
- METRIC_NAME
- QUANTILE QUANTILE `select * from inspection_summary where rule='ddl' and quantile in (0.80, 0.90, 0.99, 0.999)` `ddl 80/90/99/999 AVG_VALUE/MIN_VALUE/MAX_VALUE`

:

```
information_summary | SQL rule      select * from inspection_summary      select * from inspection_summary
where rule in ('read-link', 'ddl')  DDL
```

4.

```
hint      select /*+ time_range('2020-03-07 12:00:00', '2020-03-07 13:00:00') */ * from inspection_summary 2020-03-07
12:00:00 - 2020-03-07 13:00:00
```

```
mysql> SELECT
    t1.avg_value / t2.avg_value AS ratio,
    t1.*,
    t2.*
FROM
(
    SELECT
        /*+ time_range("2020-01-16 16:00:54.933", "2020-01-16 16:10:54.933")*/ *
        FROM inspection_summary WHERE rule='read-link'
) t1
JOIN
(
    SELECT
        /*+ time_range("2020-01-16 16:10:54.933", "2020-01-16 16:20:54.933")*/ *
        FROM inspection_summary WHERE rule='read-link'
) t2
ON t1.metrics_name = t2.metrics_name
and t1.instance = t2.instance
and t1.label = t2.label
ORDER BY
ratio DESC;
```

3.4

```
TiDB TiDB 4.0      information_schema.metrics_summary    information_schema.metrics_summary_by_label
information_schema.metrics_summary_by_label    label
```

1.

```
[ '2020-03-08 13:23:00', '2020-03-08 13:33:00' ] TiDB 3      information_schema.metrics_summary /*+
time_range() */ hint SQL
```

```
mysql> select /*+ time_range('2020-03-08 13:23:00','2020-03-08 13:33:00') */ *
       from information_schema.`METRICS_SUMMARY`
      where metrics_name like 'tidb%duration'
        and avg_value > 0
        and quantile = 0.99
     order by avg_value desc
     limit 3\G
*****
[ 1. row ]*****
METRICS_NAME | tidb_get_token_duration
QUANTILE    | 0.99
SUM_VALUE   | 8.972509
AVG_VALUE   | 0.996945
MIN_VALUE   | 0.996515
MAX_VALUE   | 0.997458
COMMENT     | The quantile of Duration (us) for getting token, it should be small until concurrency limit is reached(second)
*****
[ 2. row ]*****
METRICS_NAME | tidb_query_duration
QUANTILE    | 0.99
SUM_VALUE   | 0.269079
AVG_VALUE   | 0.007272
MIN_VALUE   | 0.000667
MAX_VALUE   | 0.01554
COMMENT     | The quantile of TiDB query durations(second)
*****
[ 3. row ]*****
METRICS_NAME | tidb_kv_request_duration
QUANTILE    | 0.99
SUM_VALUE   | 0.170232
AVG_VALUE   | 0.004601
MIN_VALUE   | 0.000975
MAX_VALUE   | 0.013
COMMENT     | The quantile of kv requests durations by store
```

```
tidb_get_token_duration COMMENT us
metrics_summary_by_label
```

```

mysql> select /*+ time_range('2020-03-08 13:23:00','2020-03-08 13:33:00') */ *
       from information_schema.`METRICS_SUMMARY_BY_LABEL`*
      where metrics_name like 'tidb%duration'
        and avg_value > 0
        and quantile = 0.99
      order by avg_value desc
     limit 10\G
*****
[ 1. row ]*****
INSTANCE | 172.16.5.40:10089
METRICS_NAME | tidb_get_token_duration
LABEL |
QUANTILE | 0.99
SUM_VALUE | 8.972509
AVG_VALUE | 0.996945
MIN_VALUE | 0.996515
MAX_VALUE | 0.997458
COMMENT | The quantile of Duration (us) for getting token, it should be small until concurrency limit is reached(second)
*****
[ 2. row ]*****
INSTANCE | 172.16.5.40:10089
METRICS_NAME | tidb_query_duration
LABEL | Select
QUANTILE | 0.99
SUM_VALUE | 0.072083
AVG_VALUE | 0.008009
MIN_VALUE | 0.007905
MAX_VALUE | 0.008241
COMMENT | The quantile of TiDB query durations(second)
*****
[ 3. row ]*****
INSTANCE | 172.16.5.40:10089
METRICS_NAME | tidb_query_duration
LABEL | Rollback
QUANTILE | 0.99
SUM_VALUE | 0.072083
AVG_VALUE | 0.008009
MIN_VALUE | 0.007905
MAX_VALUE | 0.008241
COMMENT | The quantile of TiDB query durations(second)

```

metrics_summary_by_label metrics_summary LABEL , 2,3 tidb_query_duration Select / Rollback

2.

t1 baseline

- t1 : ("2020-03-03 17:08:00", "2020-03-03 17:11:00")
- t2 : ("2020-03-03 17:18:00", "2020-03-03 17:21:00")

METRICS_NAME join /*+ time_range() */ hint

1. t1.avg_value / t2.avg_value 10

```

mysql> SELECT
    t1.avg_value / t2.avg_value AS ratio,
    t1.metrics_name,
    t1.avg_value,
    t2.avg_value,
    t2.comment
  FROM
  (
    SELECT /*+ time_range("2020-03-03 17:08:00", "2020-03-03 17:11:00")*/
    *
   FROM information_schema.metrics_summary
  ) t1
 JOIN
  (
    SELECT /*+ time_range("2020-03-03 17:18:00", "2020-03-03 17:21:00")*/
    *
   FROM information_schema.metrics_summary
  ) t2
 ON t1.metrics_name = t2.metrics_name
 ORDER BY
    ratio DESC limit 10;
+-----+-----+-----+-----+
| ratio      | metrics_name          | avg_value     | avg_value     | comment
+-----+-----+-----+-----+
| 17.6439787379 | tikv_region_average_written_bytes | 30816827.0953 | 1746591.71568 | The average rate of writing bytes to Regions per TiKV instance
| 8.88407551364 | tikv_region_average_written_keys | 108557.034612 | 12219.283193 | The average rate of written keys to Regions per TiKV instance
| 6.4105335594  | tidb_kv_write_num          | 4493.293654   | 700.923505   | The quantile of kv write times per transaction execution
| 2.99993333333 | tidb_gc_total_count        | 1.0           | 0.333341    | The total count of kv storage garbage collection time durations
| 2.66412165823 | tikv_engine_avg_seek_duration | 6569.879007  | 2466.05818   | The time consumed when executing seek operation, the unit is microsecond
| 2.66412165823 | tikv_engine_max_seek_duration | 6569.879007  | 2466.05818   | The time consumed when executing seek operation, the unit is microsecond
| 2.49994444321 | tikv_region_change          | -0.277778    | -0.111114   | The count of region change per TiKV instance
| 2.16063829787 | etcd_wal_fsync_duration    | 0.002539     | 0.001175    | The quantile time consumed of writing WAL into the persistent storage
| 2.06089264604 | node_memory_free            | 4541448192.0 | 2203631616.0 |
| 1.96749064186 | tidb_kv_write_size          | 514489.28    | 261495.159902 | The quantile of kv write size per transaction execution
+-----+-----+-----+-----+

```

- t1 tikv_region_average_written_bytes region t2 17.6
- t1 tikv_region_average_written_keys region keys t2 8.8
- t1 tidb_kv_write_size tidb_kv t2 1.96

t1 t2

1. t2.avg_value / t1.avg_value 10

```

mysql> SELECT
    t2.avg_value / t1.avg_value AS ratio,
    t1.metrics_name,
    t1.avg_value,
    t2.avg_value,
    t2.comment
  FROM
  (
    SELECT /*+ time_range("2020-03-03 17:08:00", "2020-03-03 17:11:00")*/
    *
   FROM information_schema.metrics_summary
  ) t1
 JOIN
  (
    SELECT /*+ time_range("2020-03-03 17:18:00", "2020-03-03 17:21:00")*/
    *
   FROM information_schema.metrics_summary
  ) t2
 ON t1.metrics_name = t2.metrics_name
 ORDER BY
    ratio DESC limit 10;
+-----+-----+-----+-----+
| ratio      | metrics_name          | avg_value     | avg_value     | comment
+-----+-----+-----+-----+
| 5865.59537065 | tidb_slow_query_cop_process_total_time | 0.016333 | 95.804724 | The total time of Ti
DB slow query statistics with slow query total cop process time(second) |
| 3648.74109023 | tidb_distsql_partial_scan_key_total_num | 10865.666667 | 39646004.4394 | The total num of dis
tsql partial scan key numbers
| 267.002351165 | tidb_slow_query_cop_wait_total_time | 0.003333 | 0.890008 | The total time of Ti
DB slow query statistics with slow query total cop wait time(second) |
| 192.43267836 | tikv_cop_total_response_total_size | 2515333.66667 | 484032394.445 | |
| 192.43267836 | tikv_cop_total_response_size        | 41922.227778 | 8067206.57408 | |
| 152.780296296 | tidb_distsql_scan_key_total_num    | 5304.333333 | 810397.618317 | The total num of dis
tsql scan numbers
| 126.042290167 | tidb_distsql_execution_total_time | 0.421622 | 53.142143 | The total time of di
stsql execution(second)
105.164020657	tikv_cop_scan_details	134.450733	14139.379665	
105.164020657	tikv_cop_scan_details_total	8067.043981	848362.77991	
101.635495394	tikv_cop_total_kv_cursor_operations	1070.875	108838.91113	
+-----+-----+-----+-----+

```

- t2 tidb_slow_query_cop_process_total_time tidb cop process t1 5865
- t2 tidb_distsql_partial_scan_key_total_num tidb distsql key t1 3648
- t2 tikv_cop_total_response_size tikv cop t1 192
- t2 tikv_cop_scan_detailstikv cop scan t1 105

2 t2 Cop t2 TiKV Copprocessor cop task t2

t1 ~ t2 [go-ycsb](#) t2 20 tpch tpch cop

3.5 SQL

TiDB slow-threshold 300 slow-query-file "tidb-slow.log" SQL

TiDB enable-slow-log

1.

```
# Time: 2019-08-14T09:26:59.487776265+08:00
# Txn_start_ts: 410450924122144769
# User: root@127.0.0.1
# Conn_ID: 3086
# Query_time: 1.527627037
# Parse_time: 0.000054933
# Compile_time: 0.000129729
# Process_time: 0.07 Wait_time: 0.002 Backoff_time: 0.002 Request_count: 1 Total_keys: 131073 Process_keys: 131072 Pr
ewrite_time: 0.335415029 Commit_time: 0.032175429 Get_commit_ts_time: 0.000177098 Local_latch_wait_time: 0.106869448
Write_keys: 131072 Write_size: 3538944 Prewrite_region: 1
# DB: test
# Is_internal: false
# Digest: 50a2e32d2abbd6c1764b1b7f2058d428ef2712b029282b776beb9506a365c0f1
# Stats: t:414652072816803841
# Num_cop_tasks: 1
# Cop_proc_avg: 0.07 Cop_proc_p90: 0.07 Cop_proc_max: 0.07 Cop_proc_addr: 172.16.5.87:20171
# Cop_wait_avg: 0 Cop_wait_p90: 0 Cop_wait_max: 0 Cop_wait_addr: 172.16.5.87:20171
# Mem_max: 525211
# Succ: true
# Plan_digest: e5f9d9746c756438a13c75ba3eedf601eecf555cdb7ad327d7092bdd041a83e7
# Plan: tidb_decode_plan('ZJAwCTMyXzcJMAkyMALkYXRh01RhYmx1U2Nhb182CjEJMTBfNgkxAR0AdAEY1Dp0LCByYW5nZTpblWluZiwrw5mXSw
ga2VlcCBvcmRlcjpmYWxzzSwgc3RhdHM6cHNldWRvCg==')
insert into t select * from t;
```

2.

“”

(1) Slow Query

- Time
- Query_time
- Parse_time
- Compile_time
- Digest SQL
- Stats table pseudo
- Txn_start_ts ID TiDB
- Is_internal TiDB SQL true TiDB SQL false SQL
- Index_ids ID
- Succ
- Backoff_time lockRegion tikv server is busy
- Plan_digest plan
- Plan select tidb_decode_plan('xxx...') SQL
- Query SQL Query Query

(2)

- Prewrite_time prewrite
- Commit_time commit

- Get_commit_ts_timecommit commit
- Local_latch_wait_timecommit TiDB
- Write_keys TiKV Write CF Key
- Write_size key value
- Prewrite_regionprewrite TiKV Region Region

(3)

- Memory_max TiDB byte

(4) SQL

- User
- Conn_ID ID con:3 TiDB
- DB database

(5) TiKV Coprocessor Task

- Process_time SQL TiKV TiKV Query_time
- Wait_time TiKV TiKV Coprocessor Coprocessor
- Request_count Coprocessor
- Total_keys Coprocessor key
- Process_keys Coprocessor key total_keys processed_keys MVCC processed_keys total_keys
- Cop_proc_avgcop-task
- Cop_proc_p90cop-task P90
- Cop_proc_maxcop-task
- Cop_proc_addr cop-task
- Cop_wait_avgcop-task
- Cop_wait_p90cop-task P90
- Cop_wait_maxcop-task
- Cop_wait_addr cop-task

3.

```
INFORMATION_SCHEMA.SLOW_QUERY    TiDB  Information Schema      SLOW_QUERY
TiDB 4.0  CLUSTER_SLOW_QUERY    TiDB  SLOW_QUERY  SLOW_QUERY
TiDB 4.0  SLOW_QUERY  rotate TIME 4.0
|  SLOW_QUERY TiDB
```

4. SLOW_QUERY

(1) Top N

```
Top 2  is_internal=false  TiDB
```

```
select query_time, query
from information_schema.slow_query
where is_internal = false -- TiDB SQL
order by query_time desc
limit 2;
```

```
+-----+-----+
| query_time | query
+-----+-----+
| 12.77583857 | select * from t_slim, t_wide where t_slim.c0=t_wide.c0; |
| 0.734982725 | select t0.c0, t1.c1 from t_slim t0, t_wide t1 where t0.c0=t1.c0; |
+-----+-----+
```

(2) Top N

test SQL 2

```
select query_time, query, user
from information_schema.cluster_slow_query
where is_internal = false -- TiDB SQL
  and user = "test"      --
order by query_time desc
limit 2;
```

```
+-----+-----+
| Query_time | query                                | user   |
+-----+-----+
| 0.676408014 | select t0.c0, t1.c1 from t_slim t0, t_wide t1 where t0.c0=t1.c1; | test  |
+-----+-----+
```

(3) SQL

Top N SQL SQL SQL Top N SQL

```
select query_time, query, digest
from information_schema.cluster_slow_query
where is_internal = false
order by query_time desc
limit 1;
```

```
+-----+-----+
| query_time | query                                | digest
+-----+-----+
| 0.302558006 | select * from t1 where a=1; | 4751cb6008fda383e22dacb601fde85425dc8f8cf669338d55d944bafb46a6fa |
+-----+-----+
```

SQL

```
select query, query_time
from information_schema.cluster_slow_query
where digest = "4751cb6008fda383e22dacb601fde85425dc8f8cf669338d55d944bafb46a6fa";
```

```
+-----+-----+
| query          | query_time |
+-----+-----+
| select * from t1 where a=1; | 0.302558006 |
| select * from t1 where a=2; | 0.401313532 |
+-----+-----+
```

(4) pseudo SQL

```
select query, query_time, stats
from information_schema.cluster_slow_query
where is_internal = false
and stats like '%pseudo';
```

| query | query_time | stats |
|-----------------------------|-------------|----------------------------------|
| select * from t1 where a=1; | 0.302558006 | t1:pseudo |
| select * from t1 where a=2; | 0.401313532 | t1:pseudo |
| select * from t1 where a>2; | 0.602011247 | t1:pseudo |
| select * from t1 where a>3; | 0.50077719 | t1:pseudo |
| select * from t1 join t2; | 0.931260518 | t1:407872303825682445, t2:pseudo |

(5)

SQL SQL SQL

```
select count(distinct plan_digest) as count, digest,min(query)
from cluster_slow_query
group by digest
having count>1 limit 3\G
```

```
*****[ 1. row ]*****
count | 2
digest | 17b4518fde82e32021877878bec2bb309619d384fc944106fc9c93b536e94
min(query) | SELECT DISTINCT c FROM sbtest25 WHERE id BETWEEN ? AND ? ORDER BY c [arguments: (291638, 291737)];
*****[ 2. row ]*****
count | 2
digest | 9337865f3e2ee71c1c2e740e773b6dd85f23ad00f8fa1f11a795e62e15fc9b23
min(query) | SELECT DISTINCT c FROM sbtest22 WHERE id BETWEEN ? AND ? ORDER BY c [arguments: (215420, 215519)];
*****[ 3. row ]*****
count | 2
digest | db705c89ca2dfc1d39d10e0f30f285cbbadec7e24da4f15af461b148d8ffb020
min(query) | SELECT DISTINCT c FROM sbtest11 WHERE id BETWEEN ? AND ? ORDER BY c [arguments: (303359, 303458)];
```

SQL plan

```
select min(plan),plan_digest
from cluster_slow_query where digest='17b4518fde82e32021877878bec2bb309619d384fc944106fc9c93b536e94'
group by plan_digest\G
```

```
***** 1. row *****
min(plan): Sort_6          root 100.00131380758702      sbtest.sbtest25.c:asc
  └HashAgg_10           root 100.00131380758702      group by:sbtest.sbtest25.c, funcs:firstrow(sbtest.sbt
est25.c)->sbtest.sbtest25.c
    └TableReader_15      root 100.00131380758702      data:TableRangeScan_14
      └TableScan_14       cop 100.00131380758702      table:sbtest25, range:[502791,502890], keep order:fal
se
plan_digest: 6afbbd21f60ca6c6fdf3d3cd94f7c7a49dd93c00fcf8774646da492e50e204ee
***** 2. row *****
min(plan): Sort_6          root 1                      sbtest.sbtest25.c:asc
  └HashAgg_12           root 1                      group by:sbtest.sbtest25.c, funcs:firstrow(sbtest.sbt
est25.c)->sbtest.sbtest25.c
    └TableReader_13      root 1                      data:HashAgg_8
      └HashAgg_8          cop 1                      group by:sbtest.sbtest25.c,
        └TableScan_11       cop 1.2440069558121831   table:sbtest25, range:[472745,472844], keep order:fal
se
```

(6) TIDB

```
select instance, count(*) from information_schema.cluster_slow_query where time >= "2020-03-06 00:00:00" and time < now() group by instance;
```

| instance | count(*) |
|---------------|----------|
| 0.0.0.0:10081 | 124 |
| 0.0.0.0:10080 | 119771 |

(7)

2020-03-10 13:24:00 ~ 2020-03-10 13:27:00 QPS SQL 2020-03-10 13:20:00 ~ 2020-03-10 13:23:00

```
SELECT * FROM
  (SELECT /*+ AGG_TO_COP(), HASH_AGG() */ count(*),
  min(time),
  sum(query_time) AS sum_query_time,
  sum(Process_time) AS sum_process_time,
  sum(Wait_time) AS sum_wait_time,
  sum(Commit_time),
  sum(Request_count),
  sum(process_keys),
  sum(write_keys),
  max(Cop_proc_max),
  min(query),min(prev_stmt),
  digest
  FROM information_schema.CLUSTER_SLOW_QUERY
  WHERE time >= '2020-03-10 13:24:00'
    AND time < '2020-03-10 13:27:00'
    AND Is_internal = false
  GROUP BY digest) AS t1
WHERE t1.digest NOT IN
  (SELECT /*+ AGG_TO_COP(), HASH_AGG() */ digest
  FROM information_schema.CLUSTER_SLOW_QUERY
  WHERE time >= '2020-03-10 13:20:00'
    AND time < '2020-03-10 13:23:00'
  GROUP BY digest)
ORDER BY t1.sum_query_time DESC limit 10\G
```

```
*****[ 1. row ]*****
count(*) | 200
min(time) | 2020-03-10 13:24:27.216186
sum_query_time | 50.114126194
sum_process_time | 268.351
sum_wait_time | 8.476
sum(Commit_time) | 1.044304306
sum(Request_count) | 6077
sum(process_keys) | 202871950
sum(write_keys) | 319500
max(Cop_proc_max) | 0.263
min(query) | delete from test.tcs2 limit 5000;
min(prev_stmt) |
digest | 24bd6d8a9b238086c9b8c3d240ad4ef32f79ce94cf5a468c0b8fe1eb5f8d03df
```

5. TiDB

TiDB 4.0 session tidb_slow_query_file INFORMATION_SCHEMA.SLOW_QUERY

```
set tidb_slow_query_file = "/path-to-log/tidb-slow.log"
```

TiDB 4.0 session

6. pt-query-digest TiDB

pt-query-digest TiDB

pt-query-digest 3.0.13

```
pt-query-digest --report tidb-slow.log
```

```
# 320ms user time, 20ms system time, 27.00M rss, 221.32M vsz
# Current date: Mon Mar 18 13:18:51 2019
# Hostname: localhost.localdomain
# Files: tidb-slow.log
# Overall: 1.02k total, 21 unique, 0 QPS, 0x concurrency
# Time range: 2019-03-18-12:22:16 to 2019-03-18-13:08:52
# Attribute      total     min      max      avg      95%    stddev   median
# ======      ======  ======  ======  ======  ======  ======  ======
# Exec time     218s    10ms     13s    213ms    30ms     1s    19ms
# Query size   175.37k      9    2.01k   175.89   158.58   122.36   158.58
# Commit time    46ms      2ms      7ms      3ms      7ms     1ms     3ms
# Conn ID        71       1       16      8.88    15.25     4.06    9.83
# Process keys  581.87k      2  103.15k   596.43   400.73   3.91k   400.73
# Process time    31s      1ms     10s     32ms     19ms   334ms    16ms
# Request coun   1.97k      1       10      2.02     1.96     0.33    1.96
# Total keys    636.43k      2  103.16k   652.35   793.42   3.97k   400.73
# Txn start ts  374.38E      0  16.00E   375.48P   1.25P   89.05T   1.25P
# Wait time     943ms      1ms     19ms      1ms      2ms     1ms    972us
```

•

SLOW_QUERY process_time wait_time process_time

7. admin show slow

TiDB admin show slow SQL

: TiDB

```
admin show slow recent N;
admin show slow top [internal | all] N;
```

recent N N

```
admin show slow recent 10;
```

top N internal SQL all SQL SQL

N 30 7

3 SQL

```
admin show slow top 3;
```

N N

| | |
|----------------|--------------|
| | |
| start | SQL |
| duration | SQL |
| details | |
| succ | SQL 1: 0: |
| conn_id | session ID |
| transcation_ts | commit ts |
| user | |
| db | SQL database |
| table_ids | SQL ID |
| index_ids | SQL ID |
| internal | TiDB SQL |
| digest | SQL |
| sql | SQL |

3.6 processlist

processlist TiDB

3.6.1 processlist

3 processlist

| 1. TiDB | Info | FULL | |
|--|---------------------|-------------|----------------|
| SHOW [FULL] PROCESSLIST | | | |
| 2. INFORMATION_SCHEMA.PROCESSLIST | show processlist | MEM | |
| TxnStart | | | |
| * `MEM` byte v3.0.5
* `TxnStart` v4.0.0 | | | |
| SELECT * FROM INFORMATION_SCHEMA.PROCESSLIST | | | |
| TiDB | PROCESSLIST | TiDB | |
| 3. TiDB 4.0 | CLUSTER_PROCESSLIST | TiDB | PROCESSLIST |
| PROCESSLIST | CLUSTER_PROCESSLIST | PROCESSLIST | INSTANCE |
| INSTANCE | TiDB | | |
| SELECT * FROM INFORMATION_SCHEMA.CLUSTER_PROCESSLIST | | | |
| | root | ProcessPriv | User User User |

3.6.2 KILL [TiDB]

KILL TiDB TiDB

KILL

- TiDB KILL xxx MySQL TiDB TiDB KILL TiDB xxx TiDB compatible-kill-query = true TiDB
- Ctrl+C MySQL KILL TiDB TiDB KILL KILL compatible-kill-query TiDB compatible-kill-query = true

KILL

```
SHOW PROCESSLIST;
```

```
+-----+-----+-----+-----+-----+-----+
| Id   | User    | Host     | db      | Command   | Time    | State   | Info
+-----+-----+-----+-----+-----+-----+
| 5    | root    | 127.0.0.1 | <null> | Query    | 0        | 2       | show full processlist |
+-----+-----+-----+-----+-----+-----+
```

```
KILL TiDB 5;
```

```
Query OK, 0 rows affected (0.00 sec)
```

ID 5 kill

3.6.3 PROCESSLIST

```
SHOW FULL PROCESSLIST;
```

```
+-----+-----+-----+-----+-----+-----+
| Id   | User    | Host     | db      | Command   | Time    | State   | Info
+-----+-----+-----+-----+-----+-----+
| 1    | root    | 127.0.0.1 | <null> | Sleep    | 127     | 2       | <null>
| 3    | root    | 127.0.0.1 | <null> | Sleep    | 75      | 2       | <null>
| 5    | root    | 127.0.0.1 | <null> | Query    | 0        | 2       | show full processlist |
+-----+-----+-----+-----+-----+-----+
```

- Id TiDB KILL TiDB {ID} Id
- User
- Host
- db null
- CommandSleepQuery
- Time
- StatesqlSending dataSorting for groupCreating tmp tableLocked
- Info NULL

3.6.4 INFORMATION_SCHEMA.PROCESSLIST

```
select * from INFORMATION_SCHEMA.PROCESSLIST
```

```
+-----+-----+-----+-----+-----+-----+
| ID   | USER    | HOST     | DB      | COMMAND   | TIME    | STATE   | INFO
+-----+-----+-----+-----+-----+-----+
| MEM  | TxnStart |          |          |           |         |         |
|-----+-----+-----+-----+-----+-----+
| 9    | root    | 127.0.0.1 | <null> | Query    | 0        | 2       | select * from INFORMATION_SCHEMA.PROCESSLIST |
| 0    |          |          |          |           |         |         |
| 7    | root    | 127.0.0.1 | <null> | Sleep    | 51      | 2       | <null>
| 0    |          |          |          |           |         |         |
+-----+-----+-----+-----+-----+-----+
```

```
show full processlist state Sleep select PROCESSLIST
```

```
-- Sleep
select *
from information_schema.processlist
where command != 'Sleep'
order by time desc
```

info SQL

| ID | USER | HOST | DB | COMMAND | TIME | STATE | INFO |
|----|------|-----------|--------|---------|------|-------|---|
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| 10 | root | 127.0.0.1 | <null> | Query | 0 | 2 | select * from information_schema.processlist where command != 'Sleep' order by time desc 4588 |

sql EXPLAIN ANALYZE SQL EXPLAIN ANALYZE EXPLAIN SQL

TiDB 10 info SQL

```
-- Sleep top10
select *
from information_schema.processlist
where command != 'Sleep'
order by mem desc
limit 10
```

| ID | USER | HOST | DB | COMMAND | TIME | STATE | INFO |
|----|------|-----------|--------|---------|------|-------|---|
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| 10 | root | 127.0.0.1 | <null> | Query | 0 | 2 | select * from information_schema.processlist where command != 'Sleep' order by time desc 4588 |

sql EXPLAIN ANALYZE SQL

command time ID kill ID kill concat() kill

```
-- 2 sleep kill
select concat('kill ', 'TiDB', id, ';')
from information_schema.processlist
where command != 'Sleep'
and time > 2*60
order by time desc
```

```
-- 2 sleep kill
select *
from information_schema.processlist
where command != 'Sleep'
and time > 2*60
order by time desc
```

| ID | USER | HOST | DB | COMMAND | TIME | STATE | INFO | MEM | TxnStart |
|----|------|------|----|---------|------|-------|------|-----|----------|
|----|------|------|----|---------|------|-------|------|-----|----------|

3.6.5 INFORMATION_SCHEMA.CLUSTER_PROCESSLIST

```
select * from INFORMATION_SCHEMA.CLUSTER_PROCESSLIST
```

| INSTANCE | ID | USER | HOST | DB | COMMAND | TIME | STATE | INFO |
|---|----|------|------|----|---------|------|-------|----------|
| | | | | | | | | MEM |
| | | | | | | | | TxnStart |
| 0.0.0.0:10080 9 root 127.0.0.1 <null> Sleep 1824 2 <null> | | | | | | | | |
| 0 | | | | | | | | |
| 0.0.0.0:10080 10 root 127.0.0.1 <null> Sleep 1439 2 <null> | | | | | | | | |
| 0 | | | | | | | | |
| 0.0.0.0:10081 7 root 127.0.0.1 <null> Query 0 2 select * from INFORMATION_SCHEMA.CLUSTER_PROCESSLIST 0 03-08 13:42:02.850(415143329228390401) | | | | | | | | |

TiDB 2

```
select instance, count(*)
from INFORMATION_SCHEMA.CLUSTER_PROCESSLIST
where command != 'Sleep'
and time > 2*60
group by instance;
```

| instance | count(*) |
|-------------------|----------|
| 0.0.0.0:10081 1 | |
| 0.0.0.0:10080 3 | |

`EXPLAIN ANALYZE sql`

3.6.6 MySQL

- KILL TiDB TiDB TiDB `compatible-kill-query = true`
- TiDB State TiDB GO
- TiDB show processlist MySQL show processlist ID session ID TiDB show processlist MySQL show processlist
 - TiDB tidb-server SQL TiDB MySQL tidb-servershow processlist tidb-server session session MySQL show processlist MySQL SQL

3.7 Statement Summary

2.2 TiDB Dashboard Statements Statements Statement Summary

3.7.1 Statement Summary

Statement Summary "Dynamic Performance Views" "Profile Tables" "SQL Audit" SQL

SQL TiDB 4.0 Statement Summary Tables MySQL

Statement Summary SQL TiDB

3.7.1.1 “SQL”

SQL SQL SQL

```
select * from `order` where item_id=1000;  
  
SELECT *  
FROM `order`  
WHERE item_id=1001;
```

SQL SQL SQL

SQL SQL

IndexLookup TableScan

3.7.1.2

TiDB 4.0 60

- SQL
- TiDB
- TiKV CopTask
-

slow log

```

mysql> select * from events_statements_summary_by_digest limit 1\G
***** 1. row ****
SUMMARY_BEGIN_TIME: 2020-03-04 13:00:00
SUMMARY_END_TIME: 2020-03-04 13:30:00
STMT_TYPE: select
SCHEMA_NAME: test
DIGEST: d8cc0047ec3514d418e6f425d6203966de0094f025dab14babab8f4db0947736
DIGEST_TEXT: select buyer_id , item_id from order where order_id = ?
TABLE_NAMES: test.order
INDEX_NAMES: NULL
SAMPLE_USER: root
EXEC_COUNT: 8
SUM_LATENCY: 2591978
MAX_LATENCY: 1345860
MIN_LATENCY: 135860
AVG_LATENCY: 323997
AVG_PARSE_LATENCY: 69993
MAX_PARSE_LATENCY: 78761
AVG_COMPILE_LATENCY: 137604
MAX_COMPILE_LATENCY: 881880
-----
AVG_MEM: 0
MAX_MEM: 0
AVG_AFFECTED_ROWS: 0
FIRST_SEEN: 2020-03-04 13:20:55
LAST_SEEN: 2020-03-04 13:21:24
QUERY_SAMPLE_TEXT: select buyer_id, item_id from `order` where order_id=1001221
PREV_SAMPLE_TEXT:
PLAN_DIGEST: 28ccfb38e96b6e4aab31e82d12f349eb4edcb97bed3be49cb3f73051f2cd9d2
PLAN:      Point_Get_1      root      1      table:order, handle:1001221
1 row in set (0.00 sec)

```

3.7.1.3

events_statements_summary_by_digest

TiDB events_statements_summary_by_digest_history events_statements_summary_by_digest
 SUMMARY_BEGIN_TIME SUMMARY_END_TIME

24

3.7.2

SQL Statement Summary

3.7.2.1

SQL 10ms Grafana

QUERY_SAMPLE_TEXT

```

mysql> select avg_latency, query_sample_text from events_statements_summary_by_digest where QUERY_SAMPLE_TEXT LIKE 's
elect buyer_id, item_id from `order`%\n'
***** 1. row ****
avg_latency: 202225
query_sample_text: select buyer_id, item_id from `order` where order_id=1001221
1 row in set (0.00 sec)

```

avg_latency 0.2ms 10ms

3.7.2.2

SQL

```

mysql> select sum_latency, query_sample_text, digest
->     from events_statements_summary_by_digest
->     where summary_begin_time='2020-3-1 3:00:00'
->     order by sum_latency desc limit 3\G
***** 1. row ****
sum_latency: 120663574508
query_sample_text: select buyer_id, item_id from `order` where order_id=1001221
digest: d8cc0047ec3514d418e6f425d6203966de0094f025dab14babab8f4db0947736
***** 2. row ****
sum_latency: 97252462621
query_sample_text: select count(1) from item
digest: 1626b9d694faefc4e88ae9fd5e8917e85ed26fe62dbe781eb65edad3aa939ae8
***** 3. row ****
sum_latency: 65914323442
query_sample_text: select count(1) from buyer
digest: b07c73323c511d3c18407771adb8aec865c0409f4e9a4f53baedc409fd5a1cd0
3 rows in set (0.00 sec)

```

SQL

digest SQL ID digest

3.7.2.3

SQL 10 2

SQL

```

mysql> select abnormal.avg_latency/normal.avg_latency,
->     abnormal.avg_process_time/normal.avg_process_time,
->     abnormal.avg_total_keys/normal.avg_total_keys,
->     abnormal.avg_wait_time/normal.avg_wait_time
->     from events_statements_summary_by_digest_history abnormal,
->     events_statements_summary_by_digest_history normal
->     where normal.summary_begin_time='2020-3-1 10:00:00'
->     and abnormal.summary_begin_time='2020-3-1 14:00:00'
->     and abnormal.digest=normal.digest
->     and normal.digest = 'd8cc0047ec3514d418e6f425d6203966de0094f025dab14babab8f4db0947736'\G
***** 1. row ****
abnormal.avg_latency/normal.avg_latency: 6.3433
abnormal.avg_process_time/normal.avg_process_time: 8.9993
abnormal.avg_total_keys/normal.avg_total_keys: 12.6666
abnormal.avg_wait_time/normal.avg_wait_time: 0.9883
1 row in set, 1 warning (0.00 sec)

```

avg_total_keys TiKV avg_process_time TiKV GC

3.7.3

Statement Summary

- tidb_enable_stmt_summary
- tidb_stmt_summary_refresh_interval
- tidb_stmt_summary_history_size
- max-stmt-count SQL
- max-sql-length SQL

3.7.4 FAQ

Statement Summary

3.7.4.1

Q SQL tidb_stmt_summary_history_size max-stmt-count

A Statement Summary Tables TiDB server

3.7.4.2 commit

Q TiDB commit commit

A commit prev_sample_text digest commit commit

3.7.4.3 schema_name

QSQL schema_name

A schema_name schema schema use db db table_names "{schema}.{table}" schema table_names

3.7.4.4

QStatement Summary SQL

ASysbench TPCC 2% TiDB 2% 4.0

4 TiDB

TiDB

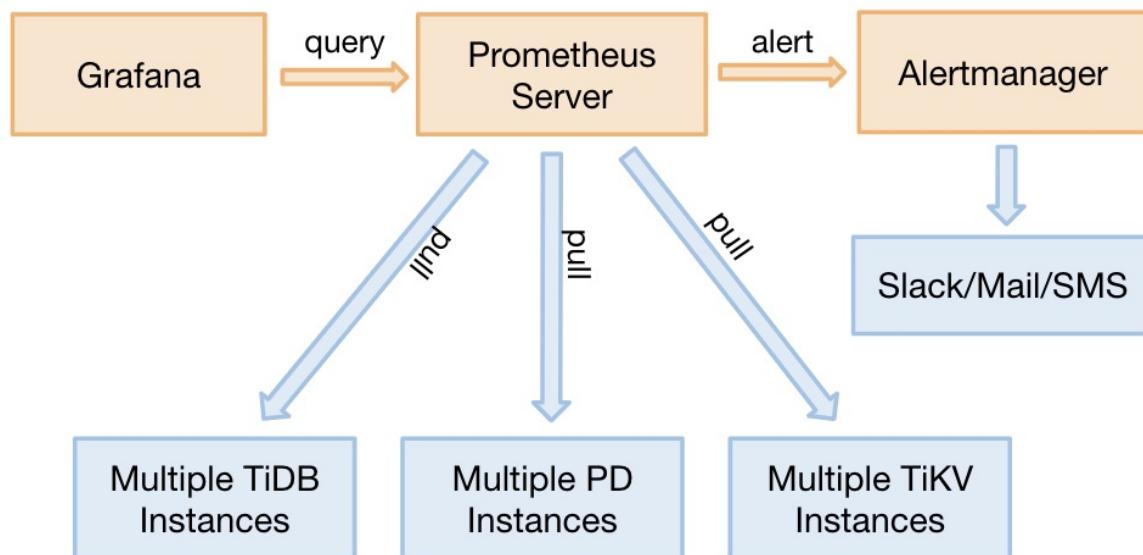
TiDB Prometheus Grafana

Prometheus TiDB

Prometheus Prometheus

Prometheus TiDB

- Prometheus Server
- Client Metric
- Alertmanager



Grafana TiDB

Grafana metric TiDB Grafana TiDB

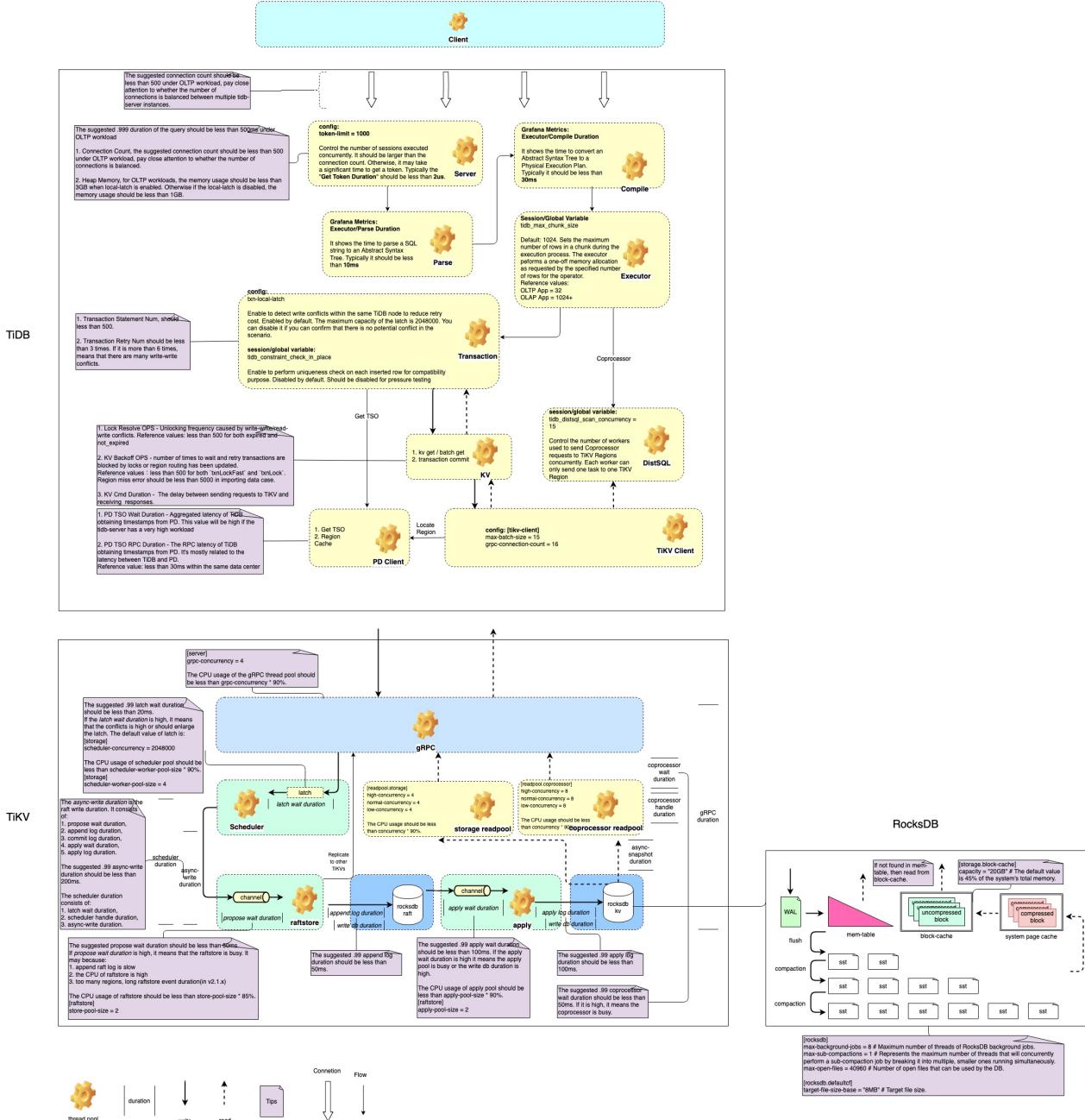


TiDB TiDB TiDB

4.1

TiDB

TiDB TiDB SQL



1. TiDB

- TiDB Server MySQL session token Connection Count TiDB TiDB 500
- session TiDB SQL Parse TiDB Executor/Parse Duration 10ms
- TiDB CompileExecutor/Compile Duration
- Executor
 - DML TiDB Transaction Commit TiKV

- TiDB DistSQL TiKV region
- TiDB TiKV KV Request/KV Request Duration 99 by type
- TiDB PD TSO PD TiDB
 - channel PD Client/PD TSO Wait Duration
 - TiDB PD PD Client/PD TSO RPC duration

2. TiKV

-
- TiKV 6 gRPCSchedulerraftstoreapplystorage-readpool coprocessor-readpool
 - gRPC TiKV
 - scheduler
 - kv get kv batch get storage-readpool
 - TiDB DistSQL Coprocessor coprocessor-readpool 4.0 storage-readpool coprocessor-readpool
 - PD region raftstore
 - scheduler key-value raftstore
 - scheduler Scheduler-/Scheduler command durationPrewrite Scheduler-prewrite
 - raft RocksDB Storage/async write duration
 - raftstore raft apply
 - raftstore Raft IO/Commit log durationraft Raft Propose/Propose wait duration
 - apply scheduler key-value RocksDB gRPC
 - apply Raft IO/Apply log duration RocksDB Raft Propose/Apply wait duration
 - storage-readpool kv get kv batch get
 - coprocessor-readpool
- TiKV RocksDB raftstore raft raft TiKV
- [8.2.1 TiKV](#)

3. RocksDB

RocskDB TiDB

- RocksDB
 - Memtable Memtable
 - Memtable Block-Cache
 - Block-Cache 45%
 - WAL Write Ahead Log logfile Memtable
 - SST Memtable SST logfile log
 - SST
 - Compaction

4.2 TiDB

CPU IOPS TiDB

- 1.
- 2.
- 3.

4.2.1

SQL TiKV Key-Value Schema

1. TiKV Key-Value
2. TiKV Key-Value
3. TiDB
4. TiDB Key-Value Schema
5. TiDB

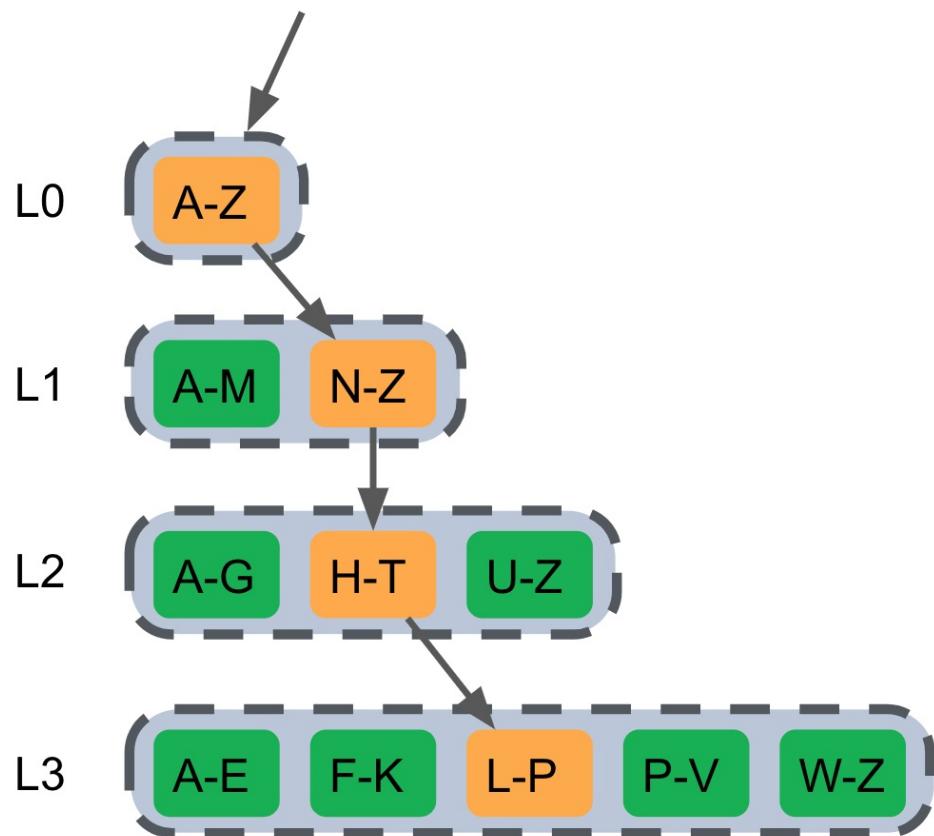
(1) TiKV Key-Value

TiKV RocksDB TiKV Query KV RocksDB SnapshotSnapshot RocksDB DB Snapshot Get KV Iterator

Point-lookup: Get(key)

“P” L0 value

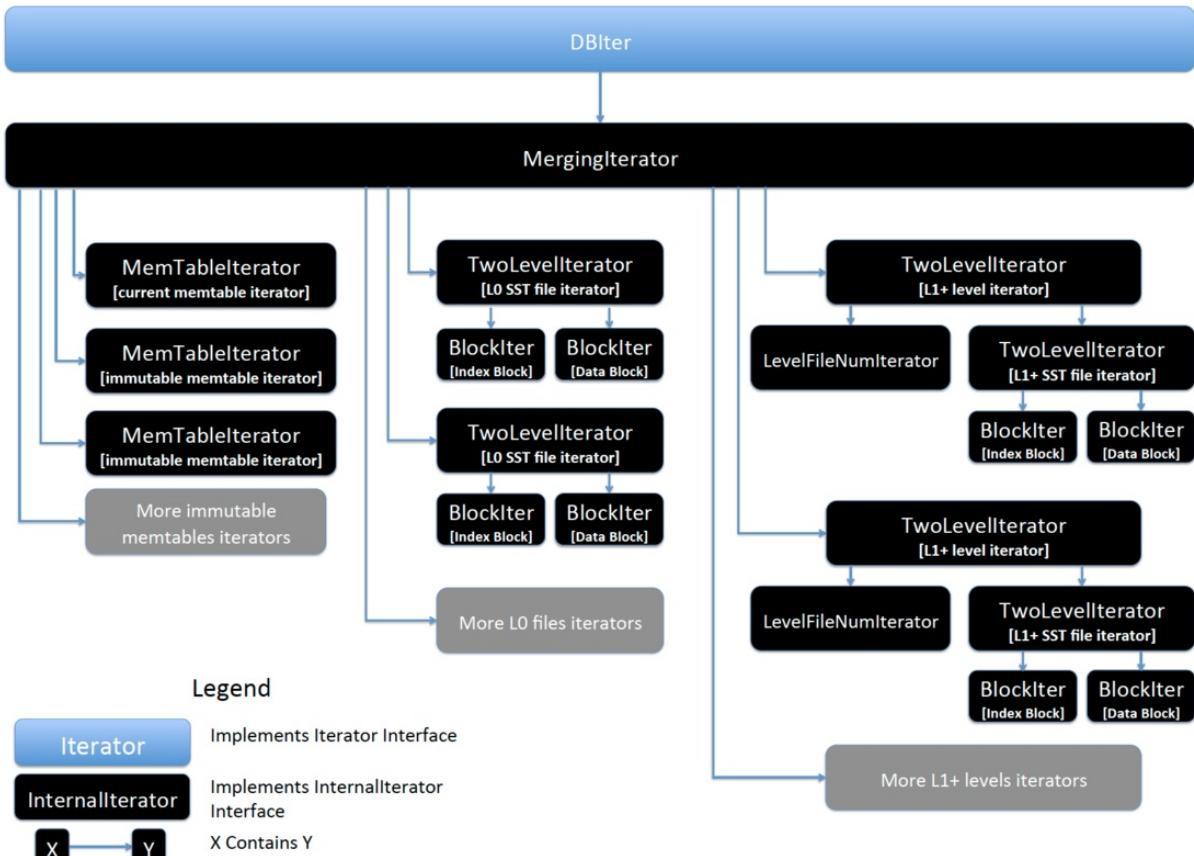
1. Memtable SST
2. key key key L0 L0
3. Bloom filter key key A [1, 10) key 11 key 5 bloom filter key key key A [1, 10) key
(1,3,5,7,9) 6 [1, 10)
4. Index Block key Block SST Block
5. Data Block value



Range-lookup: Iterator

Iterator Iterator Iterator LSM-tree

1. MemTableMutable / Immutable
2. Block (Index Block / Data Block)RocksDB TwoLevelIterator SST File first_level_iter Index Block second_level_iter Data Block



Point-lookup Range-lookup SST File Data Block point get, memtable SST KV IO SST File Data Block

IO RocksDB Block Cache Key-Value

- | | | |
|-----------------------|---------------------------|----------------|
| 1. Block Cache Memory | a. Block cache size | b. block cache |
| block cache | c. Block cache flow | block cache |
| 2. SST Files IO | d. Block cache operations | block cache |
| 3. Key CPU | a. SST read duration | SST |
| | a. RocksDB CPU | RocksDB CPU |

Key-Value Block cache size SST File IO CPU

Block cache Block cache flow Block cache operations Block cache

(2) TiKV Key-Value

TiDB gRPC kv_get / kv_batch_get / coprocessor

- gRPC message count gRPC
- gRPC message failed gRPC
- gRPC message duration gRPC
- gRPC poll CPU gRPC CPU gRPC CPU IO gRPC gRPC

TiDB 4.0 kv_get / kv_batch_get / coprocessor Storage read threads pool (kv_get / kv_batch_get) Coprocessor threads pool (coprocessor) TiDB 4.0 Unified Read Pool

Storage read threads pool

kv_get / kv_batch_get gRPC FuturePool FuturePool Pool

1. RocksDB Snapshot
2. Snapshot Point-lookup Key-Value

Snapshot Snapshot KV Storage read threads pool

- Storage async snapshot duration get RocksDB->GetSnapshot() + Snapshot->Get(key)
- Storage command total get
- Storage ReadPool CPU CPU

CPU CPU Storage async snapshot duration Snapshot Snapshot KV

Coprocessor threads pool

Coprocessor

1. TiKV coprocessor Coprocessor
2. Coprocessor Snapshot
3. Snapshot Handler 3 Handler

Coprocessor

1. Coprocessor CPU CPU
2. Wait duration + Snapshot + Handler max / .99 / .95
 - a. IO CPU CPU
 - Snapshot d. CPU CPU CPU Snapshot
3. Handle duration: Handler Handle
 - a. CPU CPU CPU
 - b. TableScan / IndexScan Snapshot Range
 - CPU Handle duration Range-lookup

- Total Request errors
- Total DAG executorsDAG executor TableScan CPU
- Total KV Cursor Operations scan key Scan keys
- Total Ops Details (Table Scan)table scan Snapshot Iterator
- Total Ops Details (Index Scan)index scan Snapshot Iterator
- Total RocksDB Perf Statistics sacn perf RocksDB operation

Unified Read Pool

TiDB 4.0 (yatp, yet another thread pool) Unified Read Pool Unified Read Pool Unified Read Pool

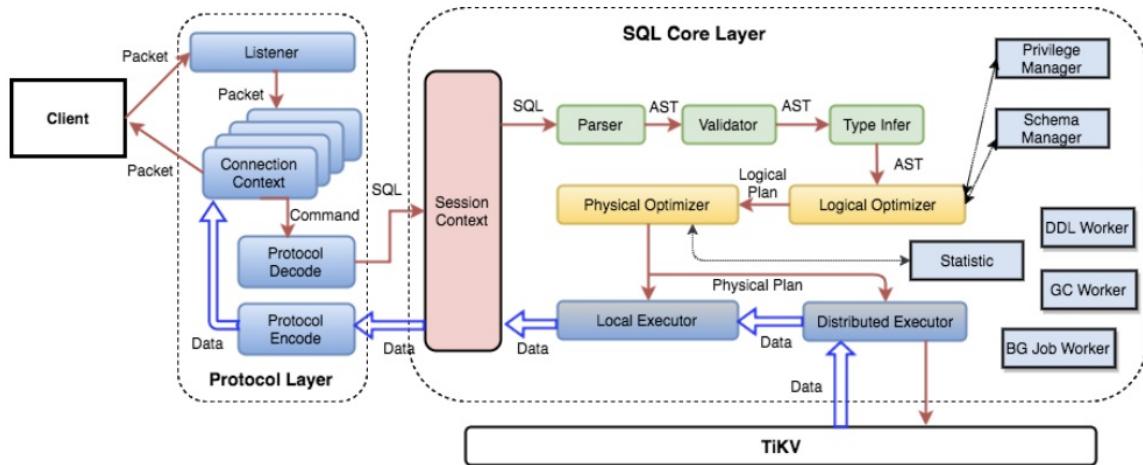
- Unified ReadPool CPU CPU
- Time used by level: level CPU
- Running tasks:

(3) TiDB

TiKV KV SQL TiKV KV

1. Socket SQL
2. Token
3. PD TSO (tsFuture tsFuture TSO)
4. Parser SQL parse AST
5. AST compile Validator / LogicalPlanOptimizer / PhysicalPlanOptimizer / Executor builder Compile duration
6. a. Executor.Open() Executor SQL Key TiKV distsql API TiKV
 - b. Executor.Next()

Executor



1. QPS
2. DurationSQL
3. Get Token Duration Token
4. Parse DurationSQL
5. Compile Duration SQL AST
6. Execution DurationSQL

1. Parse Duration / Compile Duration CPU CPU insert ... values Compile
2. Get Token Duration SQL TokenLimiter Token
3. Execution Duration Executor
4. TSO

TSO

TSO TiDB PD Client Store PD TSO

1. TSO channel channel TSO PD TSO channel batch
2. TSO

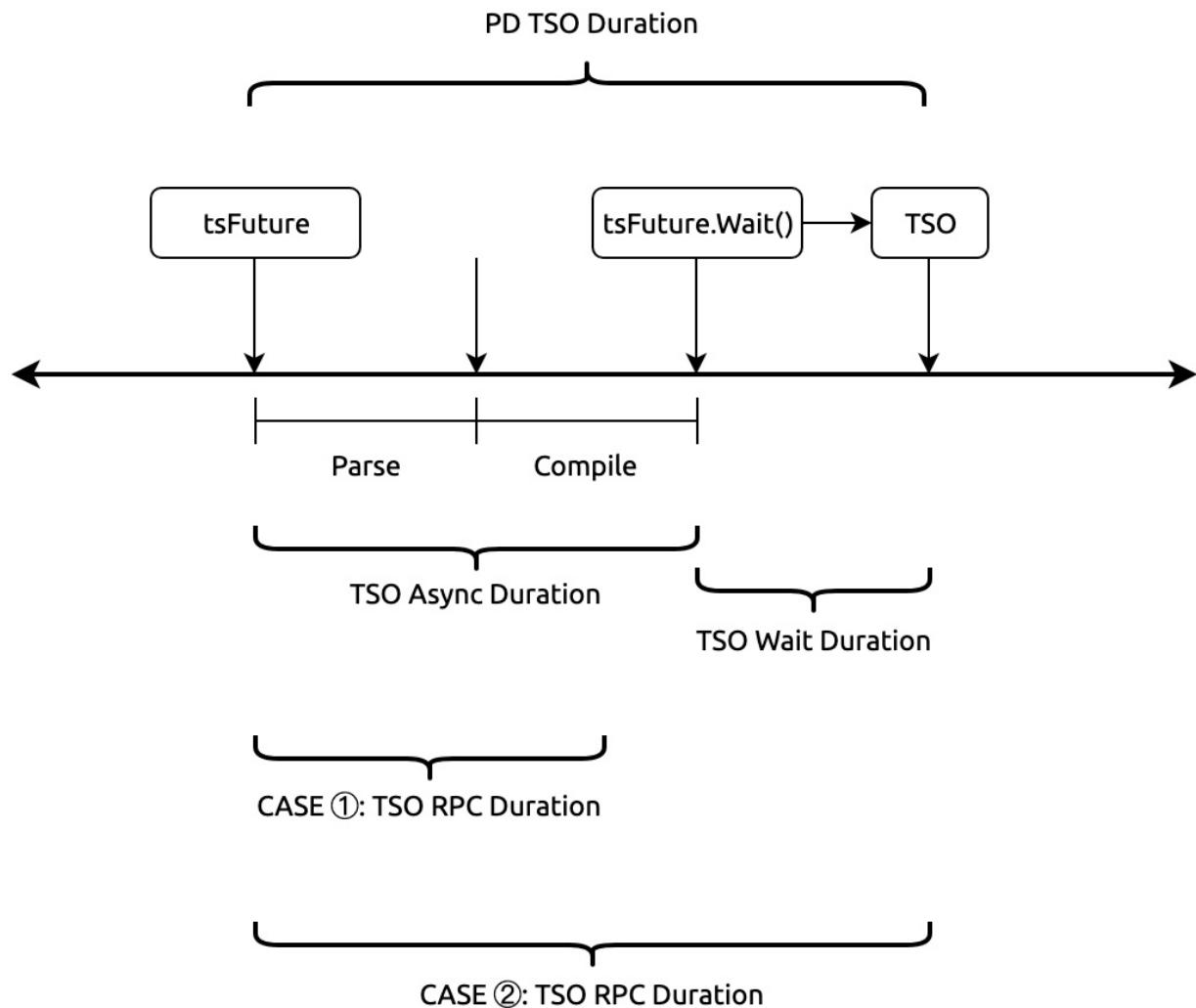
TSO

1. TSO channel GetTSAync tsFuture
2. TSO channel PD RPC TSO TSO
3. TSO tsFuture TSO tsFuture.Wait()

channel GetTSAync channel RPC RPC Duration

- PD TSO RPC Duration : PD RPC
 - TiDB PD
 - PD TSO RPC
- PD TSO Wait Duration: tsFuture Parse Compile tsFuture.Wait() PD TSO RPC

TSO



TSO RPC Duration TSO TSO SQL Parse + CompileParse + Compile tsFuture TSO TSO Async Duration =
Parse + Compile Duration

- CASE 1: TSO RPC Duration \geq Parse + Compile Duration Wait TSO RPC TSO TSO Wait Duration > 0
- CASE 2: TSO RPC Duration < Parse + Compile Duration: tsFuture.Wait() RPC TSO Wait Duration = 0

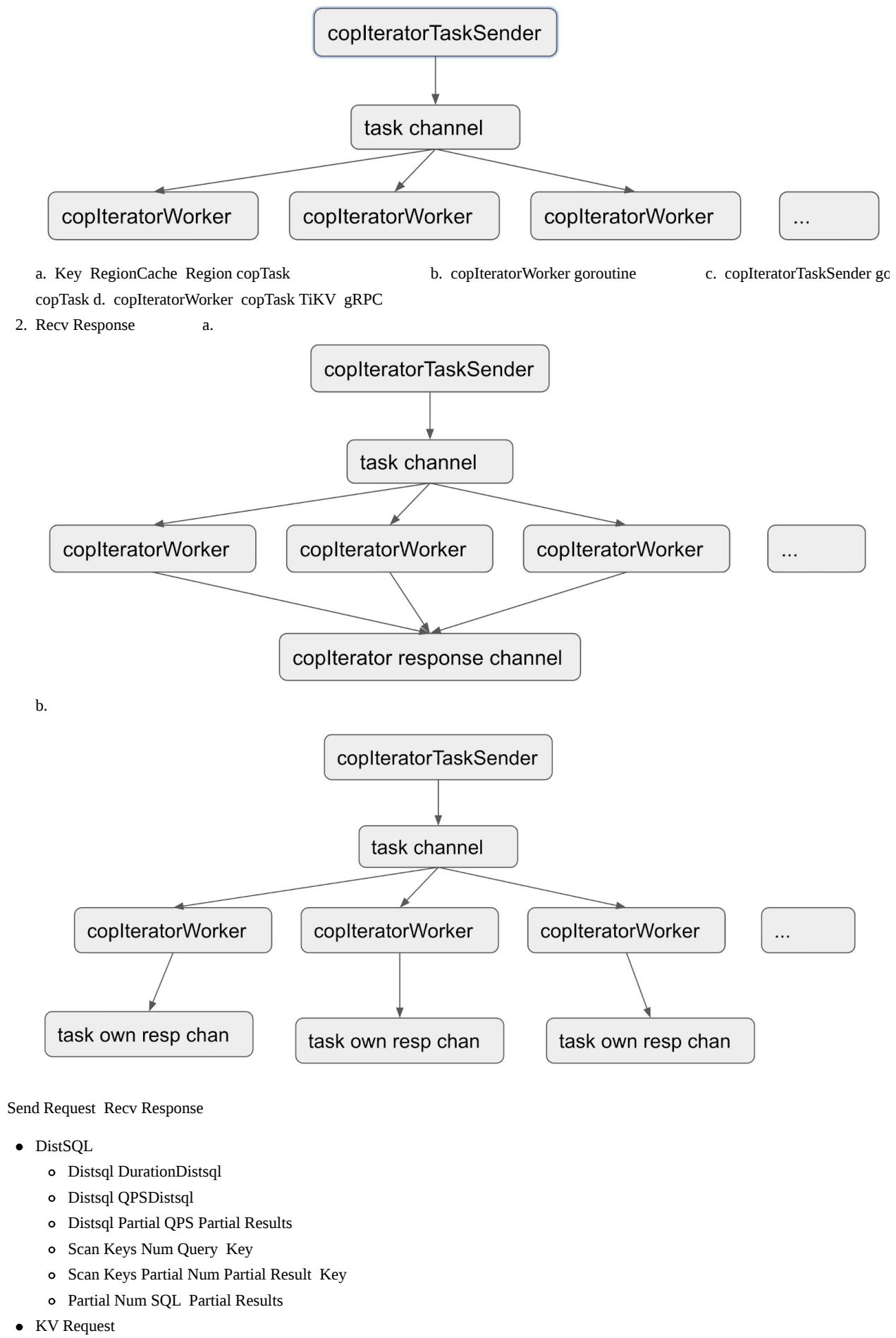
DistSQL

distsql API

1. Build Request
2. Send Request
3. Recv Response
4. Decode Response

Build Request / Decode Response Send Request / Recv Response Send Request Recv Response

1. Send Request



- KV Request Duration 999 by storeKV Request TiKV
 - KV Request Duration 999 by typeKV Request
 - KV Request OPSKV
 - Transaction
 - Transaction OPS
 - Transaction Regions Num 90 Region
 - Transaction Max Write Size Bytes
 - Transaction Max Write KV Num Key-Value
 - Load SafePoint OPS SafePoint

TiKV

KV Errors

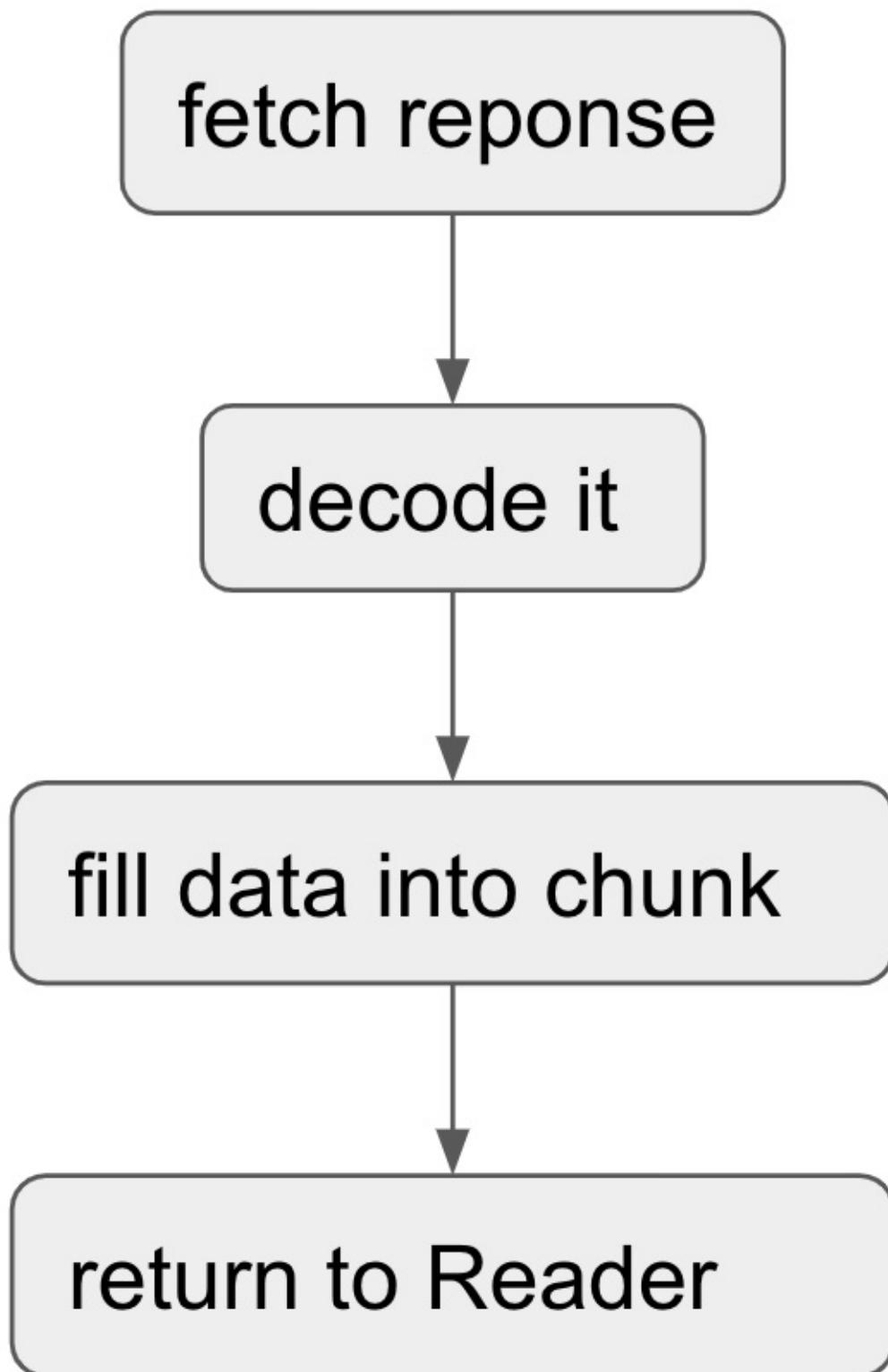
- KV Backoff Duration
 - TiClient Region Error OPSTiKV Region
 - KV Backoff OPSTiKV
 - Lock Resolve OPS
 - Other Errors OPS SafePoint

(4) TiDB Key-Value Schema

TiDB Key-Value

1. Coprocessor chunk
 2. Get / Batch Get Key-Value

Coprocessor chunk value column Get / Batch Get Key-Value encode decode



(5) TiDB

Socket

4.2.2

SQL Schema Key-Value Key-Value TiKV TiDB Executor Executor TableReaderExecutor / IndexReaderExecutor / IndexLookupExecutor Executor distsql API Executor Insert / Update / Delete 2PC

1. TiDB 2PC
2. TiKV Scheduler
3. TiKV Raftstore

(2)

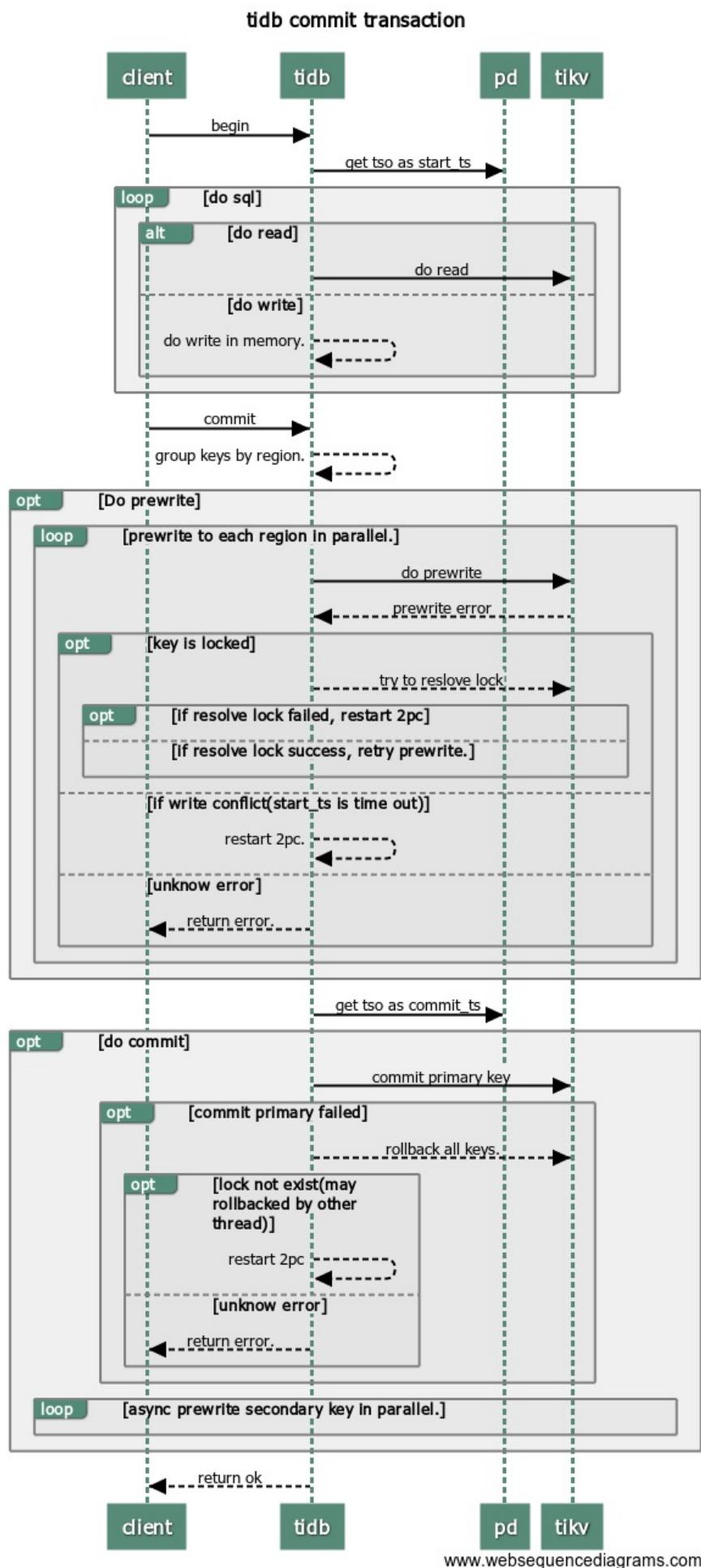
Insert / Update / Delete Key-Value In-Memory Buffer

- 1.
- 2.

Key Value Key-Value Key-Value

In-Memory Buffer Key-Value tikv client TiKV

1. Prewrite
2. Commit Transaction TSO start_ts Prewrite Commit TSO commit_ts



(3)

```
[txn-local-latches]
enabled = true
capacity = 2048000
```

Prewrite

- Local Latch Wait Duration

(4)

Distsql

- Transaction
 - Transaction OPS
 - Duration Latches Latches
 - Transaction Retry Num
 - Transaction Statement Num SQL
 - Session Retry Error OPS
 - KV Transaction OPS
 - Transaction Regions Num 90 Region
 - Transaction Max Write Size Bytes
 - Transaction Max Write KV Num Key-Value
- KV Errors
 - KV Backoff DurationKV
 - TiClient Region Error OPSTiKV Region
 - KV Backoff OPSTiKV
 - Lock Resolve OPS

TiDB 2PC TiKV

(5) TiKV Scheduler**Scheduler**

Scheduler

```
1. Latches
2. /
```

Latches TiDB Prewrite Commit Key-Value Key Scheduler

1. Snapshot
2. Snapshot PrewriteCommit Resolve Lock

Prewrite Commit

- | | | |
|---|-------------------|-----------------------|
| 1. Prewrite a. Keyload lock key key KeyIsLocked | b. key | c. Key Key Key raftst |
| 2. Commit a. Keyload lock lock key Prewrite TxnLockNotFound Value | b. Key-Value raft | |

1. Scheduler CPU
2. Command
- 3.
4. Snapshot
5. Load lock
- 6.
7. Scheduler

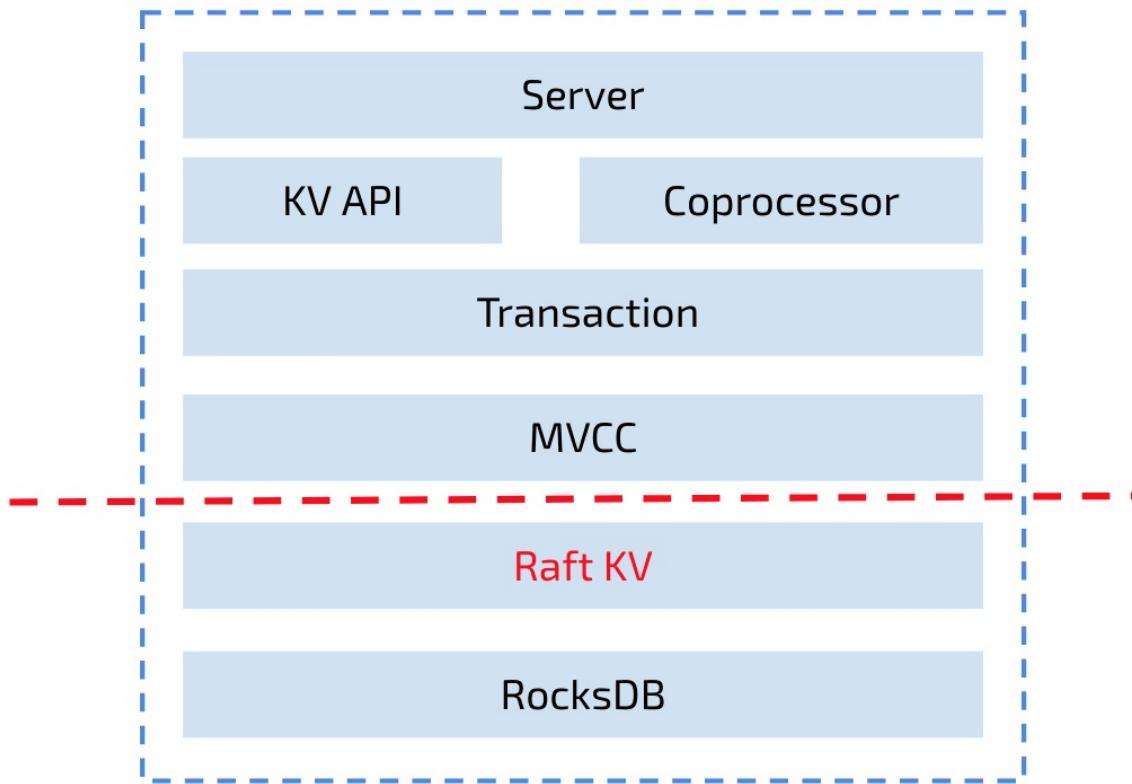
Scheduler

- gRPC
 - gRPC duration TiKV Prewrite / Commit TiKV gRPC duration TiDB KV duration gRPC duration TiDB KV duration TiDB TiKV TiDB TiKV TiDB TiKV Scheduler Duration gRPC duration gRPC Scheduler Scheduler
- Scheduler
 - Scheduler stage total
 - Scheduler priority commands
 - Scheduler pending commands TiKV pending
 - Scheduler duration latch-wait duration + async-snapshot duration + load lock (all keys total) + async-write duration
 - Scheduler worker CPU Scheduler worker CPU
 - Scheduler latch wait durationlatch wait
 - key key region key prewritecommit acquire-pessimistic-lock raft latch
 - Scheduler keys readcommit key
 - Scheduler keys writtencommit key
 - Scheduler scan details commit CF key
 - Scheduler scan details [lock] commit lock CF key
 - Scheduler scan details [write] commit write CF key
 - Scheduler scan details [default] commit default CF key
- Storage
 - Storage async-snapshot duration raft snapshot snapshot
 - Storage async-write durationraft Raftstore

Schduler Snapshot Raftstore Scheduler Snapshot Scheduler worker CPU CPU Raftstore

(6) TiKV Raftstore

Raftstore TiKV Region Raftstore Raft Region Key-Value RocksDB



Raftstore Write Raft Log Apply Raft Log

1. IO RocksDB

- Raft RocksDB Raft
- KV RocksDB Key-Value

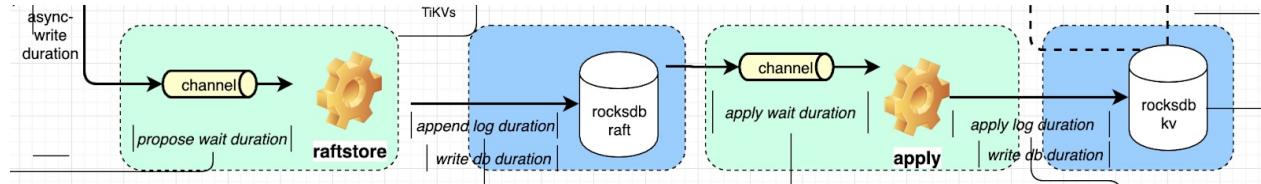
2. CPU

- raft
- apply

3. Networkleader follower

PeerMsg::RaftCommand Key-Value RaftCommand Router raft raft Raft Raft RocksDB Peerapply KV

RocksDB



TiKV dashboard Raft IO / Raft process / Raft messages / Raft propose

• Thread CPU

- Raft store CPUPraft CPU 80%
- Async apply CPUPapply CPU 90%

• Raft propose

- Raft proposals per ready mio tick Region proposal
- Raft read/write proposals proposal
- Propose wait duration raftstore raftstore append log raftstore
- Apply wait duration committed raft log apply apply apply log apply apply CPU
- Raft log speedpeer propose

- Raft IO
 - Apply log durationRaft apply raft log apply log kvdb
 - Append log durationRaft append raft log raft log raftdb 50ms Raft RocksDB
 - Commit log duration: Raft
- Raft message
 - Sent messages per server TiKV Raft
 - Flush messages per server TiKV Raft
 - Receive messages per server TiKV Raft
 - Messages Raft
 - VoteRaft
 - Raft dropped messages Raft

(7) RocksDB Titan

TiKV RocksDB TiKV RocksDB

- Raft RocksDB: Raft Log
- KV RocksDB: KV Region

RocksDB LSM-tree level

1. Memtable WAL
 - Memtable size CF memtable
 - WAL sync operationssync WAL
 - WAL sync durationsync WAL
 - Write operationswrite
 - Write durationwrite
 - Write flow
2. Memtable Immutable Memtable level (level0)
3. Compaction
 - Compaction operationsCompaction flush
 - Compaction durationCompaction flush
 - Compaction flowCompaction
 - Compaction pending bytes Compaction

level0 / Memtable / compaction RocksDB Write Stall

- Write stall duration write stall 0
- Stall conditions changed of each CF CF stall

Titan 1 8 Titan LSM-tree Key-Value Value Blob Titan LSM-tree RocksDB Blob

- Blob file size: Blob
- Blob seek duration: Blob seek
- Blob get duration: Blob get
- Blob bytes/keys flow: Blob

Titan Blob GC

- Blob GC file: Blob GC
- Blob GC duration: Blob GC
- Blob GC bytes/keys flow: Blob GC

4.4 Prometheus

Prometheus TiDB Prometheus Prometheus TiDB Prometheus alertmanager

4.4.1 Prometheus

TiDB Prometheus Grafana

Prometheus prometheus server prometheus server prometheus server grafana Prometheus alertmanager pushgateway exporters

prometheus server MySQL zabbix zabbix prometheus server prometheus server prometheus server pushgateway

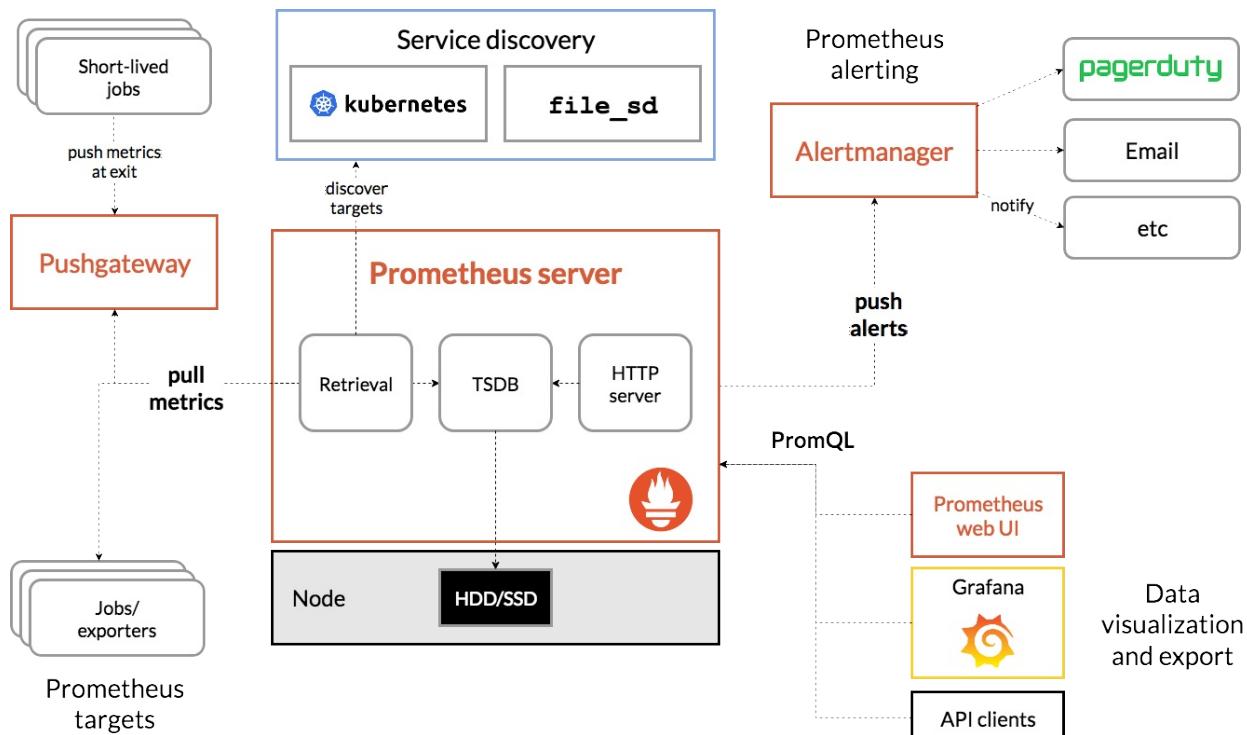
Prometheus exporter exporters

Prometheus exporters TiDB; exporter snmp_exporter; pushgateway exporter exporter ceph

kubernetes Prometheus Prometheus

4.4.2

Prometheus



Prometheus prometheus server Prometheus

Alertmanger prometheus server webhook

Prometheus exporter agent Prometheus TiDB exporter

PromQL Prometheus prometheus server web UI PromQL PromQL grafana API

Prometheus exporters service discovery kubernetes node,pod,service

```
exporter service discovery push pushgatewaypushgateway prometheus server exporterprometheus server exporters
pushgateway
```

4.4.3

Prometheus kubernetes Prometheus docker

prometheus --config.file=conf/prometheus.yml

systemd TiDB prometheus

1. Prometheus server

prometheus server yaml --config.file prometheus.yml

```
global:
  scrape_interval: 15s
  scrape_timeout: 10s
  external_labels:
    monitor: 'codelab-monitor'

rule_files:
  - rules/centos7.rules.yml
  - rules/mariadb.rules.yml

alerting:
  alertmanagers:
    - static_configs:
      - targets:
          - 21.129.127.3:9093

scrape_configs:
  - job_name: 'prometheus'
    scrape_interval: 5s
    static_configs:
      - targets: ['localhost:9090']

  - job_name: 'node'
    file_sd_configs:
      - files:
          - conf.d/centos.yml
```

- global:
- scrape_interval: 15
- scrape_timeout: 10
- external_labels: federation, remote storage, Alertmanager
- rule_files: alertmanger
- alerting: alertmanager alertmanager
- scrape_configs
- job_name job job
- static_configs:
- targets instance)
- file_sd_configs:

2.

```

groups:
- name: alert.rules
  rules:
    - alert: InstanceDown
      expr: up == 0
      for: 1s
      labels:
        level: emergency
      annotations:
        summary: ""
        description: " ", service: "
  
```

- groups
- name:
- alert
- expr
- for: client
- labels
- annotations:

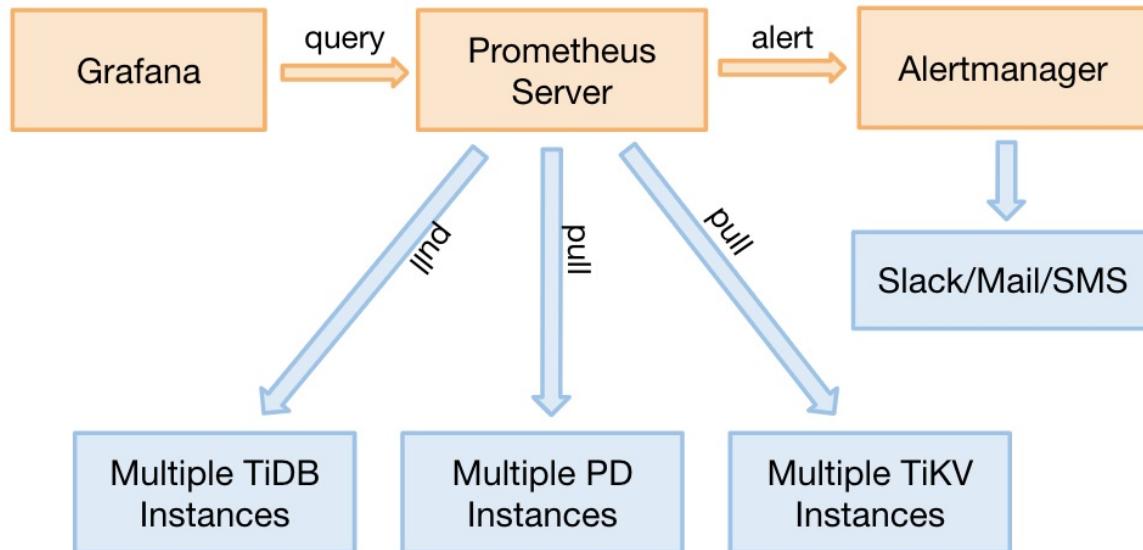
4.4.4 Prometheus TiDB

Prometheus TiDB Prometheus PromQL TiDB

1. TiDB Prometheus

TiDB Prometheus 2.1 TiDB TiDB pushgateway prometheus server pushgateway 2.1 TiDB
prometheus server Prometheus pushgateway grafana alertmanager altermanager

Metrics



2. Prometheus PromQL TiDB

PromQL(Prometheus Query Language) Prometheus PromQL TiDB

(1)

Prometheus 4

- Instant vector - ;
- Range vector - ;
- Scalar - ;
- String -

(2) web UI

web UI (<http://prometheus-server:9090/graph>) up{instance="21.129.14.103:2998"}

The screenshot shows the Prometheus web interface. At the top, there's a navigation bar with links for Prometheus, Alerts, Graph, Status, Help, and Classic UI. Below the navigation bar, there are two checked checkboxes: 'Enable query history' and 'Use local time'. A search bar contains the query 'up{instance="21.129.14.103:2998"}'. To the right of the search bar is a blue 'Execute' button. Under the search bar, there are tabs for 'Table' and 'Graph', with 'Table' currently selected. Below the tabs is a section titled 'Evaluation time' with arrows for navigating between time points. The main content area displays the query results: 'up{alert_lev="0", instance="21.129.14.103:2998", job="hadoop", project="dtl", service="hadoop"}'. To the right of this result is the number '1'. At the bottom right of the results area is a link 'Remove Panel'. At the very bottom left is a blue 'Add Panel' button.

(3)

- up: up Prometheus 1 0 0
- instance, job, project, service, alert_lev instance job prometheus.yaml project, service, alert_lev instance prometheus target
- 1

3. Instant vector

Instant vector

- :server_query_total
- {} :server_query_total{job="tikv"}
- :server_query_total{job=~"tikv|tidb"}
- :server_query_total{job!~"tikv|pd"}
- :up{tidb=~".+"} tidb up
- :tikv_engine_bytes_written{instance="21.129.14.104:21910"}/1024/1024 > 500

4. Range vector

Range vector instance vector []

- s - seconds
- m - minutes
- h - hours
- d - days
- w - weeks
- y - years

TiDB QPS 172.16.4.51:10080 TiDB QPS



5. offset

offset TiDB

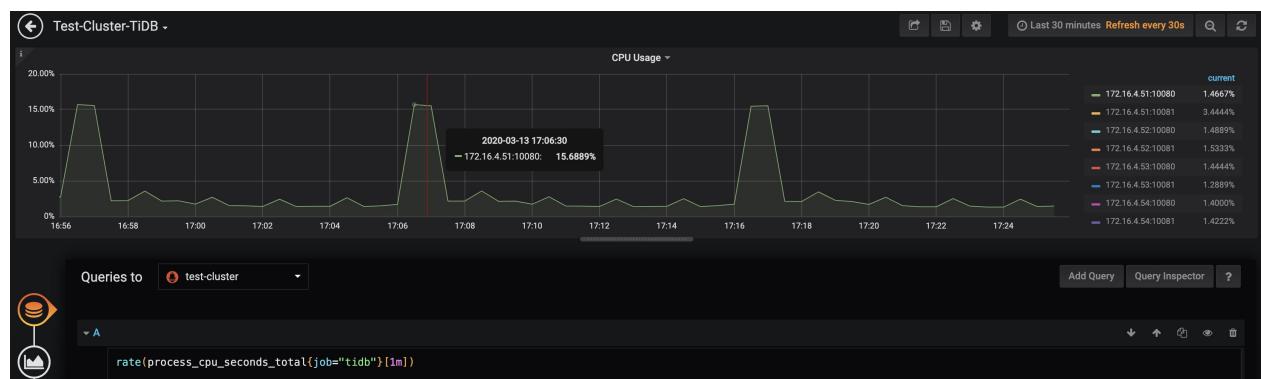
```
sum((tidb_server_query_total{result="OK"} offset 1d))
```

4.4.5 TiDB

TiDB

rate **irate**

```
counter rate() rate() irate() TiDB CPU rate(process_cpu_seconds_total{job="tidb"})[1m]
```



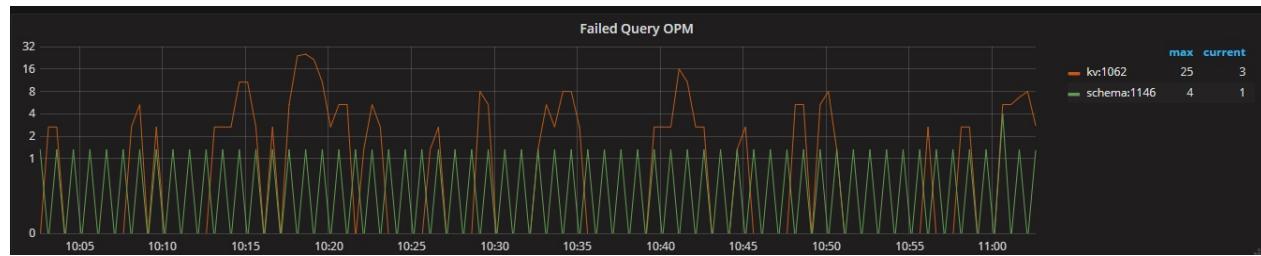
sum **avg**

```
sum avg sum(tikv_store_size_bytes{instance=~"$instance"}) by (instance) TiKV
```



increase

```
increase sum(increase(tidb_server_execute_error_total[1m])) by (type) type 1 Failed Query OPM
```



histogram_quantile

```
histogram_quantile histogram_quantile(phi float, b instant-vector) phi 0 1 0.95 200 200 95% tidb-server 99%
```

```
histogram_quantile(0.99, sum(rate(tidb_server_handle_query_duration_seconds_bucket[1m])) by (le, instance))
```



4.4.6 alertmanager TiDB

TiDB Prometheus tidb-ansible Prometheus tidb-ansible/roles/prometheus/files/tidb.rules.yml

1. TiDB

TiDB

level: emergency TiDB_schema_error TiDB Lease Schema TiDB TiKV Region TiKV TiKV

```
- alert: TiDB_schema_error
  expr: increase(tidb_session_schema_lease_error_total{type="outdated"}[15m]) > 0
  for: 1m
  labels:
    env: ENV_LABELS_ENV
    level: emergency
    expr: increase(tidb_session_schema_lease_error_total{type="outdated"}[15m]) > 0
  annotations:
    description: 'cluster: ENV_LABELS_ENV, instance: , values:'
    value: ''
    summary: TiDB schema error
```

level: critical tidb-server TiDB panic panic tidb-server OOM

```
- alert: TiDB_server_panic_total
  expr: increase(ti_db_server_panic_total[10m]) > 0
  for: 1m
  labels:
    env: ENV_LABELS_ENV
    level: critical
    expr: increase(ti_db_server_panic_total[10m]) > 0
  annotations:
    description: 'cluster: ENV_LABELS_ENV, instance: , values:'
    value: ''
    summary: TiDB server panic total
```

level: warning ti_db-server ti_db-server 10GB Join

```
- alert: TiDB_memory_abnormal
  expr: go_memstats_heap_inuse_bytes{job="ti_db"} > 1e+10
  for: 1m
  labels:
    env: ENV_LABELS_ENV
    level: warning
    expr: go_memstats_heap_inuse_bytes{job="ti_db"} > 1e+10
  annotations:
    description: 'cluster: ENV_LABELS_ENV, instance: , values:'
    value: ''
    summary: TiDB heap memory usage is over 10 GB
```

TiDB TiDB

2. TiDB alertmanager

Prometheus webhook alertmanger

TiDB alertmanager prometheus server ti_db-ansible TiDB alertmanager ti_db-ansible/conf/alertmanager.yml

```
routes:
- match:
  env: test-cluster
  level: emergency
  receiver: ti_db-emergency
  group_by: [alertname, cluster, service]
```

- match
- env level Prometheus
- receiver name
- group_by

```
receivers:  
- name: 'tidb-emergency'  
  webhook_configs:  
  - url: 'xxxx'  
    wechat_configs:  
    - corp_id: 'xxxxx'  
      to_party: 'xxx'  
      agent_id: 'xxxx'  
      api_url: 'https://qyapi.weixin.qq.com/cgi-bin/'  
      api_secret: 'xxxxxx'
```

- name receiver
- webhook_configs webhook
- wechat_configs
- webhook

5.1 GC

Oracle MySQL binlog NewSQL TiDB SQL client driver/TiDB

5.1.1

TiDB MVCC/

TiDB GCGarbage Collection 10 GC TiDB safe point 10 TiDB safe point

TiDB GC mysql.tidb SQL

```
select VARIABLE_NAME, VARIABLE_VALUE from mysql.tidb where VARIABLE_NAME like 'tikv_gc%';

+-----+-----+
| VARIABLE_NAME | VARIABLE_VALUE
+-----+
| tikv_gc_leader_uuid | 5afd54a0ea40005
| tikv_gc_leader_desc | host:tidb-cluster-tidb-0, pid:215, start at 2019-07-15 11:09:14.029668932 +0000 UTC m=+0
.463731223 |
| tikv_gc_leader_lease | 20190715-12:12:14 +0000
| tikv_gc_enable | true
| tikv_gc_run_interval | 10m0s
| tikv_gc_life_time | 10m0s
| tikv_gc_last_run_time | 20190715-12:09:14 +0000
| tikv_gc_safe_point | 20190715-11:59:14 +0000
| tikv_gc_auto_concurrency | true
| tikv_gc_mode | distributed
+-----+
-----+
13 rows in set (0.00 sec)
```

GC

```
update mysql.tidb set VARIABLE_VALUE="24h" where VARIABLE_NAME="tikv_gc_life_time";
```

```
mysql.tidbGCTiDB GC GC
.tikv_gc_leader_uuidtikv_gc_leader_desc tikv_gc_leader_lease GC leader
.tikv_gc_last_run_time GC
.tikv_gc_safe_point GC safe point
.tikv_gc_life_time:
.tikv_gc_safe_point: safePoint safePoint snapshot safePoint GC
```

5.1.2

```
:tidb_snapshot Session Set TSOTSO PD ; "2016-10-08 16:45:26.999""2016-10-08 16:45:26" TiDB
Snapshot Snapshot
```

```
: TiDB PD PD Snapshot tidb_snapshot TiDB Server PD Server PD
Session Set tidb_snapshot ""
```

5.1.3

1.

```
create table t (c int);
Query OK, 0 rows affected (0.01 sec)

insert into t values (1), (2), (3);
Query OK, 3 rows affected (0.00 sec)
```

2.

```
select * from t;
+---+
| c |
+---+
| 1 |
| 2 |
| 3 |
+---+
3 rows in set (0.00 sec)
```

3.

```
select now();
+-----+
| now()           |
+-----+
| 2016-10-08 16:45:26 |
+-----+
1 row in set (0.00 sec)
```

4.

```
update t set c=22 where c=2;
Query OK, 1 row affected (0.00 sec)
```

5.

```
select * from t;
+---+
| c |
+---+
| 1 |
| 22 |
| 3 |
+---+
3 rows in set (0.00 sec)
```

6. session scope

```
set @@tidb_snapshot="2016-10-08 16:45:26";
Query OK, 0 rows affected (0.00 sec)
```

```
update tidb_snapshot @@ @@@@ @
```

update

```
select * from t;
+---+
| c |
+---+
| 1 |
| 2 |
| 3 |
+---+
3 rows in set (0.00 sec)
```

7.

```
set @@tidb_snapshot="";
Query OK, 0 rows affected (0.00 sec)

select * from t;
+---+
| c |
+---+
| 1 |
| 22 |
| 3 |
+---+
3 rows in set (0.00 sec)
```

```
tidb_snapshot @@ @@@@ @
```

5.1.4 /

/

gc

1GC GC

```
update mysql.tidb set VARIABLE_VALUE="24h" where VARIABLE_NAME="tikv_gc_life_time";
```

2

```
set @@tidb_snapshot="2016-10-08 16:45:26";
create table t_20161008 like t;
```

3

```
insert into t_20161008 select * from t where c=2;
```

4

5

6GC

```
update mysql.tidb set VARIABLE_VALUE="10m0s" where VARIABLE_NAME="tikv_gc_life_time"
```


5.2 Recover/Flashback

DBA DBA / TiDB

5.2.1 Recover

TiDB mysql.gc_delete_range TiDB GC Worker mysql.gc_delete_range GC lifetime key

RECOVER TABLE GC Worker mysql.gc_delete_range TiDB TiDB CREATE TABLE RECOVER TABLE
DDL

5.2.2

```
MySQL [test]> show tables;

+-----+
| Tables_in_test |
+-----+
| t1           |
| t3           |
+-----+

**2 rows in set (0.00 sec)**

MySQL [test]> create table t2 like t1;

**Query OK, 0 rows affected (0.10 sec)**

MySQL [test]> insert into t2 select * from t1;

**Query OK, 524288 rows affected (17.10 sec)**

**Records: 524288  Duplicates: 0  Warnings: 0**

MySQL [test]> show tables;

+-----+
| Tables_in_test |
+-----+
| t1           |
| t2           |
| t3           |
+-----+

**3 rows in set (0.00 sec)**
```

DDL t2

```

MySQL [test]> drop table t2;

**Query OK, 0 rows affected (0.24 sec)**

MySQL [test]> show tables;

+-----+
| Tables_in_test |
+-----+
| t1           |
| t3           |
+-----+

**2 rows in set (0.00 sec)**

MySQL [test]> select * from t2 limit 1;

ERROR 1146 (42S02): Table 'test.t2' doesn't exist

```

recover

```

MySQL [test]> recover table t2;

**Query OK, 0 rows affected (1.17 sec)**

MySQL [test]> show tables;

+-----+
| Tables_in_test |
+-----+
| t1           |
| t2           |
| t3           |
+-----+

**3 rows in set (0.00 sec)**

MySQL [test]> select count(1) from t2;

+-----+
| count(1) |
+-----+
| 524288   |
+-----+

**1 row in set (0.55 sec)**

```

table t2 gc

```

MySQL [test]> recover table t2;

ERROR 8055 (HY000): snapshot is older than GC safe point 2020-03-08 15:06:19 +0800 CST

```

ddl jobs job id

```
MySQL [test]> ADMIN SHOW DDL JOBS;
```

| JOB_ID | DB_NAME | TABLE_NAME | JOB_TYPE | SCHEMA_STATE | SCHEMA_ID | TABLE_ID | ROW_COUNT | START_TIME |
|--------|---------|------------|--------------|--------------|-----------|----------|-----------|--|
| | | | | END_TIME | STATE | | | |
| 73 | test | t2 | drop table | none | 1 | 71 | 0 | 2020-03-08 15:38:47.276 +0800 CST synced |
| 72 | test | t2 | create table | public | 1 | 71 | 0 | 2020-03-08 15:38:37.726 +0800 CST synced |
| 70 | test | t2 | drop table | none | 1 | 66 | 0 | 2020-03-08 15:38:20.776 +0800 CST synced |

DDL 2 drop table t2 recover table t2

```
MySQL [test]> recover table by job 70;
**Query OK, 0 rows affected (1.15 sec)**
```

5.2.3

recover table

- DDL truncate table delete
- GC GC
- binlog recover table MySQL GC apply recover table GC

5.2.4 Flashback

Flashback Recover table recover recover truncate table recover

```

MySQL [test]> show tables;

+-----+
| Tables_in_test |
+-----+
| t1           |
| t2           |
| t3           |
+-----+

**3 rows in set (0.00 sec)**

MySQL [test]> truncate table t2;

**Query OK, 0 rows affected (0.11 sec)**

MySQL [test]> flashback table t2 to t4;

**Query OK, 0 rows affected (1.16 sec)**

MySQL [test]> show tables;

+-----+
| Tables_in_test |
+-----+
| t1           |
| t2           |
| t3           |
| t4           |
+-----+

**4 rows in set (0.00 sec)**

MySQL [test]> select count(1) from t2;

+-----+
| count(1) |
+-----+
|      0   |
+-----+

**1 row in set (0.01 sec)**

MySQL [test]> select count(1) from t4;

+-----+
| count(1) |
+-----+
|  524288  |
+-----+

**1 row in set (0.39 sec)**

```

flashback ddl job id

5.3

5.3.1

TiDB 3 Region 3 Raft Leader Raft 3 2 region 2 2 3

TiDB

- TiDB TiKV
- TiDB

Region Region TiDB

5.3.2

Region

- Region Region
 - Region 1 Region Peer Raft Log
 - Region Region 1 Region Region
 - Region Region Region IDRegion
- - Region ID 5 min DDL
 -

5.3.3 Region

Region PD Region Region Region

- PD
 - - pd-ctl config get region-schedule-limitreplica-schedule-limitleader-schedule-limitmerge-schedule-limit 4
 - pd-ctl config set 4 0
 - 4
- Region
 - pd-ctl Region ID 145 pd-ctl -u <endpoint> -d region --jq='.regions[] | {id: .id, peer_stores: [.peers[].store_id] | select(length as \$total | map(if .==(1,4,5) then . else empty end) | length>=\$total-length) }'
 - Region 2 Store TiKV
 - Region Region Peer Region tikv-ctl --db /path/to/tikv-data/db unsafe-recover remove-fail-stores -s <s1,s2> -r <r1,r2,r3>
 - Region Peer tikv-ctl --db /path/to/tikv-data/db unsafe-recover remove-fail-stores -s <s1,s2> --all-regions
 - Region Leader
- Region
 - PD TiKV pd-ctl Leader Region pd-ctl -u <endpoint> -d region --jq '.regions[]|select(has("leader")|not)|{id: .id, peer_stores: [.peers[].store_id]}'
 - Region Unavailable Store TiKV tikv-ctl --db /path/to/tikv-data/db recreate-region --pd <endpoint> -r <region_id>

6 TiDB Operator

Kubernetes TiDB

6.1

6.1.1

Kubernetes

docker-wrapper

```
git clone https://github.com/silenceshell/docker-wrapper.git
cd docker-wrapper/
# k8s.gcr.io
./docker-wrapper.py pull k8s.gcr.io/kube-apiserver:v1.16.0
-- pull k8s.gcr.io/kube-apiserver:v1.16.0 from gcr.azk8s.cn/google-containers/kube-apiserver:v1.16.0 instead --
...
-- pull k8s.gcr.io/kube-apiserver:v1.16.0 done --
```

Kubernetes Kubernetes image registry TiDB helm chart image

6.1.2 PD Running TiKV

Pod Running Pod

PD Pod pd_start_script.sh Service DNS discovery discovery Service DNS PD members

Pod

```
# pods
kubectl get pods -n <namespace>
# events
kubectl describe -n <namespace> pods <pd pod-name>
# , nslookup
kubectl logs -n <namespace> pods <pd pod-name>
#
tkctl debug <pd-name>
```

- DNS

```
# DNS
kubectl exec -it <discovery pod name> -n=<namespace> -- nslookup <service dns>
# headless service dns
$ kubectl exec -it tidb-cluster-discovery-6b7f8d9954-48ngh -n=tidb -- nslookup tidb-cluster-pd-peer.tidb.svc
Name:      tidb-cluster-pd-peer.tidb.svc
Address 1: 10.244.2.19 tidb-cluster-pd-0.tidb-cluster-pd-peer.tidb.svc.cluster.local
Address 2: 10.244.1.13 10-244-1-13.tidb-cluster-pd.tidb.svc.cluster.local
Address 3: 10.244.3.11 10-244-3-11.tidb-cluster-pd.tidb.svc.cluster.local
# cluser ip service dns
$ kubectl exec -it tidb-cluster-discovery-6b7f8d9954-48ngh -n=tidb -- nslookup tidb-cluster-prometheus.tidb.svc
Name:      tidb-cluster-prometheus.tidb.svc
Address 1: 10.107.103.198 tidb-cluster-prometheus.tidb.svc.cluster.local
```

DNS CoreDNS

```
# nslookup: can't resolve 'tidb-cluster-prometheus.tidb.svc1': Name does not resolve
command terminated with exit code
# CoreDNS
kubectl get pods -n kube-system | grep core
coredns-9d85f5447-642rs       1/1     Running   1          13h
coredns-9d85f5447-8swr2       1/1     Running   1          13h
# CoreDNS DNS
# pd pod /etc/resov nameserver
kubectl exec -it tidb-cluster-pd-0 -n=tidb -- cat /etc/resolv.conf
```

- IP

1.5 Pod

6.1.3 CrashLoopBackOff

Pod CrashLoopBackOff Pod CrashLoopBackOff Pod

- Pod

```
# kubectl -n <namespace> logs -f <pod-name>
# Pod CrashLoopBackOff
kubectl -n <namespace> logs -p <pod-name>
```

[tidb-server](#) [tikv-server](#) [pd-server](#)

- TiKV Pod “cluster id mismatch”

TiKV Pod TiKV Pod PV local volume provisioner PVC

TiKV PV TiKV Pod PVCTiKV Pod PV Kubernetes PV local volume provisioner PVC PV
reclaimPolicy PV

- Node ulimit

| | | | |
|------|------|--------|----------|
| root | soft | nofile | 10000000 |
| root | hard | nofile | 10000000 |
| root | soft | core | 1048576 |
| root | hard | core | 1048576 |

TiKV ulimit Kubernetes /etc/security/limits.conf ulimit

- pd pod CrashLoopBackOff

```
pd pod CrashLoopBackOff pod
```

```
waiting for discovery service to return start args...
```

Kubernetes discovery service:

```
kubectl get service -n {your namespace} | grep discovery //get ip port
telnet {ip} {port}
```

discovery

```
kubectl delete po {clusternamespace}-discovery-xxxxxxxx-xxxxx -n {your namespace}
```

discovery pod

```
kubectl logs -f {clusternamespace}-discovery-xxxxxxxx-xxxxx -n {your namespace}
```

```
E1218 15:49:00.395064      1 mux.go:58] failed to discover: {cluster-name}-pd-3.{cluster-name}-pd-peer.{cluster-name}.svc:2380, Get http://{cluster-name}-pd.{cluster-name}:2379/pd/api/v1/members: dial tcp x.x.x.x:2379: connect: connection refused
```

pd service

```
wget -c http://{cluster-name}-pd.{cluster-name}:2379/pd/api/v1/members
```

pd pod

```
wget -c http://{ip of pd's pod}:2379/pd/api/v1/members
```

wget wget Kubernetes kube-proxy kube-proxy Kubernetes kube-proxy kube-proxy

nginx

```
kubectl apply -f https://raw.githubusercontent.com/kubernetes/website/master/content/en/examples/service/networking/un-my-nginx.yaml
```

service

```
kubectl expose deployment/my-nginx
```

service kube-proxy kube-proxy Kubernetes Kubernetes

- pd pod CrashLoopBackOff

chart pd replicas 3 pd pod 5 3 running 3 join CrashLoopBackoff

| NAME | READY | STATUS | RESTARTS | AGE |
|--|-------|------------------|----------|-----|
| tidbcluster-...-i-discovery-68d7c75556-956kw | 1/1 | Running | 1 | 15d |
| tidbcluster-...-i-monitor-84cc9cc5-m6rm1 | 3/3 | Running | 3 | 15d |
| tidbcluster-...-i-pd-0 | 1/1 | Running | 1 | 15d |
| tidbcluster-...-i-pd-1 | 1/1 | Running | 15 | 12d |
| tidbcluster-...-i-pd-2 | 0/1 | CrashLoopBackoff | 283 | 24h |
| tidbcluster-...-i-pd-3 | 0/1 | CrashLoopBackoff | 308 | 26h |
| tidbcluster-...-i-pd-4 | 0/1 | CrashLoopBackoff | 3032 | 11d |
| tidbcluster-...-i-tidb-0 | 2/2 | Running | 0 | 15d |
| tidbcluster-...-i-tidb-1 | 2/2 | Running | 0 | 15d |
| tidbcluster-...-i-tidb-initializer-p8r21 | 0/1 | Completed | 129 | 29d |
| tidbcluster-...-i-tikv-0 | 1/1 | Running | 0 | 15d |
| tidbcluster-...-i-tikv-1 | 1/1 | Running | 0 | 15d |
| tidbcluster-...-i-tikv-2 | 1/1 | Running | 0 | 15d |

PD pod

```
[~]# kubectl logs --tail=200 pd-2
Name: pd-2
Address 1: 10.42.40.80
nslookup domain
starting pd-server
/pd-server --data-dir=/var/lib/pd --name=
--pd-2 --peer-urls=http://0.0.0.0:2380 --advertise-peer-urls=http://tidbcluster-...-pd-2.tidbcluster...
--pd-peer.tidbcluster-...-pd-2.svc:2380 --client-urls=http://0.0.0.0:2379 --advertise-client-urls=http://tidbcluster-...-pd-2.tidbcluster...
--join-urls=http://tidbcluster-...-pd-2.tidbcluster...
--pd-peer.tidbcluster-...-pd-2.svc:2379 --config=/etc/pd/pd.toml
--join=tidbcluster-...-pd-2.tidbcluster...
--pd-peer.tidbcluster-...-pd-2.svc:2380,http://tidbcluster-...-pd-2.tidbcluster...
--pd-0.tidbcluster...
--pd-peer.tidbcluster-...-pd-2.svc:2380
[2020/02/06 16:50:46.071 +08:00] [INFO] [util.go:59] ["welcome to Placement Driver (PD)"]
[2020/02/06 16:50:46.071 +08:00] [INFO] [util.go:60] [PD] [release-version=v3.0.5]
[2020/02/06 16:50:46.071 +08:00] [INFO] [util.go:61] [PD] [git-hash=4934a651d2d428411a9610a5b6b9d7156a439355]
[2020/02/06 16:50:46.071 +08:00] [INFO] [util.go:62] [PD] [git-branch=HEAD]
[2020/02/06 16:50:46.071 +08:00] [INFO] [util.go:63] [PD] [utc-build-time="2019-10-25 03:24:57"]
[2020/02/06 16:50:46.071 +08:00] [INFO] [metricutil.go:80] [{"disable Prometheus push client"}]
[2020/02/06 16:50:46.071 +08:00] [ERROR] [join.go:212] ["failed to open directory"] [error="open /var/lib/pd/member: no such file or directory"]
2020/02/06 16:50:46.071 grpcllog.go:45: [info] parsed scheme: "endpoint"
2020/02/06 16:50:46.071 grpcllog.go:45: [info] ccResolverWrapper: sending new addresses to cc: {[http://tidbcluster-...-pd-3.tidbcluster...
--pd-peer.tidbcluster-...-pd-3.svc:2380 0 <nil>} {http://tidbcluster-...-pd-1.tidbcluster...
--pd-peer.tidbcluster-...-pd-1.svc:2380 0 <nil>} {http://tidbcluster-...-pd-0.tidbcluster...
--pd-peer.tidbcluster-...-pd-0.svc:2380 0 <nil>}}
2020/02/06 16:50:46.079 grpcllog.go:60: [warning] grpc: addrConn.createTransport failed to connect to {http://tidbcluster-...-pd-3.tidbcluster...
--pd-peer.tidbcluster-...-pd-3.svc:2380 0 <nil>}. Err:connection error: desc = "transport: Error while dialing dial tcp: lookup tidbcluster-wuhanwuli-pd-3.tidbcluster...
--pd-peer.tidbcluster-...-pd-3.svc on 10.43.0.10:53: no such host"
Reconnecting...
[2020/02/06 16:50:46.081 +08:00] [FATAL] [main.go:91] ["join meet error"] [error="there is a member that has not joined successfully"]
[stack=github.com/pingcap/log.Fatal\n\t/home/jenkins/agent/workspace/release_tidb_3.0/go/pkg/mod/github.com/pingcap/log@v0.0.0-20190715063458-479153f07ebd/global.go:59\nmain.main\n\t/home/jenkins/agent/workspace/release_tidb_3.0/go/src/github.com/pingcap/pd/dpd-server/main.go:91\nruntime.main\n\t/usr/local/go/src/runtime/proc.go:200"]
2020/02/06 16:50:46.081 grpcllog.go:60: [warning] Failed to dial http://tidbcluster-...-pd-3.tidbcluster...
--pd-peer.tidbcluster-...-pd-3.svc:2380: context canceled; please retry.
```

pod PV Retain

provisioner

```
I0206 09:47:40.383224 1 cache.go:64] Updated pv "local-pv-863ba702" to cache
I0206 09:47:40.383244 1 cache.go:64] Updated pv "local-pv-ca3501f4" to cache
I0206 09:51:42.745063 1 cache.go:64] Updated pv "local-pv-863ba702" to cache
I0206 09:51:52.419372 1 deleter.go:195] Start cleanup for pv local-pv-863ba702
I0206 09:51:52.419505 1 deleter.go:266] Deleting PV file volume "local-pv-863ba702" contents at hostpath "/mnt/disks/tikv1", mountpath "/mnt/disks/tikv1"
I0206 09:52:02.420739 1 deleter.go:165] Deleting pv local-pv-863ba702 after successful cleanup
I0206 09:52:02.430763 1 cache.go:64] Updated pv "local-pv-863ba702" to cache
I0206 09:52:02.439564 1 cache.go:73] Deleted pv "local-pv-863ba702" from cache
I0206 09:52:12.433960 1 discovery.go:299] Found new volume at host path "/mnt/disks/tikv1" with capacity 3935254495232, creating Local PV "local-pv-863ba702" required volumeMode "Filesystem"
I0206 09:52:12.443227 1 discovery.go:332] Created PV "local-pv-863ba702" for volume at "/mnt/disks/tikv1"
I0206 09:52:12.443520 1 cache.go:55] Added pv "local-pv-863ba702" to cache
I0206 09:52:12.448426 1 cache.go:64] Updated pv "local-pv-863ba702" to cache
```

PV pd kube-apiserver PV kube-apiserver

--v=4

pd pd-recover

6.1.4 Pending

```
kubectl get pod
NAME READY STATUS RESTARTS AGE
pod-pvc-pv 0/1 Pending 0 4s
```

Pod Pending

- PD/TiKV/Monitor Pod PVC StorageClass PV
- Kubernetes Pod CPU
- PD TiKV Replicas tidb-scheduler

- kubectl describe Pending

```
kubectl describe pod -n <namespace> <pod-name>
kubectl describe pod pod-pvc-pv
Name:          pod-pvc-pv
...
Status:        Pending
...
Warning FailedScheduling 21s (x4 over 4m28s) default-scheduler 0/1 nodes are available: 1 node(s) didn't find available persistent volumes to bind.
```

-

CPU Kubernetes

- StorageClass

values.yaml storageClassName StorageClass helm upgrade Statefulset PVC StorageClass

```
kubectl get storageclass
```

- Local PV

PV Local PV [PV](#)

6.1.5 Pod

TiDB Pod Pod Headless Service TiDB Operator PD Service service-name PD

Pod Pod

1. Service Headless Service Endpoints

```
kubectl -n <namespace> get endpoints <release-name>-pd
kubectl -n <namespace> get endpoints <release-name>-tidb
kubectl -n <namespace> get endpoints <release-name>-pd-peer
kubectl -n <namespace> get endpoints <release-name>-tikv-peer
kubectl -n <namespace> get endpoints <release-name>-tidb-peer
```

ENDPOINTS cluster_ip:port Pod kube-controller-manager

1. Pod Network Namespace

```
tkctl debug -n <namespace> <pod-name>
```

shell dig DNS DNS [Kubernetes DNS](#)

```
dig <HOSTNAME>
```

ping IP IP dig ClusterIP:

```
ping <TARGET_IP>
```

ping Kubernetes ping telnet

```
telnet <TARGET_IP> <TARGET_PORT>
```

telnet Pod

```
#  
kubectl -n <namespace> get po <pod-name> -ojson | jq '.spec.containers[].ports[].containerPort'  
  
#  
# PD, 2379  
kubectl -n <namespace> -it exec <pod-name> -- cat /etc/pd/pd.toml | grep client-urls  
# TiKV, 20160  
kubectl -n <namespace> -it exec <pod-name> -- cat /etc/tikv/tikv.toml | grep addr  
# TiDB, 4000  
kubectl -n <namespace> -it exec <pod-name> -- cat /etc/tidb/tidb.toml | grep port
```

6.1.6 TiDB

TiDB TiDB

Running

```
kubectl get po -n <namespace>
```

TiDB

```
kubectl logs -f <tidb-pod-name> -n <namespace> -c tidb
```

1. NodePort TiDB node service domain clusterIP TiDB serviceName clusterIP Kubernetes

- node
- TiDB service externalTrafficPolicy Local Local TiDB Pod node IP

2. service domain clusterIP TiDB TiDB :4000 PodIP TiDB service domain clusterIP PodIP

- DNS

```
kubectl get po -n kube-system -l k8s-app=kube-dns  
dig <tidb-service-domain>
```

- node kube-proxy

```
kubectl get po -n kube-system -l k8s-app=kube-proxy
```

- node iptables TiDB

```
iptables-save -t nat |grep <clusterIP>
```

- endpoint
- 3. PodIP TiDB Pod
 - node route
 -
 - [Pod](#)

6.2

6.2.1 TiKV PD

6.2.2 TiKV Store Tombstone

TiKV Pod Pod Running TiKV Store Store UP TiKV TiKV Store Tombstone

Store Pod TiKV Pod Store Offline Pod Store

1. TiKV Store

```
kubectl get -n <namespace> tidbcluster <release-name> -ojson | jq '.status.tikv.stores'
```

1. TiKV Pod

```
kubectl get -n <namespace> po -l app.kubernetes.io/component=tikv
```

1. Store Pod TiKV Pod Store Offline Pod Store

- i. PD

```
kubectl port-forward -n <namespace> svc/<cluster-name>-pd <local-port>:2379 &>/tmp/portforward-pd.log &
```

- ii. Store

```
curl -X POST http://127.0.0.1:2379/pd/api/v1/store/<store-id>/state?state=Up
```

2. TiKV Pod lastHeartbeatTime Store Tombstone Pod PV

- i. Store PV reclaimPolicy Delete

```
kubectl patch $(kubectl get pv -l app.kubernetes.io/instance=<release-name>,tidb.pingcap.com/store-id=<store-id> -o name) -p '{"spec":{"persistentVolumeReclaimPolicy":"Delete"}}'
```

ii. Pod PVC

```
kubectl delete -n <namespace> pvc tikv-<pod-name> --wait=false
```

iii. Pod Pod

```
kubectl delete -n <namespace> pod <pod-name>
```

Pod Store

6.2.3 TiDB

(Load Balancer) TiDB TiDB

TiDB tcp-keep-alive TCP keepalive

- Kubernetes [kubelet](#) --allowed-unsafe-sysctls=net.* kubelet TiDB

```
tidb:
...
podSecurityContext:
  sysctls:
    - name: net.ipv4.tcp_keepalive_time
      value: "300"
```

- Kubernetes [kubelet](#) --allowed-unsafe-sysctls=net.* TiDB

```
tidb:
  annotations:
    tidb.pingcap.com/sysctl-init: "true"
  podSecurityContext:
    sysctls:
      - name: net.ipv4.tcp_keepalive_time
        value: "300"
    ...

```

TiDB Operator 1.1

6.3

Pod CrashLoopBackoff Pod kubectl exec tkctl debugTiDB in Kubernetes PD/TiKV/TiDB Pod Pod Crash
kubectl exec tkctl debug Pod

Pod CrashLoopBackoff

```
$ kubectl get pods -n tidbcluster1
NAME                      READY   STATUS    RESTARTS   AGE
demo-discovery-5c78d6bcd8-5ttcl  1/1    Running   0          20h
demo-monitor-6ddc6d6674-kcmhh   3/3    Running   0          3d22h
demo-pd-0                   1/1    Running   0          3d22h
demo-pd-1                   0/1    CrashLoopBackOff  911       3d22h
demo-pd-2                   1/1    Running   10         3d22h
demo-pd-3                   0/1    CrashLoopBackOff  932       3d22h
demo-pd-4                   1/1    Running   0          3d22h
demo-pd-5                   1/1    Running   0          3d22h
demo-tidb-0                 2/2    Running   1          3d22h
demo-tidb-1                 2/2    Running   0          3d22h
demo-tikv-0                 1/1    Running   0          3d22h
demo-tikv-1                 1/1    Running   0          3d22h
demo-tikv-2                 1/1    Running   0          3d22h
demo-tikv-3                 1/1    Running   0          3d22h
demo-tikv-4                 1/1    Running   0          3d22h
demo-tikv-5                 1/1    Running   0          3d22h
```

Pod Annotation

```
kubectl annotate pod <pod-name> -n <namespace> runmode=debug
# kubectl annotate pod demo-pd-1 -n tidbcluster1 runmode=debug
```

Pod Annotation Pod Running

```
watch kubectl get pod <pod-name> -n <namespace>
# watch kubectl get pod demo-pd-1 -n tidbcluster1
```

Every 2.0s: kubectl get pod demo-pd-1 -n tidbcluster1

| NAME | READY | STATUS | RESTARTS | AGE |
|-----------|-------|---------|----------|-------|
| demo-pd-1 | 1/1 | Running | 912 | 3d22h |

kubectl exec

```
# kubectl logs pod
$ kubectl logs demo-pd-1 -n tidbcluster1
entering debug mode.
# pod
kubectl exec -it <pod-name> -n <namespace> -- /bin/sh
# kubectl exec -it demo-pd-1 -n tidbcluster1 -- /bin/sh
```

Pod

```
kubectl delete pod <pod-name> -n <namespace>
# kubectl delete pod demo-pd-1 -n tidbcluster1
```

```
pod "demo-pd-1" deleted
```

Pod

1

“TiDB” TiDB

1. TiDB

20 “ SQL ”(RDBMS)KV SQL NoSQL NoSQL SQL -> NoSQL -> No only
SQL

SQL NewSQL TiDB NewSQL MySQL TiDB TiDB SQL ACID TiDB

TiDB TiDB

- MySQL
- Raft
-
-

2.

OLTP

MySQL MySQL NewSQL MySQL

Hash Sharding DB DDL Join / /

TiDB MySQL TiDB KV SQL DDLPercolator ACID Region Multi-Raft

| | | TiDB |
|----------------------|--|------|
| | | |
| | | |
| | | |
| (JOIN/ GROUP BY/...) | | |
| | | |
| | | |
| | | |
| ID | | |
| | | |

TiDB MySQL MySQL TiDB TiDB 4.0 K8s TiKV Pod Region

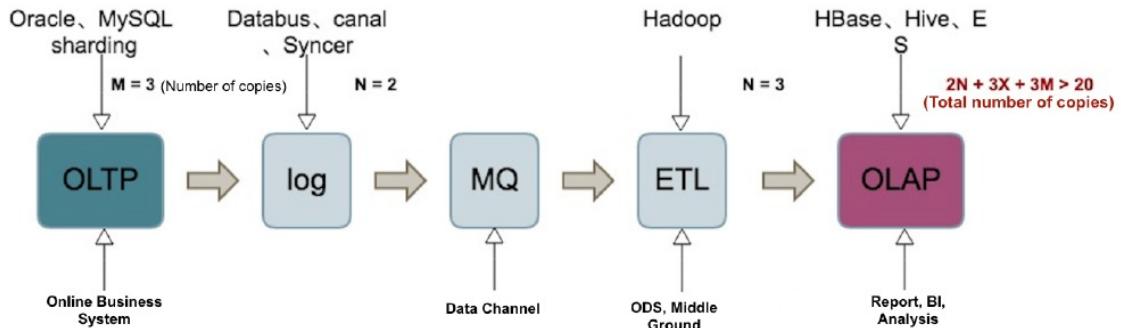
- TiDB

<https://pingcap.com/cases-cn/user-case-zhihu/>

- TiDB at

<https://pingcap.com/cases-cn/user-case-fengchao/>

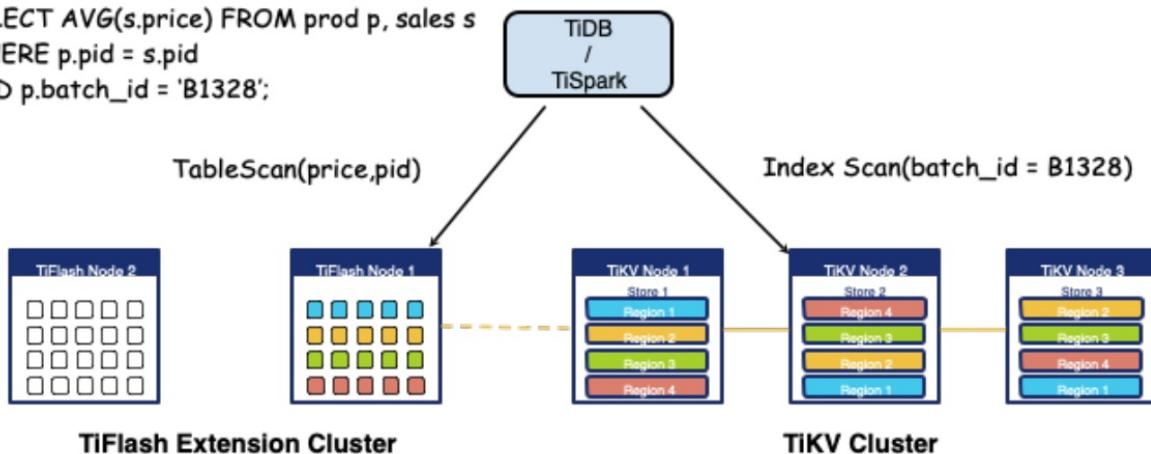
ETL



TiDB 4.0 3.0 TiSpark Spark TiDB TiKV Spark “”

TiDB 4.0 TiFlash Raft learner Engine tag TiFlash TiKV TiFlash TiDB HTAP

**SELECT AVG(s.price) FROM prod p, sales s
WHERE p.pid = s.pid
AND p.batch_id = 'B1328';**



Raft learner TiFlash Oracle DataGuard TiDB HTAP

- TiDB / TiSpark

<https://pingcap.com/cases-cn/user-case-yiguo/>

ABC A

- Oracle Extended RAC
- A B Oracle GoldenGate
-

TiDB TiDB Server Web TiKV Server PD Server Raft TiKV Server Region Multi-Raft DC / Zone /
Rack / Host Raft Leader Reject IO PD Server Raft 3

TiDB Raft TiDB 4.0 Follower Read

- TiDB

<https://pingcap.com/cases-cn/user-case-beijing-bank/>

- TiDB

<https://pingcap.com/cases-cn/user-case-webank/>

3.

TiDB TiDB

<https://pingcap.com/cases-cn/>

2

TiDB

2.1 TiDB server

TP tidb-server tidb-server cpu OLAP 16GB

- OLAP
- OLAP
- mydumper

2.2 PD

PD , SSD PD PD PD TiDB PD

2.3 TiKV

TiKV TiKV TiKV TiKV (max-replicas) location-label TiKV /

2.4

4-8GB 2-4 CPU 16GB 8 CPU 8GB 4 CPU Prometheus/Alertmanager/Grafana TiDB

2.5 Pump

- CPU4
- 8GB
- SSD
-

Pump

2.6 Drainer

- CPU4
- 8GB
- Drainer SSD
-

Drainer Drainer

2.7

mydumper/myloader

- CPU16
- 32GB
- SSD
-

64GB

Sysbench TPC-C TPC-H Sysbench TPC-C OLTP

3.1 Sysbench

TiDB MySQL OLTP OLAP TiDB MySQL sysbench

3.1.1

- sysbench

```
mkdir -p /tmp/sysbench
cd /tmp/sysbench
wget https://github.com/akopytov/sysbench/archive/1.0.14.tar.gz
yum -y install make automake libtool pkgconfig libaio-devel
yum -y install mariadb-devel
./autogen.sh
./configure
make -j
make install
sysbench --version
```

-

| | | | |
|-----------|--|---|----------------------|
| | | | |
| TiDB & PD | CPU2*E5-2650 v4@2.20GHz
128G
2*800G 3*1.6T SSD
2 * bond-1 | 3 | TiDB PD ext4 |
| TiKV | CPU2*E5-2650 v4@2.20GHz
256G
2*480G 4*1.92T NVMe SSD
2 * bond-1 | 3 | TiKV ext4PCIe |
| monitor | 8 32G800G | 3 | Grafana + Prometheus |

-

| | |
|-----------|---|
| | |
| | Redhat 7.4 |
| TiDB | TiDB-v3.0.5 |
| TiDB & PD | "TiDB" 2 tidb-server + 1 pd-server |
| TiKV | "TiKV" 4 tikv-server |
| TiDB | performance:
max-procs: 24 |
| TiKV | readpool:
coprocessor:
high-concurrency: 8
normal-concurrency: 8
low-concurrency: 8
storage:
block-cache:
capacity: "32GB" |

3.1.2

1.

- sysbench

```
mysql-host=192.168.xxx.xxx
mysql-port=4000
mysql-user=sysbench
mysql-password=*****
mysql-db=test
time=60
threads=16
report-interval=10
db-driver=mysql
```

- sysbench

```
--threads=8 8
--report-interval=10 10
--rand-type=uniform uniform ()gaussian ()special ()pareto ()
--time=120 120
--events=0 0 0
--percentile=99 95% 1% 99%
```

- sysbench

```
sysbench 1.0.14 (using bundled LuaJIT 2.1.0-beta2)
Running the test with following options:
Number of threads: 16
Report intermediate results every 10 second(s)
Initializing random number generator from current time
Initializing worker threads...
Threads started!
# 10 tps95%
[ 10s ] thds: 16 tps: 21532.38 qps: 21532.38 (r/w/o: 21532.38/0.00/0.00) lat (ms,95%): 1.04 err/s: 0.00 reconn/s: 0.0
0
[ 20s ] thds: 16 tps: 21617.20 qps: 21617.20 (r/w/o: 21617.20/0.00/0.00) lat (ms,95%): 1.01 err/s: 0.00 reconn/s: 0.0
0
[ 30s ] thds: 16 tps: 21550.98 qps: 21550.98 (r/w/o: 21550.98/0.00/0.00) lat (ms,95%): 1.03 err/s: 0.00 reconn/s: 0.0
0
[ 40s ] thds: 16 tps: 21544.16 qps: 21544.16 (r/w/o: 21544.16/0.00/0.00) lat (ms,95%): 1.01 err/s: 0.00 reconn/s: 0.0
0
[ 50s ] thds: 16 tps: 21639.76 qps: 21639.76 (r/w/o: 21639.76/0.00/0.00) lat (ms,95%): 0.99 err/s: 0.00 reconn/s: 0.0
0
[ 60s ] thds: 16 tps: 21597.56 qps: 21597.56 (r/w/o: 21597.56/0.00/0.00) lat (ms,95%): 1.01 err/s: 0.00 reconn/s: 0.0
0
SQL statistics:
queries performed:
    read:          1294886          #
    write:          0          #
    other:          0          # (COMMIT )
    total:          1294886          #
transactions:          1294886 (21579.74 per sec.)          # ( )
queries:          1294886 (21579.74 per sec.)          # ( )
ignored errors:          0      (0.00 per sec.)          # ( )
reconnects:          0      (0.00 per sec.)          # ( )
General statistics:
    total time:          60.0029s          #
    total number of events:          1294886          #
Latency (ms):
    min:          0.36          #
    avg:          0.74          #
    max:          8.90          #
    95th percentile:          1.01          # 95%
    sum:          959137.19          #
Threads fairness:
    events (avg/stddev):          80930.3750/440.48          # 80930.3750 event 440.48
    execution time (avg/stddev):          59.9461/0.00          # 59.9 0
```

- ```
sysbench --config-file=sysbench-thread-16.cfg oltp_point_select --tables=32 --table-size=10000000 prepare
```
- ```
block cache sbtest1 SQL
```

```
SELECT COUNT(pad) FROM sbtest1 USE INDEX(k_1);
```
- analyze sbtest* sbtest1


```
ANALYZE TABLE sbtest1;
```

2.

- Point select


```
sysbench --config-file=sysbench-thread-16.cfg oltp_point_select --tables=32 --table-size=10000000 run
```
- Update index


```
sysbench --config-file=sysbench-thread-16.cfg oltp_update_index --tables=32 --table-size=10000000 run
```
- Read-only

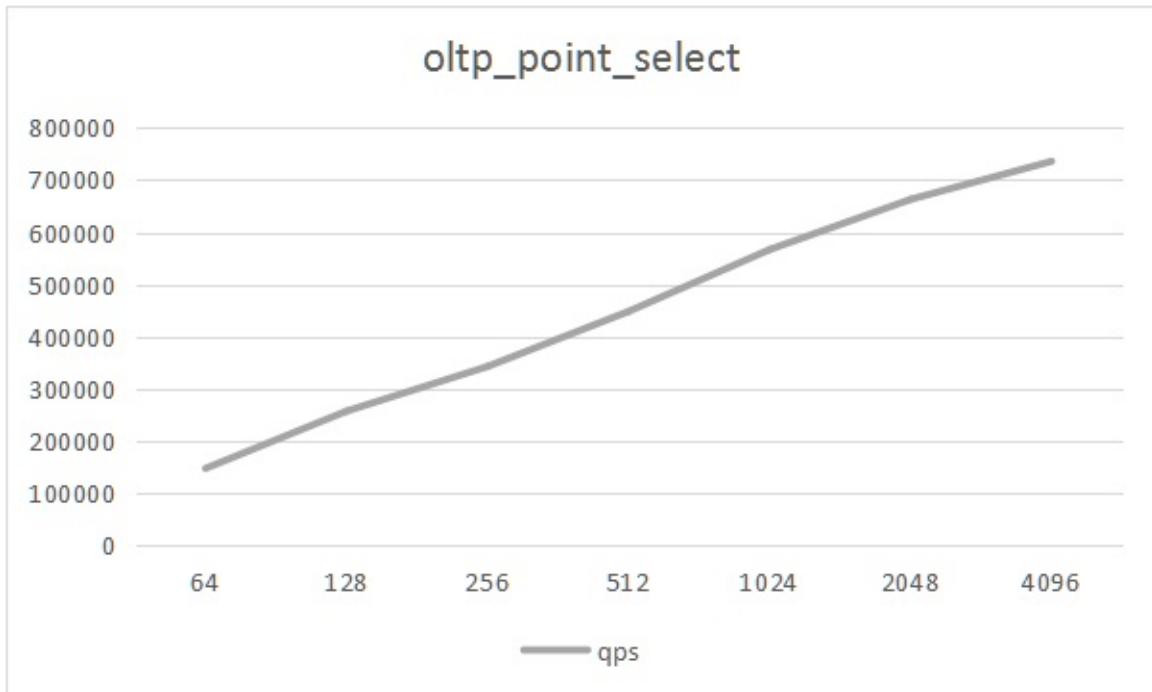

```
sysbench --config-file=sysbench-thread-16.cfg oltp_read_only --tables=32 --table-size=10000000 run
```
- Write-only


```
sysbench --config-file=sysbench-thread-16.cfg oltp_write_only --tables=32 --table-size=10000000 run
```
- Read-Write


```
sysbench --config-file=sysbench-thread-16.cfg oltp_read_write --tables=32 --table-size=10000000 run
```

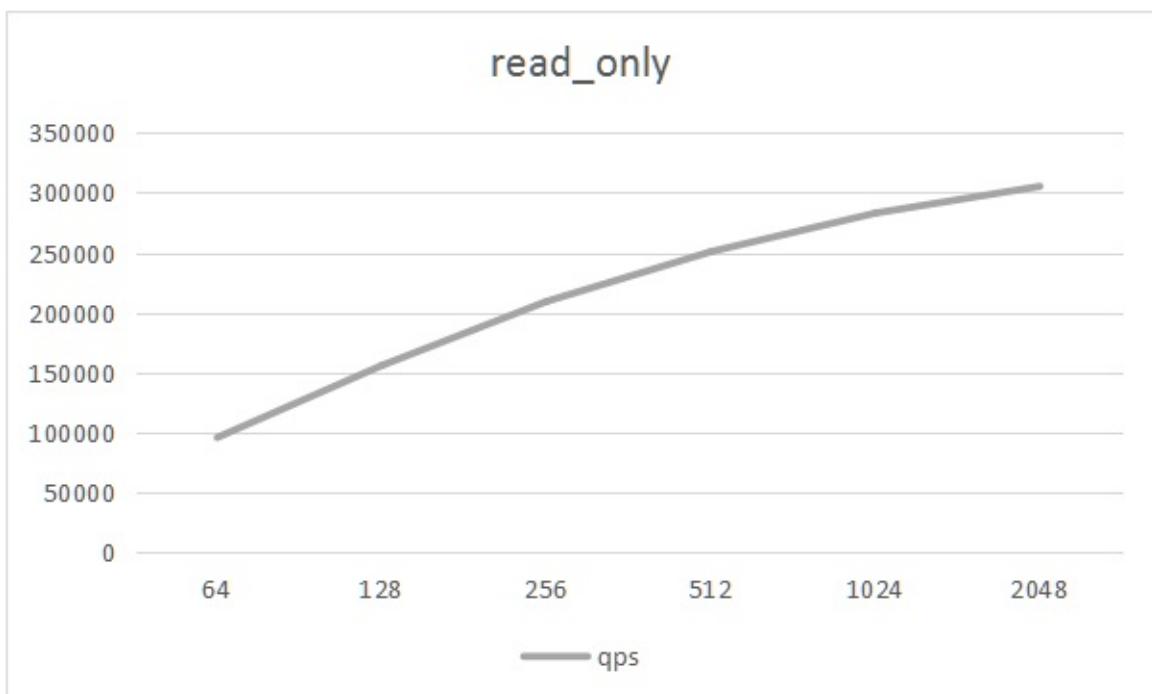
3.

- 32 10MB tidb-server sysbench
- oltp_point_select
- | type | thread | tps | qps | min latency | avg latency | 95th latency | max latency |
|--------------|---------------|------------|------------|--------------------|--------------------|---------------------|--------------------|
| point_select | 64 | 148098.00 | 148098.00 | 0.26 | 0.43 | 0.52 | 276.54 |
| point_select | 128 | 257760.00 | 257760.00 | 0.27 | 0.50 | 0.65 | 261.51 |
| point_select | 256 | 343215.00 | 343215.00 | 0.28 | 0.75 | 1.89 | 253.23 |
| point_select | 512 | 448683.00 | 448683.00 | 0.29 | 1.14 | 3.55 | 290.85 |
| point_select | 1024 | 567063.00 | 567063.00 | 0.30 | 1.80 | 5.57 | 70.21 |
| point_select | 2048 | 663217.00 | 663217.00 | 0.29 | 3.08 | 8.90 | 330.19 |
| point_select | 4096 | 736094.00 | 736094.00 | 0.33 | 5.55 | 15.00 | 431.72 |



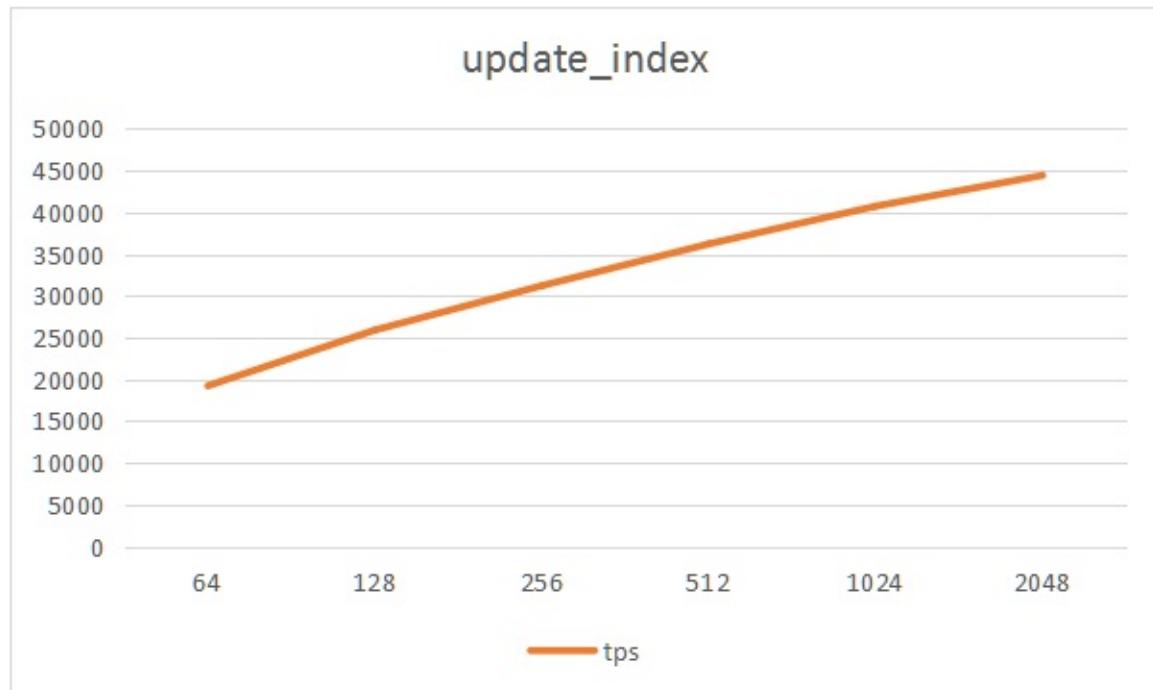
- read_only

| type | thread | tps | qps | min latency | avg latency | 95th latency | max latency |
|-----------|--------|----------|-----------|-------------|-------------|--------------|-------------|
| read_only | 64 | 5984.48 | 95751.60 | 7.87 | 10.69 | 14.21 | 85.24 |
| read_only | 128 | 9741.39 | 155862.00 | 7.64 | 13.14 | 18.28 | 236.37 |
| read_only | 256 | 13080.20 | 209284.00 | 9.22 | 19.56 | 28.16 | 99.79 |
| read_only | 512 | 15678.40 | 250854.00 | 10.40 | 32.62 | 49.34 | 115.78 |
| read_only | 1024 | 17691.40 | 283063.00 | 10.87 | 57.73 | 87.56 | 378.12 |
| read_only | 2048 | 19086.60 | 305386.00 | 7.68 | 107.12 | 164.45 | 710.91 |



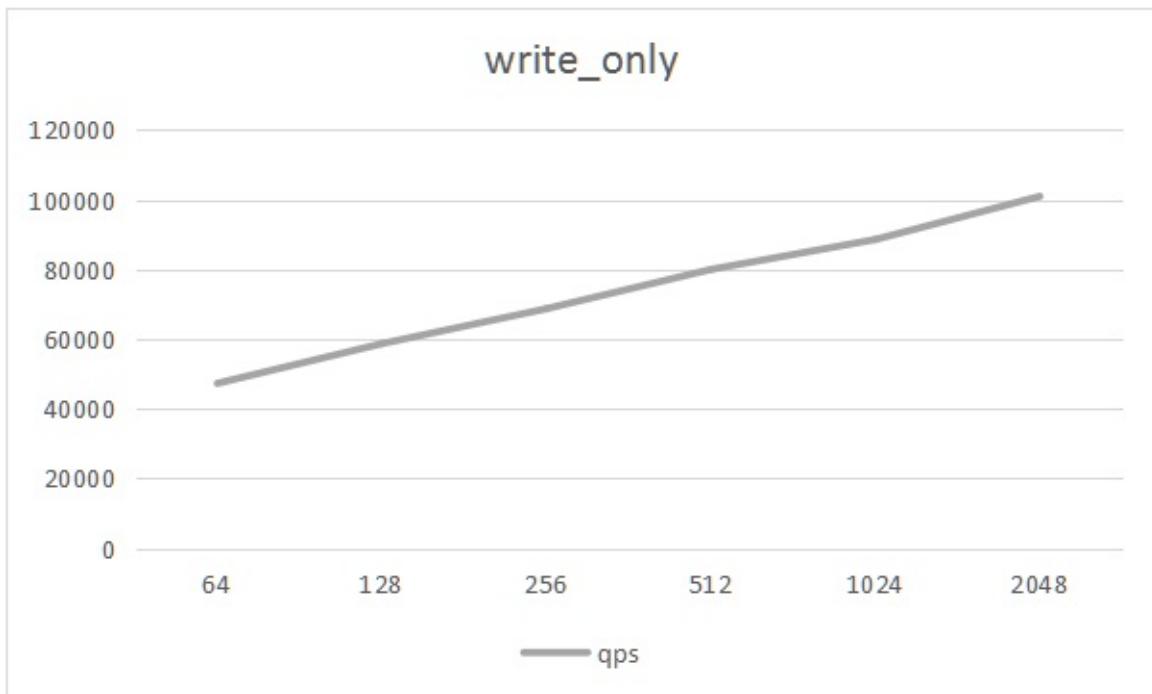
- oltp_update_index

| type | thread | tps | qps | min latency | avg latency | 95th latency | max latency |
|--------------|--------|----------|----------|-------------|-------------|--------------|-------------|
| update_index | 64 | 19232.10 | 19232.10 | 1.75 | 3.33 | 4.74 | 274.86 |
| update_index | 128 | 25898.20 | 25898.20 | 1.67 | 4.94 | 7.98 | 330.88 |
| update_index | 256 | 31214.00 | 31214.00 | 1.67 | 8.20 | 14.73 | 5189.46 |
| update_index | 512 | 36213.50 | 36213.50 | 1.74 | 14.13 | 27.66 | 5487.91 |
| update_index | 1024 | 40731.20 | 40731.20 | 1.74 | 25.12 | 52.89 | 7395.50 |
| update_index | 2048 | 44423.50 | 44423.50 | 1.77 | 46.04 | 99.33 | 5563.36 |



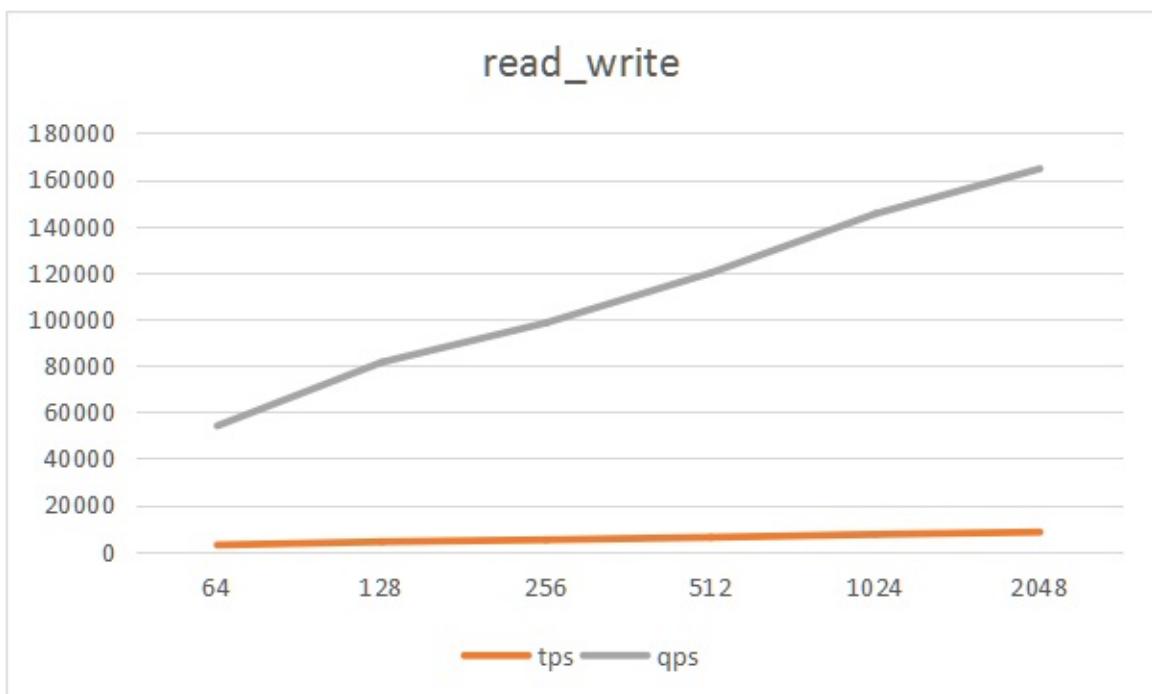
- write_only

| type | thread | tps | qps | min latency | avg latency | 95th latency | max latency |
|------------|--------|----------|-----------|-------------|-------------|--------------|-------------|
| write_only | 64 | 7882.92 | 47297.50 | 3.05 | 8.12 | 12.52 | 341.78 |
| write_only | 128 | 9780.01 | 58680.10 | 3.07 | 13.08 | 21.50 | 504.41 |
| write_only | 256 | 11450.20 | 68701.20 | 3.12 | 22.34 | 36.89 | 6874.97 |
| write_only | 512 | 13330.00 | 79979.20 | 3.04 | 38.39 | 65.65 | 6316.33 |
| write_only | 1024 | 14761.20 | 88567.40 | 3.30 | 68.39 | 118.92 | 5426.65 |
| write_only | 2048 | 16825.20 | 100951.00 | 3.25 | 121.50 | 223.34 | 5551.31 |



- **read_write**

| type | thread | tps | qps | min latency | avg latency | 95th latency | max latency |
|------------|--------|---------|-----------|-------------|-------------|--------------|-------------|
| read_write | 64 | 2698.01 | 53960.20 | 13.91 | 23.72 | 29.72 | 321.56 |
| read_write | 128 | 4066.40 | 81328.10 | 12.19 | 31.47 | 42.85 | 411.31 |
| read_write | 256 | 4915.23 | 98304.50 | 12.94 | 52.06 | 70.55 | 626.57 |
| read_write | 512 | 5988.96 | 119779.00 | 12.99 | 85.42 | 121.08 | 5023.20 |
| read_write | 1024 | 7260.19 | 145204.00 | 13.25 | 140.67 | 196.89 | 5767.52 |
| read_write | 2048 | 8228.84 | 164577.00 | 13.96 | 248.19 | 376.49 | 5475.98 |



3.1.3

TiDB MySQL TiDB MySQL MySQL TiDB

TiDB MySQL TiDB region KV region region TiDB TiKV MySQL

3.2 TPC-C

TiDB

[TPC-C](#)

1. TPC-C

TPC TPC-C Transaction Processing Performance Council OLTP TPC-C TPC-C

BenchmarkSQL TPCC PostgreSQL MySQL Oracle TiDB

2. BenchmarkSQL

TPC-C OLTP OLTP

- NewOrder –
- Payment –
- OrderStatus –
- Delivery –
- StockLevel –

TPC-C Benchmark ITEM 10 WAREHOUSE W

- STOCK W * 10 10
- DISTRICT W * 10 10
- CUSTOMER W * 10 * 3000 3000
- HISTORY W * 10 * 3000
- ORDER W * 10 * 3000 3000 900 NEW-ORDER 5 ~ 15 ORDER-LINE

TPC-C tpmC Transactions per Minute MQ Th Max Qualified Throughput Transactions NewOrder Transaction

3. TiDB

1000 warehouse 3

3 1 TiDB1 PD 1 TiKV

| | |
|------|---|
| OS | Linux (CentOS 7.3.1611) |
| CPU | 40 vCPUs, Intel(R) Xeon(R) CPU E5-2630 v4 @ 2.20GHz |
| RAM | 128GB |
| DISK | Optane 500GB SSD |

CPU NUMA taskset lscpu NUMA node

```
NUMA node0 CPU(s):    0,2,4,6,8,10,12,14,16,18,20,22,24,26,28,30,32,34,36,38
NUMA node1 CPU(s):    1,3,5,7,9,11,13,15,17,19,21,23,25,27,29,31,33,35,37,39
```

TiDB

```
nohup taskset -c 0,2,4,6,8,10,12,14,16,18,20,22,24,26,28,30,32,34,36,38 bin/tidb-server && \
nohup taskset -c 1,3,5,7,9,11,13,15,17,19,21,23,25,27,29,31,33,35,37,39 bin/tidb-server
```

HAproxy TiDB node nbproc CPU

| | TIDB | TIKV | PD |
|-------|-------------|-------------|-----------|
| node1 | 1 | 1 | 1 |
| node2 | 1 | 1 | 1 |
| node3 | 1 | 1 | 1 |

4. TiDB

1

```
[log]
level = "error"
```

2 CPU TiDB CPU

```
performance:
# Max CPUs to use, 0 use number of CPUs in the machine.
max-procs: 20
```

3 TiDB prepared plan cache

```
prepared_plan_cache:
  enabled: true
```

4 TiKV 16

```
tikv_client:
# Max gRPC connections that will be established with each tikv-server.
grpc-connection-count: 4
```

5

```
txn_local_latches:
# Enable local latches for transactions. Enable it when
# there are lots of conflicts between transactions.
enabled: true
```

5. TiKV

1 TiKV

```
global:
  log-level = "error"
```

2 sync-logTiKV Raft raftstore sync-log

```
[raftstore]
sync-log = false
```

3 TiKV RocksDB block cache 20 GB TiKV block cache

```
[storage.block-cache]
capacity = "10GB"
```

3 Grafana [TiKV] TiKV

6. BenchmarkSQL

[benchmarksql/run/props.mysql](#)

```
conn=jdbc:mysql://{{HAPROXY-HOST}}:{{HAPROXY-PORT}}/tpcc?useSSL=false&useServerPrepStmts=true&useConfigs=maxPerformance
warehouses=1000 # 1000 warehouse
terminals=500 # 500
loadWorkers=32 #
```

7.

TPC-C

1 TiDB-Server

```
create database tpcc
```

2 shell BenchmarkSQL

```
cd run && \
./runSQL.sh props.mysql sql.mysql/tableCreates.sql && \
./runSQL.sh props.mysql sql.mysql/indexCreates.sql
```

3

1 BenchmarkSQL

```
./runLoader.sh props.mysql
```

2 TiDB Lightning warehouse 1000 warehouse BenchmarkSQL csv TiDB Lightning Lightning csv

a BenchmarkSQL warehouse csv 77 MB csv

[benchmarksql/run/props.mysql](#)

```
fileLocation=/home/user/csv/ # csv
```

Lightning csv Lightning {database}.{table}.csv

```
fileLocation=/home/user/csv/tpcc. # csv + database
```

csv tpcc.bmsql_warehouse.csv Lightning

b csv

```
./runLoader.sh props.mysql
```

c inventory.ini

IP

```
[importer_server]
IS1 ansible_host=172.16.5.34 deploy_dir=/data2/is1 tikv_importer_port=13323 import_dir=/data2/import

[lightning_server]
LS1 ansible_host=172.16.5.34 deploy_dir=/data2/ls1 tidb_lightning_pprof_port=23323 data_source_dir=/home/user/csv
```

d conf/tidb-lightning.yml

```
mydumper:
  no-schema: true
csv:
  separator: ','
  delimiter: ''
  header: false
  not-null: false
  'null': 'NULL'
  backslash-escape: true
  trim-last-separator: false
```

e Lightning Importer

```
ansible-playbook deploy.yml --tags=lightning
```

f

- Lightning Importer
- Importer scripts/start_importer.sh Importer
- Lightning scripts/start_lightning.sh

ansible Lightning sql.common/test.sql | SQL

8.

BenchmarkSQL

```
nohup ./runBenchmark.sh props.mysql &> test.log &
```

test.log

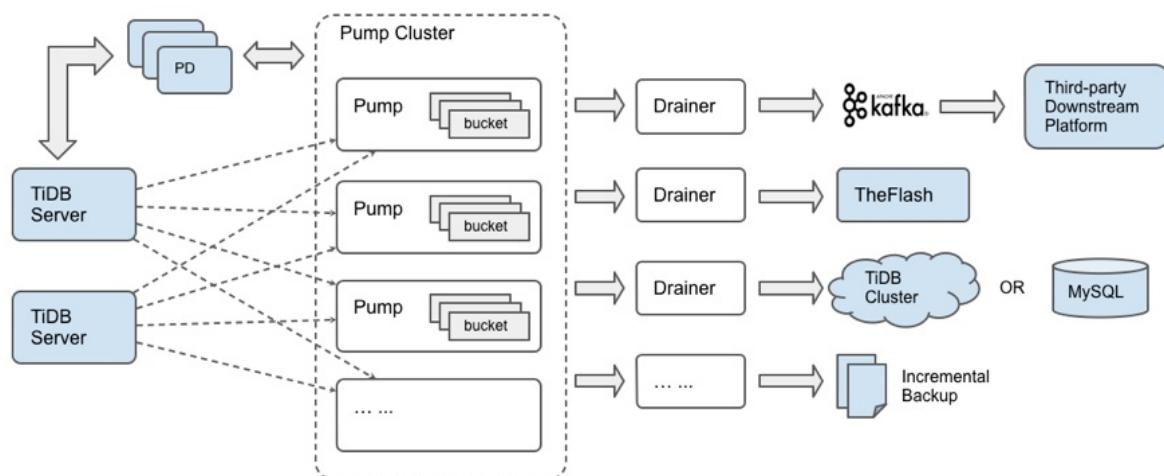
```
07:09:53,455 [Thread-351] INFO jTPCC : Term-00, Measured tpmC (NewOrders) = 77373.25
07:09:53,455 [Thread-351] INFO jTPCC : Term-00, Measured tpmTOTAL = 171959.88
07:09:53,455 [Thread-351] INFO jTPCC : Term-00, Session Start      = 2019-03-21 07:07:52
07:09:53,456 [Thread-351] INFO jTPCC : Term-00, Session End        = 2019-03-21 07:09:53
07:09:53,456 [Thread-351] INFO jTPCC : Term-00, Transaction Count = 345240
```

tpmC

TiDB-Binlog

TiDB Binlog TiDB binlog MySQL MySQL binlog TiDB Binlog TiDB TiDB binlog TiDB TiDB
TiDB Binlog

TiDB Binlog

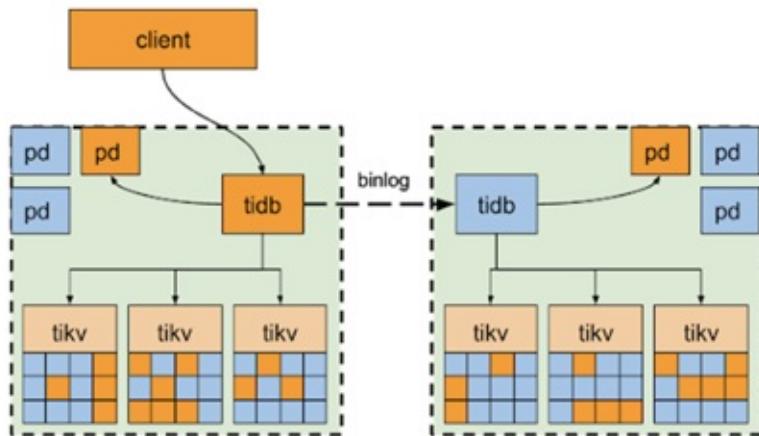


1. Pump Pump
2. TiDB Pump Client binlog Pump Pump TiDB
3. Pump kv binlog binlog
4. binlog Pump Pump
5. Drainer Pump binlog

TiDB-Binlog T+1

TiDB binlog MySQL master/slave

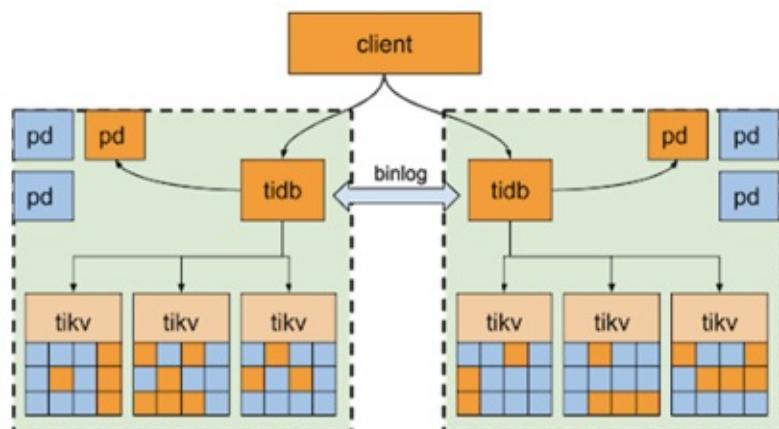
-
-



binlog

• *

TiDB binlog



•

Pump Drainer Intel x86-64 64

| | | CPU | | |
|---------|---|-----|---------------------|-----|
| Pump | 3 | 8+ | SSD, 200 GB+ | 16G |
| Drainer | 1 | 8+ | SAS, 100 GB+ binlog | 16G |

•

300M

1. Pump
 - i. tidb-ansible/inventory.ini (TiDB)
 - i. enable_binlog = True TiDB binlog ````

binlog trigger

enable_binlog = True

2. pump_servers IP

Binlog Part

[pump_servers] 172.16.10.72 172.16.10.73 172.16.10.74

3. Pump 7 tidb-ansible/conf/pump.ymlTiDB 3.0.2 tidb-ansible/conf/pump-cluster.yml gc

global:

an integer value to control the expiry date of the binlog data, which indicates for how long (in days) the binlog data would be stored

must be bigger than 0

gc: 7

4. binlog[](<https://pingcap.com/docs-cn/stable/how-to/deploy/orchestrated/ansible#%E8%B0%83%E6%95%B4%E9%83%A8%E7%BD%B2%E7%9B%AE%E5%BD%95>) Pump

Binlog Part

[pump_servers] pump1 ansible_host=172.16.10.72 deploy_dir=/data1/pump pump2 ansible_host=172.16.10.73 deploy_dir=/data2/pump pump3 ansible_host=172.16.10.74 deploy_dir=/data3/pump

2. Pump TiDB

1. pump_servers node_exporters

```
$ ansible-playbook deploy.yml --tags=pump -l ${pump1_ip},${pump2_ip},[$alias1_name,$alias2_name]
```

2. pump_servers

```
$ ansible-playbook start.yml --tags=pump
```

3. tidb_servers

```
$ ansible-playbook rolling_update.yml --tags=tidb
```

4.

```
$ ansible-playbook rolling_update_monitor.yml --tags=prometheus
```

3. Pump

```
binlogctl Pump pd-urls PD State online Pump
```

```
$ cd /home/tidb/tidb-ansible && resources/bin/binlogctl -pd-urls=http://172.16.10.72:2379 -cmd pumps INFO[0000] pump: {NodeID: ip-172-16-10-72:8250, Addr: 172.16.10.72:8250, State: online, MaxCommitTS: 403051525690884099, UpdateTime: 2018-12-25 14:23:37 +0800 CST} INFO[0000] pump: {NodeID: ip-172-16-10-73:8250, Addr: 172.16.10.73:8250, State: online, MaxCommitTS: 403051525703991299, UpdateTime: 2018-12-25 14:23:36 +0800 CST} INFO[0000] pump: {NodeID: ip-172-16-10-74:8250, Addr: 172.16.10.74:8250, State: online, MaxCommitTS: 403051525717360643, UpdateTime: 2018-12-25 14:23:35 +0800 CST}
```

2. TSO

```
Drainer initial_commit_ts initial_commit_ts
```

```
$ cd /home/tidb/tidb-ansible && resources/bin/binlogctl -pd-urls=http://127.0.0.1:2379 -cmd generate_meta INFO[0000] [pd] create pd client with endpoints [http://192.168.199.118:32379] INFO[0000] [pd] leader switches to: http://192.168.199.118:32379, previous: INFO[0000] [pd] init cluster id 6569368151110378289 2018/06/21 11:24:47 meta.go:117: [info] meta: & {CommitTS:400962745252184065}
```

```
meta: &{CommitTS:400962745252184065} CommitTS
```

3. mydumper

4. loader

5. Drainer Drainer

```
1. tidb-ansible/inventory.ini drainer_servers IPinitial_commit_ts initial_commit_ts Drainer
```

[drainer_servers]

```
drainer_mysql ansible_host=172.16.10.71 initial_commit_ts="402899541671542785"
```

2.

```
cd /home/tidb/tidb-ansible/conf && cp drainer.toml drainer_mysql_drainer.toml && vi drainer_mysql_drainer.toml
```

```
```
[syncer]
downstream storage, equal to --dest-db-type
Valid values are "mysql", "file", "tidb", "kafka".
db-type = "mysql"
the downstream MySQL protocol database
[syncer.to]
host = "172.16.10.72"
user = "root"
password = "123456"
port = 3306
```

### 1. Drainer

```
ansible-playbook deploy_drainer.yml
```

### 2. Drainer

```
ansible-playbook start_drainer.yml
```

Drainer

sync-diff-inspector PingCAP sync-diff-inspector

TiDB-Binlog TiDB TiDB MySQL Drainer PB TiDB-Binlog PingCAP

## 4.1

### 4.1.1 Raft

1.Raft TiDB PD TiKV Raft Raft

- Raft Raft leader
- Raft Raft Group leader Raft Group quorum leader

2.Raft TiDB Raft Raft Group n Raft Group

- n Raft Group (n-1)/2
- n Raft Group n/2 -1

3. Raft Group

- 1 1 4
- n/2 leader
- leader

4. Raft

- 1 3
- 1 3
- 1 3
- 1 3

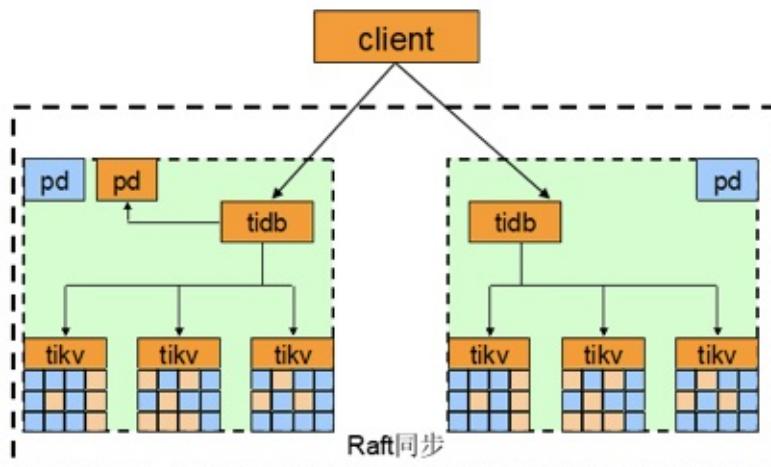
Raft 3 Raft on-premise 3 2 3

Raft 2:1 3 ACID

TiKV Raft Raft

### 4.1.2

TiDB Raft



3 6 Raft 3

- 
- 
- 
- RPO Recovery Point Objective 0 RTO Recovery Time Objective

### 4.1.3

Raft Raft Group quorum quorum Raft Group TiKV Raft Group quorum 2:1 Raft leader quorum 3 quorum 2

## 1. TiKV

TiKV custom quorum quorum  $Q1(n) = \text{floor}((n + 1) / 2) + 1$  quorum quorum 3 leader follower

(1)

TiKV max-idle-time region raft group leader max-idle-time Oracle DG // region raft group leader  
quorum  $Q2(n) = \text{floor}(n / 2) + 1$  2 term Raft log gap quorum  $Q1(n)$  PD

(2)

5 & 3 5%

## 2. TiDBPD

TiDB tidb-server PD 2:1 quorum PD pd-recover

3.

max-idle-time

### 4.1.4

(1)

max-idle-time max-idle-time region max-idle-time

(2)

TiKV max-idle-time max-idle-time max-store-down-time PD PD

(3)

TiKV region pd-recover PD tikv-ctl

## 4. RPO RTO

- RPO = 0
  - 0
- RTO
  - RTO max-idle-time
  - RTO max-idle-time

- RTO PD DBA TiKV DBA

# 1.

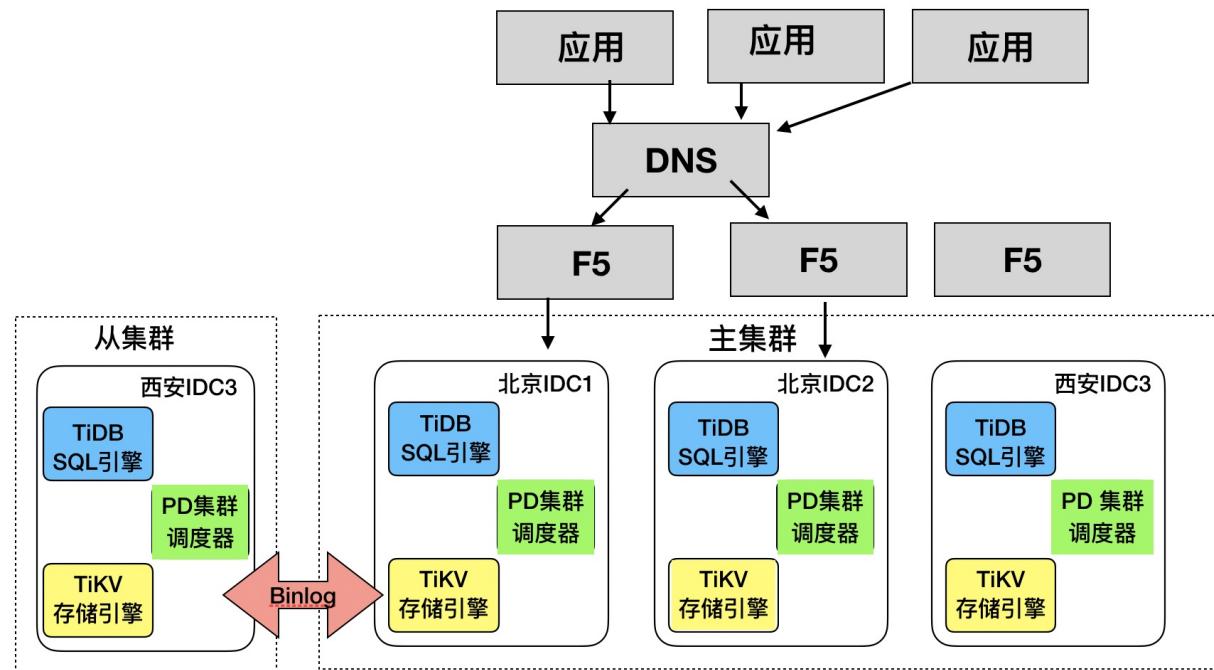
TiDB Raft Region leader PD leader

## 1.1

TiDB IDC1 IDC2 IDC3 3ms ISP 20ms

1

- TiDB-Binlog
- IDC1 IDC2 IDC3
- 5 IDC1 IDC2 2 IDC3 1 TiKV Label
- TiDB-Binlog



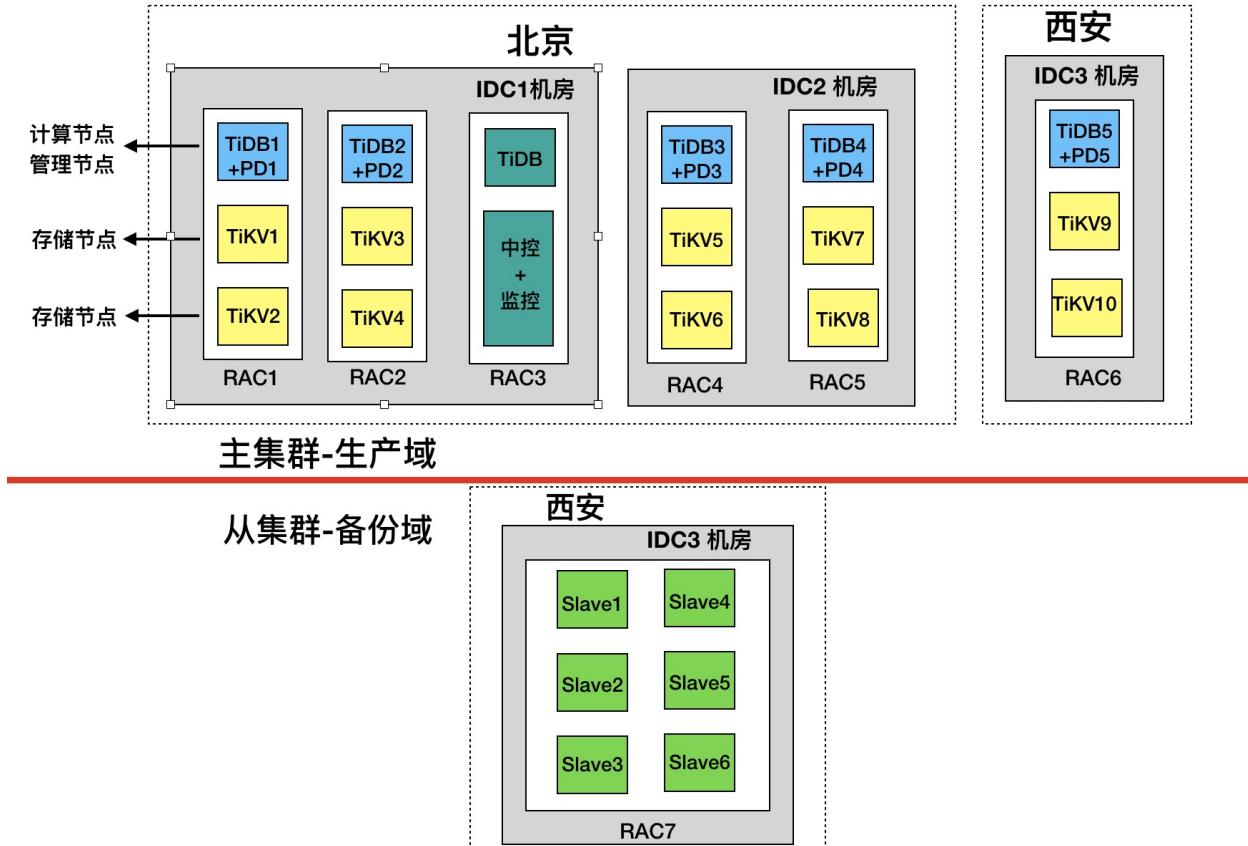
1

PD region leader region leader

- 
- 
- 

Raft Raft

## 1.2



2

1. 2 IDC1 IDC2 IDC1 RAC1RAC2RAC3 IDC2 RAC4RAC5 IDC3 RAC6RAC7 RAC7 TiDB T+1
2. RAC1 TiDBPD TiKV TiKV 2 TiKV tikv-server TiKV PCIe SSD TiKV RAC2RAC4RAC5RAC6
3. RAC3 TiDB-Server + TiDB-Server + TiDB-AnsiblePrometheusGrafana
4. 2 TiKV 3 TiDB PD
5. Mydumper Dranier Drainer PB

## 2.

### 2.1 inventory

inventory.ini

```
TiDB Cluster Part
[tidb_servers]
TiDB-10 ansible_host=10.63.10.10 deploy_dir=/data/tidb_cluster/tidb
TiDB-11 ansible_host=10.63.10.11 deploy_dir=/data/tidb_cluster/tidb
TiDB-12 ansible_host=10.63.10.12 deploy_dir=/data/tidb_cluster/tidb
TiDB-13 ansible_host=10.63.10.13 deploy_dir=/data/tidb_cluster/tidb
TiDB-14 ansible_host=10.63.10.14 deploy_dir=/data/tidb_cluster/tidb

[tikv_servers]
TiKV-30 ansible_host=10.63.10.30 deploy_dir=/data/tidb_cluster/tikv tikv_port=20171 labels="dc=1,rack=1,zone=1,host=30"
TiKV-31 ansible_host=10.63.10.31 deploy_dir=/data/tidb_cluster/tikv tikv_port=20171 labels="dc=1,rack=2,zone=1"
```

```

=2,host=31"
TiKV-32 ansible_host=10.63.10.32 deploy_dir=/data/tidb_cluster/tikv tikv_port=20171 labels="dc=2, rack=3, zone
=3,host=32"
TiKV-33 ansible_host=10.63.10.33 deploy_dir=/data/tidb_cluster/tikv tikv_port=20171 labels="dc=2, rack=4, zone
=4,host=33"
TiKV-34 ansible_host=10.63.10.34 deploy_dir=/data/tidb_cluster/tikv tikv_port=20171 labels="dc=3, rack=5, zone
=5,host=34"

[pd_servers]
PD-10 ansible_host=10.63.10.10 deploy_dir=/data/tidb_cluster/pd
PD-11 ansible_host=10.63.10.11 deploy_dir=/data/tidb_cluster/pd
PD-12 ansible_host=10.63.10.12 deploy_dir=/data/tidb_cluster/pd
PD-13 ansible_host=10.63.10.13 deploy_dir=/data/tidb_cluster/pd
PD-14 ansible_host=10.63.10.14 deploy_dir=/data/tidb_cluster/pd

[spark_master]

[spark_slaves]

[lightning_server]

[importer_server]

Monitoring Part
prometheus and pushgateway servers
[monitoring_servers]
proth-60 ansible_host=10.63.10.60 prometheus_port=8380 deploy_dir=/data/tidb_cluster/prometheus

[grafana_servers]
graf-60 ansible_host=10.63.10.60 grafana_port=8690 grafana_collector_port=8691 deploy_dir=/data/tidb_cluster/
grafana

node_exporter and blackbox_exporter servers
[monitored_servers]
10.63.10.10
10.63.10.11
10.63.10.12
10.63.10.13
10.63.10.14
10.63.10.30
10.63.10.31
10.63.10.32
10.63.10.33
10.63.10.34

[alertmanager_servers]
alertm ansible_host=10.63.10.60 deploy_dir=/data/tidb_cluster/alertmanager

[kafka_exporter_servers]

Binlog Part
[pump_servers]
pump-10 ansible_host=10.63.10.10 deploy_dir=/data/tidb_cluster/pump
pump-11 ansible_host=10.63.10.11 deploy_dir=/data/tidb_cluster/pump
pump-12 ansible_host=10.63.10.12 deploy_dir=/data/tidb_cluster/pump
pump-13 ansible_host=10.63.10.13 deploy_dir=/data/tidb_cluster/pump

[drainer_servers]

Group variables
[pd_servers:vars]
location_labels = ["dc", "rack", "zone", "host"]

Global variables
[all:vars]
deploy_dir = /data/tidb_cluster/tidb

Connection
ssh via normal user
ansible_user = tidb

cluster_name = test

```

```

tidb_version = v3.0.5

process supervision, [systemd, supervise]
process_supervision = systemd

timezone = Asia/Shanghai

enable_firewalld = False
check NTP service
enable_ntpd = False
set_hostname = False

binlog trigger
enable_binlog = True

kafka cluster address for monitoring, example:

zookeeper address of kafka cluster for monitoring, example:
zookeeper_addrs = "192.168.0.11:2181,192.168.0.12:2181,192.168.0.13:2181"

enable TLS authentication in the TiDB cluster
enable_tls = False

KV mode
deploy_without_tidb = False

wait for region replication complete before start tidb-server.
wait_replication = True

Optional: Set if you already have a alertmanager server.
Format: alertmanager_host:alertmanager_port
alertmanager_target = ""

grafana_admin_user = ""
grafana_admin_password = ""

Collect diagnosis
collect_log_recent_hours = 2
enable_bandwidth_limit = True
default: 10Mb/s, unit: Kbit/s
collect_bandwidth_limit = 10000

```

## 2.2 inventory

inventory.ini TiDB ansible-playbook -l

### 1.TiDB Servers

```

[tidb_servers]
TiDB-10 ansible_host=10.63.10.10 deploy_dir=/data/tidb_cluster/tidb
TiDB-11 ansible_host=10.63.10.11 deploy_dir=/data/tidb_cluster/tidb

```

### 2.TiKV Servers tikv label PD

```

[tikv_servers]
TiKV-30 ansible_host=10.63.10.30 deploy_dir=/data/tidb_cluster/tikv1 tikv_port=20171 labels="dc=1,rack=1,zone=1,host=30"
TiKV-31 ansible_host=10.63.10.31 deploy_dir=/data/tidb_cluster/tikv1 tikv_port=20171 labels="dc=1,rack=2,zone=2,host=31"

```

### 3.PD PD TiKV

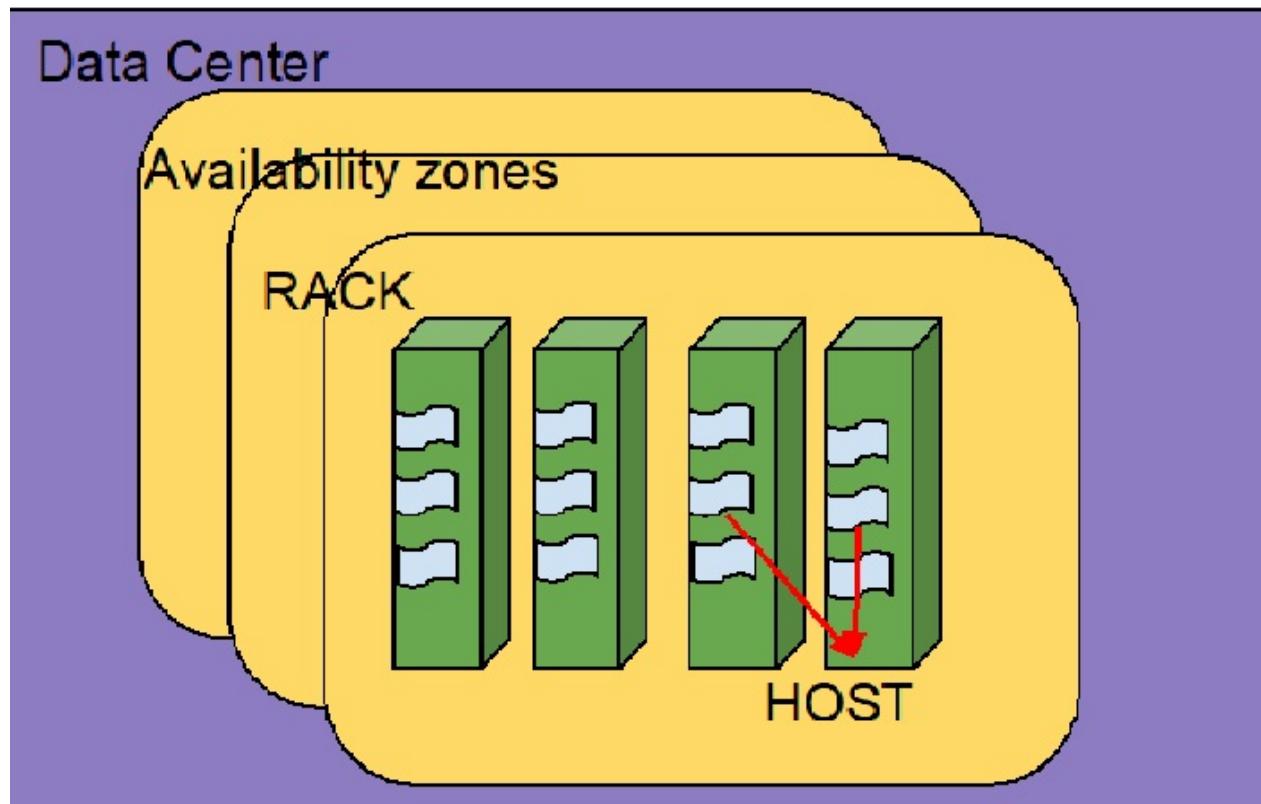
```

[pd_servers:vars]
location_labels = ["dc", "rack", "zone", "host"]

```

## 2.3 Labels

Labels DCAZRACKHOST



3 label

## 2.4

- tikv.yml

/tidb-ansible/conf/tikv.yml

```
+ block-cache-size
TiKV
```

capacity = MEM\_TOTAL \* 0.5 / TiKV :

```
...
Storage:
 block-cache:
 capacity: "1G"
...
```

- grpc gRPC

```
server:
 grpc-compression-type: gzip
```

- pd.yml /tidb-ansible/conf/pd.yml

PD balance PD PD score store leader Region Region size PD balance

```
```
schedule:
  tolerant-size-ratio: 20.0
```
```

- DC3 TiKV /tikv/conf/tikv.toml

TiKV raft DC3 TiKV DC3 TiKV

```
```
raftstore:
  raft-min-election-timeout-ticks= 1000
  raft-max-election-timeout-ticks= 1020
```
```

- PD control

- TiKV raft 5

```
config set max-replicas 5
```

- raft leader raft leader TiDB raft leader

```
config set label-property reject-leader dc dc3
```

- PD PD leader PD () PD

```
member leader_priority PD-10 5
member leader_priority PD-11 5
member leader_priority PD-12 5
member leader_priority PD-13 5
member leader_priority PD-14 1
```

## 4.4 AWS AZ TiDB

AWS (region)Availability Zones AZ TiDB

### 4.4.1

TiDB EC2 AWS TiDB AZ

#### 1.

PD AWS AZ TiKV region AWS AZ

- PD AZ PD
- max-replicas 3

AZ PD region AWS AZ Northern Virginia AZ AWS AZ

- PD AZ (2:2:1)
- max-replicas 5

max-replicas=3 AZ EC2

#### 2. (location labels)

TiKV

- 
- 
- (logical group)
- AZ

10 TiKV AZ 1a:1b:1c=2:4;4max-replicas=5

- PDlocation-labels = ["az","zone","host"]
- TiKV

<b>az</b>	<b>zone</b>	<b>host</b>
az=1a	zone=1	host=1a_1
az=1a	zone=1	host=1a_2
az=1b	zone=2	host=1b_1
az=1b	zone=2	host=1b_2
az=1b	zone=3	host=1b_3
az=1b	zone=3	host=1b_4
az=1c	zone=4	host=1c_1
az=1c	zone=4	host=1c_2
az=1c	zone=5	host=1c_3
az=1c	zone=5	host=1c_4

zone PD Region Peer AZ(3 AZ ,5 ) az.zone ( 5 zone5 ) az.zone Region peer zone az.zone 5

zone AZ az.zone.host Region peers

### 3. (placement groups)

[Placement Groups](#) Partition Placement Groups Spread Placement Groups placement groups

## 4.

AWS

(Recovery Time Objective, RTO)

### 4.4.2 AZ

AWS AZ AZ 1 AZ 23

AZ (1a, 1b, 1c) AZ (1a) AZ

#### Reject region leaders

AZ region leader

```
$ pd-ctl --pd="http://pd-url:2379"
» config set label-property reject-leader az 1a
» config show label-property
{
 "reject-leader": [
 {
 "key": "az",
 "value": "1a"
 }
]
}
```

#### 1. PD leader

AZ PD leader AZ PD leader 1a PD leader\_priority 3 1b 1c PD leader\_priority 5 leader\_priority

```

» member leader_priority pd_1a_1 3
Success!

» member leader_priority pd_1c_1 5
Success!

» member leader_priority pd_1b_1 5
Success!

» member
{
...
"members": [
 {
 "name": "pd_1a_1",
...
 "leader_priority": 3
 },
 {
 "name": "pd_1b_1",
...
 "leader_priority": 5
 },
...
 {
 "name": "pd_1c_1",
...
 "leader_priority": 5
 }
],
"leader": {
...
},
"etcd_leader": {
...
}
}

```

## 2. TiDB server

TiDB server AZ AZ AZ TiDB 1a1b 1c 1a TiDB 1b 1c1b 1c —— TiDB AZ

### 4.4.3

#### 1.

Adjust the alerts if you know for sure that triggered alerts are false alarms due to AWS hardware. For our deployment on AWS, alerts related to disk latency had to be adjusted in 2 ways:

- Increase “for” duration from 1m to 5m so that intermittent performance degradations for short period are ignored
- Increase the threshold so that lower hardware performance compared to the on-prem deployment is accounted for

AWS AWS

- “for” 1m 5m
- 

#### 2.

EC2 20

- TiDB
- EC2 —— TiDB



## 5

IBM Db2MongoDBMySQLOracle DRDSSQL ServerTiDBTiDB

- IBM Db2 TiDB (CDC)
- MongoDB TiD
- DM MySQL TiDB
- Oracle TiDB (OGG)
- DM MySQL TiDB
- DM MySQL TiDB
- SQLServer TiDB (DataX)
- SQL Server TiDB (yutong)
- TiDB TiDB (DataX)

## 5.1 DM MySQL TiDB

DM (Data Migration) MySQL MariaDB TiDB

DM

## 5.1.1 DM MySQL TiDB

MySQL TiDB

8 DM 1.0.2

- [DM](#)
- [DM](#)
- [DM](#)
- 
- [task](#)
- 
- 
- [tips](#)

### 1. DM

- &
- SQL
- 
- 

### 2. DM

- - 5.5 < MySQL < 8.0
  - MariaDB >= 10.1.2
- TiDB parser DDL
- `sql_model`
- binlog binlog\_format=ROW
- 

### 3. DM

DM 5

- DM-masterDM
- DM-woker
  - DM-worker MySQL MariaDB slave
  - DM DM-worker MySQL MariaDB
  - DM-worker
  - Dumper
  - Loader dumper TiDB
  - Syncer relay log binlog event event SQL SQL TiDB
- dmctlDM DM-Master
- task DM-worker DM-worker
- Prometheus

## 4.

- | DM-Master |  | 172.16.10.71 | 8261 |
|-----------|--|--------------|------|
| DM-worker |  | 172.16.10.72 | 8262 |
| MariaDB   |  | 172.16.10.81 | 3306 |
| TiDB      |  | 172.16.10.83 | 4000 |
- MySQL book session TiDB
  - mysql,information\_schema,percona,performance\_schema
  - droptruncate
  - book draft
- DMTiDBinventory.ini
  - [dm\_worker\_servers]
    - 1.server\_id MySQL TiDB
    - 2.source\_id task
    - 3.mysql\_password dmctl task
      - eg: dm-ansible/resources/bin/dmctl -encrypt
    - 3.enable\_gtid GTID MySQL GTID
    - 4.

```
[dm_worker_servers]
dm-worker1 ansible_host=1.1.1.1 source_id="mariadb-01" server_id=101 mysql_host=172.16.10.81 mysql_user=tidbdm mysql_password="encrytpwd" mysql_port=3306
```
- MySQL
  - REPLICATION SLAVE
    - REPLICATION CLIENT
    - RELOAD
    - SELECT
- TiDB
- task
  - (binlog position) binlog
  - Binlog
    - book
    - 
    -

## 5. task

task DM-Worker 9

- 

```
name: "taskname" # task
task-mode: all #
meta-schema: "dm_meta" # checkpoint
remove-meta: false # checkpoint
```

- target-database TiDB DM-Worker

- mysql-instances MySQL source-id

```
source-id: "mariadb-01" # dm-worker source-id
route-rules: ["book-route-rules-schema", "book-route-rules"] #
filter-rules: ["book-filter-1"] # binlog event
black-white-list: "bookblack" #

mydumper-config-name: "global" # mydumper
mydumper-thread: 4

loader-config-name: "global" # loader
loader-thread: 8

syncer-config-name: "global" # syncer
syncer-thread: 9
```

- routes

```
book-route-rules-schema: # mysql-instances
 schema-pattern: "book"
 target-schema: "book"
book-route-rules:
 schema-pattern: "book"
 table-pattern: "session"
 target-schema: "book"
 target-table: "session"
```

- filtersBinlog Event

```
book-filter-1:
 schema-pattern: "book"
 table-pattern: "session"
 events: ["truncate table", "drop table"]
 action: Ignore
```

- black-white-list

```
bookblack:
 do-dbs: ["~^book.*"]
 ignore-dbs: ["mysql", "performance_schema", "percona", "information_schema"]
 ignore-tables:
 - db-name: "book.*"
 tbl-name: "draft"
```

- mydumpers

```
global:
 mydumper-path: "./bin/mydumper"
 threads: 4
 chunk-filename-size: 64
 skip-tz-utc: true
```

- loaders

```
global:
 pool-size: 8
 dir: "./dumped_data"
```

- syncers

```
global:
 worker-count: 8
 batch: 100
```

## 6.

task dmctl

- dmctl DM-master task

```
./dmctl -master-addr 172.16.10.71:8261
```

- 

```
» check-task task-path
{
 "result": true,
 "msg": "check pass!!!"
}
```

- 

```
» start-task task-path
{
 "result": true,
 "msg": "",
 "workers": [
 {
 "result": true,
 "worker": "172.16.10.72:8262",
 "msg": ""
 }
]
}
```

- result true worker binlog

```
query-status taskname
```

- 

```
query-error taskname
```

- 

```
start-task taskname
```

- 

## 7.

- task
- resume-task
- SQL
- [Data Migration](#)

- task

```
"msg": "[code=44003:class=schema-tracker:scope=downstream:level=high] current pos (mysql-bin.000010, 814332497): failed to create table for `db_1`.`tb_1` in schema tracker: [types:1067]Invalid default value for 'expire_time'
```

- db\_1.tb\_1

```
expire_time datetime NOT NULL DEFAULT '0000-00-00 00:00:00'
```

- 
- sql\_mode datetime '0000-00-00 00:00:00'

```
TiDB> show variables like 'sql_mode';
+-----+-----+
| Variable_name | Value |
+-----+-----+
| sql_mode | STRICT_TRANS_TABLES,NO_ENGINE_SUBSTITUTION |
+-----+-----+
```

- sql\_mode task

## 8. Tips

- task MySQL
- task meta binlog
- 
- task [DM Portal](#)
- DM-worker Rely log

## 5.1.2 DM MySQL TiDB

### 5.1.2.1 DM

DRDS DM TiDB

- 1.
2. DBA
3. TiDB

ID DM

dm loader binlog

#### 1.

IP				
192.168.128.131	centos7.3 x86_64	MySQL5.7	3306	root/password
192.168.128.131	centos7.3 x86_64	MySQL5.7	3307	root/password
192.168.128.131	centos7.3 x86_64	MySQL5.7	3308	root/password
192.168.128.132	centos7.3 x86_64	dm-master/dmctl		root/password
192.168.128.133	centos7.3 x86_64	dm-worker	dm-worker	root/password
192.168.206.28	centos7.3 x86_64	TiDB	4000	root/password

#### 2.

**root 192.168.128.132**

```
[tidb@dmmaster ~]# yum -y install epel-release git curl sshpass
[tidb@dmmaster ~]# yum -y install python-pip
```

**tidb SSH**

1 tidb

```
[tidb@dmmaster ~]# useradd -m -d /home/tidb tidb
```

2 tidb

```
[tidb@dmmaster ~]# echo "password" | passwd --stdin tidb
```

3 tidb sudo

```
[tidb@dmmaster ~]# echo "tidb ALL=(ALL) NOPASSWD: ALL" >>/etc/sudoers
```

4 tidb home SSH

```
[tidb@dmmaster ~]# su - tidb
[tidb@dmmaster ~]$ ssh-keygen -t rsa
```

### tidb DM-Ansible

```
[tidb@dmmaster ~]$ wget https://download.pingcap.org/dm-ansible-v1.0.2.tar.gz
```

### DM-Ansible

```
[tidb@dmmaster ~]$ tar -xf dm-ansible-v1.0.2.tar.gz
[tidb@dmmaster ~]$ mv dm-ansible-v1.0.2.tar.gz dm-ansible
[tidb@dmmaster ~]$ cd /home/tidb/dm-ansible
[tidb@dmmaster dm-ansible]$ sudo pip install --upgrade pip
[tidb@dmmaster dm-ansible]$ sudo pip install -r ./requirements.txt
```

#### : ssh sudo

```
[tidb@dmmaster dm-ansible]$ cat hosts.ini
[servers]
192.168.128.132
192.168.128.133
[all:vars]
username = tidb
ansible_ssh_port = 22
ntp_server = ntp.aliyun.com
```

dm-worker tidb root password

```
[tidb@dmmaster dm-ansible]$ ansible-playbook -i hosts.ini create_users.yml -u root -k
SSH password:
PLAY [all]

TASK [create user]

changed: [192.168.128.133]
TASK [set authorized key]

changed: [192.168.128.133]
TASK [update sudoers file] *****
changed: [192.168.128.133]
PLAY RECAP

192.168.128.133 : ok=3 changed=3 unreachable=0 failed=0
```

### DM

```
[tidb@dmmaster dm-ansible]$ ansible-playbook local_prepare.yml
PLAY [do local preparation] *****
TASK [download : Stop if ansible version is too low, make sure that the Ansible version is Ansible 2.5.0 or later, otherwise a compatibility issue occurs.]

ok: [localhost] => {
 "changed": false,
 "msg": "All assertions passed"
}

localhost : ok=13 changed=5 unreachable=0 failed=0
```

### MySQL TiDB

```
root@localhost >grant Reload,Replication slave, Replication client,select on *.* to tidb@'%' IDENTIFIED by 'tidb@2020';
';
```

**dmctl**

```
[tidb@dmmaster bin]$ dmctl -encrypt tidb@2020
BXTTVvKeWhXgAefaFRNoN0BS4XjZ85uZByE=
```

**3. dm-worker****inventory.ini**

dm-master dm-worker dm-worker

```
[tidb@dmmaster dm-ansible]$ cat inventory.ini
DM modules
[dm_master_servers]
dm_master ansible_host=192.168.128.132
[dm_worker_servers]
dm_worker3306 ansible_host=192.168.128.133 deploy_dir=/data/mysql3306 dm_worker_port=13306 source_id="mysql3306" server_id=13306 mysql_host=192.168.128.131 mysql_user=tidb mysql_password=BXTTVvKeWhXgAefaFRNoN0BS4XjZ85uZByE mysql_port=3306
dm_worker3307 ansible_host=192.168.128.133 deploy_dir=/data/mysql3307 dm_worker_port=13307 source_id="mysql3307" server_id=13307 mysql_host=192.168.128.131 mysql_user=tidb mysql_password=BXTTVvKeWhXgAefaFRNoN0BS4XjZ85uZByE mysql_port=3307
dm_worker3308 ansible_host=192.168.128.133 deploy_dir=/data/mysql3308 dm_worker_port=13308 source_id="mysql3308" server_id=13308 mysql_host=192.168.128.131 mysql_user=tidb mysql_password=BXTTVvKeWhXgAefaFRNoN0BS4XjZ85uZByE mysql_port=3308
[dm_portal_servers]
dm_portal ansible_host=192.168.128.132
Monitoring modules
[prometheus_servers]
prometheus ansible_host=192.168.128.132
[grafana_servers]
grafana ansible_host=192.168.128.132
[alertmanager_servers]
alertmanager ansible_host=192.168.128.132
Global variables
[all:vars]
cluster_name = dm-cluster
ansible_user = tidb
ansible_port = 5622
dm_version = v1.0.2
deploy_dir = /home/tidb/deploy
grafana_admin_user = "admin"
grafana_admin_password = "admin"
```

```
inventory.ini
[dm_master_servers] dm-master dm-master
[dm_worker_servers] dm-worker dm-worker
----dm_worker3306 dmansible-playbook -l
----ansible_host dm-worker
----dm_worker_port dm-worker
----deploy_dir dm-worker
----source_id dm-worker source-id
----mysql_host MySQL
----mysql_user MySQL
----mysql_port MySQL
----mysql_password MySQL(dmctl,2.2)
```

**dm-worker**

dm-worker:

```
[tidb@dmmaster dm-ansible]$ ansible-playbook deploy.yml
PLAY RECAP ****
alertmanager : ok=13 changed=7 unreachable=0 failed=0
dm_master : ok=13 changed=8 unreachable=0 failed=0
dm_portal : ok=12 changed=5 unreachable=0 failed=0
dm_worker3306: ok=14 changed=2 unreachable=0 failed=0
dm_worker3307: ok=14 changed=2 unreachable=0 failed=0
dm_worker3308: ok=14 changed=2 unreachable=0 failed=0
grafana : ok=17 changed=10 unreachable=0 failed=0
localhost : ok=4 changed=3 unreachable=0 failed=0
prometheus : ok=15 changed=13 unreachable=0 failed=0
#
#
```

dm-worker:

```
[tidb@dmmaster dm-ansible]$ ansible-playbook start.yml
PLAY RECAP ****
alertmanager : ok=10 changed=1 unreachable=0 failed=0
dm_master : ok=10 changed=1 unreachable=0 failed=0
dm_portal : ok=9 changed=1 unreachable=0 failed=0
dm_worker3306: ok=11 changed=1 unreachable=0 failed=0
dm_worker3307: ok=11 changed=1 unreachable=0 failed=0
dm_worker3308: ok=11 changed=1 unreachable=0 failed=0
grafana : ok=13 changed=1 unreachable=0 failed=0
localhost : ok=4 changed=0 unreachable=0 failed=0
prometheus : ok=13 changed=4 unreachable=0 failed=0
dm binlog dm
```

## 4. & task

### TiDB

shard_db01	shard_tb01	merge_db	merge_tb
shard_db02	shard_tb02		
shard_db03	shard_tb03		
shard_db04	shard_tb04		
shard_db05	shard_tb05		
shard_db06	shard_tb06		

```

CREATE TABLE shard_tb01~06 (
 id bigint(20) NOT NULL AUTO_INCREMENT COMMENT 'ID',
 uid bigint(20) NOT NULL COMMENT 'ID',
 uname varchar(10) NOT NULL DEFAULT '' COMMENT '',
 gender tinyint(1) NOT NULL DEFAULT '0' COMMENT ' 0-1-',
 shard varchar(50) NOT NULL DEFAULT '' COMMENT '',
 mobile varchar(15) NOT NULL DEFAULT '' COMMENT '',
 PRIMARY KEY (id)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COMMENT='';
INSERT INTO shard_db01.shard_tb01
(uid,uname,gender,shard,mobile) VALUES
('10001','tb01001','0','shard_db01_tb01','136*****17'), ('10002','tb01002','1','shard_db01_tb01','136*****17');
INSERT INTO shard_db02.shard_tb02
(uid,uname,gender,shard,mobile) VALUES ('20001','tb02001','1','shard_db02_tb02','136*****33'), ('20002','tb02002','0',
,'shard_db02_tb02','139*****63');
INSERT INTO shard_db03.shard_tb03
(uid,uname,gender,shard,mobile) VALUES ('30001','tb03001','0','shard_db03_tb03','135*****73'),
('30002','tb03002','0','shard_db03_tb03','139*****46');

INSERT INTO shard_db04.shard_tb04
(uid,uname,gender,shard,mobile) VALUES ('40001','tb04001','0','shard_db04_tb04','137*****91'), ('40002','tb04002','1'
,'shard_db04_tb04','138*****91');
INSERT INTO shard_db05.shard_tb05
(uid,uname,gender,shard,mobile) VALUES ('50001','tb05001','1','shard_db05_tb05','158*****96'), ('50002','tb05002','0'
,'shard_db05_tb05','188*****92');
INSERT INTO shard_db06.shard_tb06
(uid,uname,gender,shard,mobile) VALUES ('60001','tb06001','1','shard_db06_tb06','178*****98'), ('60002','tb06002','1'
,'shard_db06_tb06','175*****31');

```

### task yaml

```

[tidb@tidb-dm-4-0-95 task]$ cat shardmysql_to_tidb.yaml
name: "shard_to_tidb" #task
is-sharding: true #
task-mode: "all" # full-incremental-all-
meta-schema: "tidb_meta" #
remove-meta: false
target-database:
 host: "192.168.206.28" # TiDB IP
 port: 4000 #TiDB
 user: "root" #TiDB
 password: "vLnqQt44rNFHSxA" # dmctl
mysql-instances:
 -
 source-id: "mysql3306" # inventory.ini source-id
 route-rules: ["rt000","rt001"] #
 filter-rules: ["ymdd-filter-rule"] #
 mydumper-config-name: "global"
 loader-config-name: "global"
 syncer-config-name: "global"
 black-white-list: "br01" #
 column-mapping-rules: ["cm001"] #

 -
 source-id: "mysql3307"
 route-rules: ["rt000","rt001"]
 filter-rules: ["ymdd-filter-rule"]
 mydumper-config-name: "global"
 loader-config-name: "global"
 syncer-config-name: "global"
 black-white-list: "br01"
 column-mapping-rules: ["cm002"]

 -
 source-id: "mysql3308"
 route-rules: ["rt000","rt001"]
 filter-rules: ["ymdd-filter-rule"]
 mydumper-config-name: "global"
 loader-config-name: "global"
 syncer-config-name: "global"
 black-white-list: "br01"
 column-mapping-rules: ["cm003"]

```

```

filters:
 ymdd-filter-rule:
 schema-pattern: "shard_db *"
 # 2 binlog
 events: ["truncate table","delete","drop table","drop database"]
 action: Ignore

routes:
 rt000:
 # shard_db* merge_db
 schema-pattern: "shard_db*"
 target-schema: "merge_db"
 rt001:
 # shard_tb* merge_db merge_tb
 schema-pattern: "shard_db*"
 table-pattern: "shard_tb??"
 target-schema: "merge_db"
 target-table: "merge_tb"

 #

 #

column-mappings:
 cm001:
 schema-pattern: "shard_db*"
 table-pattern: "shard_tb??"
 expression: "partition id"
 source-column: "id"
 target-column: "id"
 arguments: ["1","shard_db","shard_tb"]
 cm002:
 schema-pattern: "shard_db*"
 table-pattern: "shard_tb??"
 expression: "partition id"
 source-column: "id"
 target-column: "id"
 arguments: ["2","shard_db","shard_tb"]
 cm003:
 schema-pattern: "shard_db*"
 table-pattern: "shard_tb??"
 expression: "partition id"
 source-column: "id"
 target-column: "id"
 arguments: ["3","shard_db","shard_tb"]

 #
 black-white-list:
 br01:
 do-dbs: ["~shard_db*"] #
 #
 ignore-dbs: ["mysql","performance_schema","information_schema"]
 #
 ignore-tables:
 - db-name: "~shard_db*"
 tbl-name: "~txc_undo_log*"

 #
 mydumpers:
 global:
 threads:
 chunk-filename-size: 64
 skip-tz-utc: true
 extra-args: " --no-locks "
 loaders:
 global:
 pool-size: 64
 dir: "./dumped_data"
 syncers:
 global:
 worker-count: 6
 batch: 1000

```

```
[tidb@dmworker dmctl]$ dmctl -master-addr 192.168.128.132:8261
» start-task shard_to_tidb.yaml
» query-status
{
 "result": true,
 "msg": "",
 "tasks": [
 {
 "taskName": "shard_to_tidb",
 "taskStatus": "Running",
 "workers": [
 "192.168.128.133:53306",
 "192.168.128.133:53307",
 "192.168.128.133:53308"
]
 }
]
}
#
```

**TiDB dm-worker**

```
[root@dmworker dumped_data.shard_to_tidb]# ll |awk '{print $NF}'
metadata
shard_db01-schema-create.sql
shard_db01.shard_tb01-schema.sql
shard_db01.shard_tb01.sql
shard_db02-schema-create.sql
shard_db02.shard_tb02-schema.sql
shard_db02.shard_tb02.sql
```

```
[root@dmworker dumped_data.shard_to_tidb]# ls | grep -v schema | xargs rm -f
[root@dmworker dumped_data.shard_to_tidb]# ls
shard_db01.shard_tb01-schema.sql shard_db02.shard_tb02-schema.sql
shard_db02-schema-create.sql
```

**merge\_db merge\_tb**

对象 merge\_tb @merge\_db (tidb...)

开始事务 备注 筛选 排序 导入 导出

id	uid	uname	gender	shard	mobile	
580981944116838401	10001	tb01001		0 shard_db01_tb01	136	17
580981944116838402	10002	tb01002		1 shard_db01_tb01	136	17
585503135930253313	20001	tb02001		1 shard_db02_tb02	136	33
585503135930253314	20002	tb02002		0 shard_db02_tb02	139	63
1166485080047091713	30001	tb03001		0 shard_db03_tb03	135	73
1166485080047091714	30002	tb03002		0 shard_db03_tb03	139	46
1171006271860506625	40001	tb04001		0 shard_db04_tb04	137	91
1171006271860506626	40002	tb04002		1 shard_db04_tb04	138	91
1751988215977345025	50001	tb05001		1 shard_db05_tb05	158	96
1751988215977345026	50002	tb05002		0 shard_db05_tb05	188	92
1756509407790759937	60001	tb06001		1 shard_db06_tb06	178	98
1756509407790759938	60002	tb06002		1 shard_db06_tb06	175	31

## 5.1.2.2 DM

### 1. dm-worker

#### deploy.yml

```
#inventory.ini
[tidb@dmworker dm-ansible]$ ansible-playbook deploy.yml
-1 mysql3306 dm-worker
[tidb@dmworker dm-ansible]$ ansible-playbook deploy.yml -l mysql3306
--tags deploy.yml dm-worker
[tidb@dmworker dm-ansible]$ ansible-playbook deploy.yml --tags=dm-worker
```

#### dm-worker

```
dm MySQL binlog
MySQL Slave_IO_Running
[tidb@dmworker dm-ansible]$ ansible-playbook start.yml
dm MySQL binlog
[tidb@dmworker dm-ansible]$ ansible-playbook stop.yml
dm
[tidb@dmworker dm-ansible]$ ansible-playbook rolling_update.yml
*yml-l--tags
```

### 2. dm-worker task

#### task dmctl help

```

[tidb@dmworker dm-ansible]$ dmctl -master-addr 192.168.128.132:8261
»help
```

#### start-task task MySQL Slave\_SQL\_Running

```
[tidb@dmworker dm-ansible]$ dmctl -master-addr 192.168.128.132:8261
task.yaml dm-worker
»start-task shard_to_tidb.yaml
shard_to_tidb.yaml dm-worker
shard_to_tidb.yaml 192.168.128.133:53307:
»start-task -w '192.168.128.133:53307' shard_to_tidb.yaml
```

#### stop-task MySQL Slave\_SQL\_Running

```
[tidb@dmworker dm-ansible]$ dmctl -master-addr 192.168.128.132:8261
shard_to_tidb.yaml dm-worker
#-w [-w IP:PORT]
»start-task shard_to_tidb
```

#### query-status

```
[tidb@dmworker dm-ansible]$ dmctl -master-addr 192.168.128.132:8261
shard_to_tidb
»query-status shard_to_tidb
```

#### query-error

```
[tidb@dmworker dm-ansible]$ dmctl -master-addr 192.168.128.132:8261
shard_to_tidb
»query-error shard_to_tidb
```

**skip\_sql SQL**

```
binlog (failedBinlogPosition)
query-error shard_to_tidb
binlog
sql-skip --worker=192.168.128.133:53307 --binlog-pos=mysql-bin|000001.000003:737983 shard_to_tidb
#
resume-task --worker=192.168.128.133:53307 shard_to_tidb
#
query-error shard_to_tidb
```

DBA RDBMS NoSQL

TiDB RDBMS NoSQL MySQL TiDB TiDB TiDB TiDB

TiDB OGG Oracle TiDB

## 5.2.1

### 1.

Oracle GoldenGate 12.3.0.1.4 for Oracle on Linux x86-64

Oracle GoldenGate 12.3.0.1.5 for MySQL-compatible DBs on Linux x86-64

Oracle 11G Oracle OGG TiDB

Oracle ZHS16GBK

### 2.

#### 2.1.

Oracle TiDB Navicat

#### 2.2.

OGG initial load Oracle TiDB Oracle initial load Oracle

#### 2.3.

initial load logOGG handlecolisions

## 3. Oracle

- 
- Force logging
- ENABLE\_GOLDENGATE\_REPLICATION true (11.2.0.4)
- - 
  - Schema
  -
- - DBA

## 4. TiDB

- set tidb\_constraint\_check\_in\_place = 1;

  | TiDB commit insert abend OGG session

- lower-case-table-names = 1

  | OGG 1tidb

## 5. OGG

## DDL

- Oracle

```
create table account (
 account_number number(10,0),
 account_balance decimal(38,2),
 account_trans_ts timestamp(6),
 account_trans_type varchar2(30),
 primary key (account_number)
 using index
);
```

- TiDB

```
create table account (
 account_number int,
 account_balance decimal(38,2),
 account_trans_ts timestamp,
 account_trans_type varchar(30),
 primary key (account_number));
```

## 5.2.2

### 1. OGG - Oracle

```
$unzip V975837-01.zip
$ ls -l
total 1201820
drwxr-xr-x 3 oracle oinstall 4096 Apr 16 2018 fbo_ggs_Linux_x64_shiphome
-rw-r--r-- 1 oracle oinstall 1396 May 10 2018 OGG-12.3.0.1.4-README.txt
-rw-r--r-- 1 oracle oinstall 293566 May 10 2018 OGG_WinUnix_Rel_Notes_12.3.0.1.4.pdf
```

(1)

```
$vi ./fbo_ggs_Linux_x64_shiphome/Disk1/response/oggcore.rsp
#####
Copyright(c) Oracle Corporation 2017. All rights reserved.
##
Specify values for the variables listed below to customize
your installation.
##
Each variable is associated with a comment. The comment
can help to populate the variables with the appropriate
values.
##
IMPORTANT NOTE: This file should be secured to have read
permission only by the oracle user or an administrator who
own this installation to protect any sensitive input values.
##
#####
#-----
Do not change the following system generated value.
#-----
oracle.install.responseFileVersion=/oracle/install/rspfmt_ogginstall_response_schema_v12_1_2
#####
##
```

```

Oracle GoldenGate installation option and details
##
#####
#-----#
Specify the installation option.
Specify ORA12c for installing Oracle GoldenGate for Oracle Database 12c and
ORA11g for installing Oracle GoldenGate for Oracle Database 11g
#-----
INSTALL_OPTION=ORA11g
Oracle 11.2.0.4 ORA11G 12c ORA12c
#-----
Specify a location to install Oracle GoldenGate
#-----
SOFTWARE_LOCATION=/home/oracle/ogg12.3
##
#-----
Specify true to start the manager after installation.
#-----
START_MANAGER=false
mgr false
#-----
Specify a free port within the valid range for the manager process.
Required only if START_MANAGER is true.
#-----
MANAGER_PORT=
#-----
Specify the location of the Oracle Database.
Required only if START_MANAGER is true.
#-----
DATABASE_LOCATION=

#####
Specify details to Create inventory for Oracle installs
Required only for the first Oracle product install on a system.

#####
#-----#
Specify the location which holds the install inventory files.
This is an optional parameter if installing on
Windows based Operating System.
#-----
INVENTORY_LOCATION=
#-----
Unix group to be set for the inventory directory.
This parameter is not applicable if installing on
Windows based Operating System.
#-----
UNIX_GROUP_NAME=

```

(2)

fbo\_ggs\_Linux\_x64\_shiphome/Disk1/runInstaller

```

$./fbo_ggs_Linux_x64_shiphome/Disk1/runInstaller -silent -responseFile /home/oracle/oggsoft/fbo_ggs_Linux_x64_shiphom
e/Disk1/response/oggcore.rsp
response

```

successful

The installation of Oracle GoldenGate Core was successful

OGG

```

$ ls -l
total 225044
-rwxr-xr-x 1 oracle oinstall 426 Oct 15 2010 bcpfmt.tpl

```

```

-rwxr-xr-x 1 oracle oinstall 1725 Oct 15 2010 bcrypt.txt
-rwxrwxr-x 1 oracle oinstall 1612776 Apr 15 2018 cachefiledump
drwxr-xr-x 4 oracle oinstall 4096 Jul 28 14:45 cfgtoollogs
-rwxrwxr-x 1 oracle oinstall 3563576 Apr 15 2018 checkprm
-rw-rw-r-- 1 oracle oinstall 1021 Apr 15 2018 chkpt_ora_create.sql
-rwxrwxr-x 1 oracle oinstall 3379568 Apr 15 2018 convchk
-rwxrwxr-x 1 oracle oinstall 4716080 Apr 15 2018 convprm
drwxr-xr-x 2 oracle oinstall 4096 Jul 28 14:45 crypto
-rwxr-xr-x 1 oracle oinstall 159 Oct 15 2010 db2cntl.tpl
-rwxrwxr-x 1 oracle oinstall 9696 Apr 15 2018 db_upgrade
-rw-rw-r-- 1 oracle oinstall 455 Apr 15 2018 ddl_cleartrace.sql
-rw-rw-r-- 1 oracle oinstall 8414 Apr 15 2018 ddl_create.sql
-rw-rw-r-- 1 oracle oinstall 3176 Apr 15 2018 ddl_ddl2file.sql
-rw-rw-r-- 1 oracle oinstall 90 Apr 15 2018 ddl_disable.sql
-rw-rw-r-- 1 oracle oinstall 88 Apr 15 2018 ddl_enable.sql
-rw-rw-r-- 1 oracle oinstall 2036 Apr 15 2018 ddl_filter.sql
-rw-rw-r-- 1 oracle oinstall 12220 Apr 15 2018 ddl_ora10.sql
-rw-rw-r-- 1 oracle oinstall 1725 Apr 15 2018 ddl_ora10upCommon.sql
-rw-rw-r-- 1 oracle oinstall 13539 Apr 15 2018 ddl_ora11.sql
-rw-rw-r-- 1 oracle oinstall 12564 Apr 15 2018 ddl_ora9.sql
-rw-rw-r-- 1 oracle oinstall 216 Apr 15 2018 ddl_pin.sql
-rw-rw-r-- 1 oracle oinstall 3184 Apr 15 2018 ddl_remove.sql
-rw-rw-r-- 1 oracle oinstall 1 Apr 15 2018 ddl_session1.sql
-rw-rw-r-- 1 oracle oinstall 629 Apr 15 2018 ddl_session.sql
-rw-rw-r-- 1 oracle oinstall 287877 Apr 15 2018 ddl_setup.sql
-rw-rw-r-- 1 oracle oinstall 8401 Apr 15 2018 ddl_status.sql
-rw-rw-r-- 1 oracle oinstall 2122 Apr 15 2018 ddl_staymetadata_off.sql
-rw-rw-r-- 1 oracle oinstall 2118 Apr 15 2018 ddl_staymetadata_on.sql
-rw-rw-r-- 1 oracle oinstall 2186 Apr 15 2018 ddl_tracelevel.sql
-rw-rw-r-- 1 oracle oinstall 2133 Apr 15 2018 ddl_trace_off.sql
-rw-rw-r-- 1 oracle oinstall 2383 Apr 15 2018 ddl_trace_on.sql
-rwxrwxr-x 1 oracle oinstall 5037440 Apr 15 2018 defgen
drwxr-xr-x 2 oracle oinstall 4096 Jul 28 14:45 deinstall
-rw-rw-r-- 1 oracle oinstall 882 Apr 15 2018 demo_more_ora_create.sql
-rw-rw-r-- 1 oracle oinstall 649 Apr 15 2018 demo_more_ora_insert.sql
-rw-rw-r-- 1 oracle oinstall 583 Apr 15 2018 demo_ora_create.sql
-rw-rw-r-- 1 oracle oinstall 504 Apr 15 2018 demo_ora_insert.sql
-rw-rw-r-- 1 oracle oinstall 3597 Apr 15 2018 demo_ora_lob_create.sql
-rw-rw-r-- 1 oracle oinstall 1943 Apr 15 2018 demo_ora_misc.sql
-rw-rw-r-- 1 oracle oinstall 1056 Apr 15 2018 demo_ora_pk_befores_create.sql
-rw-rw-r-- 1 oracle oinstall 1013 Apr 15 2018 demo_ora_pk_befores_insert.sql
-rw-rw-r-- 1 oracle oinstall 2305 Apr 15 2018 demo_ora_pk_befores_updates.sql
drwxr-xr-x 3 oracle oinstall 4096 Jul 28 14:45 diagnostics
drwxr-xr-x 3 oracle oinstall 4096 Jul 28 14:45 diretc
drwxr-xr-x 2 oracle oinstall 4096 Jul 28 14:45 dirout
drwxr-xr-x 4 oracle oinstall 4096 Jul 28 14:45 dirsca
-rwxrwxr-x 1 oracle oinstall 4272400 Apr 15 2018 emsclnt
-rwxrwx--- 1 oracle oinstall 12545040 Apr 15 2018 extract
-rwxr-xr-x 1 oracle oinstall 1968 Oct 15 2010 freeBSD.txt
-rwxrwxr-x 1 oracle oinstall 4280528 Apr 15 2018 ggcmd
-rwxr-xr-x 1 oracle oinstall 2303056 Apr 15 2018 ggMessage.dat
-rwxr-xr-x 1 oracle oinstall 49675440 Apr 15 2018 ggparam.dat
-rwxrwx--- 1 oracle oinstall 9340192 Apr 15 2018 ggsci
drwxr-xr-x 2 oracle oinstall 4096 Jul 28 14:45 healthcheck
-rwxr-xr-x 1 oracle oinstall 299451 Nov 16 2017 help.txt
drwxr-xr-x 3 oracle oinstall 4096 Jul 28 14:45 install
drwxr-x--- 12 oracle oinstall 4096 Jul 28 14:45 inventory
drwxr-xr-x 7 oracle oinstall 4096 Jul 28 14:45 jdk
-rwxrwxr-x 1 oracle oinstall 144496 Apr 15 2018 keygen
-rw-rw-r-- 1 oracle oinstall 56 Apr 15 2018 label.sql
-rwxrwx--- 1 oracle oinstall 102840 Apr 15 2018 libantlr3c.so
-rwxrwxr-x 1 oracle oinstall 12312 Apr 15 2018 libboost_system-mt.so.1.58.0
-rwxrwx--- 1 oracle oinstall 2190856 Apr 15 2018 libdb-6.1.so
-rwxrwx--- 1 oracle oinstall 2198136 Apr 15 2018 libglog.so
-rwxrwx--- 1 oracle oinstall 10524064 Apr 15 2018 libgnnzitp.so
-rwxrwx--- 1 oracle oinstall 21993240 Apr 15 2018 libggparam.so
-rwxrwx--- 1 oracle oinstall 210264 Apr 15 2018 libggperf.so
-rwxrwx--- 1 oracle oinstall 779352 Apr 15 2018 libggrepo.so
-rwxrwx--- 1 oracle oinstall 1108760 Apr 15 2018 libggssl.so
-rwxrwx--- 1 oracle oinstall 125624 Apr 15 2018 libgutil.so
-rwxrwxr-x 1 oracle oinstall 29764432 Apr 15 2018 libicudata.so.56
-rwxrwxr-x 1 oracle oinstall 2909360 Apr 15 2018 libicui18n.so.56

```

```

-rwxrwxr-x 1 oracle oinstall 1995808 Apr 15 2018 libicuuc.so.56
-rwxrwx--- 1 oracle oinstall 86960 Apr 15 2018 liblmbdb.so
-rwxrwxr-x 1 oracle oinstall 175136 Apr 15 2018 libPocoCrypto.so.48
-rwxrwxr-x 1 oracle oinstall 3079760 Apr 15 2018 libPocoFoundation.so.48
-rwxrwxr-x 1 oracle oinstall 373232 Apr 15 2018 libPocoJSON.so.48
-rwxrwxr-x 1 oracle oinstall 1326504 Apr 15 2018 libPocoNet.so.48
-rwxrwxr-x 1 oracle oinstall 350064 Apr 15 2018 libPocoNetSSL.so.48
-rwxrwxr-x 1 oracle oinstall 520864 Apr 15 2018 libPocoUtil.so.48
-rwxrwxr-x 1 oracle oinstall 680288 Apr 15 2018 libPocoXML.so.48
-rwxrwx--- 1 oracle oinstall 1115360 Apr 15 2018 libudt.so
-rwxrwx--- 1 oracle oinstall 4782344 Apr 15 2018 libxerces-c-3.1.so
-rwxrwxr-x 1 oracle oinstall 5025840 Apr 15 2018 logdump
-rw-rw-r-- 1 oracle oinstall 1553 Apr 15 2018 marker_remove.sql
-rw-rw-r-- 1 oracle oinstall 3309 Apr 15 2018 marker_setup.sql
-rw-rw-r-- 1 oracle oinstall 675 Apr 15 2018 marker_status.sql
-rwxrwxr-x 1 oracle oinstall 6570480 Apr 15 2018 mgr
-rwxr-xr-x 1 oracle oinstall 41643 Jun 30 2017 notices.txt
-rwxrwxr-x 1 oracle oinstall 1661024 Apr 15 2018 oggerr
drwxr-xr-x 12 oracle oinstall 4096 Jul 28 14:45 OPatch
-rw-r----- 1 oracle oinstall 59 May 30 13:35 oraInst.loc
drwxr-xr-x 8 oracle oinstall 4096 Jul 28 14:45 oui
-rw-rw-r-- 1 oracle oinstall 3146 Apr 15 2018 params.sql
-rwxrwxr-x 1 oracle oinstall 11524576 Apr 15 2018 pmsrvr
-rwxr-xr-x 1 oracle oinstall 1272 Dec 28 2010 prvtclkm.plb
-rwxr-xr-x 1 oracle oinstall 9487 May 27 2015 prvtlmpg.plb
-rw-rw-r-- 1 oracle oinstall 2724 Apr 15 2018 prvtlmpg_uninstall.sql
-rw-rw-r-- 1 oracle oinstall 1532 Apr 15 2018 remove_seq.sql
-rwxrwx--- 1 oracle oinstall 10873240 Apr 15 2018 replicat
-rwxrwxr-x 1 oracle oinstall 1656864 Apr 15 2018 retrace
-rw-rw-r-- 1 oracle oinstall 3187 Apr 15 2018 role_setup.sql
-rw-rw-r-- 1 oracle oinstall 35254 Apr 15 2018 sequence.sql
-rwxrwxr-x 1 oracle oinstall 4659736 Apr 15 2018 server
-rwxr-xr-x 1 oracle oinstall 4917 Jan 5 2017 SQLDataTypes.h
-rwxr-xr-x 1 oracle oinstall 248 Oct 15 2010 sqldr.tpl
drwxr-xr-x 3 oracle oinstall 4096 Jul 28 14:45 srvm
-rwxrwxr-x 1 oracle oinstall 759 Oct 15 2010 tcperrs
-rwxr-xr-x 1 oracle oinstall 37877 Apr 16 2016 ucharset.h
-rw-rw-r-- 1 oracle oinstall 7341 Apr 15 2018 ulg.sql
drwxr-xr-x 7 oracle oinstall 4096 Jul 28 14:45 UserExitExamples
-rwxr-xr-x 1 oracle oinstall 32987 Jun 2 2017 usrdecs.h
-rwxr-xr-x 1 oracle oinstall 1033 Oct 19 2016 zlib.txt

```

## (3) .bash\_profile

```
export LD_LIBRARY_PATH=$ORACLE_HOME/lib
```

## (4)

```
$./ggsci
Oracle GoldenGate Command Interpreter for Oracle
Version 12.3.0.1.4 OGGCORE_12.3.0.1.0_PLATFORMS_180415.0359_F80
Linux, x64, 64bit (optimized), Oracle 11g on Apr 15 2018 21:16:09
Operating system character set identified as UTF-8.
Copyright (C) 1995, 2018, Oracle and/or its affiliates. All rights reserved.

GGSCI (hostname) 1> create subdirs
Creating subdirectories under current directory /home/oracle/ogg12.3
Parameter file /home/oracle/ogg12.3/dirprm: created.
Report file /home/oracle/ogg12.3/dir rpt: created.
Checkpoint file /home/oracle/ogg12.3/dir chk: created.
Process status files /home/oracle/ogg12.3/dir pcs: created.
SQL script files /home/oracle/ogg12.3/dir sql: created.
Database definitions files /home/oracle/ogg12.3/dir def: created.
Extract data files /home/oracle/ogg12.3/dir ddat: created.
Temporary files /home/oracle/ogg12.3/dir tmp: created.
Credential store files /home/oracle/ogg12.3/dir crd: created.
Masterkey wallet files /home/oracle/ogg12.3/dir wlt: created.
Dump files /home/oracle/ogg12.3/dir dmp: created.
```

## Oracle OGG

**2. OGG - TiDB**

(1)

```
$ unzip V978711-01.zip
Archive: V978711-01.zip
 inflating: ggs_Linux_x64(MySQL)_64bit.tar
 inflating: OGG-12.3.0.1-README.txt
 inflating: OGG_WinUnix_Rel_Notes_12.3.0.1.5.pdf
```

(2) ggs\_Linux\_x64(MySQL)\_64bit.tar .

```
$tar xvf ./ggs_Linux_x64(MySQL)_64bit.tar -C /home/tidb/ogg12.3
```

(3) ggsci

```
$./ggsci
Oracle GoldenGate Command Interpreter for MySQL
Version 12.3.0.1.5 OGGCORE_12.3.0.1.0_PLATFORMS_180501.2124
Linux, x64, 64bit (optimized), MySQL Enterprise on May 2 2018 10:58:16
Operating system character set identified as UTF-8.
Copyright (C) 1995, 2018, Oracle and/or its affiliates. All rights reserved.

GGSCI (hostname) 1> create subdirs
Creating subdirectories under current directory /home/tidb/ogg12.3
Parameter file /home/tidb/ogg12.3/dirprm: created.
Report file /home/tidb/ogg12.3/dir rpt: created.
Checkpoint file /home/tidb/ogg12.3/dir chk: created.
Process status files /home/tidb/ogg12.3/dir pcs: created.
SQL script files /home/tidb/ogg12.3/dir sql: created.
Database definitions files /home/tidb/ogg12.3/dir def: created.
Extract data files /home/tidb/ogg12.3/dir ddat: created.
Temporary files /home/tidb/ogg12.3/dir tmp: created.
Credential store files /home/tidb/ogg12.3/dir crd: created.
Masterkey wallet files /home/tidb/ogg12.3/dir wlt: created.
Dump files /home/tidb/ogg12.3/dir dmp: created.
```

## OGG TiDB (OGG For MySQL)

### 3. TiDB

(1) lower-case-table-names 1

```
$grep lower-case-table-names tidb.toml
lower-case-table-names = 1
```

(2)

```
MySQL [(none)]> show variables like '%lower%';
+-----+-----+
| Variable_name | Value |
+-----+-----+
| lower_case_table_names | 1 |
| lower_case_file_system | 1 |
+-----+-----+
MySQL [(none)]> show variables like '%place%' ;
+-----+-----+
| Variable_name | Value |
+-----+-----+
| tidb_constraint_check_in_place | 0 |
+-----+-----+
```

(3)

```
MySQL [scott]> create user 'tidb' identified by 'tidb';
Query OK, 1 row affected (0.01 sec)
MySQL [scott]> GRANT ALL PRIVILEGES ON scott.* to oggadmin ;
Query OK, 0 rows affected (0.01 sec)
MySQL [scott]> show create user tidb;
+-----+
| CREATE USER for tidb@% |
+-----+
| CREATE USER 'tidb'@'%' IDENTIFIED WITH 'mysql_native_password' AS '465D123EE55795DBDBDAE36AFD3DCD9C429B718A' REQUIRE NONE PASSWORD EXPIRE DEFAULT ACCOUNT UNLOCK |
+-----+
1 row in set (0.00 sec)
```

(4)

### 4. Oracle

4.1

```
SQL> archive log list;
```

(1)

```
SQL> shutdown immediate;
```

(2) mount

```
SQL> startup mount;
```

(3)

```
SQL> alter database archivelog;
```

(4)

```
SQL> alter database open;
```

(5)

```
SQL> archive log list;
```

4.2

(1)

```
select log_mode, supplemental_log_data_min, force_logging from v$database;
```

(2) force logging

```
nolog
```

```
SQL> ALTER DATABASE FORCE LOGGING;
```

(3)

```
SQL> ALTER DATABASE ADD SUPPLEMENTAL LOG DATA;
SQL> ALTER SYSTEM SWITCH LOGFILE;
```

(4) OGG

```
SQL> alter system set enable_goldengate_replication=true;
```

4.3 ogg

goldengate

```
SQL> create tablespace goldengate datafile '/opt/oracle/data/goldengate001.dbf' size 4G;
SQL> create user goldengate identified by goldengate default tablespace goldengate;
SQL> grant dba to goldengate;
```

## 5. OGG

(1) ggsic

```
./ggscli
create subdirs
add credentialstore
alter credentialstore add user goldengate,password goldengate
```

(2) ogg goldengate

```
dblogin useralias goldengate
```

## 6. schema

schema schema

## (1) 11.2.0.2 Oracle Patch 13794550

```
oracle schema
```

ERROR OGG-06522 Cannot verify existence of table function that is required to ADD schema level supplemental logging, failed.

## (2) GGSCI

- 1) dblogin userid goldengate, password goldengate

- 2) dblogin useridalias goldengate

## (3) schema

```
ADD SCHEMATRANSACTIONAL schema ALLCOLS
```

## 7.

```
schema
```

## (1) GGSCI

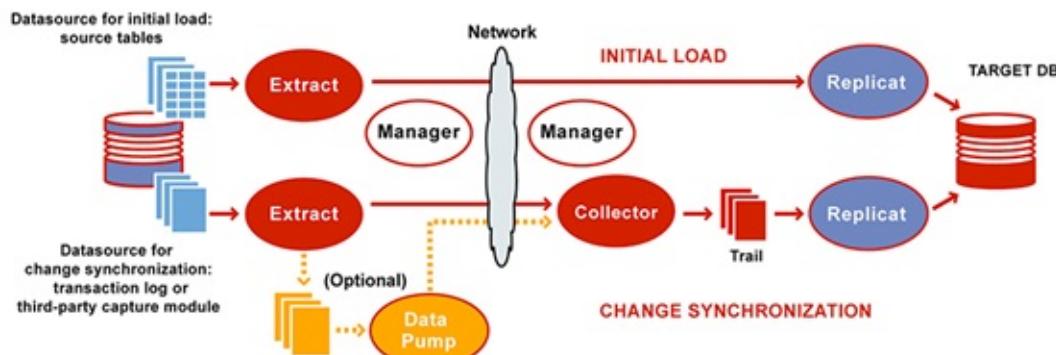
```
dblogin userid goldengate, password goldengate
```

## (2) schema

```
ADD TRANSACTIONAL schema.tablename NOKEY
```

## 5.2.3 OGG

### 1. OGG



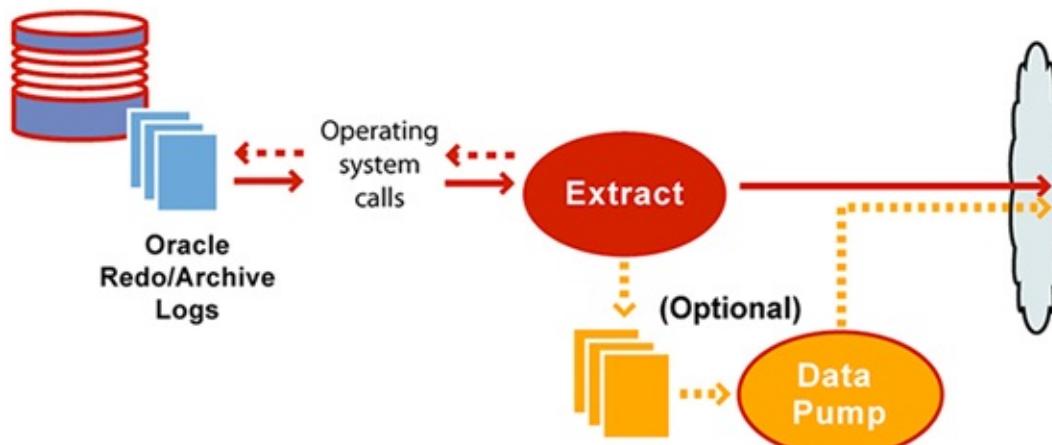
- INITIAL LOAD
- CHANGE SYNCHRONIZATION trail oracle oracle redo log archive log trail

/opt/oracle/11.2.0	ORACLE_HOME
ora	Oracle sid
ext_ora	
tdpm_ora	
rep_ora	
./dirdat/sk	
/Data/tidb/ogg/dirdat/sk	Replicat

## 2.

ADG

Extract pump



Extract redo archive log Data pump

## 2.1 MGR

Manager OGG ExtractPumpReplicat ExtractPumpReplicat Manager

(1)

```

GGSCI (b-db-ps-055) 5> view param mgr

port 9001
DYNAMICPORTLIST 9001-9020
--AUTORESTART ER *,RETRIES 5,WAITMINUTES 7
PURGEOLDEXTRACTS ./dirdat/*,usecheckpoints, minkeepdays 2
LAGREPORTHOURS 1
LAGINFOMINUTES 30
LAGCRITICALMINUTES 45

```

port	
DYNAMICPORTLIST	OGG OGG
AUTORESTART	
PURGEOLDEXTRACTS	trail 2 trail
LAGREPORTHOURS	lag 1
LAGINFOMINUTES	

(2) mgr

```
./ggsci
start mgr
```

## 2.2 Extract

Extract OGG Extract

(1)

```
GGSCI (b-db-ps-055) 2> edit param ext_ora

EXTRACT ext_ora
SETENV (ORACLE_HOME="/opt/oracle/11.2.0")
setenv (NLS_LANG="AMERICAN_AMERICA.AL32UTF8")
setenv (ORACLE_SID="ora")
useridalias goldengate
GETTRUNCATES
REPORTCOUNT EVERY 1 MINUTES, RATE
DISCARDFILE ./dir rpt/ext_ora.dsc,APPEND,MEGABYTES 1000
WARNLONGTRANS 2h,CHECKINTERVAL 10m
EXTTRAIL ./dirdat/sk
TRANLOGOPTIONS EXCLUDEUSER goldengate
TRANLOGOPTIONS MINEFROMACTIVEVG
DBOPTIONS ALLOWUNUSEDCOLUMN
DYNAMICRESOLUTION
FETCHOPTIONS FETCHKUPDATECOLS,INCONSISTENTROW ALLOW
ddl include mapped objname hr.*
getupdatebefores
nocompressdeletes
nocompressupdates
table hr.table_name;
```

|||:----:----| SETENV ||| useridalias ||| GETTRUNCATES | truncate || REPORTCOUNT |||  
 WARNLONGTRANS | OGG || EXTTRAIL | trail || TRANLOGOPTIONS EXCLUDEUSER |||  
 TRANLOGOPTIONS MINEFROMACTIVEVG | ADG || DBOPTIONS ALLOWUNUSEDCOLUMN |||  
 DYNAMICRESOLUTION | OGG OGG OGG OGG ID  
 FETCHKUPDATECOLS,INCONSISTENTROW ALLOW | row id () ALLOW |

(2)

```
./ggsci
add extract ext_ora,TRANLOG, begin now
add ExtTrail ./dirdat/sk, Extract ext_ora, Megabytes 50
```

## 2.3 Pump

Pump Extract Pump Extract Trail Trail

```
GGSCI (b-db-ps-055) 4> edit param tdpm_ora

EXTRACT tdpm_ora
RMTHOST 31.***.***.***, MGRPORT 9001, compress
PASSTHRU
RMTTRAIL /Data/tidb/ogg/dirdat/sk
DYNAMICRESOLUTION
--table
table hr.table_name;
```

| | |:----|----| RMTHOST | IP MGR || PASSTHRU | OGG |

(1) pump

```
./ggsci
add extract tdpm_ora, extrailsource ./dirdat/sk
add rmttrail /Data/tidb/ogg/dirdat/sk, extract tdpm_ora, megabytes 100
```

### 3.

## 3.1 MGR

(1)

```
GGSCI (w-db-ps-082) 2> edit param mgr

port 9001
DYNAMICPORTLIST 9001-9120
ACCESSRULE, PROG *, IPADDR 31.*.*.* , ALLOW
--AUTORESTART ER *, RETRIES 5, WAITMINUTES 7
PURGEOLDEXTRACTS ./dirdat/*, usecheckpoints, minkeepdays 10
LAGREPORTHOURS 1
LAGINFOMINUTES 30
LAGCRITICALMINUTES 45
```

(2) mgr

```
./ggsci
start mgr
```

## 3.2 Replicat

Replicat Extract DDL

(1) checkpoint table

```
dblogin sourcedb tidb@31.***.***.***:4000 userid username password user_password
add checkpointtable tidb.checkpoint_table
```

(2) replicat

```
GGSCI (w-db-ps-082) 4> edit param rep_ora

replicat rep_ora
targetdb tidb@31.***.***.***:4000 userid username password user_password
SQLEXEC "set tidb_constraint_check_in_place=1"
handlecollisions
MAXTRANSOPS 10000
discardfile /Data/tidb/ogg/dirrpt/reporadsc, purge
map hr.table_name, target tidb.table_name, keycols(ID);
```

```
| | |:----|----| SQLEXEC | tidb_constraint_check_in_place 0 INSERT OGG 1 || handlecollisions | replicat
replicat insertrepliact () |
```

(3) replicat

```
add replicat rep_ora,exttrail /Data/tidb/ogg/dirdat/sk,checkpointtable tidb.checkpoint_table
```

## 4.

### 4.1

(1)

```
GGSCI (w-db-ps-082) 4>edit param init_ora

extract init_ora
setenv (NLS_LANG=AMERICAN_AMERICA.AL32UTF8)
setenv (ORACLE_SID="ora")
useridalias goldengate
rmthost 31.***.***.***,mgrport 9001
rmttask replicat,group rnit_ora
table hr.table_name;
```

rmttask replicat,group	replicat

(2)

```
./ggsci
add extract init_ora,sourceistable
```

### 4.2

(1)

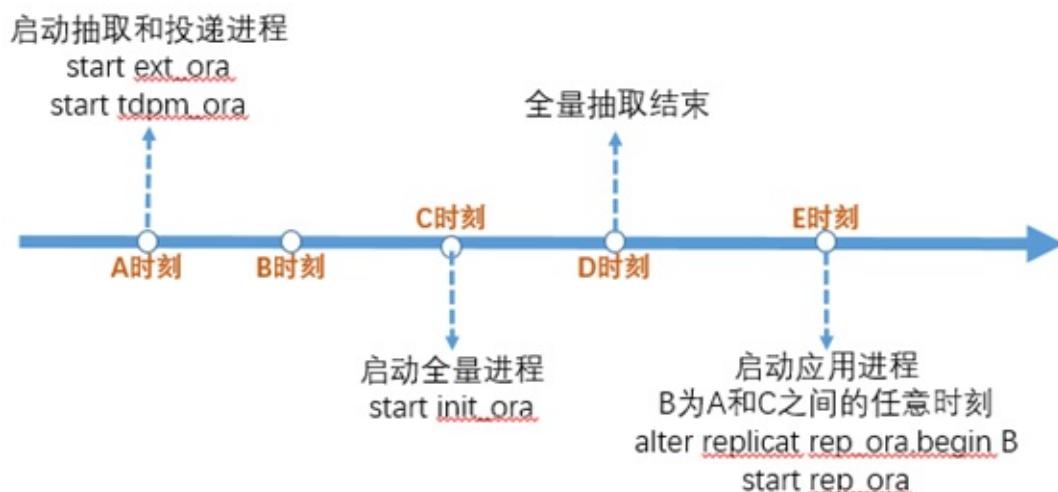
```
GGSCI (w-db-ps-082) 4>edit param rnit_ora

replicat rnit_ora
SETENV (NLS_LANG = AMERICAN_AMERICA.AL32UTF8)
targetdb tidb@31.***.***.***:4000 userid username password user_password
discardfile /Data/tidb/ogg/dir rpt/rnit_ora.dsc,purge
map hr.table_name, target tidb.hr.table_name;
```

(2)

```
./ggsci
add replicat rnit_ora,specialrun
```

## 5.



1. MGR
2. “A” trail
3. “C” “A”
4. “D”
5. “E” replicat repliact “C” “A” “handlecollisions”

## 5.2.4 OGG

OGG

```
SQL> CREATE TABLE hr.table_name
 2 (
 3 ID NUMBER NOT NULL,
 4 CREATED_AT DATE DEFAULT sysdate NOT NULL,
 5 CREATED_BY VARCHAR2(128 BYTE) DEFAULT 'SYS' NOT NULL,
 6 UPDATED_AT DATE DEFAULT sysdate NOT NULL,
 7 UPDATED_BY VARCHAR2(128 BYTE) DEFAULT 'SYS' NOT NULL
 8)
 9 ;
SQL> ALTER TABLE hr.table_name ADD CONSTRAINT PK_table_name;
```

1.

1.1

(1)

```
SQL> insert into hr.table_name(id) values(1);
1 row created.

SQL> commit;

Commit complete.
```

(3) OGG

```

Logdump 12 >ghdr on
Logdump 13 >detail on
Logdump 14 >detail data

Logdump 18 >n

Hdr-Ind : E (x45) Partition : . (x0c)
UndoFlag : . (x00) BeforeAfter: A (x41)
RecLength : 81 (x0051) IO Time : 2020/03/06 17:26:30.000.000
IOType : 5 (x05) OrigNode : 255 (xff)
TransInd : . (x03) FormatType: R (x52)
SyskeyLen : 0 (x00) Incomplete: . (x00)
AuditRBA : 3539 AuditPos : 31052304
Continued : N (x00) RecCount : 1 (x01)

2020/03/06 17:26:30.000.000 Insert Len 81 RBA 2014
Name: HR.TABLE_NAME (TDR Index: 1)
After Image: Partition 12 G s
0000 0005 0000 0001 3100 0100 1500 0032 3032 302d |1.....2020-
3033 2d30 363a 3137 3a32 363a 3238 0002 0007 0000 | 03-06:17:26:28.....
0003 5359 5300 0300 1500 0032 3032 302d 3033 2d30 | ..SYS.....2020-03-0
363a 3137 3a32 363a 3238 0004 0007 0000 0003 5359 | 6:17:26:28.....SY
53 | S
Column 0 (x0000), Len 5 (x0005) |1
0000 0001 31
Column 1 (x0001), Len 21 (x0015) 3a32 | ..2020-03-06:17:26:2
0000 3230 3230 2d30 332d 3036 3a31 373a 3236 38 | 8
Column 2 (x0002), Len 7 (x0007) 0000 0003 5359 53 |SYS
Column 3 (x0003), Len 21 (x0015) 0000 3230 3230 2d30 332d 3036 3a31 373a 3236 3a32 | ..2020-03-06:17:26:2
38 | 8
Column 4 (x0004), Len 7 (x0007) 0000 0003 5359 53 |SYS

```

**IOType** 3 delete5 insert10 updatefull >record15 updatecompressed record

**Continued** Y N lobblob > varchar Y N

**BeforeAfter** update before image( B ) >after image A insert after images

delete > before images

**FormatType** F >RF fetched from databaseR readable in transaction log

trail

1.2

(1) update

```

SQL> update hr.table_name set updated_by='aaaaaaaa';
1 row updated.

SQL> commit;

```

(2) OGG

```
Logdump 19 >n

Hdr-Ind : E (x45) Partition : . (x0c)
UndoFlag : . (x00) BeforeAfter: B (x42)
Reclength : 19 (x0013) IO Time : 2020/03/06 17:26:36.000.000
IOType : 15 (x0f) OrigNode : 255 (xff)
TransInd : . (x00) FormatType: R (x52)
SyskeyLen : 0 (x00) Incomplete: . (x00)
AuditRBA : 3539 AuditPos : 31055376
Continued : N (x00) RecCount : 1 (x01)

2020/03/06 17:26:36.000.000 FieldComp Len 19 RBA 2207
Name: HR.TABLE_NAME (TDR Index: 1)
Before Image: Partition 12 G b
0000 0004 ffff 0000 0004 0007 0000 0003 5359 53 |SYS
Column 0 (x0000), Len 4 (x0004)
ffff 0000 |
Column 4 (x0004), Len 7 (x0007)
0000 0003 5359 53 |SYS
Logdump 20 >n

Hdr-Ind : E (x45) Partition : . (x0c)
UndoFlag : . (x00) BeforeAfter: A (x41)
Reclength : 23 (x0017) IO Time : 2020/03/06 17:26:36.000.000
IOType : 15 (x0f) OrigNode : 255 (xff)
TransInd : . (x02) FormatType: R (x52)
SyskeyLen : 0 (x00) Incomplete: . (x00)
AuditRBA : 3539 AuditPos : 31055376
Continued : N (x00) RecCount : 1 (x01)

2020/03/06 17:26:36.000.000 FieldComp Len 23 RBA 2334
Name: HR.TABLE_NAME (TDR Index: 1)
After Image: Partition 12 G e
0000 0004 ffff 0000 0004 000b 0000 0007 6161 6161 |aaaa
6161 61 | aaa
Column 0 (x0000), Len 4 (x0004)
ffff 0000 |
Column 4 (x0004), Len 11 (x000b)
0000 0007 6161 6161 6161 61 |aaaaaaaa
```

trail

replicat "handlecollisions" insert

## 2.

### (1) OGG

```
GGSCI (1-db-ps-005) 3> dblogin userid goldengate, password goldengate
Successfully logged into database.
```

### (2)

```
GGSCI (1-db-ps-005 as goldengate@ogg) 4> add trandata hr.table_name,nokey

2020-03-06 21:13:51 WARNING OGG-01387 Table HR.TABLE_NAME has no valid key columns, added unconditional supplemental log group for all table columns.

Logging of supplemental redo data enabled for table HR.TABLE_NAME.
TRANSACTION for scheduling columns has been added on table 'HR.TABLE_NAME'.
TRANSACTION for instantiation CSN has been added on table 'HR.TABLE_NAME'.
```

### (3)

```
GGSCI (1-db-ps-005 as goldengate@ogg) 7> info trandata hr.table_name
Logging of supplemental redo log data is enabled for table HR.TABLE_NAME.
Columns supplementally logged for table HR.TABLE_NAME: ALL.
Prepared CSN for table HR.TABLE_NAME: 57187843
```

ogg

(4)

```
SQL> update hr.table_name set updated_by='bbb';
1 row updated.

SQL> commit;

Commit complete.
```

(5) ogg

Logdump 23 &gt;n

---

Hdr-Ind : E (x45) Partition : . (x0c)  
 UndoFlag : . (x00) BeforeAfter: B (x42)  
 RecLength : 85 (x055) IO Time : 2020/03/06 21:22:21.000.000  
 IOType : 15 (x0f) OrigNode : 255 (xff)  
 TransInd : . (x00) FormatType : R (x52)  
 SyskeyLen : 0 (x00) Incomplete : . (x00)  
 AuditRBA : 3539 AuditPos : 45237264  
 Continued : N (x00) RecCount : 1 (x01)

2020/03/06 21:22:21.000.000 FieldComp Len 85 RBA 2437  
 Name: HR.TABLE\_NAME (TDR Index: 1)  
 Before Image: Partition 12 G b  
 0000 0005 0000 0001 3100 0100 1500 0032 3032 302d | .....1.....2020-  
 3033 2d30 363a 3137 3a32 363a 3238 0002 0007 0000 | 03-06:17:26:28.....  
 0003 5359 5300 0300 1500 0032 3032 302d 3033 2d30 | ..SYS.....2020-03-0  
 363a 3137 3a32 363a 3238 0004 000b 0000 0007 6161 | 6:17:26:28.....aa  
 6161 6161 61 | aaaaaa  
 Column 0 (x0000), Len 5 (x0005)  
 0000 0001 31 | ....1  
 Column 1 (x0001), Len 21 (x0015)  
 0000 3230 3230 2d30 332d 3036 3a31 373a 3236 3a32 | ..2020-03-06:17:26:2  
 38 | 8  
 Column 2 (x0002), Len 7 (x0007)  
 0000 0003 5359 53 | ....SYS  
 Column 3 (x0003), Len 21 (x0015)  
 0000 3230 3230 2d30 332d 3036 3a31 373a 3236 3a32 | ..2020-03-06:17:26:2  
 38 | 8  
 Column 4 (x0004), Len 11 (x000b)  
 0000 0007 6161 6161 6161 61 | ....aaaaaaaaaa

Logdump 24 &gt;n

---

Hdr-Ind : E (x45) Partition : . (x0c)  
 UndoFlag : . (x00) BeforeAfter: A (x41)  
 RecLength : 81 (x051) IO Time : 2020/03/06 21:22:21.000.000  
 IOType : 15 (x0f) OrigNode : 255 (xff)  
 TransInd : . (x02) FormatType : R (x52)  
 SyskeyLen : 0 (x00) Incomplete : . (x00)  
 AuditRBA : 3539 AuditPos : 45237264  
 Continued : N (x00) RecCount : 1 (x01)

2020/03/06 21:22:21.000.000 FieldComp Len 81 RBA 2630  
 Name: HR.TABLE\_NAME (TDR Index: 1)  
 After Image: Partition 12 G e  
 0000 0005 0000 0001 3100 0100 1500 0032 3032 302d | .....1.....2020-  
 3033 2d30 363a 3137 3a32 363a 3238 0002 0007 0000 | 03-06:17:26:28.....  
 0003 5359 5300 0300 1500 0032 3032 302d 3033 2d30 | ..SYS.....2020-03-0  
 363a 3137 3a32 363a 3238 0004 0007 0000 0003 6262 | 6:17:26:28.....bb  
 62 | b  
 Column 0 (x0000), Len 5 (x0005)  
 0000 0001 31 | ....1  
 Column 1 (x0001), Len 21 (x0015)  
 0000 3230 3230 2d30 332d 3036 3a31 373a 3236 3a32 | ..2020-03-06:17:26:2  
 38 | 8  
 Column 2 (x0002), Len 7 (x0007)  
 0000 0003 5359 53 | ....SYS  
 Column 3 (x0003), Len 21 (x0015)  
 0000 3230 3230 2d30 332d 3036 3a31 373a 3236 3a32 | ..2020-03-06:17:26:2  
 38 | 8  
 Column 4 (x0004), Len 7 (x0007)  
 0000 0003 6262 62 | ....bbb

## 5.2.5

### 1.

```
MAXTRANSOPS 100
```

TiDB OGG 100

```
SQLEXEC "set tidb_constraint_check_in_place = 1"
```

OGG TiDB set tidb\_constraint\_check\_in\_place = 1 abend

## 2.

GBK report

```
2019-07-28 21:01:10 INFO OGG-02243 Opened trail file /home/tidb/ogg12.3/dirdat/p1000000002 at 2019-07-28 21:01:10.232896.

2019-07-28 21:01:10 INFO OGG-03506 The source database character set, as determined from the trail file, is zhs16gbk.

2019-07-28 21:05:57 INFO OGG-06505 MAP resolved (entry scott.account): MAP "SCOTT"."ACCOUNT", TARGET "scott"."account".

2019-07-28 21:05:57 INFO OGG-02756 The definition for table SCOTT.ACOUNT is obtained from the trail file.

2019-07-28 21:05:57 INFO OGG-06511 Using following columns in default map by name: account_number, account_balance, account_trans_ts, account_trans_type.

2019-07-28 21:05:57 INFO OGG-06510 Using the following key columns for target table scott.account: account_number.

2019-07-28 21:05:57 INFO OGG-03010 Performing implicit conversion of column data from character set zhs16gbk to UTF-8.
```

## 3.

OGG OGG TiDB

## 4.

TiDB OGG replicat map

## 5. replicat

OGG RANGE replicat

```
filter(@RANGE (, ,));
```

```

EXTRACT initext1

userid oggadmin ,password oggadmin

RMTHOST target-ogg, MGRPORT 7909

RMTTASK replicat, GROUP initrep1

TABLE scott.account, filter(@RANGE(1,2,account)) ;

EXTRACT initext2

userid oggadmin ,password oggadmin

RMTHOST target-ogg, MGRPORT 7909

RMTTASK replicat, GROUP initrep2

TABLE scott.account, filter(@RANGE(2,2,account)) ;

```

## 5.2.6

### 1. OGG-01201 Error reported by MGR : Access denied

This is due to a new security restriction in GoldenGate 12.2. In order to allow access from a remote system the ACCESSRULE parameter must be put into the manager parameter file on the target in order to allow access from the source.

```
mgr ip
```

```

(1) mgr

edit params mgr

(2)

ACCESSRULE, PROG *, IPADDR *, ALLOW

```

### 2. replicat invalid time format

Oracle date TIDB timestamp Oracle TIDB timestamp TIDB 1970-01-01 00:00:01.000000 2038-01-19  
03:14:07.999999Oracle 2038

TIDB timestamp datetime OGG replicat abended discard REPERROR (default,discard) OGG

### 3. extract ORA-01801: date format is too long for internal buffer

Oracle OCI Oracle Oracle select \* Oracle

to\_chardate,'yyyy-mm-dd hh24:mi:ss'Oracle '0000-00-00 00:00:00'

```

1
create table t_bak as select * from t where to_char(date1,'yyyy-mm-dd hh24:mi:ss')= '0000-00-00 00:00:00';
2:
delete from t where id in (select id from t_bak);
commit;

```

o have read ##

**permission only by the oracle user or an administrator who own this installation to protect any sensitive input values.**

#

#

---

**Do not change the following system generated value.**

---

oracle.install.responseFileVersion=/oracle/install/rspfmt\_ogginstall\_response\_schema\_v12\_1\_2

#

#

**Oracle GoldenGate installation option and details**

#

#

---

**Specify the installation option.**

**Specify ORA12c for installing Oracle GoldenGate for Oracle Database 12c and**

## **ORA11g for installing Oracle GoldenGate for Oracle Database 11g**

---

INSTALL\_OPTION=ORA11g

**Oracle 11.2.0.4 ORA11G 12c ORA12c**

---

**Specify a location to install Oracle GoldenGate**

---

SOFTWARE\_LOCATION=/home/oracle/ogg12.3

---

**Specify true to start the manager after installation.**

---

START\_MANAGER=false

**mgr false**

---

**Specify a free port within the valid range for the manager process.**

**Required only if START\_MANAGER is true.**

---

MANAGER\_PORT=

---

**Specify the location of the Oracle Database.**

**Required only if START\_MANAGER is true.**

---

DATABASE\_LOCATION=

#

#

**Specify details to Create inventory for Oracle installs**

**Required only for the first Oracle product install on a system.**

#

#

---

**Specify the location which holds the install inventory files.**

**This is an optional parameter if installing on**

**Windows based Operating System.**

---

INVENTORY\_LOCATION=

---

**Unix group to be set for the inventory directory.**

# This parameter is not applicable if installing on Windows based Operating System.

---

UNIX\_GROUP\_NAME=

```
(2)

fbo_ggs_Linux_x64_shiphome/Disk1/runInstaller

```
bash
$ ./fbo_ggs_Linux_x64_shiphome/Disk1/runInstaller -silent -responseFile /home/oracle/oggsoft/fbo_ggs_Linux_x64_shiphome/Disk1/response/oggcore.rsp
# response
```

successful

The installation of Oracle GoldenGate Core was successful

OGG

```
$ ls -l
total 225044
-rwxr-xr-x 1 oracle oinstall      426 Oct 15  2010 bcpfmt.tpl
-rwxr-xr-x 1 oracle oinstall    1725 Oct 15  2010 bcrypt.txt
-rwxrwxr-x 1 oracle oinstall 1612776 Apr 15  2018 cachefiledump
drwxr-xr-x 4 oracle oinstall     4096 Jul 28 14:45 cfgtoollogs
-rwxrwxr-x 1 oracle oinstall 3563576 Apr 15  2018 checkprm
-rw-rw-r-- 1 oracle oinstall     1021 Apr 15  2018 ckpt_ora_create.sql
-rwxrwxr-x 1 oracle oinstall 3379568 Apr 15  2018 convchk
-rwxrwxr-x 1 oracle oinstall 4716080 Apr 15  2018 convprm
drwxr-xr-x 2 oracle oinstall     4096 Jul 28 14:45 crypto
-rwxr-xr-x 1 oracle oinstall     159 Oct 15  2010 db2cntl.tpl
-rwxrwxr-x 1 oracle oinstall    9696 Apr 15  2018 db_upgrade
-rw-rw-r-- 1 oracle oinstall     455 Apr 15  2018 ddl_cleartrace.sql
-rw-rw-r-- 1 oracle oinstall    8414 Apr 15  2018 ddl_create.sql
-rw-rw-r-- 1 oracle oinstall    3176 Apr 15  2018 ddl_ddl2file.sql
-rw-rw-r-- 1 oracle oinstall     90 Apr 15  2018 ddl_disable.sql
-rw-rw-r-- 1 oracle oinstall     88 Apr 15  2018 ddl_enable.sql
-rw-rw-r-- 1 oracle oinstall    2036 Apr 15  2018 ddl_filter.sql
-rw-rw-r-- 1 oracle oinstall   12220 Apr 15  2018 ddl_ora10.sql
-rw-rw-r-- 1 oracle oinstall   1725 Apr 15  2018 ddl_ora10upCommon.sql
-rw-rw-r-- 1 oracle oinstall  13539 Apr 15  2018 ddl_ora11.sql
-rw-rw-r-- 1 oracle oinstall  12564 Apr 15  2018 ddl_ora9.sql
-rw-rw-r-- 1 oracle oinstall     216 Apr 15  2018 ddl_pin.sql
-rw-rw-r-- 1 oracle oinstall   3184 Apr 15  2018 ddl_remove.sql
-rw-rw-r-- 1 oracle oinstall      1 Apr 15  2018 ddl_session1.sql
-rw-rw-r-- 1 oracle oinstall    629 Apr 15  2018 ddl_session.sql
-rw-rw-r-- 1 oracle oinstall  287877 Apr 15  2018 ddl_setup.sql
-rw-rw-r-- 1 oracle oinstall   8401 Apr 15  2018 ddl_status.sql
-rw-rw-r-- 1 oracle oinstall   2122 Apr 15  2018 ddl_staymetadata_off.sql
-rw-rw-r-- 1 oracle oinstall   2118 Apr 15  2018 ddl_staymetadata_on.sql
-rw-rw-r-- 1 oracle oinstall   2186 Apr 15  2018 ddl_tracelevel.sql
-rw-rw-r-- 1 oracle oinstall   2133 Apr 15  2018 ddl_trace_off.sql
-rw-rw-r-- 1 oracle oinstall   2383 Apr 15  2018 ddl_trace_on.sql
-rwxrwxr-x 1 oracle oinstall 5037440 Apr 15  2018 defgen
drwxr-xr-x 2 oracle oinstall     4096 Jul 28 14:45 deinstall
-rw-rw-r-- 1 oracle oinstall    882 Apr 15  2018 demo_more_ora_create.sql
-rw-rw-r-- 1 oracle oinstall    649 Apr 15  2018 demo_more_ora_insert.sql
-rw-rw-r-- 1 oracle oinstall    583 Apr 15  2018 demo_ora_create.sql
```

```

-rw-rw-r-- 1 oracle oinstall      504 Apr 15 2018 demo_ora_insert.sql
-rw-rw-r-- 1 oracle oinstall    3597 Apr 15 2018 demo_ora_lob_create.sql
-rw-rw-r-- 1 oracle oinstall   1943 Apr 15 2018 demo_ora_misc.sql
-rw-rw-r-- 1 oracle oinstall   1056 Apr 15 2018 demo_ora_pk_befores_create.sql
-rw-rw-r-- 1 oracle oinstall   1013 Apr 15 2018 demo_ora_pk_befores_insert.sql
-rw-rw-r-- 1 oracle oinstall  2305 Apr 15 2018 demo_ora_pk_befores_updates.sql
drwxr-xr-x 3 oracle oinstall    4096 Jul 28 14:45 diagnostics
drwxr-xr-x 3 oracle oinstall    4096 Jul 28 14:45 diretc
drwxr-xr-x 2 oracle oinstall    4096 Jul 28 14:45 dirout
drwxr-xr-x 4 oracle oinstall    4096 Jul 28 14:45 dirsca
-rwxrwxr-x 1 oracle oinstall 4272400 Apr 15 2018 emsclnt
-rwxrwx--- 1 oracle oinstall 12545040 Apr 15 2018 extract
-rwxr-xr-x 1 oracle oinstall    1968 Oct 15 2010 freeBSD.txt
-rwxrwxr-x 1 oracle oinstall  4280528 Apr 15 2018 ggcmd
-rwxr-xr-x 1 oracle oinstall  2303056 Apr 15 2018 ggMessage.dat
-rwxr-xr-x 1 oracle oinstall 49675440 Apr 15 2018 ggparam.dat
-rwxrwx--- 1 oracle oinstall  9340192 Apr 15 2018 ggsci
drwxr-xr-x 2 oracle oinstall    4096 Jul 28 14:45 healthcheck
-rwxr-xr-x 1 oracle oinstall  299451 Nov 16 2017 help.txt
drwxr-xr-x 3 oracle oinstall    4096 Jul 28 14:45 install
drwxr-x--- 12 oracle oinstall   4096 Jul 28 14:45 inventory
drwxr-xr-x 7 oracle oinstall    4096 Jul 28 14:45 jdk
-rwxrwxr-x 1 oracle oinstall 144496 Apr 15 2018 keygen
-rw-rw-r-- 1 oracle oinstall     56 Apr 15 2018 label.sql
-rwxrwx--- 1 oracle oinstall 102840 Apr 15 2018 libantlr3c.so
-rwxrwxr-x 1 oracle oinstall  12312 Apr 15 2018 libboost_system-mt.so.1.58.0
-rwxrwx--- 1 oracle oinstall 2190856 Apr 15 2018 libdb-6.1.so
-rwxrwx--- 1 oracle oinstall 2198136 Apr 15 2018 libglog.so
-rwxrwx--- 1 oracle oinstall 10524064 Apr 15 2018 libgnnzip.so
-rwxrwx--- 1 oracle oinstall 21993240 Apr 15 2018 libggparam.so
-rwxrwx--- 1 oracle oinstall 210264 Apr 15 2018 libggperf.so
-rwxrwx--- 1 oracle oinstall 779352 Apr 15 2018 libggrepo.so
-rwxrwx--- 1 oracle oinstall 1108760 Apr 15 2018 libgssl.so
-rwxrwx--- 1 oracle oinstall 125624 Apr 15 2018 libggutil.so
-rwxrwxr-x 1 oracle oinstall 29764432 Apr 15 2018 libicudata.so.56
-rwxrwxr-x 1 oracle oinstall 2909360 Apr 15 2018 libicui18n.so.56
-rwxrwxr-x 1 oracle oinstall 1995808 Apr 15 2018 libicuu.so.56
-rwxrwx--- 1 oracle oinstall 86960 Apr 15 2018 liblmbd.so
-rwxrwxr-x 1 oracle oinstall 175136 Apr 15 2018 libPocoCrypto.so.48
-rwxrwxr-x 1 oracle oinstall 3079760 Apr 15 2018 libPocoFoundation.so.48
-rwxrwxr-x 1 oracle oinstall 373232 Apr 15 2018 libPocoJSON.so.48
-rwxrwxr-x 1 oracle oinstall 1326504 Apr 15 2018 libPocoNet.so.48
-rwxrwxr-x 1 oracle oinstall 350064 Apr 15 2018 libPocoNetSSL.so.48
-rwxrwxr-x 1 oracle oinstall 520864 Apr 15 2018 libPocoUtil.so.48
-rwxrwxr-x 1 oracle oinstall 680288 Apr 15 2018 libPocoXML.so.48
-rwxrwx--- 1 oracle oinstall 1115360 Apr 15 2018 libudt.so
-rwxrwx--- 1 oracle oinstall 4782344 Apr 15 2018 libxml2.so
-rwxrwxr-x 1 oracle oinstall 5025840 Apr 15 2018 logdump
-rw-rw-r-- 1 oracle oinstall    1553 Apr 15 2018 marker_remove.sql
-rw-rw-r-- 1 oracle oinstall    3309 Apr 15 2018 marker_setup.sql
-rw-rw-r-- 1 oracle oinstall     675 Apr 15 2018 marker_status.sql
-rwxrwxr-x 1 oracle oinstall 6570480 Apr 15 2018 mgr
-rwxr-xr-x 1 oracle oinstall 41643 Jun 30 2017 notices.txt
-rwxrwxr-x 1 oracle oinstall 1661024 Apr 15 2018 oggerr
drwxr-xr-x 12 oracle oinstall   4096 Jul 28 14:45 OPatch
-rw-r----- 1 oracle oinstall     59 May 30 13:35 oraInst.loc
drwxr-xr-x 8 oracle oinstall    4096 Jul 28 14:45 oui
-rw-rw-r-- 1 oracle oinstall    3146 Apr 15 2018 params.sql
-rwxrwxr-x 1 oracle oinstall 11524576 Apr 15 2018 pmsrvr
-rwxr-xr-x 1 oracle oinstall    1272 Dec 28 2010 prvtclkm.plb
-rwxr-xr-x 1 oracle oinstall    9487 May 27 2015 prvtlmpg.plb
-rw-rw-r-- 1 oracle oinstall    2724 Apr 15 2018 prvtlmpg_uninstall.sql
-rw-rw-r-- 1 oracle oinstall    1532 Apr 15 2018 remove_seq.sql
-rwxrwx--- 1 oracle oinstall 10873240 Apr 15 2018 replicat
-rwxrwxr-x 1 oracle oinstall 1656864 Apr 15 2018 retrace
-rw-rw-r-- 1 oracle oinstall    3187 Apr 15 2018 role_setup.sql
-rw-rw-r-- 1 oracle oinstall 35254 Apr 15 2018 sequence.sql
-rwxrwxr-x 1 oracle oinstall 4659736 Apr 15 2018 server
-rwxr-xr-x 1 oracle oinstall    4917 Jan  5 2017 SQLDataTypes.h
-rwxr-xr-x 1 oracle oinstall    248 Oct 15 2010 sqlldr.tpl
drwxr-xr-x 3 oracle oinstall    4096 Jul 28 14:45 srvm
-rwxrwxr-x 1 oracle oinstall    759 Oct 15 2010 tcper

```

```
-rwxr-xr-x 1 oracle oinstall 37877 Apr 16 2016 ucharset.h
-rw-rw-r-- 1 oracle oinstall 7341 Apr 15 2018 ulg.sql
drwxr-xr-x 7 oracle oinstall 4096 Jul 28 14:45 UserExitExamples
-rwxr-xr-x 1 oracle oinstall 32987 Jun 2 2017 usrdecs.h
-rwxr-xr-x 1 oracle oinstall 1033 Oct 19 2016 zlib.txt
```

(3) .bash_profile

```
export LD_LIBRARY_PATH=$ORACLE_HOME/lib
```

(4)

```
$ ./ggsci
Oracle GoldenGate Command Interpreter for Oracle
Version 12.3.0.1.4 OGGCORE_12.3.0.1.0_PLATFORMS_180415.0359_FBO
Linux, x64, 64bit (optimized), Oracle 11g on Apr 15 2018 21:16:09
Operating system character set identified as UTF-8.
Copyright (C) 1995, 2018, Oracle and/or its affiliates. All rights reserved.
```

```
GGSCI ( hostname ) 1> create subdirs
Creating subdirectories under current directory /home/oracle/ogg12.3
Parameter file          /home/oracle/ogg12.3/dirprm: created.
Report file             /home/oracle/ogg12.3/dir rpt: created.
Checkpoint file         /home/oracle/ogg12.3/dir chk: created.
Process status files   /home/oracle/ogg12.3/dir p cs: created.
SQL script files        /home/oracle/ogg12.3/dir sql: created.
Database definitions files /home/oracle/ogg12.3/dir def: created.
Extract data files     /home/oracle/ogg12.3/dir d dat: created.
Temporary files         /home/oracle/ogg12.3/dir tmp: created.
Credential store files /home/oracle/ogg12.3/dir crd: created.
Masterkey wallet files /home/oracle/ogg12.3/dir wl t: created.
Dump files              /home/oracle/ogg12.3/dir dmp: created.
```

Oracle OGG

2.2 OGG - TiDB

(1)

```
$ unzip V978711-01.zip
Archive: V978711-01.zip
  inflating: ggs_Linux_x64_SQLite_64bit.tar
  inflating: OGG-12.3.0.1-README.txt
  inflating: OGG_WinUnix_Rel_Notes_12.3.0.1.5.pdf
```

(2) ggs_Linux_x64_SQLite_64bit.tar .

```
$tar xvf ./ggs_Linux_x64_SQLite_64bit.tar -C /home/tidb/ogg12.3
```

(3) ggsci

```
$ ./ggsci
Oracle GoldenGate Command Interpreter for MySQL
Version 12.3.0.1.5 OGGCORE_12.3.0.1.0_PLATFORMS_180501.2124
Linux, x64, 64bit (optimized), MySQL Enterprise on May 2 2018 10:58:16
Operating system character set identified as UTF-8.
Copyright (C) 1995, 2018, Oracle and/or its affiliates. All rights reserved.

GGSCI ( hostname ) 1> create subdirs
Creating subdirectories under current directory /home/tidb/ogg12.3
Parameter file          /home/tidb/ogg12.3/dirprm: created.
Report file             /home/tidb/ogg12.3/dir rpt: created.
Checkpoint file         /home/tidb/ogg12.3/dir chk: created.
Process status files   /home/tidb/ogg12.3/dir pcs: created.
SQL script files       /home/tidb/ogg12.3/dir sql: created.
Database definitions files /home/tidb/ogg12.3/dir def: created.
Extract data files     /home/tidb/ogg12.3/dir dta: created.
Temporary files        /home/tidb/ogg12.3/dir tmp: created.
Credential store files /home/tidb/ogg12.3/dir crd: created.
Masterkey wallet files /home/tidb/ogg12.3/dir wlt: created.
Dump files              /home/tidb/ogg12.3/dir dmp: created.
```

OGG TiDB (OGG For MySQL)

2.3TiDB

(1) lower-case-table-names 1

```
$grep lower-case-table-names tidb.toml
lower-case-table-names = 1
```

(2)

```
MySQL [(none)]> show variables like '%lower%';
+-----+-----+
| Variable_name      | Value |
+-----+-----+
| lower_case_table_names | 1    |
| lower_case_file_system | 1    |
+-----+-----+
MySQL [(none)]> show variables like '%place%';
+-----+-----+
| Variable_name      | Value |
+-----+-----+
| tidb_constraint_check_in_place | 0    |
+-----+-----+
```

(3)

```
MySQL [scott]> create user 'tidb' identified by 'tidb';
Query OK, 1 row affected (0.01 sec)
MySQL [scott]> GRANT ALL PRIVILEGES ON scott.* to oggadmin ;
Query OK, 0 rows affected (0.01 sec)
MySQL [scott]> show create user tidb;
+-----+
| CREATE USER for tidb@%                                |
+-----+
| CREATE USER 'tidb'@'%' IDENTIFIED WITH 'mysql_native_password' AS '*465D123EE55795DBDBDAE36AFD3DCD9C429B718A' REQUIRE NONE PASSWORD EXPIRE DEFAULT ACCOUNT UNLOCK |
+-----+
1 row in set (0.00 sec)
```

(4)

2.4 Oracle

2.4.1

```
SQL> archive log list;
```

(1)

```
SQL> shutdown immediate;
```

(2) mount

```
SQL> startup mount;
```

(3)

```
SQL> alter database archivelog;
```

(4)

```
SQL> alter database open;
```

(5)

```
SQL> archive log list;
```

2.4.2

(1)

```
select log_mode, supplemental_log_data_min, force_logging from v$database;
```

(2) force logging

```
nolog
```

```
SQL> ALTER DATABASE FORCE LOGGING;
```

(3)

```
SQL> ALTER DATABASE ADD SUPPLEMENTAL LOG DATA;
SQL> ALTER SYSTEM SWITCH LOGFILE;
```

(4) OGG

```
SQL> alter system set enable_goldengate_replication=true;
```

2.4.3 ogg

goldengate

```
SQL> create tablespace goldengate datafile '/opt/oracle/data/goldengate001.dbf' size 4G;
SQL> create user goldengate identified by goldengate default tablespace goldengate;
SQL> grant dba to goldengate;
```

2.5OGG

(1) ggsic

```
./ggsic
create subdirs
add credentialstore
alter credentialstore add user goldengate,password goldengate
```

(2) ogg goldengate

```
dblogin useridalias goldengate
```

2.6 schema

schema schema

(1) 11.2.0.2 Oracle Patch 13794550

oracle schema

ERROR OGG-06522 Cannot verify existence of table function that is required to ADD schema level supplemental logging.
failed.

(2) GGSCI

- 1) dblogin userid goldengate, password goldengate
- 2) dblogin useridalias goldengate

(3) schema

ADD SCHEMATRANSACTION schema ALLCOLS

2,7

schema

(1) GGSCI

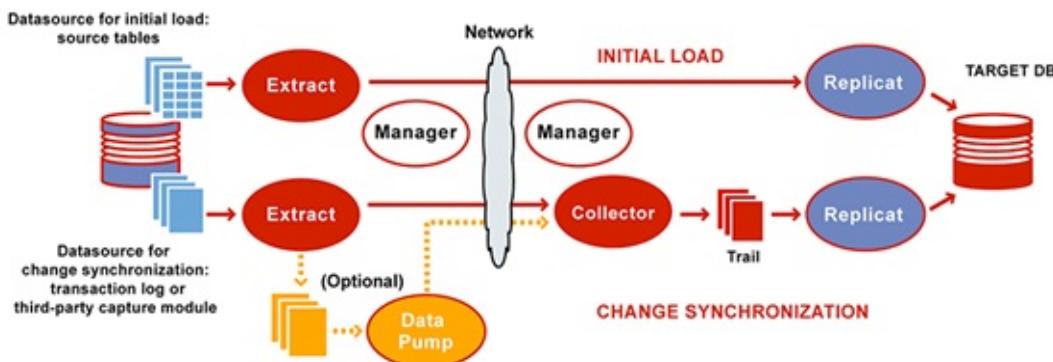
```
dblogin userid goldengate, password goldengate
```

(2) schema

```
ADD TRANSACTION schema.tablename NOKEY
```

OGG

3.1OGG



1. INITIAL LOAD

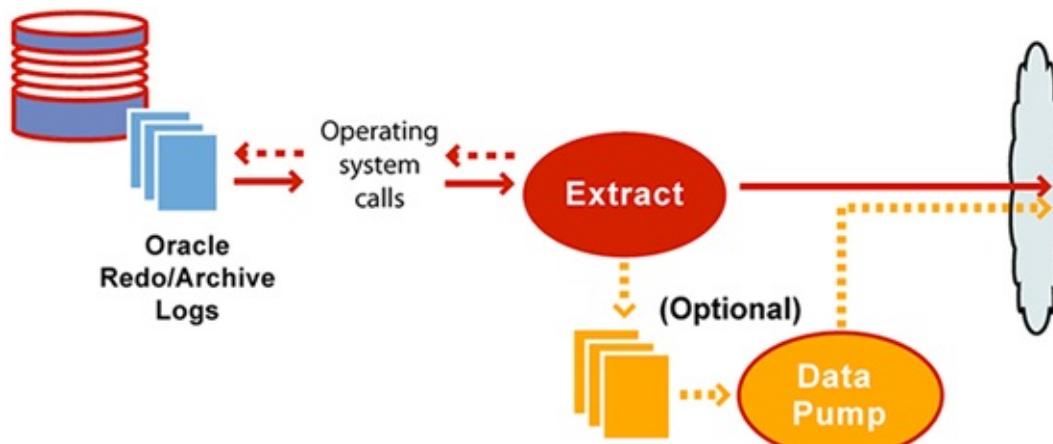
2. CHANGE SYNCHRONIZATION trail oracle oracle redo log archive log trail

| | |
|--------------------------|-------------|
| | |
| /opt/oracle/11.2.0 | ORACLE_HOME |
| ora | Oracle sid |
| ext_ora | |
| tdpm_ora | |
| rep_ora | |
| ./dirdat/sk | |
| /Data/tidb/ogg/dirdat/sk | Replicat |

3.2

ADG

Extract pump



Extract redo archive log Data pump

3.2.1 MGR

Manager OGG ExtractPumpReplicat ExtractPumpReplicat Manager

(1)

```
GGSCI (b-db-ps-055) 5> view param mgr

port 9001
DYNAMICPORTLIST 9001-9020
--AUTORESTART ER *,RETRIES 5,WAITMINUTES 7
PURGEOLDEXTRACTS ./dirdat/*,usecheckpoints, minkeepdays 2
LAGREPORTHOURS 1
LAGINFOMINUTES 30
LAGCRITICALMINUTES 45
```

| | |
|------------------|---------------|
| | |
| port | |
| DYNAMICPORTLIST | OGG OGG |
| AUTORESTART | |
| PURGEOLDEXTRACTS | trail 2 trail |
| LAGREPORTHOURS | lag 1 |
| LAGINFOMINUTES | |

(2) mgr

```
./ggscli
start mgr
```

3.2.2 Extract

Extract OGG Extract

(1)

```
GGSCI (b-db-ps-055) 2> edit param ext_ora

EXTRACT ext_ora
SETENV (ORACLE_HOME="/opt/oracle/11.2.0")
setenv (NLS_LANG="AMERICAN_AMERICA.AL32UTF8")
setenv (ORACLE_SID="ora")
useridalias goldengate
GETTRUNCATES
REPORTCOUNT EVERY 1 MINUTES, RATE
DISCARDFILE ./dirrpt/ext_ora.dsc,APPEND,MEGABYTES 1000
WARNLONGTRANS 2h,CHECKINTERVAL 10m
EXTTRAIL ./dirdat/sk
TRANLOGOPTIONS EXCLUDEUSER goldengate
TRANLOGOPTIONS MINEFROMACTIVEDG
DBOPTIONS ALLOWUNUSEDCOLUMN
DYNAMICRESOLUTION
FETCHOPTIONS FETCHPKUPDATECOLS,INCONSISTENTROW ALLOW
ddl include mapped objname hr.*
getupdatebefores
nocompressdeletes
nocompressupdates
table hr.table_name;
```

|||:----|:----| SETENV | || useridalias | || GETTRUNCATES | truncate || REPORTCOUNT ||
 WARNLONGTRANS | OGG || EXTTRAIL | trail || TRANLOGOPTIONS EXCLUDEUSER ||
 TRANLOGOPTIONS MINEFROMACTIVEDG | ADG || DBOPTIONS ALLOWUNUSEDCOLUMN ||

DYNAMICRESOLUTION | OGG OGG OGG OGG ID
FETCHPKUPDATECOLS,INCONSISTENTROW ALLOW | row id () ALLOW

(2)

```
./ggsci  
add extract ext_ora,TRANLOG, begin now  
add ExtTrail ./dirdat/sk, Extract ext_ora, Megabytes 50
```

3.2.3 Pump

Pump Extract Pump Extract Trail Trail

```
GGSCI (b-db-ps-055) 4> edit param    tdpm_ora

EXTRACT tdpm_ora
RMTHOST 31.***.***.***, MGRPORT 9001,    compress
PASSTHRU
RMTTRAIL /Data/tidb/ogg/dirdat/sk
DYNAMICRESOLUTION
--table
table hr.table_name;
```

| | |:----|:----| RMTHOST | IP MGR || PASSTHRU | OGG

(1) pump

```
./ggsci  
add extract tdpm_ora, extrailsource    ./dirdat/sk  
add rmttrail /Data/tidb/ogg/dirdat/sk,   extract tdpm_ora, megabytes 100
```

33

3.3.1 MGR

(1)

```
GGSCI (w-db-ps-082) 2> edit param mgr

port 9001
DYNAMICPORTLIST 9001-9120
ACCESSRULE, PROG *, IPADDR 31.*.*.*,, ALLOW
--AUTORESTART ER *,RETRIES 5,WAITMINUTES 7
PURGEOLDEXTRACTS ./dirdat/*,usecheckpoints, minkeepdays 10
LAGREPORTHOURS 1
LAGINFOMINUTES 30
LAGCRITICALMINUTES 45
```

(2) mgr

```
./ggsci  
start mar
```

3.3.2 Replicat

Replicat

5.3 SQL Server TiDB

5.3.1

SQL server DAU 1000000000+ SQL server



SQL Server TiDB SQL Server TiDB

5.3.2



| | |
|--|---|
| | |
| | 1. SQL Server CDC
2. SQL Server
3. ETL TiDB
4. |
| | 1. CDC TiDB
2. |
| | 1. SQL Server
2. CDC
3. TiDB |

5.3.3

1.

yugongyugong Oracle <https://github.com/alswl/yugong> SQL Server

2.

(1) **yugong**

```
# jar
wget https://github.com/alswl/yugong/releases/download/541e5f8/yugong-shaded.jar

#
git git@github.com:alswl/yugong.git
cd yugong
mvn clean package
```

(2)

```
yugong  yugong.properties  yugong.yaml
```

- yugong.properties

```
# vi yugong.properties
vi yugong.properties
```

```

#
yugong.database.source.url=jdbc:sqlserver://192.168.1.88:1433;databaseName=example
yugong.database.source.username=sa
yugong.database.source.password=sa
yugong.database.source.type=SQL_SERVER
yugong.database.source.encode=UTF-8
yugong.database.source.poolSize=200

# TiDB MySQL
yugong.database.target.url=jdbc:mysql://192.168.1.99:3306/example?autoReconnect=true
yugong.database.target.username=root
yugong.database.target.password=root
yugong.database.target.type=MYSQL
yugong.database.target.encode=UTF-8
yugong.database.target.poolSize=200
yugong.table.batchApply=true
yugong.table.onceCrawNum=1000
yugong.table.tpsLimit=0
# schema
yugong.table.ignoreSchema=false
# skip Applier Load Db failed data
yugong.table.skipApplierException=false
#
yugong.table.white=user
yugong.table.black=
#
yugong.table.concurrent.enable=true
#
yugong.table.concurrent.size=6
# retry times
yugong.table.retry.times=3
# retry interval or sleep time (ms)
yugong.table.retry.interval=1000
# MARK/FULL/INC/ALL(REC+FULL+INC)/CHECK/CLEAR
# FULL
yugong.table.mode=FULL

# yugong extractor
yugong.extractor.dump=false
yugong.extractor.concurrent.enable=true
yugong.extractor.concurrent.global=false
yugong.extractor.concurrent.size=60
yugong.extractor.noupdate.sleep=1000
yugong.extractor.noupdate.thresold=0
yugong.extractor.once=false

# yugong applier
yugong.applier.concurrent.enable=true
yugong.applier.concurrent.global=false
yugong.applier.concurrent.size=20
yugong.applier.dump=false
# stats
yugong.stat.print.interval=5
yugong.progress.print.interval=1
# alarm email
yugong.alarm.email.host=
yugong.alarm.email.username=
yugong.alarm.email.password=
yugong.alarm.email.smtp.port=
yugong.alarm.email.ssl.support=

```

- `yugong.yaml` translator

```

# yugong.yaml.sample
wget https://raw.githubusercontent.com/alswl/yugong/feature/sql-server-to-mysql-overview/yugong.yaml.sample

# yugong.yaml
mv yugong.yaml.sample yugong.yaml

```

(3) yugong

- yugong

```
# -c yugong.properties
# -y yugong.yaml ETL translator
#
java -jar yugong-shaded.jar -c yugong.properties -y yugong.yaml
```

-

```
#
2019-12-02 20:49:23.923 [main] INFO com.taobao.yugong.YugongApp - ## start the YuGong.
#
2019-12-02 20:49:24.000 [main] INFO com.taobao.yugong.controller.YuGongController - check source database connection ...
2019-12-02 20:49:24.017 [main] INFO com.taobao.yugong.controller.YuGongController - check source database is ok
2019-12-02 20:49:24.017 [main] INFO com.taobao.yugong.controller.YuGongController - check target database connection ...
...
2019-12-02 20:49:24.028 [main] INFO com.taobao.yugong.controller.YuGongController - check target database is ok
2019-12-02 20:49:24.028 [main] INFO com.taobao.yugong.controller.YuGongController - check source tables read privileges ...
#
2019-12-02 20:49:24.071 [main] INFO com.alibaba.druid.pool.DruidDataSource - {dataSource-1} initied
2019-12-02 20:49:24.277 [main] INFO com.alibaba.druid.pool.DruidDataSource - {dataSource-2} initied
2019-12-02 20:49:25.418 [main] INFO com.taobao.yugong.controller.YuGongController - check source tables is ok.
#
2019-12-02 20:49:26.933 [main] INFO com.taobao.yugong.controller.YuGongController - ## prepare start tables[1] with concurrent[6]
2019-12-02 20:49:26.959 [main] INFO com.taobao.yugong.YugongApp - ## the YuGong is running now .....
#
2019-12-02 20:51:08.801 [Thread-7] INFO com.taobao.yugong.common.stats.ProgressTracer - {::0,:0,:1,:0}
#
2019-12-02 20:51:08.801 [Thread-7] INFO com.taobao.yugong.common.stats.ProgressTracer - :[example.user]
2019-12-02 20:51:08.805 [Thread-7] INFO com.alibaba.druid.pool.DruidDataSource - {dataSource-2} closed
2019-12-02 20:51:08.806 [Thread-7] INFO com.alibaba.druid.pool.DruidDataSource - {dataSource-1} closed
2019-12-02 20:51:08.807 [Thread-7] INFO com.taobao.yugong.YugongApp - ## YuGong is down.
```

(4)

- logs

```
logs
├── example.user  #
│   ├── applier.log  #
│   ├── extractor.log  #
│   └── table.log  #
└── yugong
    └── table.log  #
positioner_data
└── example_user.dat #
```

table.log ERROR

5.3.4**1.**

SQL Server CDC SQL Server CDC CDC Change Data CaptureCDC TiDBCapture process transaction log
Change Table CDC

2.

(1) SQL Server CDC

```
-- CDC
sys.sp_cdc_enable_db;

-- example CDC
EXEC sys.sp_cdc_enable_table @source_schema = N'dbo', @source_name = N'example', @role_name = NULL;

-- SQL CDC
-- CDC
SELECT name, is_cdc_enabled FROM sys.databases WHERE is_cdc_enabled = 1;
```

(2) CDC Change Table cdc.dbo_example_CT

```
.schema cdc.dbo_example_CT
name      default nullable type      length indexed
-----  -----  -----  -----  -----
__$end_lsn    null     YES   binary     10    NO
__$operation   null     NO    int        4     NO
__$seqval      null     NO   binary     10    NO
__$start_lsn   null     NO   binary     10    YES
__$update_mask  null     YES  varbinary  128   NO
id          null     YES   int        4     NO
name         null     YES  varchar(255) 255   NO
```

id name example

(3) CDC

- example CDC

```
-- 
DECLARE @begin_time datetime, @end_time datetime, @begin_lsn binary(10), @end_lsn binary(10);

-- 
SET @begin_time = '2020-03-08 10:00:00.000';
SET @end_time   = '2020-03-08 10:10:00.000';

-- lsn
SELECT @begin_lsn = sys.fn_cdc_map_time_to_lsn('smallest greater than', @begin_time);
SELECT @end_lsn = sys.fn_cdc_map_time_to_lsn('largest less than or equal', @end_time);

-- lsn
SELECT * FROM cdc.fn_cdc_get_all_changes_dbo_example(@begin_lsn, @end_lsn, 'all');
```

-

| __\$start_lsn | __\$end_lsn | __\$seqval | __\$operation | __\$update_mask | id | name |
|----------------------|-------------|----------------------|---------------|-----------------|----|------|
| 0000dede0000019f001a | null | 0000dede0000019f0018 | 2 | 03 | 1 | AAA |
| 0000dede000001ad0004 | null | 0000dede000001ad0003 | 2 | 03 | 2 | BBB |
| 0000dede000001ba0003 | null | 0000dede000001ba0002 | 3 | 02 | 2 | BBB |
| 0000dede000001ba0003 | null | 0000dede000001ba0002 | 4 | 02 | 2 | CCC |
| 0000dede000001c10003 | null | 0000dede000001c10002 | 2 | 03 | 3 | DDD |
| 0000dede000001cc0005 | null | 0000dede000001cc0002 | 1 | 03 | 3 | DDD |

```
__$operation
•
•
•
•
```

(4) CDC

CDC SQL Change Table `__$operation` SQL TiDB

1. Kafka RabbitMQ
2. Kafka Topic Partition 1
- 3.
4. CDC `insert ignore` SQL

5.3.5

ETL

yugong ETL `CHECK FULL INC AUTO` `CHECK` yugong yugong JDBC

yugong SQL Server —— `physloc`

5 TiDB

5.3.6

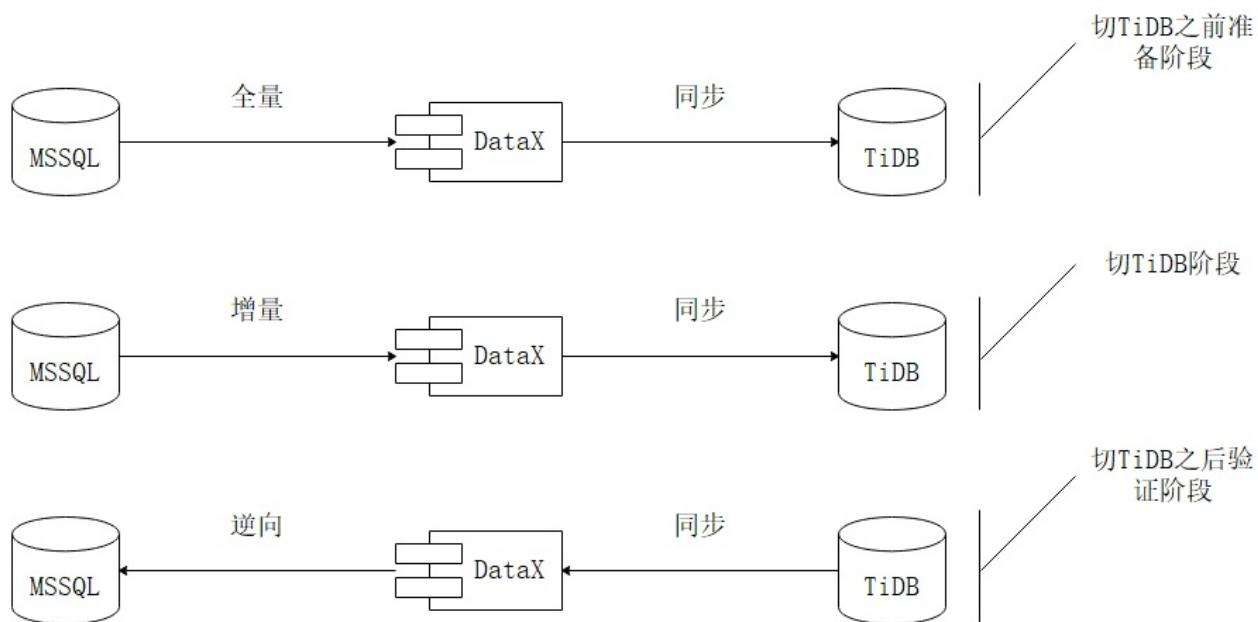
1. SQL server DBA SQL server DMV
2. SQL Server TiDBMySQL
3. SQL Server TiDBMySQL
- 4.
- 5.
- 6.

5.3.7

TiDB MySQL “” SQL Server TiDB

DataX / MySQL SQL Server Oracle PostgreSQL HDFS Hive HBase OTS ODPS DataX + GitHub

DataX TiDB



TiDB SQL Server DataX TiDB

UpdateTime Where t_sync_record

TiDB SQL Server SQL Server

DataX

```
wget http://datax.opensource.oss-cn-hangzhou.aliyuncs.com/datax.tar.gz
```

```
tar -zvxf datax.tar.gz
```

```
python {YOUR_DATAX_HOME}/bin/datax.py {YOUR_DATAX_HOME}/job/job.json
```

DataX JobJson

1Job

vi full.json

```
{
  "job": {
    "setting": {
      "speed": {
```

```

    #
    "channel": 128
}

},
"content": [{

    #SQL Server

    "reader": {

        "name": "sqlserverreader",

        "parameter": {

            #

            "username": "${srcUserName}",

            #

            "password": "${srcPassword}",

            #*

            "column": ["*"]

        }

        "connection": [{

            #

            "table": ["${tableName}"],

            jdbc

            "jdbcUrl": ["${srcUrl}"]

        }]

    }

},

#TiDB

"writer": {

    "name": "tidbwriter",

    "parameter": {

        #

        "username": "${desUserName}",

        #

        "password": "${desPassword}",

        # Replace

        "writeMode": "replace",

        #*

        "column": ["*"],

        "connection": [{


```

```

        # jdbc

        "jdbcUrl": "${desUrl}",
        #

        "table": ["${tableName}"]

    ],
    #sql-
    "preSql": [
        "replace into t_sync_record(table_name,start_date,end_date) values('@table',now(),null)",
        # sql-
        "postSql": [
            "update t_sync_record set end_date=now() where table_name='@table' "
        ]
    ]
}
}
}
}

```

2 Job

vi increase.json

```

{
    "job": {
        "setting": {
            "speed": {
                #
                "channel": 128
            }
        },
        "content": [
            {
                "#SQL Server

                "reader": {

                    "name": "sqlserverreader",
                    "parameter": {
                        #
                        "username": "${srcUserName}",
                        #
                        "password": "${srcPassword}",

```

```

#*
"column": ["*"],
"connection": [{

    #

    "table": ["${tableName}"],

    jdbc

    "jdbcUrl": ["${srcUrl}"]

}],

    #

    "where": "updateTime >= '${syncTime}'"

}

},

#TiDB

"writer": {

    "name": "tidbwriter",

    "parameter": {

        #

        "username": "${desUserName}",

        #

        "password": "${desPassword}",

        # Replace

        "writeMode": "replace",

        #*

        "column": ["*"],

        "connection": [{

            # jdbc

            "jdbcUrl": "${desUrl}",

            #

            "table": ["${tableName}"]

        }],


        #sql-


        "preSql": [

            "replace into t_sync_record(table_name,start_date,end_date) values('@table',now(),null)",

            # sql-


            "postSql": [


                "update t_sync_record set end_date=now() where table_name='@table' " ]



        ]



    }

}

```

```
    }
}
}]
}
}
```

3DataX JobShell

vi datax_execute_job.sh

```

#!/bin/bash

source /etc/profile

srcUrl="Reader Sql Server "
srcUserName="Sql Server "
srcPassword="Sql Server "

desUrl="Writer TiDB "
desUserName="TiDB "
desPassword="TiDB "

#
defaultSyncUpdateTime="2020-03-03 18:00:00.000"

# ()

sleepTime="N"

tableName="Table1,Table2, . . ."

# -1

flg=-1

while [ "$flg" -gt 0 -o "$flg" -eq -1 ]
do
#
if [ "" = "$preRunTime" ]; then
    syncTime=$defaultSyncUpdateTime
else
    syncTime=$preRunTime
fi
#
preRunTime=$(date -d +"%Y-%m-%d %T").000";
echo $syncTime
echo $preRunTime
echo $flg

python {YOUR_DATAAX_HOME}/bin/dataax.py -p "-DsyncTime='${syncTime}' -DtableName='${tableName}' -DsrcUrl='${srcUrl}' -DsrcUserName='${srcUserName}' -DsrcPassword='${srcPassword}' -DdesUrl='${desUrl}' -DdesUserName='${desUserName}' -DdesPassword='${desPassword}'" {YOUR_DATAAX_HOME}/job/increase.json

if [ -1 -lt "$flg" ]; then
    let "flg-=1"
fi
sleep $sleepTime
done

```

4 Shell

```
chmod +x datax_excute_job.sh  
nohup ./datax_excute_job.sh > info.file 2>&1 &
```

SQL Server TiDB

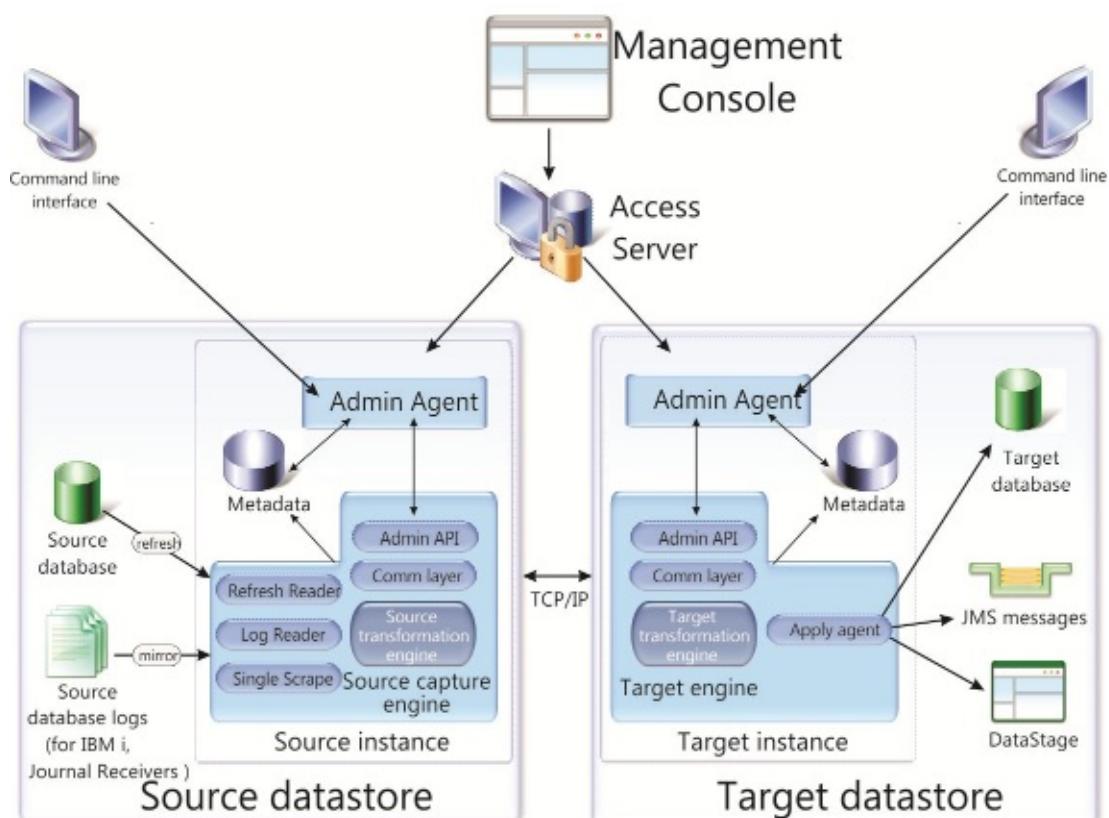
Db2 TiDB CDC

Db2 TiDB Db2 Db2 for LUW Db2 for iAS/400 Db2 for z/OS Mainframe Db2 for LUW IBM CDC OGG
Db2 for i Db2 for z/OS IBM CDC Db2 for i CDC TiDB IBM CDC TiDB

IBM Db2 InfoSphere Data Replication IIDR IIDR Change data capture (CDC) Replication Q Replication SQL
Replication Db2 for i TiDB Change data capture (CDC) Replication IBM CDC IIDR

CDC

- IIDR-CDC CDC



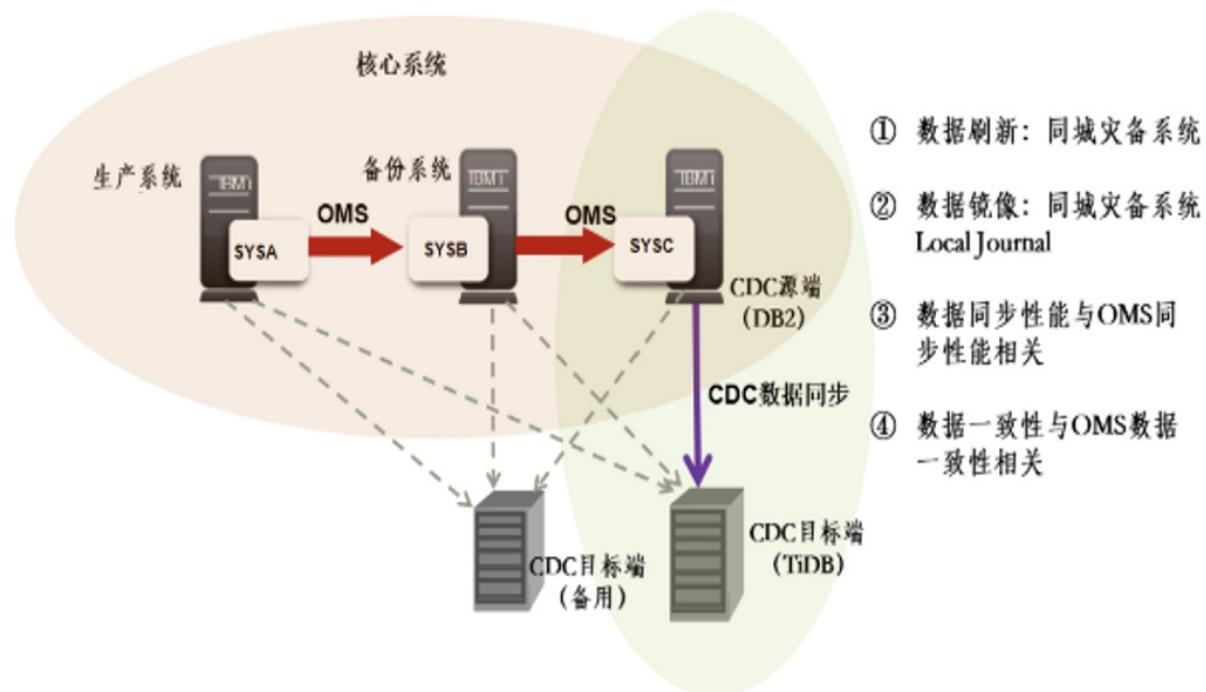
- CDC

https://www.ibm.com/support/knowledgecenter/SSTRGZ_11.4.0/com.ibm.cdc.doc.sysreq.doc/concepts/supportedsourceandtarget.html

| Supported source databases | Supported target databases and middleware applications |
|--|---|
| IBM??Db2? for Linux, UNIX and Windows (LUW) | IBM?Db2 for Linux, UNIX and Windows (LUW) |
| IBM?Db2 for i | IBM?Db2 for i |
| IBM?Db2 for z/OS? | IBM?Db2 for z/OS |
| IMS | IBM?InfoSphere? DataStage? |
| Microsoft SQL Server | Microsoft SQL Server |
| Oracle | CDC Replication Engine for FlexRep |
| Sybase?1 | Netezza? |
| Informix??1 | Oracle |
| Db2 on Cloud?(formerly dashDB? for Transactions)?2 | Sybase?1 |
| VSAM | Informix?1 |
| PostgreSQL | CDC Replication Engine for Event Server?1 |
| Db2 Warehouse on Cloud?(formerly dashDB for Analytics)?3 | IBM Cloudant? |
| Db2 Warehouse?(formerly dashDB Local)?3 | Apache? Hadoop |
| MySQL | Apache Kafka |
| | Db2 Warehouse on Cloud?(formerly dashDB for Analytics) |
| | Db2 on Cloud?(formerly dashDB for Transactions)?2 |
| | Db2 Warehouse?(formerly dashDB Local) |
| | IBM MQ for z/OS (using Classic CDC for z/OS) |
| | Teradata |
| | Microsoft Azure SQL Database |
| | Microsoft Azure SQL Database Managed Instance |
| | IBM Integrated Analytics System |

- TiDB CDC Replication Engine for FlexRep JDBC TiDB MySQL MySQL JDBC driver TiDB
- CDC datastage type FlexRep JDBC driver TiDB

Db2 for i TiDB

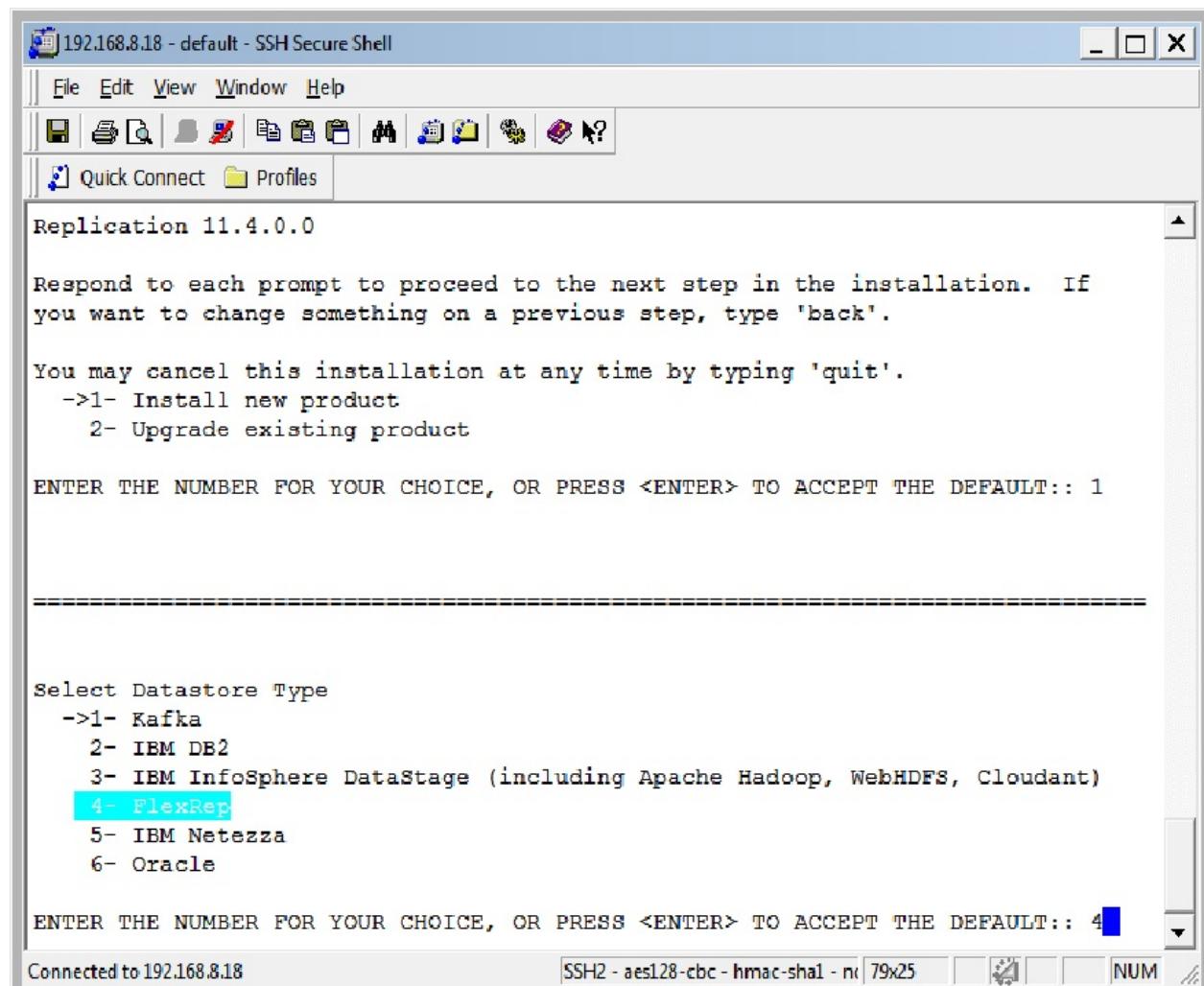


IBM OMS IIIDR TiDB IIIDR

IIIDR IIIDR TiDB

- IIIDR FlexRep

TiDB IIIDR FlexRep



- MySQL

IIIDR MySQL JDBC TiDB MySQL

192.168.8.18 - default - SSH Secure Shell

File Edit View Window Help

Quick Connect Profiles

NEW INSTANCE: tstcdc >> JDBC DRIVERS

Please refer to documentation for the list of supported drivers and their associated configuration values/parameters.

1. Add Driver

Enter your selection:1

Enter the location of the new JDBC driver file:/etc/mysql-connector-java-5.1.36-bin.jar

Connected to 192.168.8.18 SSH2 - aes128-cbc - hmac-sha1 - nc | 79x25 NUM

- IIDR

| | |
|---|---------------------|
| convert_not_nullable_column | true |
| events_max_retain | 10000 |
| global_conversion_not_possible_warning | false |
| global_max_batch_size | 25 |
| global_shutdown_after_no_heartbeat_reponse_minute | 10 |
| Implicit_transformation_warning | true |
| jdbc_refresh_commit_after_max_operation | 4000 |
| Mirror_commit_after_max_operations | 4000 |
| Mirror_global_disk_quota_gb | 9223372036854775807 |
| Mirror_interim_commit_thresholds | 100 |
| Userexit_max_lob_size_kb | 2097151 |
| Mirror_commit_on_transaction_boundary | False |

-

IIDR TiDB Db2 for i TiDB

| Db2 | TiDB |
|------------|--------------|
| L | date |
| T | time |
| Z | timestamp(6) |
| A | varchar |
| P | Decimal |
| S | decimal |
| O | varchar |

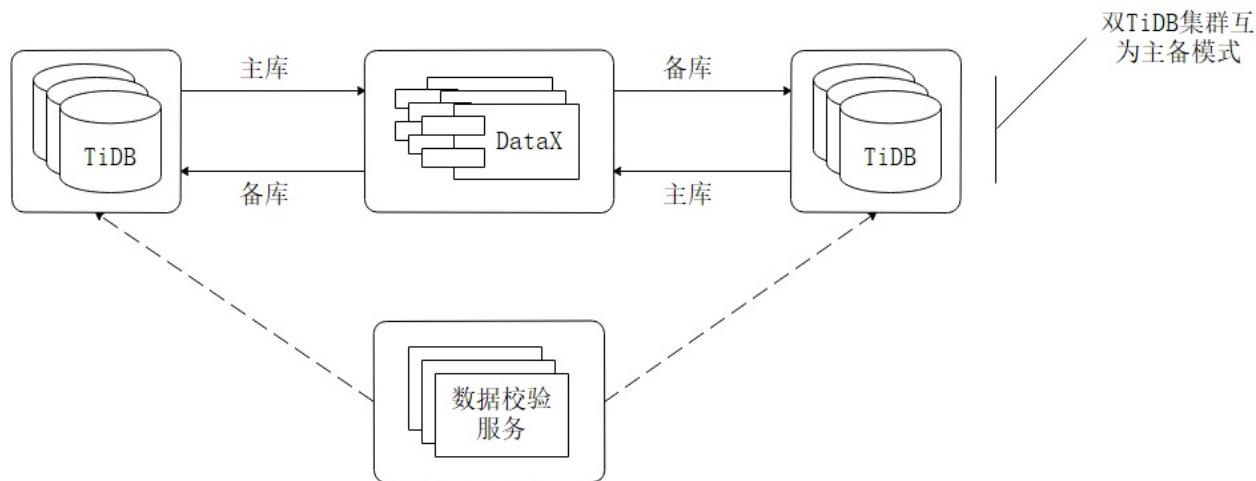
- CDC Db2 Db2 TiDB loadalter table xxx activate not logged initially
- CDC datastage type FlexRep TiDB

IIDR Db2 for i TiDB IIDR TiDB

TiDB TiDB TiDB

DataX TiDB TiDB 4.0 CDC

TiDB



DataX

```
wget http://datax.opensource.oss-cn-hangzhou.aliyuncs.com/datax.tar.gz
```

```
tar -zxvf datax.tar.gz
```

```
python {YOUR_DATAx_HOME}/bin/datax.py {YOUR_DATAx_HOME}/job/job.json
```

Job

vi increase.json

```
{
  "job": {
    "setting": {
      "speed": {
        "channel": 128 #Channel
      }
    },
    "content": [
      {
        "reader": {
          "name": "tidbreader",

```

```

"parameter": {
    "username": "${srcUserName}",
    "password": "${srcPassword}",
    "column": ["*"],
    "connection": [
        {
            "table": ["${tableName}"],
            "jdbcUrl": ["${srcUrl}"]
        }
    ],
    "where": "updateTime >= '${syncTime}'"
},
"writer": {
    "name": "tidbwriter",
    "parameter": {
        "username": "${desUserName}",
        "password": "${desPassword}",
        "writeMode": "replace",
        "column": ["*"],
        "connection": [
            {
                "jdbcUrl": "${desUrl}",
                "table": ["${tableName}"]
            }
        ],
        "preSql": [
            "replace into t_sync_record(table_name,start_date,end_date) values('@table',now(),null)"
        ],
        "postSql": [
            "update t_sync_record set end_date=now() where table_name='@table' "
        ]
    }
}
}

```

DataX JobShell

vi datax_execute_job.sh

```

#!/bin/bash

source /etc/profile

srcUrl="Reader SourceTiDB"

srcUserName=""

srcPassword=""

desUrl="Writer DestTiDB"

desUserName=""

desPassword=""

# 

defaultSyncUpdateTime="2020-03-03 18:00:00.000"

#()

sleepTime="N"

tableName="Table1,Table2, . . ."

#-1

flg=-1

while [ "$flg" -gt 0 -o "$flg" -eq -1 ]

do

# 

if [ "" = "$preRunTime" ]; then

syncTime=$defaultSyncUpdateTime

else

syncTime=$preRunTime

fi

# 

preRunTime=$(date -d +"%Y-%m-%d %T").000";

echo $syncTime

echo $preRunTime

echo $flg

python {YOUR_DATAx_HOME}/bin/datamax.py -p "-DsyncTime='${syncTime}' -DtableName='${tableName}' -DsrcUrl='${srcUrl}' -DsrcUserName='${srcUserName}' -DsrcPassword='${srcPassword}' -DdesUrl='${desUrl}' -DdesUserName='${desUserName}' -DdesPassword='${desPassword}'" {YOUR_DATAx_HOME}/job/increase.json

if [ -1 -lt "$flg" ]; then

let "flg-=1"

fi

sleep $sleepTime

done

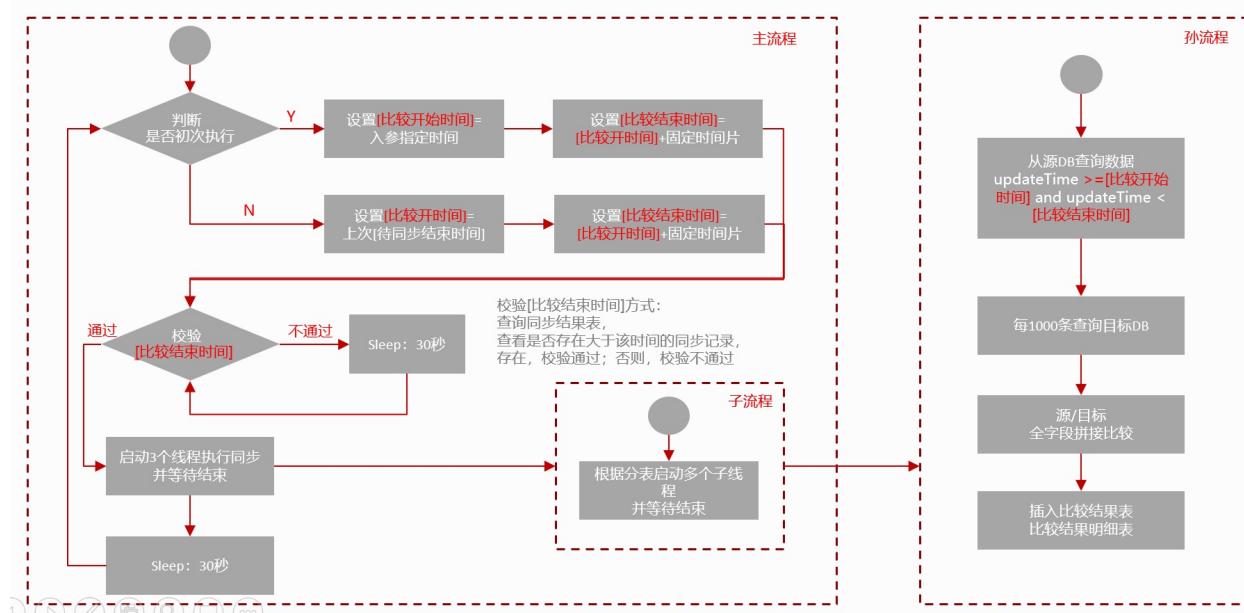
```

Shell

```
chmod +x datax_execute_job.sh
nohup ./datax_execute_job.sh > log.file 2>&1 &
```

DataX Job Shell Supervisor TiDB

TiDB TiDB



DataX TiDB

5.7 MongoDB TiDB

5.7.1

MongoDB MongoDB

MongoS Opolog MySQL

MongoS AP MongoDB Kafka ClickHouse

MongoDB TiDB

5.7.2 TiDB MongoDB

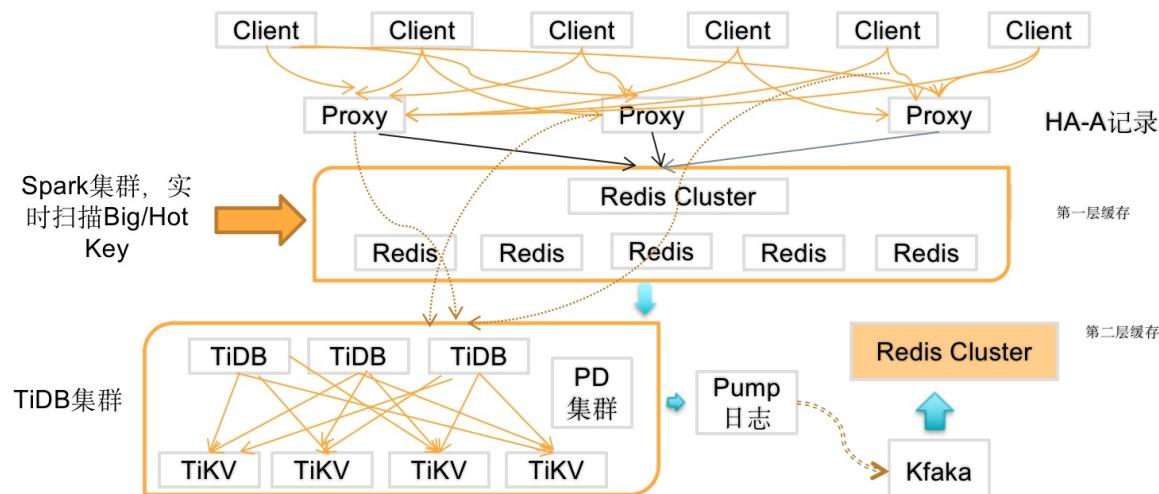
| MongoDB | TiDB |
|---|----------------------|
| 1.
2. json
3.
4. Sharded cluster | 1.
2.
3.
4. |

5.7.3 TiDB

Redis Cluster MongoDB TiDB Redis TiDB

Web Service Proxy Redis Cluster TiDB Transaction too large Redis Cluster TiDB TiDB TiDB-binlog Kafka Redis Cluster

Proxy TiDB Redis Cluster key-value Redis TiDB SQL TiDB Spark Big Key Hot Key
Redis LFU key-value 5 key-value Spark 35 TiDB Redis Cluster



Redis TiDB

sleep

5.7.4

1.

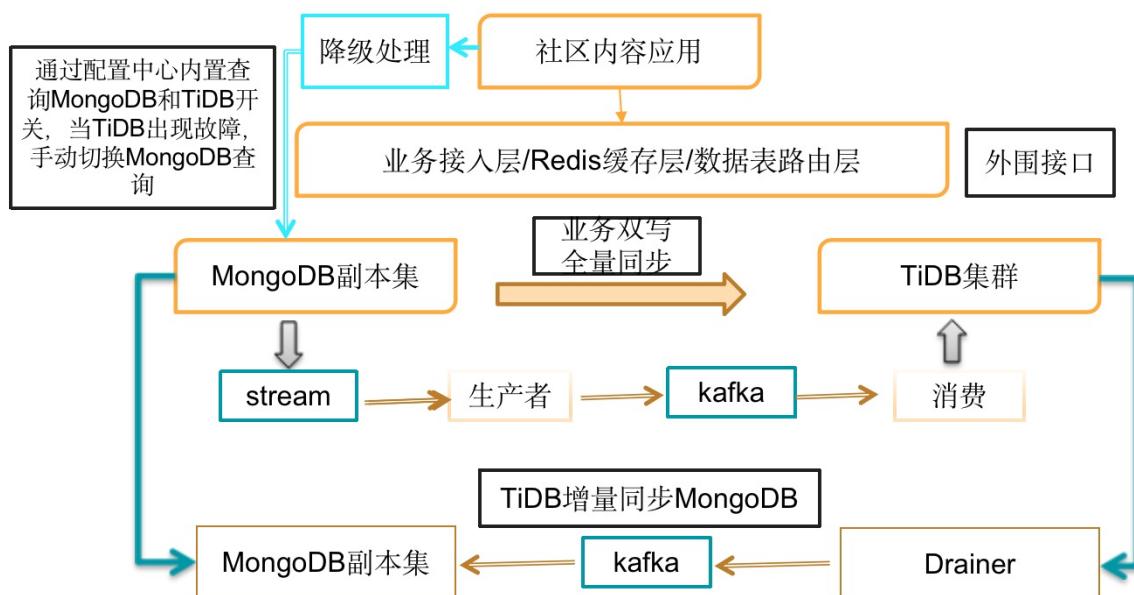
•

- o MongoDB MongoDB _id TiDB ID

- uuid snowflakes
- PK PK TiDB rowid INSERT Region SHARD_ROW_ID_BITS rowid Region RPC CPU
- v4.0 TiDB Auto Random Key Auto Random Key int
- MongoDB TiDB MongoDB TiDB DATEDATETIME TIMESTAMP TIME YEAR
-
-
- Web Service
- watch
- uid uid

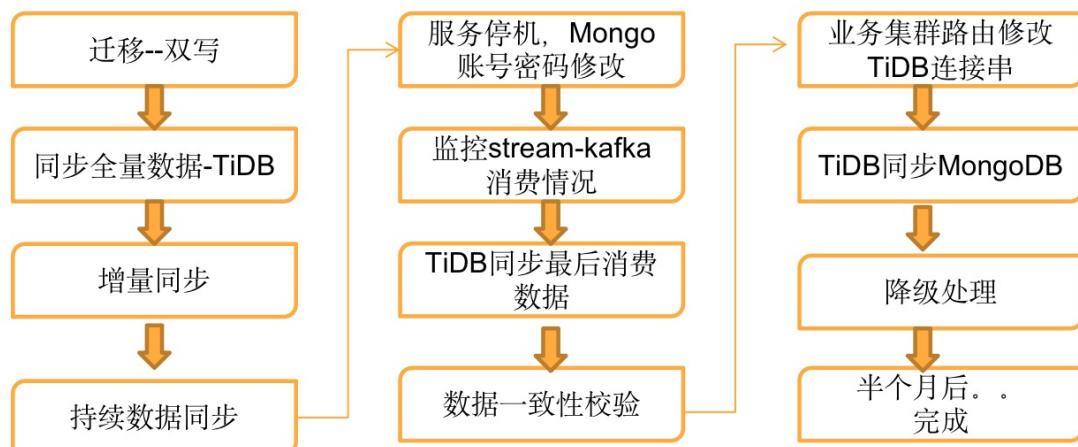
2.

MongoDB java Redis Cluster MongoDB Change Stream Kafka TiDB TiDB TiDB TiDB-binlog
 Kafka MongoDB TiDB MongoDB TiDB MongoDB TiDB MongoDB MongoDB



MongoDB v3.6 Change Stream watch MongoCursor

3.



(1)

| | | |
|---------------|----------------------------|-----------------------|
| max(uid) TiDB | max(uid) TiDB MongoDB TiDB | max(uid) DML TiDB DML |
|---------------|----------------------------|-----------------------|

(2) Redis

20 uid Redis Cluster Redis log

```

127.0.0.1:6379> KEYS *
1) "migration:data:st"
127.0.0.1:6379> HGETALL "migration:comminfo:st"
1) "2186200681228325_2382488577771540"
2) "2188513952998466"
3) "2859668199691429_93967041042064455"
4) "2861166835008519"
5) "2487557684037633_2612301467395085"
6) "2488649487696938"
7) "2612301467395085_2657794730464475"
8) "2613352273347597"
9) "2382488577771540_2438002819885096"
10) "2383933609133067"
11) "2708706397606993_2859668199691429"
12) "2709891834446963"
13) "2657794730464475_2708706397606993"
14) "2658901430273111"
15) "2094211406200848_2186200681228325"
16) "2096628556585988"
17) "2438002819885096_2487557684037633"
18) "2439044447011868"
  
```

(3)

| | | | | | | | | |
|------------------|---------------|------------------|--------|--------|--------|---------|---------------|-------|
| UpdateTime Kafka | Change Stream | UpdatedTime TiDB | insert | update | delete | MongoDB | Change Stream | Kafka |
|------------------|---------------|------------------|--------|--------|--------|---------|---------------|-------|

(4)

(5)

(6)

- TiDB MongoDB TiDB TiDB TiDB TiDB
- TiDB MongoDB TiDB TiDB
- TiDB TiDB-binlog Kafka MongoDB

5.7.5

1.

```

MongoDB  Change Stream Kafka TiDB Redis      uid  Kafka      uid  Redis  uid
127.0.0.1:6379> KEYS *
1) "migration:data:st"
127.0.0.1:6379> HGETALL "migration:comminfo:st"
 1) "2186200681228325_2382488577771540"
 2) "2188513952998466"
 3) "2859668199691429_93967041042064455"
 4) "2861166835008519"
 5) "2487557684037633_2612301467395085"
 6) "2488649487696938"
 7) "2612301467395085_2657794730464475"
 8) "2613352273347597"
 9) "2382488577771540_2438002819885096"
10) "2383933609133067"
11) "2708706397606993_2859668199691429"
12) "2709891834446963"
13) "2657794730464475_2708706397606993"
14) "2658901430273111"
15) "2094211406200848_2186200681228325"
16) "2096628556585988"
17) "2438002819885096_2487557684037633"
18) "2439044447011868"

```

2. MongoDB metadata

Metadata MongoDB TiDB MongoDB MongoDB Metadata Metadata TiDB

3. TiDB

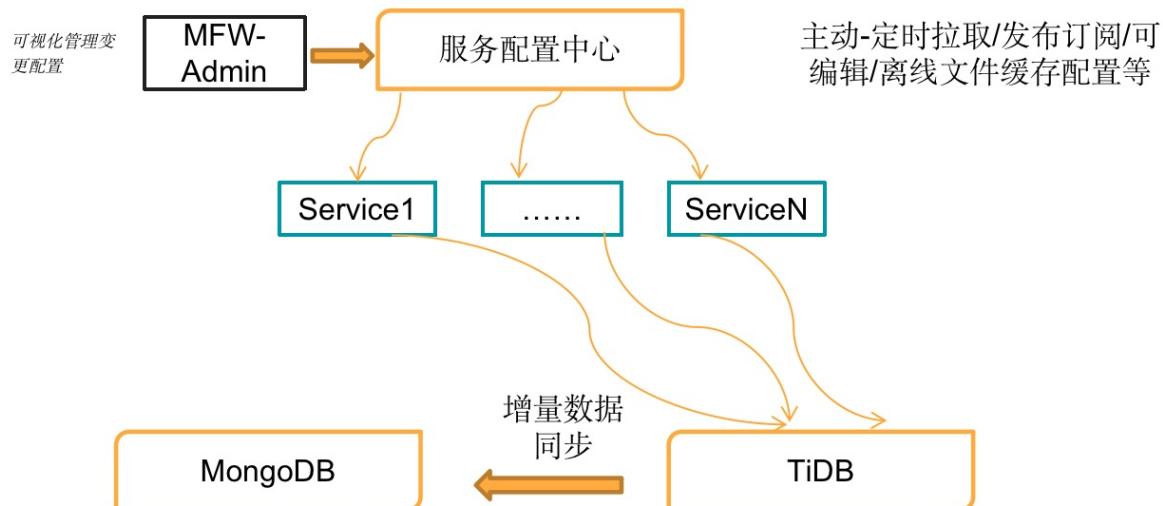
TiDB MongoDB TiDB SQL

5.7.6 -

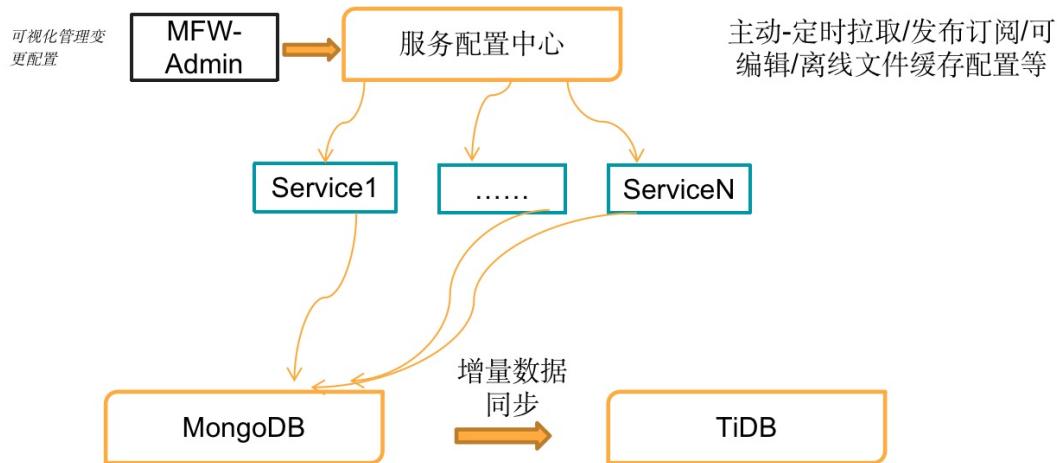
TiDB MongoDB

1.

TiDB TiDB TiDB-binlog Kafka MongoDB MongoDB



TiDB Web Service MongoDB MongoDB Change Stream Kafka TiDB



1.

“” RPC “” RPC

RPC RPC ”“

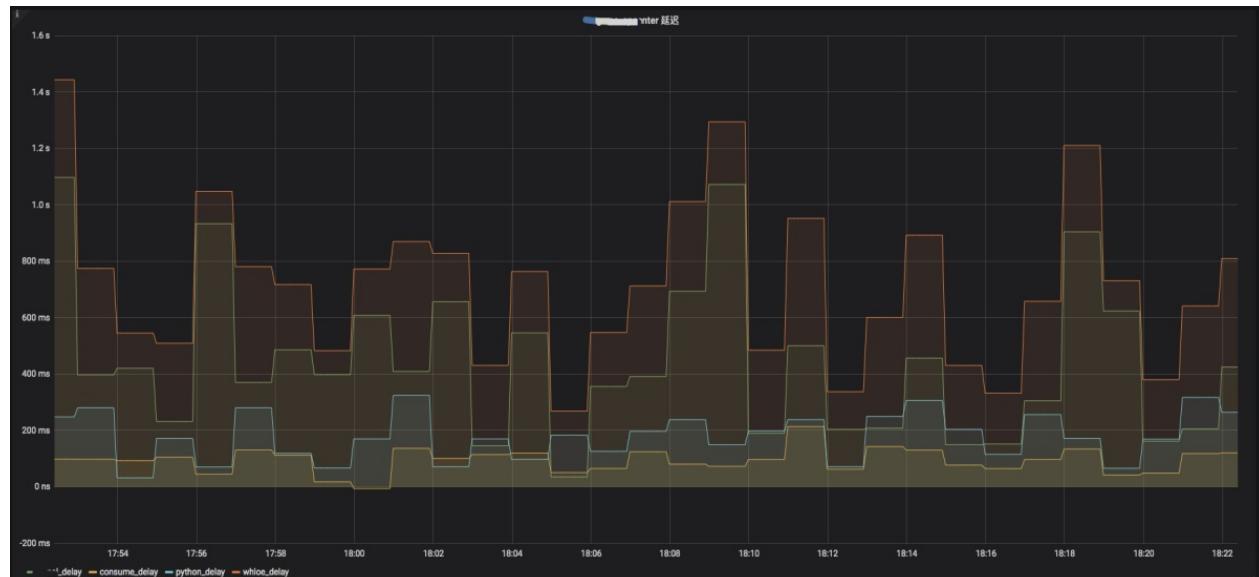
```

#
circuitBreaker.requestVolumeThreshold    // 20
circuitBreaker.errorThresholdPercentage // 50%
circuitBreaker.sleepWindowInMilliseconds// 5000 5s
#
20 50% 5s

```

“”Web Service RPC “”“”

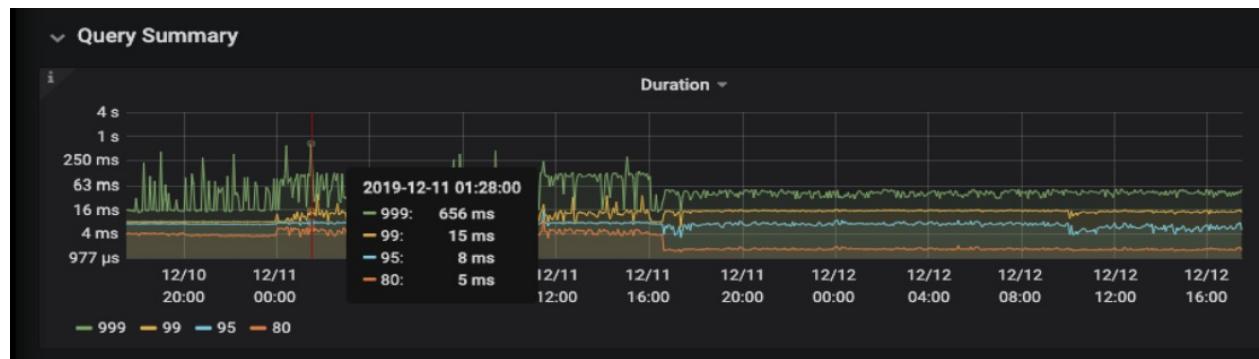
MongoDB TiDB TiDB MongoDB TiDB Kafka



5.7.7 SQL

MongoDB SQL TP 99% 15ms 99.9% 700ms sum group by TP 99% 15ms 99.9% 48ms
SQL

•



5.7.8

MongoDB TiDB MongoDB change stream TiDB TiDB-binlog Kafka Updatetime RedisMongoDBTiDB
MongoDB TiDB

6.1

TiDB TiDB

-
- [TiDB](#)
-
-
- [hint](#)
- [SQL](#)
- [TiDB + TiSpark](#)

6.1.1

3.0.8 TiDB TiDB

-
-

-
-
- commit

6.1.1.1

- Key Key
- Key

TiDB commit

1. TiDB
2. TCC) TiDB

(, ,), , ,

1. , TiDB v3 , (redis),
2. .. TiDB (kafka partition),
- 3.

6.1.1.2

TiDB (100500)

1. TiDB OOM
- 2.
- 3.

TiDB

1. SQL 5000
2. 6MB
3. 300,000
4. 100MB

6.1.1.3

prewrite TiKV TiDB

- TiDB TiDB TiKV
- TiKV prewrite TiDB TiDB TiDB TiKV

TiDB

```
txn-local-latches    enable    false
capacity Hash slot 2 slot 32 Bytes 1024000
```

capacity key hash hash capacity hash

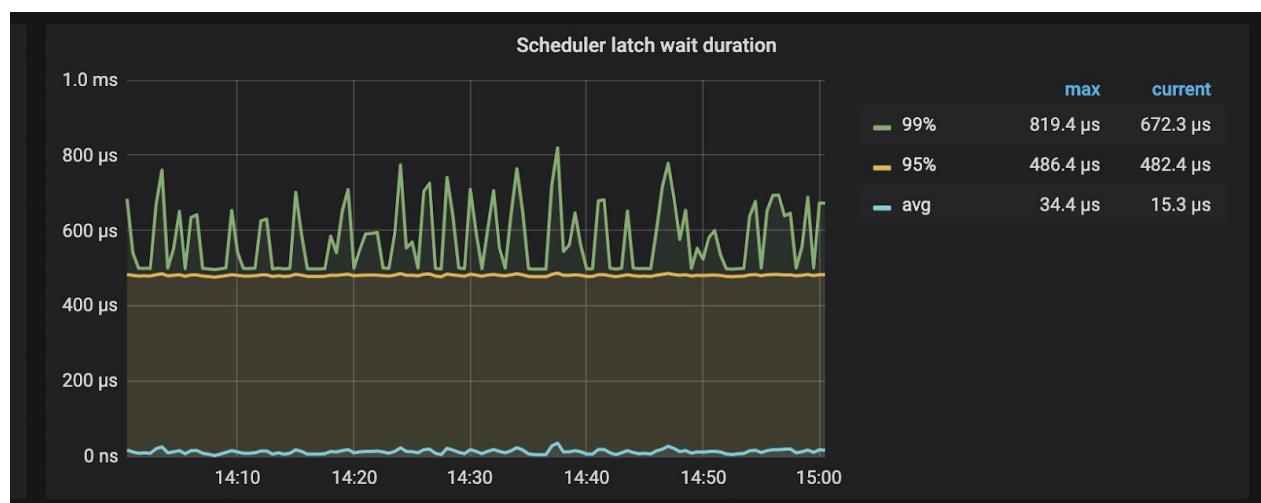
- capacity
- capacity

TiDB

TiKV TiKV hash

```
scheduler-concurrency scheduler key key hash 2048000
```

TiKV latch



wait duration

6.1.1.4 tidb

commit Commit MySQL SQL commit TiDB

- [tidb_disable_txn_auto_retry](#) on
- [tidb_retry_limit](#)

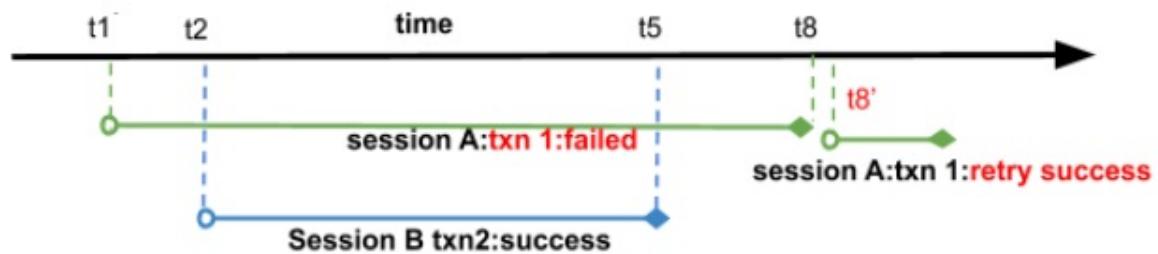
1. start_ts
2. SQL
- 3.

`tidb_disable_txn_auto_retry = off` lost update TiDB tso, region TiDB `tidb_disable_txn_auto_retry = off`

| | Session A | Session B |
|----|--|-------------------------|
| t1 | MySQL [test]>
begin;
<i>Query OK, 0 rows affected (0.00 sec)</i> | |
| | | MySQL [test]>
begin; |

| | | | | | | | | | | |
|----|---|--|---------|-----------|--|------|---------|------|--|--|
| | | Query OK, 0 rows affected (0.00 sec) | | | | | | | | |
| t3 | | MySQL [test]>
update tidb set
status=0 where
id=1;
Query OK, 1 row affected (0.01 sec)
Rows matched: 1
Changed: 1
Warnings: 0 | | | | | | | | |
| t4 | MySQL [test]>
select from tidb
where id=1;
**+-----+---+
-----+
\\ | id \\ | name \\ | status \\ | +----
---+--
---+
+----
---+
\\ | 1 \\ | tikv \\ | 1 \\ | +----
---+--
---+
+----
---+
1
row
in set
(0.01
sec)* | |
| t5 | | commit;
Query OK, 0 rows affected (0.01 sec) | | | | | | | | |
| t6 | update tidb set
name='pd' where
id =1 and
status=1;
Query OK, 1 row affected (0.00 sec)
Rows matched: 1
Changed: 1
Warnings: 0 | | | | | | | | | |
| t7 | MySQL [test]>
select from tidb
where id=1;
**+-----+---+
-----+
\\ | id \\ | name \\ | status \\ | +----
---+--
---+
+----
---+
\\ | 1 \\ | pd \\ | 1 \\ | +----
---+--
---+
+----
---+
1
row
in set
(0.01
sec)* | |
| t8 | MySQL [test]>
commit;
Query OK, 0 rows affected (0.01 sec) | | | | | | | | | |
| t9 | MySQL [test]>
select from tidb
where id=1;
**+-----+---+
-----+
\\ | id \\ | name \\ | status \\ | +----
---+--
---+
+----
---+
\\ | 1 \\ | tikv \\ | 0 \\ | +----
---+--
---+
+----
---+
1
row
in set
(0.00
sec)* | |

case:



- session B t2 2t5 session A 1 2 2
- 1 2
- session A 1 tidb 1
 - start_ts t8'
 - update tidb set name='pd' where id =1 and status=1
 - t8'
 -
- tidb 1
- session A

start_ts ReadRepeatable

SQL TiDB

TiDB TiDB

4.6 TiDB

2.1 TiDB

4.6.1

>

TiDB TiDB

6MB

300,000

100MB

SQL 5000

PingCAP -

TiDB RocksDB Key-Value 6MB 100MB 30W

4.6.2 30W

30W TiKV Key-Value

Key-Value

| Key | Value | Flag |
|-----|-------|------|
| | | |

- **Insert**

TiKV

(1)

| Key: PK + TSO | Value: Fields | Flag: Put |
|---------------|---------------|-----------|
| | | |

(2)

| Key: Index (UK) + TSO | Value: PK | Flag: Put |
|-----------------------|-----------|-----------|
| | | |

(3)

| Key: Index + PK + TSO | Value: Null | Flag: Put |
|-----------------------|-------------|-----------|
| | | |

Insert 30W 1 + ()

- **Delete**

TiKV RocksDB Flag = Del

(1)

| Key: PK + TSO | Value: Null | Flag: Del |
|---------------|-------------|-----------|
| | | |

(2)

| | | |
|------------------------------|--------------------|------------------|
| Key: Index (UK) + TSO | Value: Null | Flag: Del |
|------------------------------|--------------------|------------------|

(3)

| | | |
|------------------------------|--------------------|------------------|
| Key: Index + PK + TSO | Value: Null | Flag: Del |
|------------------------------|--------------------|------------------|

Delete 30W (1 + 0)

- **Update**

(1)

| | | |
|----------------------|----------------------|------------------|
| Key: PK + TSO | Value: Fields | Flag: Put |
|----------------------|----------------------|------------------|

30W 30W

(1)

| | | |
|----------------------|----------------------|------------------|
| Key: PK + TSO | Value: Fields | Flag: Put |
|----------------------|----------------------|------------------|

(2)

| | | |
|------------------------------|--------------------|------------------|
| Key: Index (UK) + TSO | Value: Null | Flag: Del |
|------------------------------|--------------------|------------------|

| | | |
|------------------------------|------------------|------------------|
| Key: Index (UK) + TSO | Value: PK | Flag: Put |
|------------------------------|------------------|------------------|

(3)

| | | |
|------------------------------|--------------------|------------------|
| Key: Index + PK + TSO | Value: Null | Flag: Del |
|------------------------------|--------------------|------------------|

| | | |
|------------------------------|--------------------|------------------|
| Key: Index + PK + TSO | Value: Null | Flag: Put |
|------------------------------|--------------------|------------------|

30W (1 + * 2)

PK Key

(1)

| | | |
|----------------------|--------------------|------------------|
| Key: PK + TSO | Value: Null | Flag: Del |
|----------------------|--------------------|------------------|

| | | |
|----------------------|----------------------|------------------|
| Key: PK + TSO | Value: Fields | Flag: Put |
|----------------------|----------------------|------------------|

(2)

| | | |
|------------------------------|------------------|------------------|
| Key: Index (UK) + TSO | Value: PK | Flag: Put |
|------------------------------|------------------|------------------|

(3)

| | | |
|------------------------------|--------------------|------------------|
| Key: Index + PK + TSO | Value: Null | Flag: Del |
| Key: Index + PK + TSO | Value: Null | Flag: Put |

30W ((1 +) * 2 +) Update 1 KV 2 KV Key-Value Key UK Value PKUpdate PK Key Put kv
 Key PK Update Del kvPut kv

4.6.3 30W

||| |:----:|:----:| | Insert | 30W/(1+Idx_Count) || Delete | 30W/(1+Idx_Count) || Update_On_PK | 30W/((1+Non_UK)*2+UK*1)
 Update_non_PK | 30W/(1+Involved_Idx_Count*2) |

```
CREATE TABLE t1 (
  id int(11) NOT NULL AUTO_INCREMENT,
  name char(10) CHARSET utf8mb4 COLLATE utf8mb4_bin DEFAULT NULL,
  age int(11) DEFAULT NULL,
  PRIMARY KEY (id),
  Key idx_name (name)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_bin
```

1

| | | |
|----------------|-------------------|------|
| | | |
| Insert | 30W/(1+1) | 15W |
| Delete | 30W/(1+1) | 15W |
| Update_On_id | 30W/((1+1)*2 + 0) | 7.5W |
| Update_On_name | 30W/(1+1*2) | 10W |
| Update_On_age | 30W/(1+0*2) | 30W |

TiDB int int

```
CREATE TABLE t1 (
  id varchar(11) NOT NULL,
  name char(10) CHARSET utf8mb4 COLLATE utf8mb4_bin DEFAULT NULL,
  age int(11) DEFAULT NULL,
  PRIMARY KEY (id),
  Key idx_name (name)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_bin
```

1 () 1

| | | |
|----------------|-------------|-----|
| | | |
| Insert | 30W/(1+2) | 10W |
| Delete | 30W/(1+2) | 10W |
| Update_On_id | 30W/(1+1*2) | 10W |
| Update_On_name | 30W/(1+1*2) | 10W |
| Update_On_age | 30W/(1+0*2) | 30W |

4.6.4

TiDB

(1) stmt-count-limit 5000

StmtCountLimit limits the max count of statement inside a transaction.

5000 SQL statement TiDB

(2) Insert Into Select

ERROR 1105 (HY000): BatchInsert failed with error: [try again later]: con:3877 **txn takes too much time**, start: 405023027269206017, commit: 405023312534306817

max-txn-time-use gc_life_time - 10s 590

PingCAP GitHub <https://github.com/pingcap/TiDB/blob/master/config/config.toml.example#L240>

The max time a Txn may use (in seconds) from its startTS to commitTS. # We use it to guarantee GC worker will not influence any active txn. Please make sure that this# Value is less than gc_life_time - 10s.

gc_life_time - 10s gc + TiDB Batch

4.6.5

Batch

- tidb_batch_insert

: SESSION : 0 Autocommit True Batch Batch

- tidb_batch_delete

: SESSION : 0 Autocommit True Batch Batch

- tidb_dml_batch_size

: SESSION : 20000 / Batch tidb_batch_insert tidb_batch_delete 20k

Update limit Batch Update

Batch DBA Batch 100 TiDB

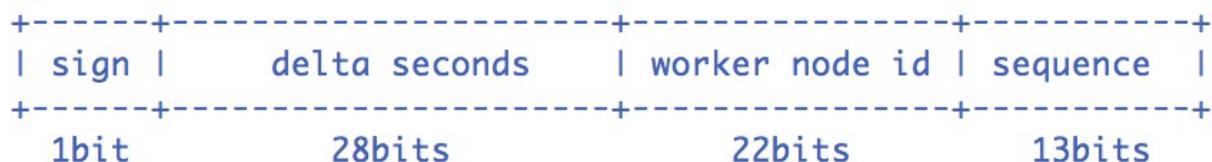
6.1.3

SequenceTiDB 4.0

Snowflake

Snowflake Twitter ID uid-generator leaf uid-generator

uid-generator 64 ID



- sign 1 0 ID
- delta seconds 28 ("2016-05-20") 28 8.7
- worker node id 22 id ID Zookeeper worker node id22 420
- sequence 13 13 8192

Snowflake

- delta seconds ,
- delta seconds , 28 44
- delta seconds
- worker node id 500 TiDB worker node id TiDB 3 500 / 3 = 166 worker node id Snowflake

ID | | | | :---- | :----- || SEQ_NAME | varchar(128) | | MAX_ID | bigint(20) | | STEP | int(11) | |

ID TiDB IDbit-reverse ID

SQL / MySQL limit offset

```

begin;
update sbtest1 set pad='new_value' where id in (select id from sbtest1 order by id limit 0,10000);
commit;
begin;
update sbtest1 set pad='new_value' where id in (select id from sbtest1 order by id limit 10000,10000);
commit;
begin;
update sbtest1 set pad='new_value' where id in (select id from sbtest1 order by id limit 20000,10000);
commit;

```

SQL / TiKV row_number()

200tmp_loan

```

MySQL [demo]> desc tmp_loan;
+-----+-----+-----+-----+
| Field | Type | Null | Key | Default | Extra |
+-----+-----+-----+-----+
serialno	int(11)	NO	PRI	NULL	
name	varchar(40)	NO			
businesssum	int(10)	NO		0	
+-----+-----+-----+-----+

```

200

```

MySQL [demo]> select count(1) from tmp_loan;
+-----+
| count(1) |
+-----+
| 1998985 |
+-----+
MySQL [demo]> select * from tmp_loan limit 10;
+-----+-----+-----+
| serialno | name | businesssum |
+-----+-----+-----+
200000000		10000
200000001		10000
200000002		10000
200000003		10000
200000004		10000
200000005		10000
200000006		10000
200000007		10000
200000008		10000
200000009		10000
+-----+-----+-----+

```

MOD

```

select serialno from tmp_table where MOD(substring(serialno,-3),${ThreadNums}) = ${ThreadId} order by serialno;
${ThreadNums} ${ThreadId}

```

```
MySQL [demo]> select serialno from tmp_loan where MOD(substring(serialno,-3),17) = 1 order by serialno;
+-----+
| serialno |
+-----+
| 200000001 |
| 200000018 |
| 200000035 |
| 200000052 |
| 200000069 |
| 200000086 |
| .... |
+-----+
117942 rows in set (1.407 sec)
```

1712 RabbitMQ sql ResultSet Mysql TiDB GC lift time 10min ResultSet GC life time is shorter than transaction duration ResultSet ResultSet sql mod

row_number()

```
MySQL [demo]> SELECT min(t.serialno) AS start_key, max(t.serialno) AS end_key, count(*) AS page_size FROM ( SELECT *, row_number () over (ORDER BY serialno) AS row_num FROM tmp_loan ) t GROUP BY floor((t.row_num - 1) / 50000) ORDER BY start_key;
+-----+-----+-----+
| start_key | end_key | page_size |
+-----+-----+-----+
200000000	200050001	50000
200050002	200100007	50000
200100008	200150008	50000
200150009	200200013	50000
200200014	200250017	50000
....	....	....
201900019	201950018	50000
201950019	201999003	48985
+-----+-----+-----+
40 rows in set (1.51 sec)
```

serialno between start_key and end_key row_number() mod TiKV / row_number()

```
MySQL [demo]> select serialno from tmp_loan where serialno BETWEEN 200050002 and 200100007;
+-----+
| serialno |
+-----+
| 200050002 |
| 200050003 |
| 200050004 |
| 200050005 |
| 200050006 |
| .... |
+-----+
50000 rows in set (0.070 sec)
```

businesssum () row_number()

```
MySQL [demo]> update tmp_loan set businesssum='6666' where serialno BETWEEN 200000000 and 200050001;
Query OK, 50000 rows affected (0.89 sec)
Rows matched: 50000  Changed: 50000  Warnings: 0
MySQL [demo]> select * from tmp_loan ORDER BY serialno limit 10;
+-----+-----+-----+
| serialno | name      | businesssum |
+-----+-----+-----+
200000000		6666
200000001		6666
200000002		6666
200000003		6666
200000004		6666
200000005		6666
200000006		6666
200000007		6666
200000008		6666
200000009		6666
+-----+-----+-----+
```

6.1.5 Hint

hint TiDB mysql use indexforce indexignore index TiDB Optimizer Hints MySQL 5.7 comment / +
TIDB_XX(t1, t2) / TiDBhint

| Hint | |
|-------------------------------------|-------------------------------------|
| USE INDEX | Index Hint: Choose Index |
| FORCE INDEX | Index Hint: Choose Index |
| IGNORE INDEX | Index Hint: Ignore Index |
| /+ <i>TIDB_INLJ(t)</i> / | Join Hint: Nested Index Lookup Join |
| /+ <i>TIDB_HJ(t)</i> / | Join Hint: Hash Join |
| /+ <i>TIDB_SMJ(t)</i> / | Join Hint: Merge Join |
| /+ <i>MAX_EXECUTION_TIME(num)</i> / | Execution Time Limit |

6.1.5.1 USE INDEX, FORCE INDEX, IGNORE INDEX

MySQL , USE INDEX

use/force index SQL

```
mysql> explain select * from t;
+-----+-----+-----+-----+
| id      | estRows | task      | access object | operator info      |
+-----+-----+-----+-----+
| TableReader_5    | 8193.00 | root      |                | data:TableFullScan_4 |
| └─TableFullScan_4 | 8193.00 | cop[tikv] | table:t      | keep order:false   |
+-----+-----+-----+-----+
2 rows in set (0.00 sec)

mysql> explain select * from t use index(idx_1);
+-----+-----+-----+-----+
| id      | estRows | task      | access object        | operator info      |
+-----+-----+-----+-----+
IndexLookUp_6	8193.00	root		
└─IndexFullScan_4(Build)	8193.00	cop[tikv]	table:t, index:idx_1(a)	keep order:false
└─TableRowIDScan_5(Probe)	8193.00	cop[tikv]	table:t	keep order:false
+-----+-----+-----+-----+
3 rows in set (0.00 sec)

mysql> explain select * from t force index(idx_1);
+-----+-----+-----+-----+
| id      | estRows | task      | access object        | operator info      |
+-----+-----+-----+-----+
IndexLookUp_6	8193.00	root		
└─IndexFullScan_4(Build)	8193.00	cop[tikv]	table:t, index:idx_1(a)	keep order:false
└─TableRowIDScan_5(Probe)	8193.00	cop[tikv]	table:t	keep order:false
+-----+-----+-----+-----+
3 rows in set (0.00 sec)
```

ignore index SQL

```

mysql> explain select a from t where a=2;
+-----+-----+-----+-----+
| id      | estRows | task      | access object           | operator info      |
+-----+-----+-----+-----+
| IndexReader_6    | 1.00    | root      |                   | index:IndexRangeScan_5   |
| └─IndexRangeScan_5 | 1.00    | cop[tikv] | table:t, index:idx_1(a) | range:[2,2], keep order:false |
+-----+-----+-----+-----+
2 rows in set (0.00 sec)

mysql> explain select a from t ignore index(idx_1) where a=2 ;
+-----+-----+-----+-----+
| id      | estRows | task      | access object           | operator info      |
+-----+-----+-----+-----+
TableReader_7	1.00	root		data:Selection_6
└─Selection_6	1.00	cop[tikv]	eq(test.t.a, 2)	
└─TableFullScan_5	8193.00	cop[tikv]	table:t	keep order:false
+-----+-----+-----+-----+
3 rows in set (0.00 sec)

```

MySQL , TiDB use index force index

use index tidbanalyze table use index

6.1.5.2 MAX_EXECUTION_TIME(N)

SELECT MAX_EXECUTION_TIME(N) N

1

```
SELECT /*+ MAX_EXECUTION_TIME(1000) */ * FROM t1
```

max_execution_time

MAX_EXECUTION_TIME bug tidb . OLTP 5

mysql jdbc tidb KILL tidb TiDB tidb KILL MAX_EXECUTION_TIME

6.1.5.3 JOIN HINT

TiDB Join Sort Merge JoinIndex Nested Loop JoinHash Join join [TiDB](#)

1. TIDB_SMJ(t1, t2)

```
SELECT /*+ TIDB_SMJ(t1, t2) */ * from t1t2 where t1.id = t2.id;
```

Sort Merge Join Join ,

2. TIDB_INLJ(t1, t2)

```
SELECT /*+ TIDB_INLJ(t1, t2) */ * from t1t2 where t1.id = t2.id;
```

Index Nested Loop Join Index Look Up Join WHERE 1 TIDB_INLJ() TIDB_INLJ(t1) t1

3. TIDB_HJ(t1, t2)

```
SELECT /*+ TIDB_HJ(t1, t2) */ * from t1t2 where t1.id = t2.id;
```

Hash Join t1 t2 Hash Join Inner Outer

hint / + *TIDB_STREAMAGG()* // + *TIDB_HASHAGG()* /

Hint SQL SQL SQL

6.2 SQL

SQL SQL

count estRows.

1 Delete OOM

```
MySQL [db_stat]> explain delete from t_stat where imp_date<='20200202';
+-----+-----+-----+
| id      | count    | task   | operator info          |
+-----+-----+-----+
TableReader_6	220895815.00	root	data:Selection_5
└Selection_5	220895815.00	cop	le(db_stat.t_stat.imp_date, "20200202")
└TableScan_4	220895815.00	cop	table:t_stat, range:[-inf,+inf], keep order:false
+-----+-----+-----+
3 rows in set (0.00 sec)

MySQL [db_stat]> select count(*)  from t_stat where imp_date<='20200202';
+-----+
| count(*) |
+-----+
| 184340473 |
+-----+
1 row in set (17.88 sec)
```

- TiDB OOM
- imp_date IndexScan Table ScanTiDB TiKV Coprocessor
- TiKV Coprocessor CPU
- Delete TiDB TiDB
- limit N
- 3.0 Range

2

```

MySQL [db_stat]> explain SELECT * FROM `tbl_article_check_result` `t` WHERE (articleid = '20190925A0PYT800') ORDER BY checkTime desc LIMIT 100 ;
+-----+-----+-----+
| id      | count   | task  | operator info
+-----+-----+-----+
| Projection_7    | 100.00  | root  | db_stat.t.type, db_stat.t.articleid, db_stat.t.docid, db_stat.t.version, db_stat.t.checkid, db_stat.t.checkstatus, db_stat.t.seclevel, db_stat.t.t1checkstatus, db_stat.t.t2checkstatus, db_stat.t.mdaichannel, db_stat.t.mdaisubchannel, db_stat.t.checkuser, db_stat.t.checktime, db_stat.t.addtime, db_stat.t.havegot, db_stat.t.checkcode
| └Limit_12        | 100.00  | root  | offset:0, count:100
|
|   └IndexLookup_34 | 100.00  | root  |
|
|     └IndexScan_31 | 30755.49 | cop   | table:t, index:checkTime, range:[NULL,+inf], keep order:true, desc
|
|       └Selection_33 | 100.00  | cop   | eq(db_dayu_1.t.articleid, "20190925A0PYT800")
|
|         └TableScan_32 | 30755.49 | cop   | table:tbl_article_check_result, keep order:false
+
+-----+-----+-----+
6 rows in set (0.00 sec)

```

- articleid checkTime articleid checkTime
- LIMIT 100 100 checkTime articleid checkTime articleid articleid
- Duration
- analyze table crontab analyze
- auto analyze analyze ratio
 - set global tidb_auto_analyze_ratio=0.2;
 - set global tidb_auto_analyze_start_time='00:00 +0800';
 - set global tidb_auto_analyze_end_time='06:00 +0800';
- SQL force index articleid
- 3.0 SQL create binding force index SQL
- 4.0 SQL Plan Management 1.3

3

```
MySQL [db_stat]> explain select * from t_like_list where person_id=1535538061143263;
+-----+-----+-----+
| id      | count   | task | operator info
|          |
+-----+-----+-----+
| Selection_5    | 1430690.40 | root | eq(cast(db_stat.t_like_list.person_id), 1.535538061143263e+15)
|          |
|  └TableReader_7 | 1788363.00 | root | data:TableScan_6
|          |
|    └TableScan_6 | 1788363.00 | cop  | table:t_like_list, range:[-inf,+inf], keep order:false
|          |
+-----+-----+-----+
-----+
3 rows in set (0.00 sec)
```

- person_id IndexScan

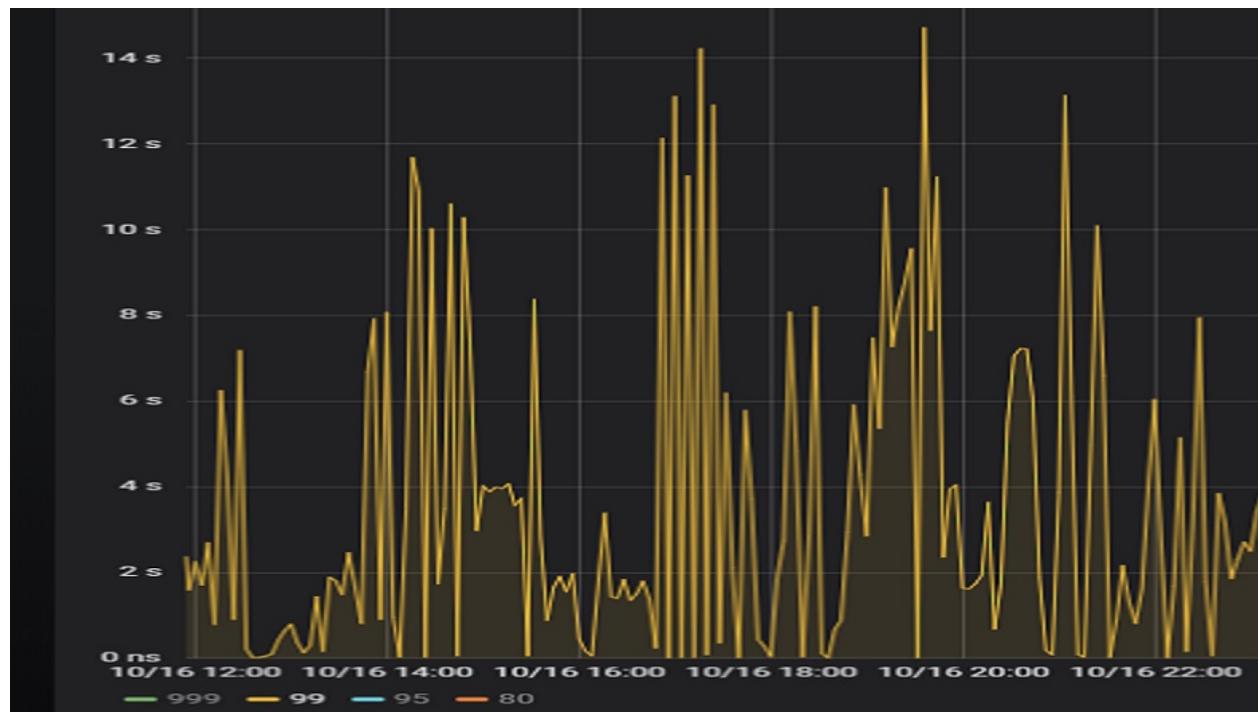
- person_id cast

- where

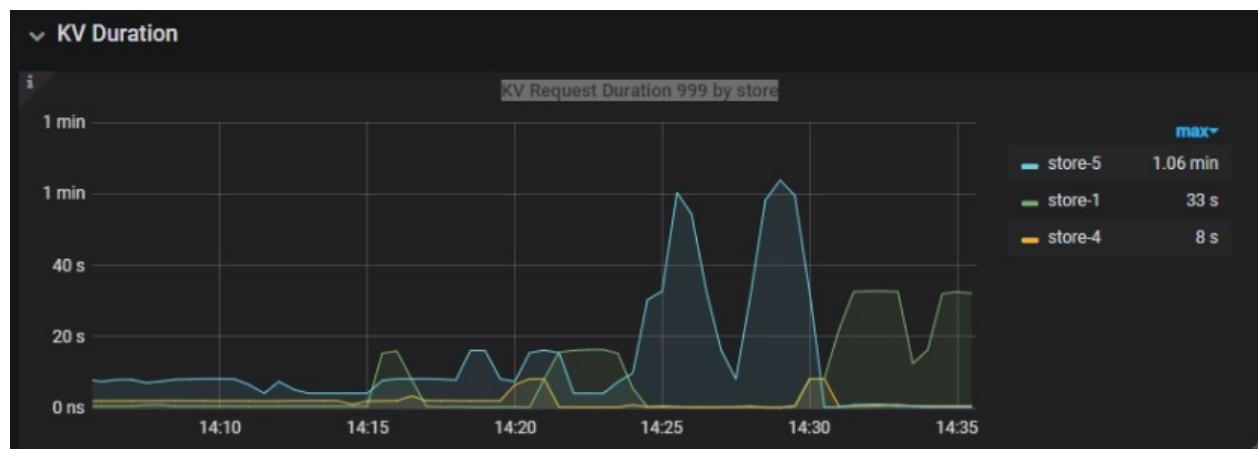
```
MySQL [db_stat]> explain select * from table:t_like_list where person_id='1535538061143263';
+-----+-----+-----+
| id      | count   | task | operator info
|          |
+-----+-----+-----+
| IndexLookUp_10 | 0.00  | root |
|          |
|  └IndexScan_8  | 0.00  | cop  | table:t_like_list, index:person_id, range:["1535538061143263", "1535538061143263"], keep order:false
|          |
|    └TableScan_9 | 0.00  | cop  | table:t_like_list, keep order:false
|          |
+-----+-----+-----+
-----+
3 rows in set (0.00 sec)
```

4 SQL

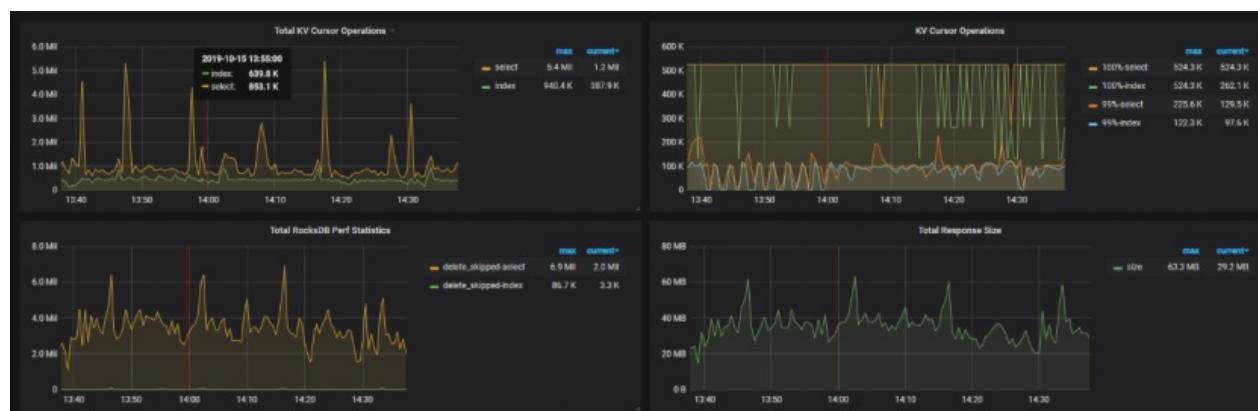
600G TiDB TiDB Query Summary - Duration p99



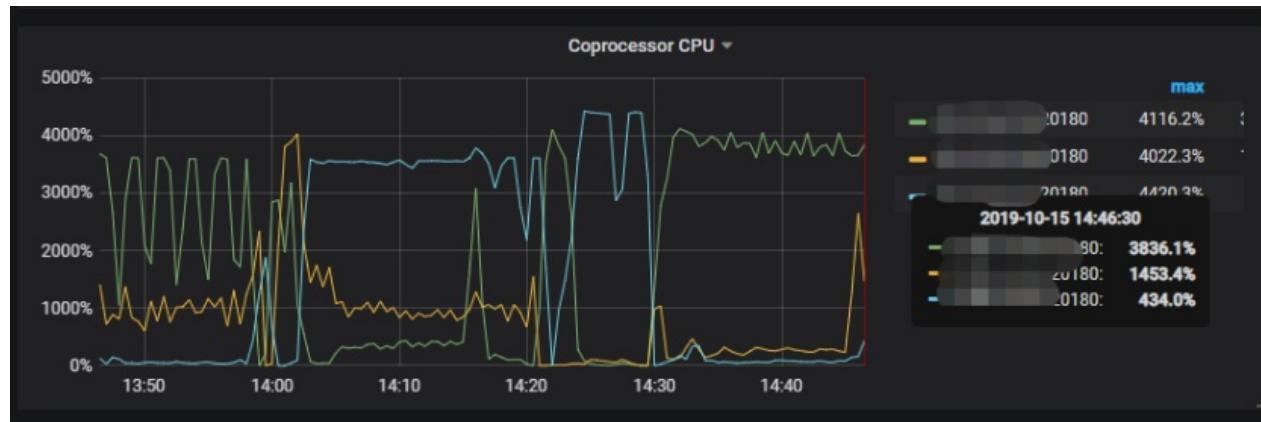
TiDB KV Duration KV Request Duration 999 by store TiKV Duration



TiKV Coprocessor Overview



Coprocessor CPU



Coprocessor CPU Duration Coprocessor CPU

pt-query-digest TiDB

```
./pt-query-digest tidb_slow_query.log > result
```

TopSQL Process keys Process time Process keys SQL Process time SQL Process keys 22.09MProcess time 51s

```
# Query 26: 0.08 QPS, 0.34x concurrency, ID 0x2AA5E845384665393DC48BB36364AC51 at byte 140924321
# This item is included in the report because it matches --outliers.
# Scores: V/M = 4.20
# Time range: 2019-10-15T12:14:48 to 2019-10-15T14:45:55
# Attribute   pct      total      min      max      avg     95%    stddev   median
# ======  ======  ======  ======  ======  ======  ======  ======  ======
# Count          0       739
# Exec time     0  3039s   300ms     28s      4s     12s      4s      3s
# Query size    0   22.37k      31      31      31      31      0      31
# Conn ID       0   559.72k     15    7.32k   775.58     4.71k   1.32k   299.03
# Cop proc avg   0   50.76      0.04     0.14     0.07     0.09     0.01     0.07
# Cop proc max   0   50.76      0.04     0.14     0.07     0.09     0.01     0.07
# Cop proc p90   0   50.76      0.04     0.14     0.07     0.09     0.01     0.07
# Cop wait avg   0   2.30k      0     24.03     3.19     8.91     3.15     2.39
# Cop wait max   0   2.30k      0     24.03     3.19     8.91     3.15     2.39
# Cop wait p90   0   2.30k      0     24.03     3.19     8.91     3.15     2.39
# Mem max        0  699.31M  968.95k  969.37k  969.01k  961.27k      0  961.27k
# Num cop task   0       739      1       1       1       1       0       1
# Process keys   0   22.09M     157   32.17k   30.60k   31.59k   6.77k   31.59k
# Process time   0      51s     35ms   138ms    69ms    91ms    15ms    68ms
# Request coun   0       739      1       1       1       1       0       1
# Total keys    0   22.43M     197   32.88k   31.08k   31.59k   6.76k   31.59k
# Txn start ts  0  263.99E  365.80P  365.81P  365.80P   1.25P      0   1.25P
# Wait time      0   2355s     1ms     24s      3s      9s      3s      3s
# String:
# Cop proc add  .245:20171 (392/53%)... 2 more
# Cop wait add  .245:20171 (392/53%)... 2 more
# DB
# Digest        769ecfd1b40ab79558c26dale5b64a2b68d65919b23...
# Is internal   false
# Stats          grp:411858... (100/13%), grp:411859... (95/12%)... 20 more
# Succ           true
```

SQL Process keys 12.68M Process time 142353s

```
# Query 4: 0.36 QPS, 14.84x concurrency, ID 0xF0ECDD453DA544E79C7D0E42506FF4BE at byte 163086266
# This item is included in the report because it matches --limit.
# Scores: V/M = 13.91
# Time range: 2019-10-15T12:13:52 to 2019-10-15T15:20:49
# Attribute      pct    total      min      max      avg     95%   stddev   median
# ====== ====== ====== ====== ====== ====== ====== ====== ======
# Count          1    3999
# Exec time     11 166406s      3s    170s     42s     80s     24s     39s
# Query size    1 515.50k     132     132     132     132      0    132
# Conn ID       0    2.22M      10    7.32k   583.33    2.06k   1.03k   258.32
# Cop proc avg  32    2.53k     0.17    1.34     0.65     0.99     0.17     0.61
# Cop proc max  35    7.15k     0.45    3.60     1.83     2.76     0.49     1.70
# Cop proc p90  34    6.44k     0.34    3.32     1.65     2.63     0.46     1.62
# Cop wait avg  1    8.10k     0.00    8.48     2.07     4.50     1.41     1.96
# Cop wait max  3    30.67k     0.00   36.06     7.85   16.81     5.21     6.98
# Cop wait p90  1    18.28k     0.00   23.68     4.68     9.83     3.04     4.50
# Mem max       39 119.23G   30.27M   30.54M   30.53M   29.99M      0   29.99M
# Num cop task 18 214.79k      51     55   55.00     54.21     0.08     54.21
# Process keys  0    12.68M   3.18k    3.36k    3.25k    3.19k   57.63   3.19k
# Process time  61 142353s      9s    74s     36s     54s     10s     33s
# Request coun 18 214.79k      51     55   55.00     54.21     0.08     54.21
# Total keys    0   13.00M   3.23k    3.50k    3.33k    3.35k   69.99   3.19k
# Txn start ts  1    1.40Z 365.80P 365.81P 365.80P   1.25P      0   1.25P
# Wait time     24 456039s      6ms    466s    114s    246s     78s    107s
# String:
# Cop proc add [REDACTED] .245:20171 (1814/45%)... 2 more
# Cop wait add [REDACTED] .245:20171 (1835/45%)... 2 more
# DB
# Digest        473bf876de8bb5210bf4b300a6b3d94b0c0210balale...
# Is internal   false
# Stats          tpl:411850... (202/5%), tpl:411850... (192/4%)... 196 more
# Succ          true
```

Process time SQL 3 slow query

- SQL1

```
select a.a_id, a.b_id, uqm.p_id from a join hsq on a.b_id=hsq.id join uqm on a.a_id=uqm.id;
```

```
id count task operator info execution info memory
Projection_9 1878.75 root
time:lm29.290313266s, loops:69, rows:69176 N/A
└─IndexJoin_13 1878.75 root inner join, inner:TableReader_12, outer key:[REDACTED], inner
key:[REDACTED] time:lm29.290100131s, loops:69, rows:69176 22.52175521850586 MB
  └─IndexJoin_31 1878.75 root inner join, inner:IndexLookUp_30, outer key:[REDACTED], inner
key:[REDACTED] time:lm11.787661278s, loops:422, rows:429136 7.81768798828125 MB
    └─TableReader_44 1503.00 root data:TableScan_43 time:1.963267ms, loops:3, rows:1503 16.3544921875 KB
      └─TableScan_43 1503.00 cop table:tpl, range:[0,+inf], keep order:false time:1ms, loops:6, rows:1503 N/A
        └─IndexLookUp_30 284.10 root time:lm35.454370598s, loops:422, rows:429136 100.359375 KB
          └─IndexScan_28 284.10 cop table:[REDACTED], index:[REDACTED], range: decided by [eq("
[REDACTED]",)], keep order:false proc max:470ms, min:36ms, p80:470ms, p95:470ms, rows:429136, iters:428, tasks:2 N/A
            └─TableScan_29 284.10 cop table:[REDACTED], keep order:false proc max:1.695s, min:30ms, p80:1.291s, p95:1.525s, rows:429136,
iters:543, tasks:27 N/A
      └─TableReader_12 1.00 root data:TableScan_11 time:lm9.184965119s, loops:100, rows:64302 N/A
        └─TableScan_11 1.00 cop table:[REDACTED], range: decided by [REDACTED], keep order:false proc max:702ms,
min:4ms, p80:490ms, p95:627ms, rows:64302, iters:165, tasks:25 N/A
```

- SQL2

```
select distinct g.abc, g.def, g.ghi, h.abcd, hi.jq from ggg g left join ggg_host gh on g.id = gh.ggg_id left join host h on gh.a_id = h.id left join a_jq hi on h.id = hi.hid where h.abcd is not null and h.abcd <> '' and hi.jq is not null and hi.jq <> '';
```

```
id count task operator info execution info memory
HashAgg_13 80296.47 root group by:dbl.g.def, dbl.g.abc, dbl.g.ghi, dbl.h.abcd, dbl.hi.jq, funcs:firstrow(dbl.g.abc),
firstrow(dbl.g.def), firstrow(dbl.g.ghi), firstrow(dbl.h.abcd), firstrow(dbl.hi.jq) time:8.646954728s, loops:347, rows:354274 N/A
└─IndexJoin_18 80296.47 root inner join, inner:IndexReader_17, outer key:dbl.hi.hid, inner key:dbl.hi.hid time:7.613316688s,
loops:507, rows:517734 257.38256072998047 MB
  └─IndexJoin_47 41156.25 root inner join, inner:TableReader_46, outer key:dbl.gh.a_id, inner key:dbl.hi.hid time:2.76087231s,
loops:656, rows:669536 438.8576126098633 MB
  | └─IndexJoin_54 41156.25 root inner join, inner:IndexReader_53, outer key:dbl.g.id, inner key:dbl.gh.ggg_id
  time:1.69100157s, loops:1189, rows:1215334 27.39978790283203 MB
  |   └─TableReader_66 32925.00 root data:TableScan_65 time:75.325365ms, loops:35, rows:32925 4.20347785949707 MB
  |     └─TableScan_66 32925.00 cop table:g, range:[0,+inf], keep order:false time:52ms, loops:37, rows:32925 N/A
  |       └─IndexReader_53 0.99 root index:Selection_52 time:1.898590556s, loops:1202, rows:1215334 15.298088073730469 MB
  |         └─Selection_52 0.99 cop not(isnull(dbl.gh.ggg_id)) proc max:664ms, min:29ms, p80:317ms, p95:664ms, rows:1215334,
iters:1218, tasks:10 N/A
  |           └─IndexScan_51 0.99 cop table:gh, index:ggg_id, a_id, range: decided by [eq(dbl.gh.ggg_id, dbl.g.id)
not(isnull(dbl.gh.a_id))], keep order:false proc max:640ms, min:29ms, p80:316ms, p95:640ms, rows:1215334, iters:1218, tasks:10 N/A
  |             └─TableReader_46 0.80 root data:Selection_45 time:3.441797722s, loops:732, rows:666902 N/A
  |               └─Selection_45 0.80 cop ne(dbl.h.abcd, ""), not(isnull(dbl.h.id)), not(isnull(dbl.h.abcd)) proc max:206ms, min:0s,
p80:10ms, p95:30ms, rows:666902, iters:4012, tasks:1487 N/A
  |                 └─TableScan_44 1.00 cop table:h, range: decided by [dbl.gh.a_id], keep order:false proc max:204ms, min:0s, p80:10ms,
p95:30ms, rows:668753, iters:4012, tasks:1487 N/A
  └─IndexReader_17 1.20 root index:Selection_16 time:20.824781527s, loops:545, rows:510094 144.6767578125 KB
    └─Selection_16 1.20 cop not(isnull(dbl.hi.hid)) proc max:876ms, min:4ms, p80:728ms, p95:789ms, rows:510094, iters:781, tasks:93
N/A
      └─IndexScan_15 1.20 cop table:hi, index:hid, jq, source, status, range: decided by [eq(dbl.hi.hid, dbl.h.id)
not(isnull(dbl.hi.jq)) ne(dbl.hi.jq, )], keep order:false proc max:875ms, min:4ms, p80:728ms, p95:789ms, rows:510094, iters:781,
tasks:93 N/A
```

- SQL3

```
select tb1.mt, tb2.name from tb2 left join tb1 on tb2.mtId=tb1.id where tb2.type=0 and (tb1.mt is not null and t
b1.mt != '') and (tb2.name is not null or tb2.name != '');
```

```
id count task operator info execution info memory
Projection_6 121393.07 root dbl.tbl.mt, dbl.tb2.name time:713.605177ms, loops:14, rows:10753 N/A
└─HashLeftJoin_11 121393.07 root inner join, inner:TableReader_26, equal:[eq(dbl.tb2.mtid, dbl.tbl.id)] time:713.544424ms,
loops:14, rows:10753 2.650096893310547 MB
  └─IndexLookup_23 121393.07 root time:621.192752ms, loops:120, rows:121269 2.6519317626953125 MB
    └─IndexScan_20 121393.07 cop table:tb2, index:type, range:[0,0], keep order:false time:69ms, loops:123, rows:121269 N/A
      └─Selection_22 121393.07 cop or(not(isnull(dbl.tb2.name)), ne(dbl.tb2.name, "")) proc max:390ms, min:9ms, p80:368ms, p95:390ms,
rows:121269, iters:166, tasks:10 N/A
        └─TableScan_21 121393.07 cop table:tb2, keep order:false proc max:387ms, min:8ms, p80:366ms, p95:387ms, rows:121269,
iters:166, tasks:10 N/A
  └─TableReader_26 41200.00 root data:Selection_25 time:171.796012ms, loops:42, rows:41200 2.2558584213256836 MB
    └─Selection_25 41200.00 cop ne(dbl.tbl.mt, ""), not(isnull(dbl.tbl.mt)) proc max:156ms, min:11ms, p80:156ms, p95:156ms,
rows:41200, iters:53, tasks:3 N/A
      └─TableScan_24 41200.00 cop table:tbl, range:[-inf,+inf], keep order:false proc max:152ms, min:11ms, p80:152ms, p95:152ms,
rows:41200, iters:53, tasks:3 N/A
```

TiDB TiKV

TiKV [slow-query] region

```
more tikv.log.2019-10-16-06\:28\:13 |grep slow-query |awk -F '[' '{print $1}' | awk '{print $6}' | sort | uniq -c |
sort -n
```

3 region

```
73 29452
140 33324
757 66625
```

region region region [slow-query] table_id start_ts TiDB table_id 1318start_ts 411837294180565013 SQL

```
more tidb-2019-10-14T16-40-51.728.log | grep '"/[1318/]"' |grep 411837294180565013
```

region split region 66625 x.x.x.x pd

```
pd-ctl -u http://x.x.x.x:2379 operator add split-region 66625
```

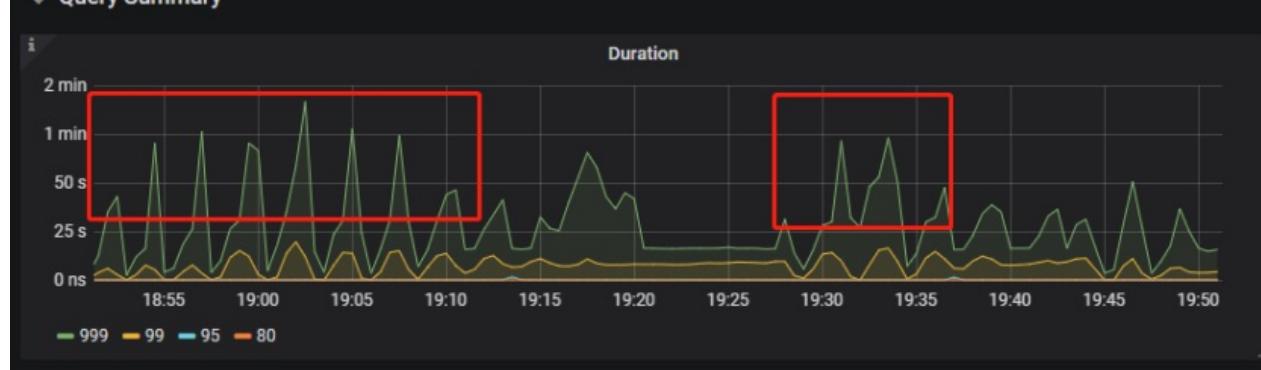
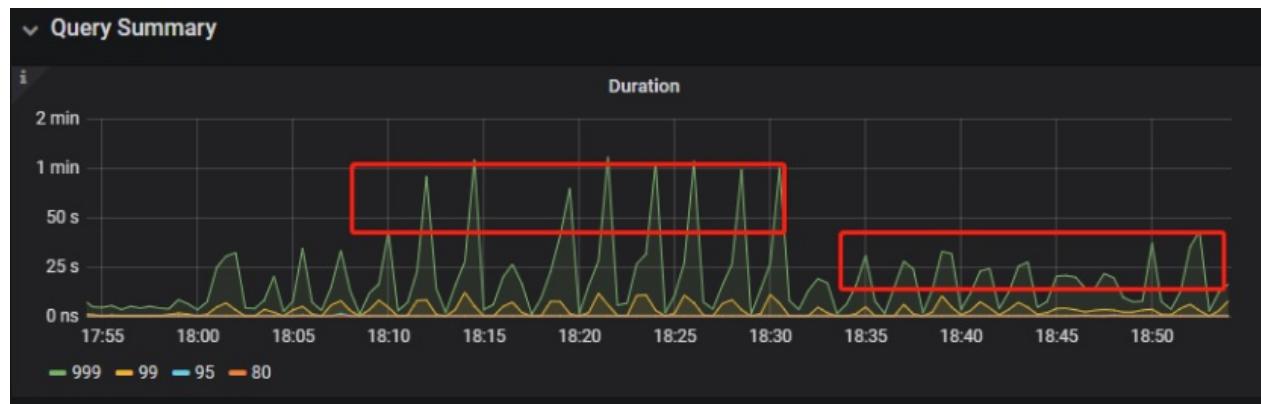
PD

```
[2019/10/16 18:22:56.223 +08:00] [INFO] [operator_controller.go:99] ["operator finish"] [region-id=30796] [operator=""
\admin-split-region (kind:admin, region:66625(1668,3), createAt:2019-10-16 18:22:55.888064898 +0800 CST m=+110918.82
3762963, startAt:2019-10-16 18:22:55.888223469 +0800 CST m=+110918.823921524, currentStep:1, steps:[split region with
policy SCAN]) finished"]
```

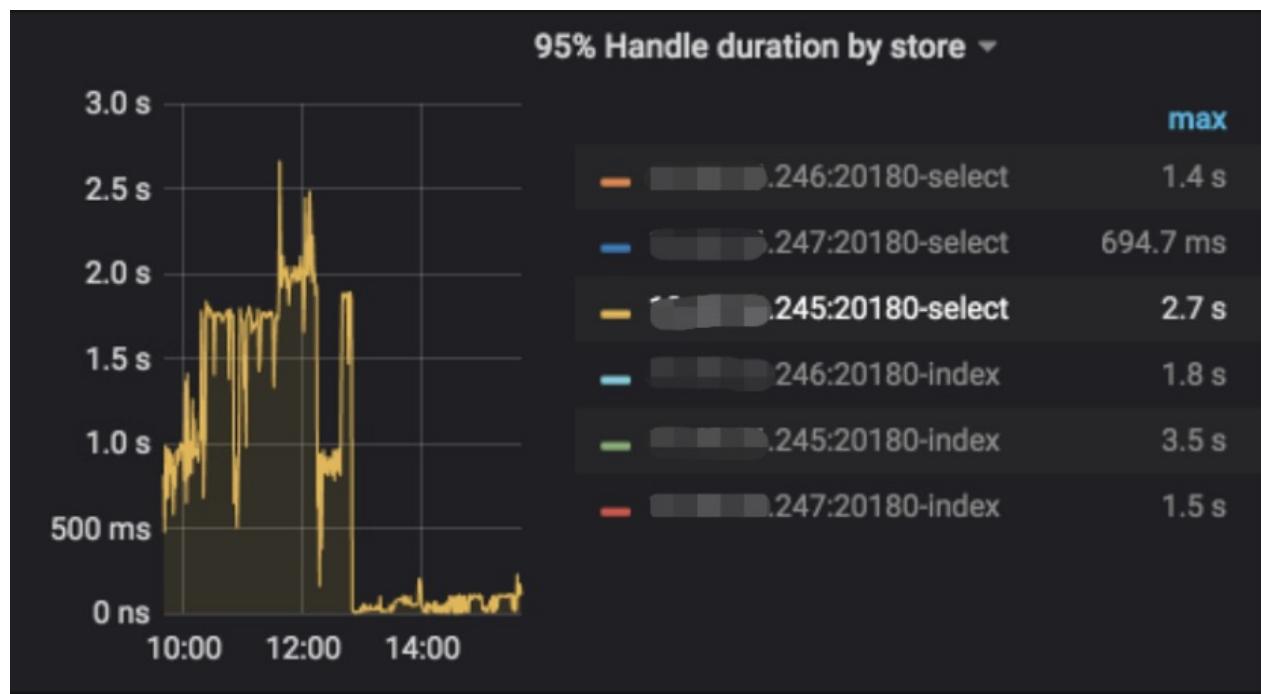
region region slow-query

```
more tikv.log.2019-10-16-06\:28\:13 |grep slow-query | grep 66625
```

66625 region region region Query Summary - Duration



```
pd-ctl -u http://x.x.x.x:2379 operator add transfer-leader 1 2 // region1 leader store2
```





SQL TiDB TiKV region region TiKV TiKV SQL 7.2

5 SQL

- SQL

- SQL

```
select count(*)
from todbs.bus_jijin_trade_record a, todbs.bus_jijin_info b
where a.fund_code=b.fund_code and a.type in ('PURCHASE','APPLY')
and a.status='CANCEL_SUCCESS' and a.pay_confirm_status = 1
and a.cancel_app_no is not null and a.id >= 177045000
and a.updated_at > date_sub(now(), interval 48 hour) ;

1 3.7s
mysql> select count(*)
-> from todbs.bus_jijin_trade_record a, todbs.bus_jijin_info b
-> where a.fund_code=b.fund_code and a.type in ('PURCHASE','APPLY')
-> and a.status='CANCEL_SUCCESS' and a.pay_confirm_status = 1
-> and a.cancel_app_no is not null and a.id >= 177045000
-> and a.updated_at > date_sub(now(), interval 48 hour) ;
+-----+
| count(*) |
+-----+
|      708 |
+-----+
1  row in set (1 min 3.77 sec)
```

-

| | | |
|------------------------|-----------|--|
| | | |
| bus_jijin_trade_record | 176384036 | PRIMARY KEY (ID) KEY idx_bus_jijin_trade_record_upd (UPDATED_AT) |
| bus_jijin_info | 6442 | PRIMARY KEY (ID) |

-

```

mysql> explain
-> select count(*)
-> from todbs.bus_jijin_trade_record a, todbs.bus_jijin_info b
-> where a.fund_code=b.fund_code and a.type in ('PURCHASE','APPLY')
-> and a.status='CANCEL_SUCCESS' and a.pay_confirm_status = 1
-> and a.cancel_app_no is not null and a.id >= 177045000
-> and a.updated_at > date_sub(now(), interval 48 hour) ;
+-----+-----+-----+
| id | count | task | operator info |
+-----+-----+-----+
StreamAgg_13	1.00	root	funcs:count(1)			
└─HashRightJoin_27	421.12	root	inner join, inner:TableReader_18, equal:[eq(a.fund_code, b.fund_code)]			
	TableReader_18	421.12	root	data:Selection_17		
		Selection_17	421.12	cop	eq(a.pay_confirm_status, 1), eq(a.status, "CANCEL_SUCCESS"), gt(a.updated_at, 2020-03-03 22:31:08), in(a.type, "PURCHASE", "APPLY"), not(isnull(a.cancel_app_no))	
			TableScan_16	145920790.55	cop	table:a, range:[177045000,+inf], keep order:false
			TableReader_37	6442.00	root	data:TableScan_36
			TableScan_36	6442.00	cop	table:b, range:[-inf,+inf], keep order:false
+-----+-----+-----+
TableScan_16TableScan_36 TiKV a b TableScan_16 1.46
Selection_17 a where
TableReader_37 b TiDB
TableReader_18 coprocessor a TiDB
HashRightJoin_27 TableReader_37 TableReader_18 hash join
StreamAgg_13

```

a (bus_jijin_trade_record) id >= 177045000 updated_at > date_sub(now(), interval 48 hour) TiDB

```

mysql> SELECT COUNT(*) FROM todbs.bus_jijin_trade_record WHERE id >= 177045000 ;
+-----+
| COUNT(*) |
+-----+
| 145917327 |
+-----+
1 row in set (16.86 sec)

mysql> SELECT COUNT(*) FROM todbs.bus_jijin_trade_record WHERE updated_at > date_sub(now(), interval 48 hour) ;
+-----+
| COUNT(*) |
+-----+
| 713682 |
+-----+

```

bus_jijin_trade_record 1.7 updated_at

hint 6.27 63s 6.3s 10

```
mysql> select count(*)
-> from tod.s.bus_jijin_trade_record a  use index(idx_bus_jijin_trade_record_upt), tod.s.bus_jijin_info b
-> where a.fund_code=b.fund_code and a.type in ('PURCHASE','APPLY')
-> and a.status='CANCEL_SUCCESS' and a.pay_confirm_status = 1
-> and a.cancel_app_no is not null and a.id >= 177045000
-> and a.updated_at > date_sub(now(), interval 48 hour) ;
+-----+
| count(*) |
+-----+
|    709 |
+-----+
1 row in set (6.27 sec)
```

hint

```
mysql> explain
-> select count(*)
-> from tod.s.bus_jijin_trade_record a  use index(idx_bus_jijin_trade_record_upt), tod.s.bus_jijin_info b
-> where a.fund_code=b.fund_code and a.type in ('PURCHASE','APPLY')
-> and a.status='CANCEL_SUCCESS' and a.pay_confirm_status = 1
-> and a.cancel_app_no is not null and a.id >= 177045000
-> and a.updated_at > date_sub(now(), interval 48 hour) ;
+-----+-----+-----+
| id      | count     | task | operator info
+-----+-----+-----+
| StreamAgg_13   | 1.00      | root | funcs:count(1)
|           |           |       |
| └HashRightJoin_24 | 421.12    | root | inner join, inner:IndexLookUp_20, equal:[eq(a.fund_code, b.fun
d_code)]
|   | IndexLookUp_20 | 421.12    | root |
|   |           |           |
|   |   | Selection_18 | 146027634.83 | cop | ge(a.id, 177045000)
|   |   |           |           |
|   |   |   | IndexScan_16 | 176388219.00 | cop | table:a, index:UPDATED_AT, range:(2020-03-03 23:05:30,+inf], k
eep order:false
|   |   |   | Selection_19 | 421.12    | cop | eq(a.pay_confirm_status, 1), eq(a.status, "CANCEL_SUCCESS"), i
n(a.type, "PURCHASE", "APPLY"), not(isnull(a.cancel_app_no))
|   |   |   | TableScan_17 | 146027634.83 | cop | table:bus_jijin_trade_record, keep order:false
|   |   |   |
|   |   | TableReader_31 | 6442.00    | root | data:TableScan_30
|   |   |           |
|   |   | TableScan_30 | 6442.00    | cop | table:b, range:[-inf,+inf], keep order:false
|   |   |
+-----+-----+-----+
```

hint updated_at 176388219

bus_jijin_trade_record

```
mysql> show stats_meta where table_name like 'bus_jijin_trade_record' and db_name like 'tod.s';
+-----+-----+-----+-----+
| Db_name | Table_name          | Update_time      | Modify_count | Row_count |
+-----+-----+-----+-----+
| tod.s | bus_jijin_trade_record | 2020-03-05 22:04:21 |      10652939 | 176381997 |
+-----+-----+-----+-----+

mysql> show stats_healthy where table_name like 'bus_jijin_trade_record' and db_name like 'tod.s';
+-----+-----+
| Db_name | Table_name          | Healthy |
+-----+-----+
| tod.s | bus_jijin_trade_record |      93 |
+-----+-----+
```

```
bus_jijin_trade_record 176381997 10652939(176381997-10652939)/176381997 *100=93
```

-

```
mysql> set tidb_build_stats_concurrency=10;
Query OK, 0 rows affected (0.00 sec)

#
mysql> analyze table todbs.bus_jijin_trade_record;
Query OK, 0 rows affected (3 min 48.74 sec)
```

- hint

```
mysql> explain select count(*)
   -> from todbs.bus_jijin_trade_record a, todbs.bus_jijin_info b
   -> where a.fund_code=b.fund_code and a.type in ('PURCHASE','APPLY')
   -> and a.status='CANCEL_SUCCESS' and a.pay_confirm_status = 1
   -> and a.cancel_app_no is not null and a.id >= 177045000
   -> and a.updated_at > date_sub(now(), interval 48 hour) ;;
+-----+-----+-----+
| id      | count    | task | operator info
+-----+-----+-----+
| StreamAgg_13          | 1.00     | root | funcs:count(1)
|           |          |       |
| └HashRightJoin_27      | 1.99     | root | inner join, inner:IndexLookUp_23, equal:[eq(a.fund_code, b.fund_code)]
|   | IndexLookUp_23        | 1.99     | root |
|   |           |          |       |
|   | Selection_21          | 626859.65 | cop  | ge(a.id, 177045000)
|   |           |          |       |
|   |   | IndexScan_19        | 757743.08 | cop  | table:a, index:UPDATED_AT, range:(2020-03-03 23:28:14,+inf], keep order:false
		Selection_22	1.99	cop	eq(a.pay_confirm_status, 1), eq(a.status, "CANCEL_SUCCESS"), in(a.type, "PURCHASE", "APPLY"), not(isnull(a.cancel_app_no))
		TableScan_20	626859.65	cop	table:bus_jijin_trade_record, keep order:false
		TableReader_37	6442.00	root	data:TableScan_36
		TableScan_36	6442.00	cop	table:b, range:[-inf,+inf], keep order:false
+-----+-----+-----+
9 rows in set (0.00 sec)
```

hint 757743

1.69s

```
mysql> select count(*)
   -> from todbs.bus_jijin_trade_record a, todbs.bus_jijin_info b
   -> where a.fund_code=b.fund_code and a.type in ('PURCHASE','APPLY')
   -> and a.status='CANCEL_SUCCESS' and a.pay_confirm_status = 1
   -> and a.cancel_app_no is not null and a.id >= 177045000
   -> and a.updated_at > date_sub(now(), interval 48 hour) ;
+-----+
| count(*) |
+-----+
|      712 |
+-----+
1 row in set (1.69 sec)
```

SQL

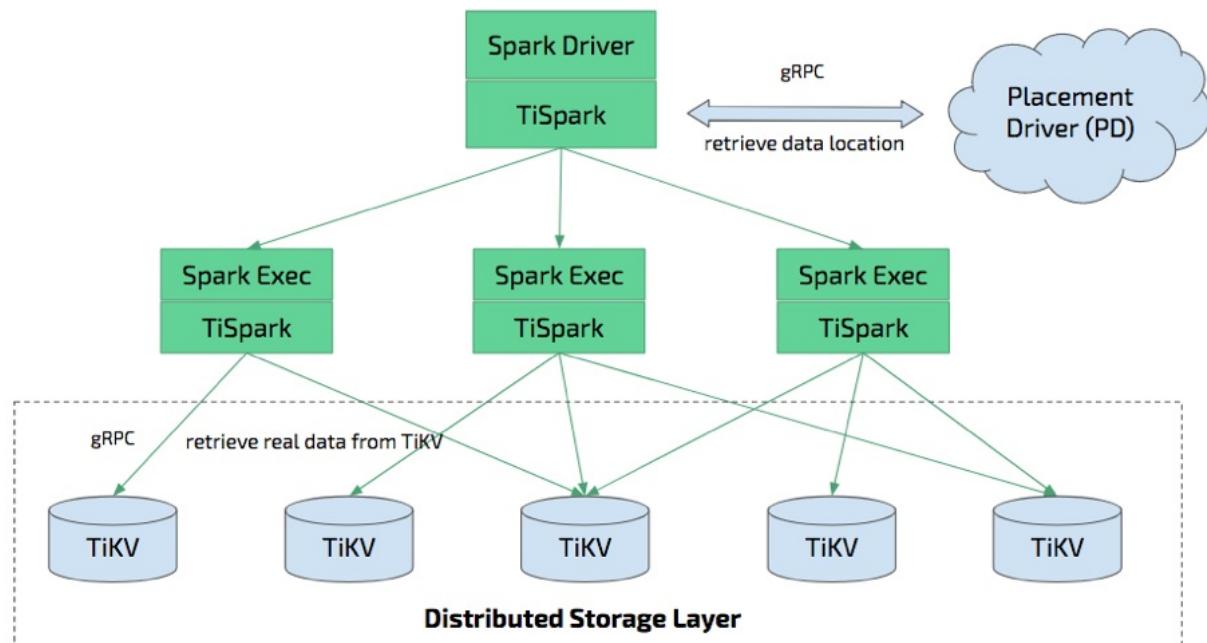
712

6.3 TiDB + TiSpark

TiSpark PingCAP OLAP Spark TiKV TiDB TiSpark Spark

6.3.1 TiSpark

TiSpark Spark SQL TiDB TiKV OLAP TiSpark



- TiSpark Spark Catalyst , Spark TiKV
- Spark SQL TiDB
- TiSpark+TiDB ETL
- TiSpark TiDB Spark TiSpark ETL TiKV

6.3.2

(1) TiSpark

TiSpark 2.1.8 Spark 2.3.0+ Spark 2.4.0+ TiDB-2.x TiDB-3.x

TiSpark GitHub

(2) Spark

TiDB TiSpark Spark Hadoop

(3) JDK

TiSpark JDK 1.8+ Scala 2.11 Spark2.0+ Scala

6.3.3 TiSpark

TiSpark YARN Mesos Standalone Spark Saprk Standalone Saprk Standalone TiSpark

(1) spark-env.sh

```
SPARK_EXECUTOR_MEMORY=10g
SPARK_EXECUTOR_CORES=5
SPARK_WORKER_MEMORY=40g
SPARK_WORKER_CORES=20
```

(2) spark-defaults.conf

```
spark.sql.extensions org.apache.spark.sql.TiExtensions
spark.tispark.pd.addresses 127.0.0.1:2379
```

PD : PD

(3) TiSpark

TiSpark Spark Spark --jars TiSpark :

```
sh spark-shell --jars $TISPARK_FOLDER/tispark-${name_with_version}.jar
```

TiSpark TiSpark jar Spark jars Spark

```
 ${SPARK_HOME}/jars
```

(4) TiSpark

Spark Master

```
cd ${SPARK_HOME}
./sbin/start-all.sh
```

master slave log log <http://spark-master-hostname:8080> Spark-Master web

6.3.4

TiSpark Spark-Shell Spark-Submit OLAP

(1) Spark-Shell

TiSpark 2.0 Spark-Shell Spark SQL TiDB

```
spark.sql("use test")
spark.sql("select count(*) from user").show
```

TiSpark 2.0

```
import org.apache.spark.sql.TiContext
val ti = new TiContext(spark, List("127.0.0.1:2379"))
ti.tidbMapDatabase("test")
```

Spark SQL 2.0 TiSpark

(2) Spark-Submit

Spark-Shell Spark-Submit TiSpark

Maven POM Spark TiSpark Spark Provided

```
<dependencies>
  <dependency>
    <groupId>com.pingcap.tispark</groupId>
    <artifactId>tispark-core</artifactId>
    <version>2.1.8-spark_${spark.version}</version>
    <scope>provided<scope>
  </dependency>

  <dependency>
    <groupId>org.apache.spark</groupId>
    <artifactId>spark-core_2.11</artifactId>
    <version>${spark.version}</version>
    <scope>provided<scope>
  </dependency>

  <dependency>
    <groupId>org.apache.spark</groupId>
    <artifactId>spark-sql_2.11</artifactId>
    <version>${spark.version}</version>
    <scope>provided<scope>
  </dependency>
</dependencies>
```

SparkSession Spark API ScalaPythonJava Java

```
SparkSession sc = SparkSession
  .builder()
  .appName("TiSpark example")
  .master("local")
  .config("spark.sql.extensions", "org.apache.spark.sql.TiExtensions")
  .config("spark.tispark.pd.addresses", "127.0.0.1:2379")
  .getOrCreate();
```

TiSpark Spark master TiSpark

TiSpark TiDB Spark JDBC Java

```

sc.read().schema(getStructType())
    .option("delimiter",true)
    .option("header",true)
    .csv(filePath)
    .withColumn("input_time", functions.current_timestamp())
    .repartition(100)
    .write().mode(SaveMode.Append)
    .jdbc(url,tableName,connProperties);

public StructType getStructType(){
    List<StructField> fields=new ArrayList<>();
    StructField transferAccount = DataTypes.createStructField("transferAccount",DataTypes.StringType,false);
    StructField receiveAccount = DataTypes.createStructField("receiveAccount",DataTypes.StringType,false);
    StructField amount = DataTypes.createStructField("amount",DataTypes.createDecimalType(10,2),false);
    fields.add(transferAccount);
    fields.add(receiveAccount);
    fields.add(amount);
    return DataTypes.createStructType(fields);
}

```

TiDB

1. Spark repartition Spark
2. TiDB rewriteBatchedStatements True JDBCOptions.JDBC_BATCH_INSERT_SIZE 150
3. isolationLevel NONE

Java

```

sc.sqlContext()
    .udf()
    .register("convert",newConvertUDF(),DataTypes.createDecimalType(10,2));
sc.sql("select transferAccount,receiveAccount,convert(amount) from tableName where receiveAccount='0001'");

```

0001 convert UDF Job Spark

SQL Spark SQL UDF UDAF SQL TiSpark SQL

delete

Spark-Submit Spark

```

cd ${SPARK_HOME}
./bin/spark-submit \
--class Analyze \
--master spark://127.0.0.1:7077 \
/home/tispark/TiSparkExample.jar

```

Spark-Submit Spark

6.3.5

TiDB TiSpark TiSpark TiSpark Spark Spark TiSpark TiSpark Spark

TiSpark

Spark-Submit Oozie Azkaban

Spark API Spark Spark API

6.4

DELETE Range DDL

6.4.1 TiDB

TiDB 2.1 3.0 4.0 Bug

6.4.1.1

Range Hash MySQL List Key

Range Range VALUES LESS THAN Range

Hash Hash Hash CREATE TABLE PARTITION BY HASH (expr) PARTITIONS num expr num

6.4.1.2

1.

Range

```
CREATE TABLE employees_attendance (
    id INT NOT NULL AUTO_INCREMENT,
    uid INT NOT NULL,
    name VARCHAR(25) NOT NULL,
    login_date date NOT NULL,
    create_time timestamp NOT NULL COMMENT '',
    type tinyint NOT NULL DEFAULT '0' COMMENT '0:1:',
    PRIMARY KEY (`id`,`login_date`),
    UNIQUE KEY `idx_attendance` (`uid`,`login_date`,`type`)
)
PARTITION BY RANGE COLUMNS(login_date) (
    PARTITION p20200306 VALUES LESS THAN ('20200307'),
    PARTITION p20200307 VALUES LESS THAN ('20200308'),
    PARTITION pmax VALUES LESS THAN MAXVALUE
);
```

login_date

```
> ERROR 1503 (HY000): A (PRIMARY KEY/UNIQUE INDEX) must include all columns in the table's partitioning function
```

2.

TiDB Hash Range

Load Data load

```
load local data infile "xxx" into t partition (p1);
```

6.4.1.3 TiDB 4.0

TiDB 4.0 Bug

- Bug
- INFORMATION_SCHEMA.PARTITION
- - Range Hash

- Range Partition Expression Expression int Partition Constant Propagate Constant Propagate Pattern Expr Range Partition
- 4.0 Hash Partition Range Partition Hash Partition Range Partition Hash Partition PointGet Hash Partition
partition by hash(id) id PointGet Hash Partition Hash Partition

6.4.2 TiDB

0

- Range TiDB 4.0 DELETE DDL SST
- Hash Hash
- SQL
-
- - /TiDB Range 3
 - Range Log
 - BIGINT UID Hash UID Hash Hash

6.4.3 TiDB

- DELETE TRUNCATE

```
ALTER TABLE employees_attendance TRUNCATE PARTITION p20200306;
```

- TiDB SQL SQL SQL//SQL

```
ALTER TABLE employees_attendance ANALYZE PARTITION p20200306;
```

- DBA TiDB UNIX_TIMESTAMP

```
ALTER TABLE ADD PARTITION p20200306 VALUES LESS THAN (UNIX_TIMESTAMP('2020-03-07'))
```

TiDB

- Range MAXVALUE Data Maintenance

- Hash Hash Range Hash

- (TiDB) UPDATE DROP

6.4.4 TiDB

1. TiDB 4.0 Fixed Bug GA

- - Table Partition

```
create table t (a int) partition by hash(a) partitions 4;
begin
insert into t values (0),(1);
select * from t where a>0;
+---+
| a |
+---+
| 1 |
| 0 | -- Bug: filter is a>0
+---+
```

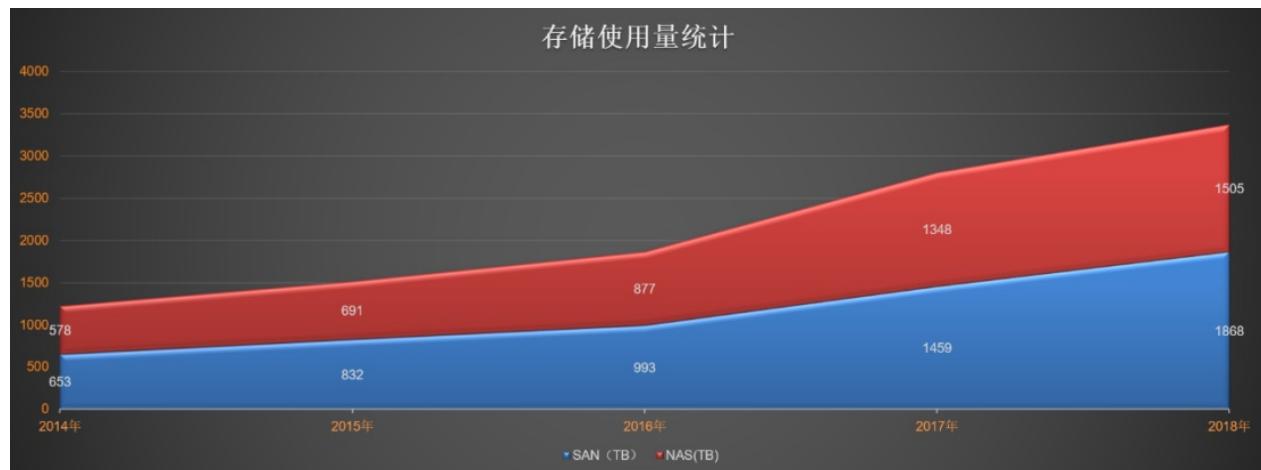
- INSERT ... ON DUPLICATE Table Partition

```
create table t1 (a int,b int,primary key(a,b))
partition by range(a)
(partition p0 values less than (100),
 partition p1 values less than (1000)
 );
insert into t1 set a=1,b=1;
insert into t1 set a=1,b=1 on duplicate key update a=1,b=1;
ERROR 1105:can not be duplicated row, due to old row not found. handle 1 not found
```

- Bug
- LIMIT TiDB Server OOM Kill
- OR Partition Partition Column CheckScalarFunction OR Partition Column TiDB

6.5 TiDB

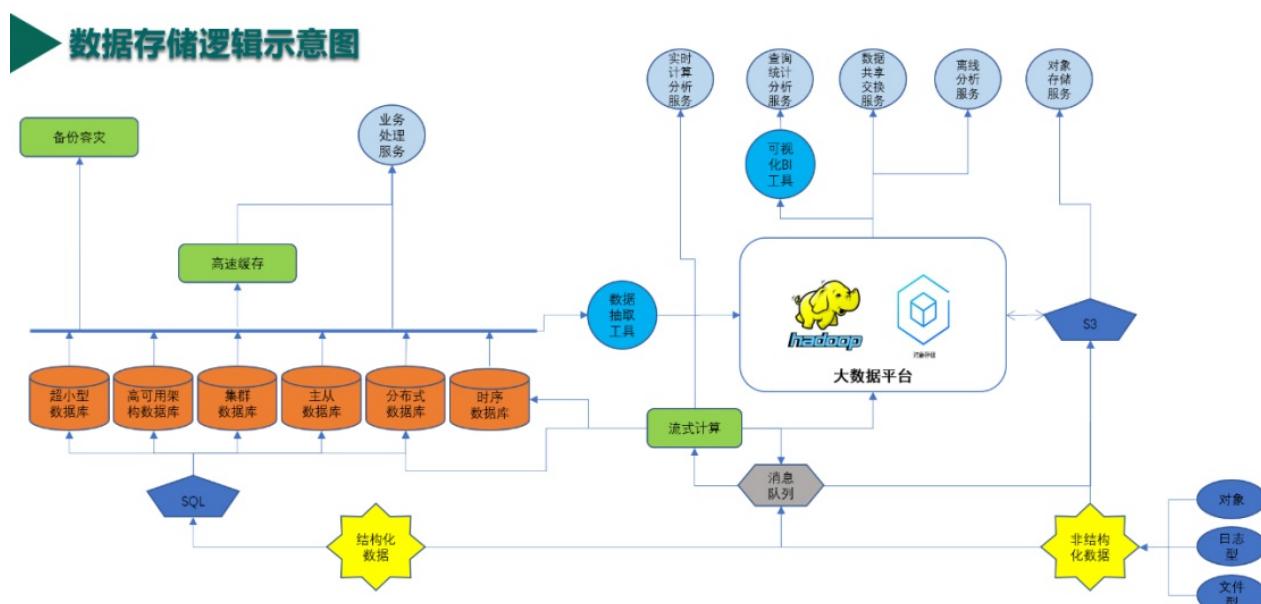
6.5.1



2014 SAN NAS 1PB 2018 3 2017 180 100

180 OracleDB2MySQLGBASE 8aMongodbHIVEHBASE 20

6.5.2



- • •

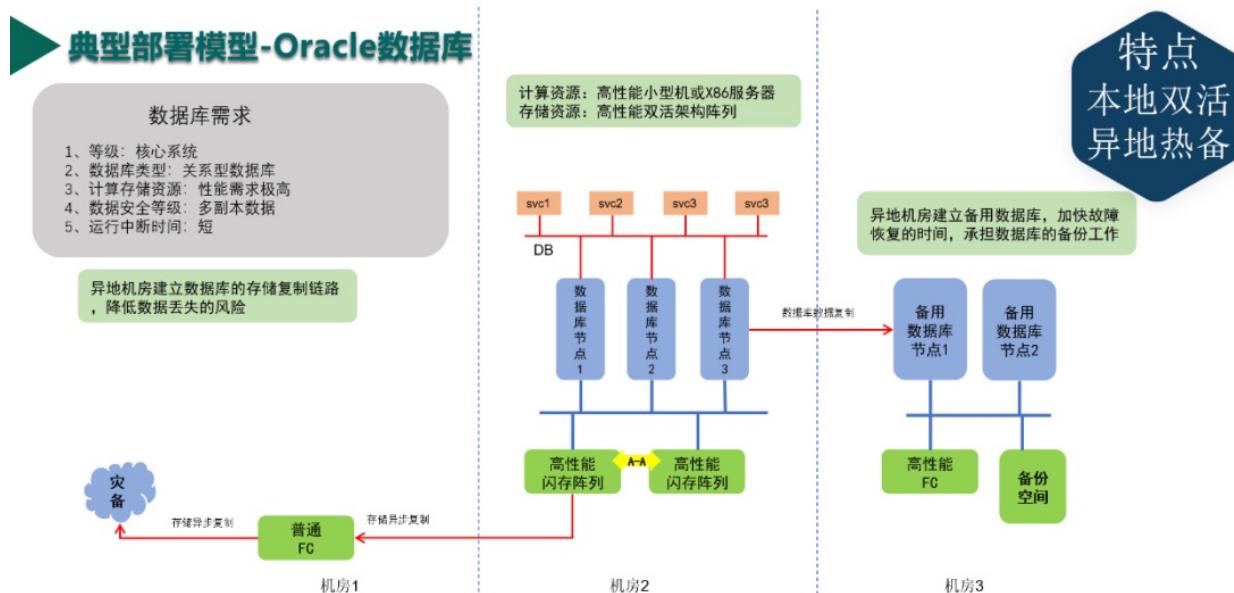
6.5.3

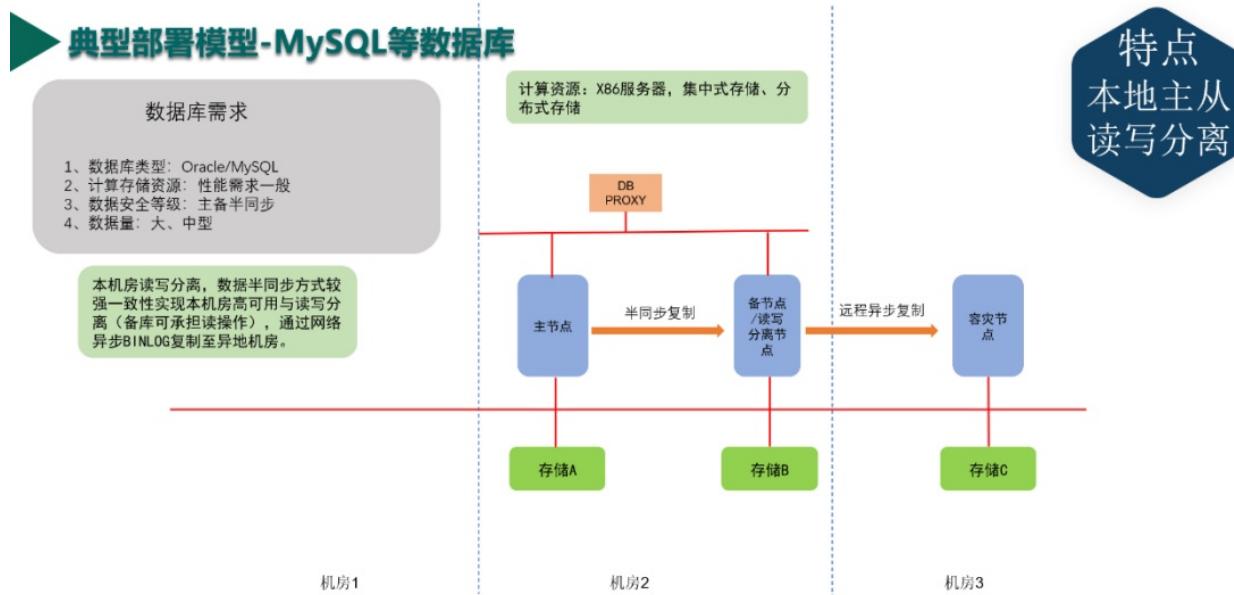
OracleMySQL Oracle MySQL

- (1) ACID ACID
 - (2)
 - (3)
 - (4) OLTP OLAP PB OLTP OLAP OLTP/OLAP
 - (5) MySQL

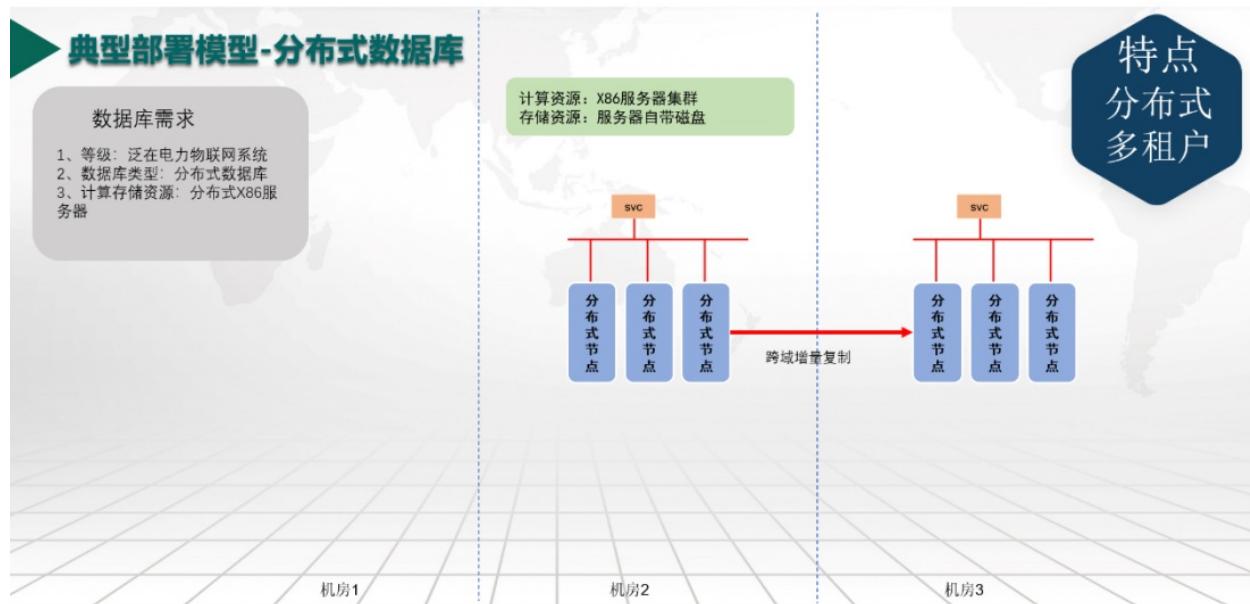
TiDB

- (1)





MySQL Oracle MySQL



TiDB

6.5.4

TiDB

6.6 TiDB HBaseESDruid

TiDB TiDB

TiDB HbaseESDruid

6.6.1. ES TiDB

ES TiDB

Kettle tableinputElasticsearch Query

ES 'bmsshakehandsinfo_202003' TiDB

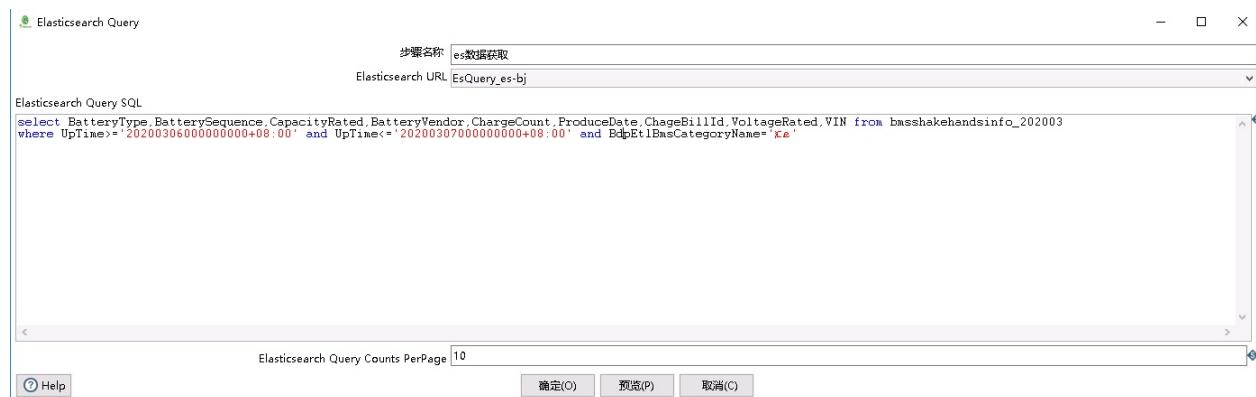
'Temp_ETL_Handshake'

1. ES

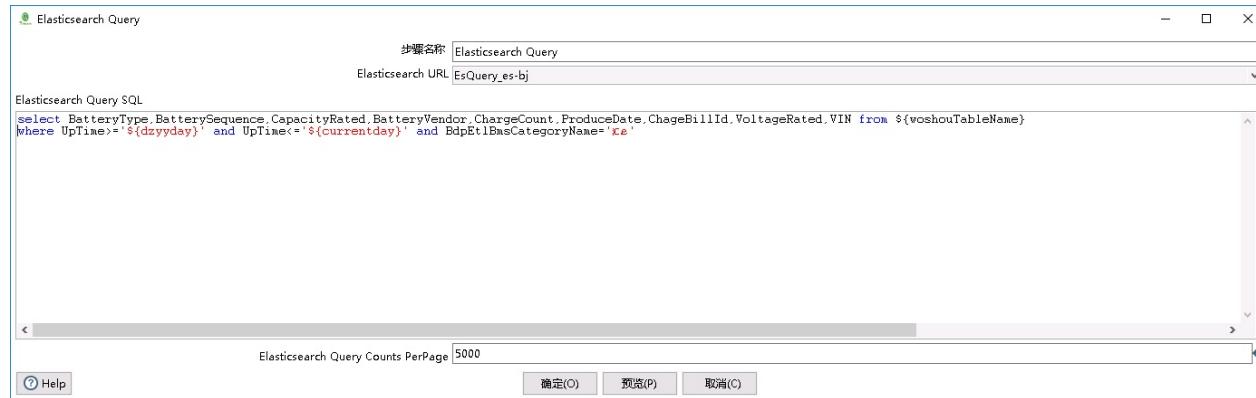
1 ES SQL ES

SQL

Kettle 2



1



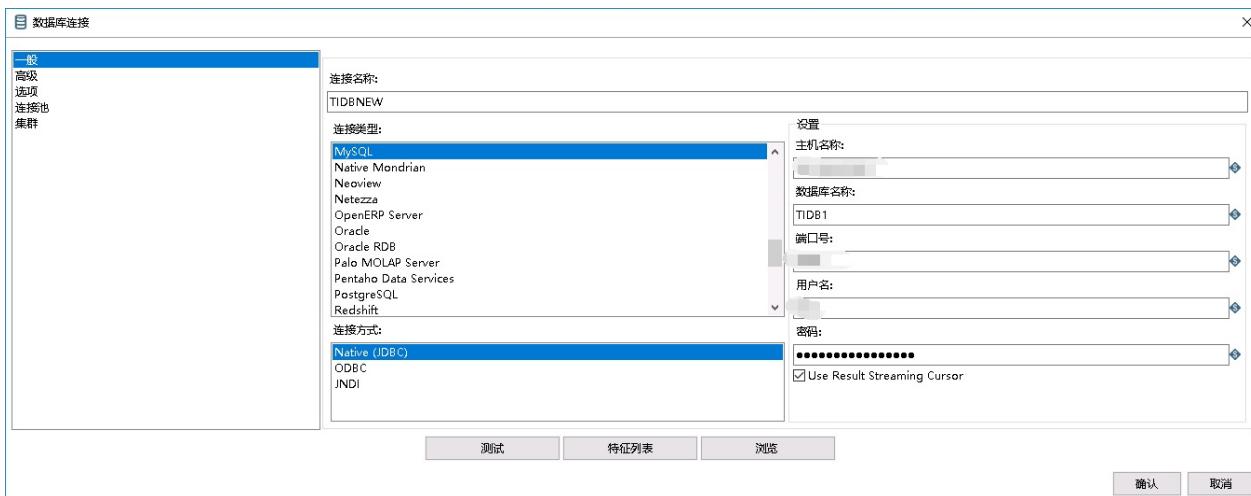
2

2. TiDB

TiDB Kettle tableout

2.1

Kettle TiDB Kettle MySQL 3



3

2.2 TiDB

TiDB 4

- 1. TiDB 60-100.
 2. TiDB
 - - 1.
 2. Kettle /, TiDB
 - on duplicate



4

3

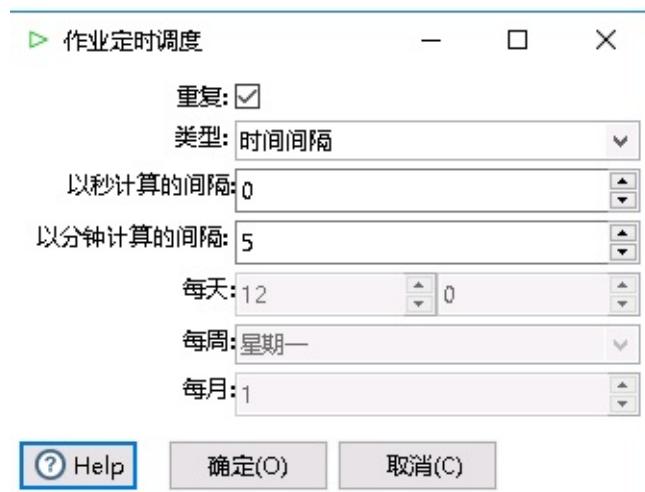
Kettle ES TiDB

5 60 60 TiDB



5

Kettle 6



6

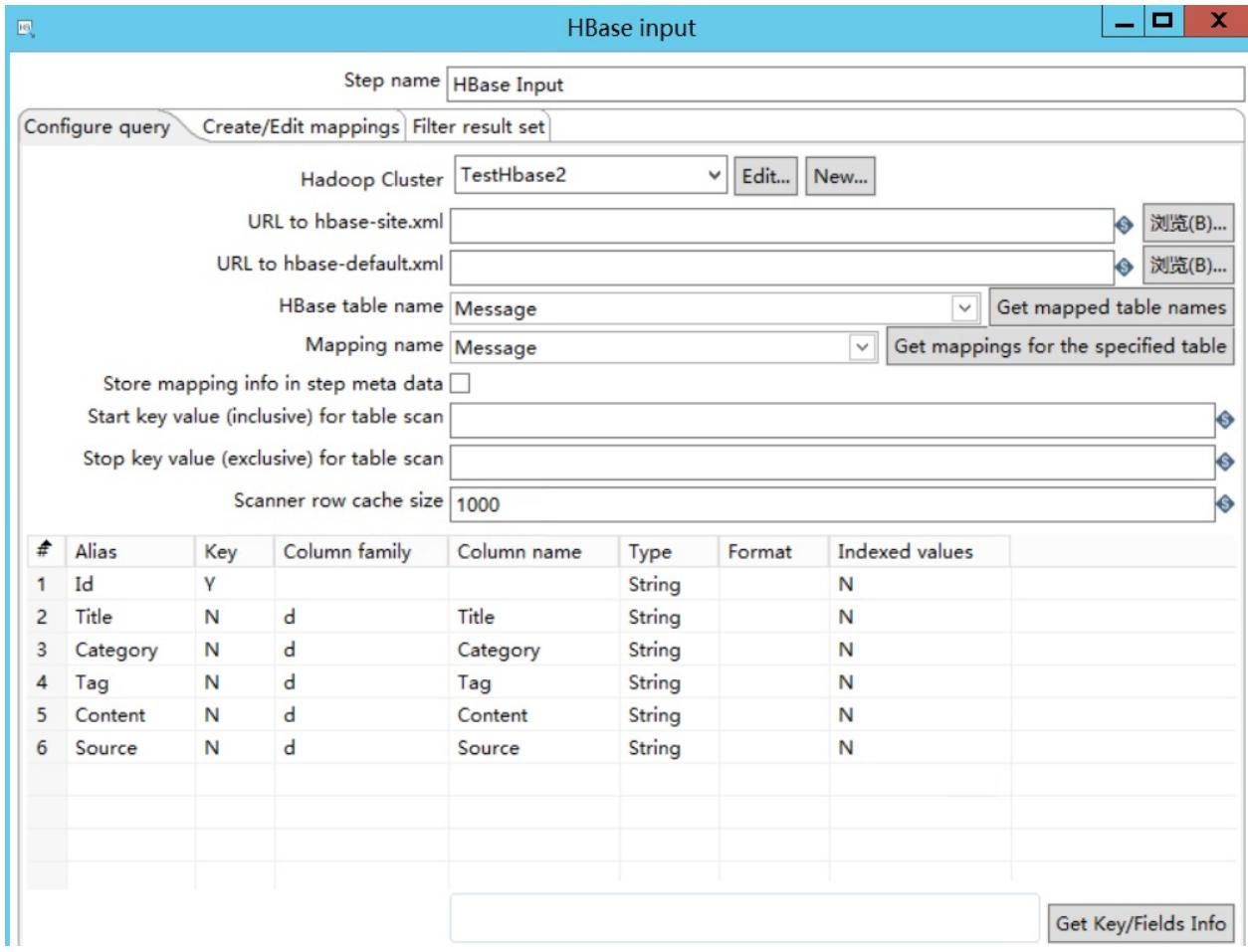
Kettle ES TiDB

6.6.2 Hbase TiDB

Hbase TiDB Kettle

HbaseInputTableOutPut TiDB TableOutPut ES TiDB

Kettle HBase input , 7, Hbase mapping



7



8

6.6.3 Druid TiDB

Druid TiDB Druid

TableOutPut TiDB TableOutPut ES TiDB

Druid Druid SQL



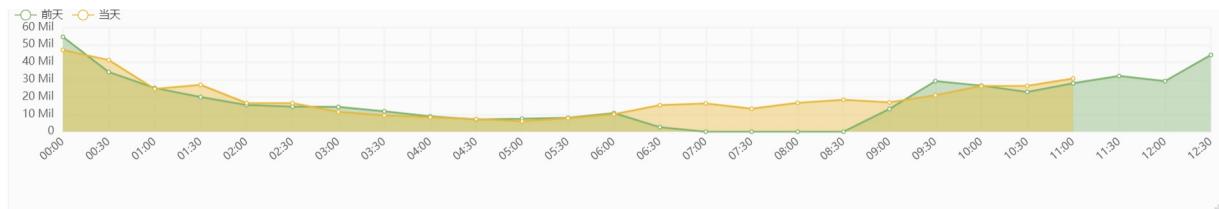
9

6.7 TiDB

TiDB HTAP TiDB Grafana Saiku .

6.7.1. TiDB Grafana

Grafana Grafana MySQL TiDB MySQL TiDB



TiDB ECharts

General Metrics Time range Time cycle Style Series Options Events Calc Cumulative SerieType Serie Detail Echarts Option Drilldown ×

A `select datagetttime`
GROUP BY datagetttime
order by time_sec asc

Format as Time series Show Help ▾

1

1 Grafana SQL Grafana

1. Grafana TiDB
2. Grafana

6.7.2 TiDB Saiku

Saiku Web OLAP

TiDB +Saiku

BI TiDB

Saiku
TiDB :TiDB Saiku Saiku SQL TiDB

6.7.2.1. Schema

Saiku Schema TiDB Schema Schema

```

▼<Schema name="kms_dt_tday_bizdate">
  <Cube name="001充电日期-年 (字段:RYear)" caption="001充电日期-年" cache="true" enabled="true">
    <Table name="kms_dt" alias="" />
    <!-- 001充电日期-年 (字段:RYear) -->
    ▼<Dimension type="StandardDimension" name="001充电日期-年" foreignKey="bizdate">
      ▼<Hierarchy name="" hasAll="true" primaryKey="tday" allLevelName="全部" allMemberCaption="全部">
        <Table name="kms_dt" />
        <Level name="充电日期-年" column="tyear" type="String" uniqueMembers="true" levelType="Regular" />
      </Hierarchy>
    </Dimension>
    <!-- 002充电日期-月 (字段:RYearMonth) -->
    ▼<Dimension type="StandardDimension" name="002充电日期-月" foreignKey="bizdate">
      ▼<Hierarchy name="" hasAll="true" primaryKey="tday" allLevelName="全部" allMemberCaption="全部">
        <Table name="kms_dt" />
        <Level name="充电日期-月" column="ymonth" type="String" uniqueMembers="true" levelType="Regular" />
      </Hierarchy>
    </Dimension>
    <!-- 003充电日期-周 (字段:周描述) 先把周注释掉, 重新修改去除UPPER()逻辑后, 再显示该维度。(通过修改sensitive配置, 已成功去除Upper)
    -->
    ▼<Dimension type="StandardDimension" name="003充电日期-周" foreignKey="bizdate">
      ▼<Hierarchy name="" hasAll="true" primaryKey="tday" allLevelName="全部" allMemberCaption="全部">
        <Table name="kms_dt" />
        <Level name="周" column="seqdesc" type="String" uniqueMembers="true" levelType="Regular" />
      </Hierarchy>
    </Dimension>
    <!-- 004充电日期-日 (字段:RBizDate) -->
    ▼<Dimension type="StandardDimension" name="004充电日期-日" foreignKey="bizdate">
      ▼<Hierarchy name="" hasAll="true" primaryKey="tday" allLevelName="全部" allMemberCaption="全部">
        <Table name="kms_dt" />
        <Level name="充电日期-日" column="tday" type="String" uniqueMembers="true" levelType="Regular" />
      </Hierarchy>
    </Dimension>
    <!-- 005启动时段 (字段:DHour) -->
    ▼<Dimension type="StandardDimension" name="005启动时段">
      ▼<Hierarchy name="" hasAll="true" allLevelName="全部" allMemberCaption="全部">
        <Level name="启动时段" column="dhour" type="String" uniqueMembers="true" levelType="Regular" />
      </Hierarchy>
    </Dimension>
    <!-- 011省 (字段:省名称) -->
  
```

2

Schema , kms_dt tday bizdate .

```

:
Schema , TiDB , Schema Saiku .

```

6.7.2.2 TiDB

Saiku 1-3	URL: TiDB	Schema: schema	jdbcdriver	TiDB MySQL MyS
Driver:com.mysql.jdbc.Driver	TiDB			

Create Data Source

Name: [REDACTED]

Connection Type: Mondrian

URL: jdbc:mysql://[REDACTED].4000/tidb1

Schema: /datasources/[REDACTED].xml

Jdbc Driver: com.mysql.jdbc.Driver

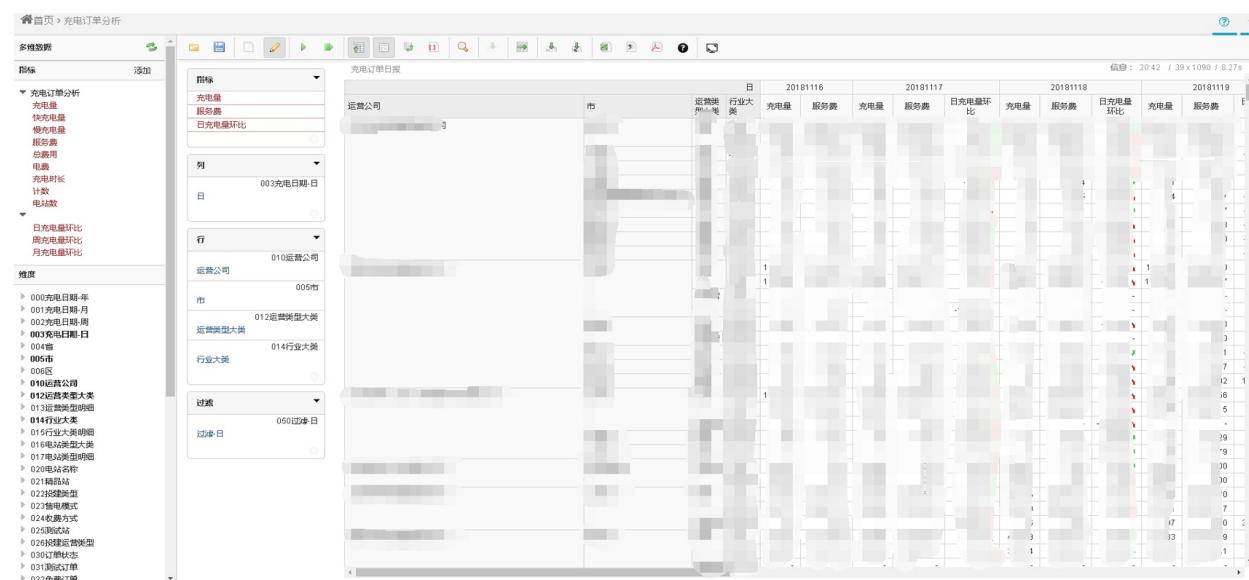
Username: [REDACTED]

Password: [REDACTED]

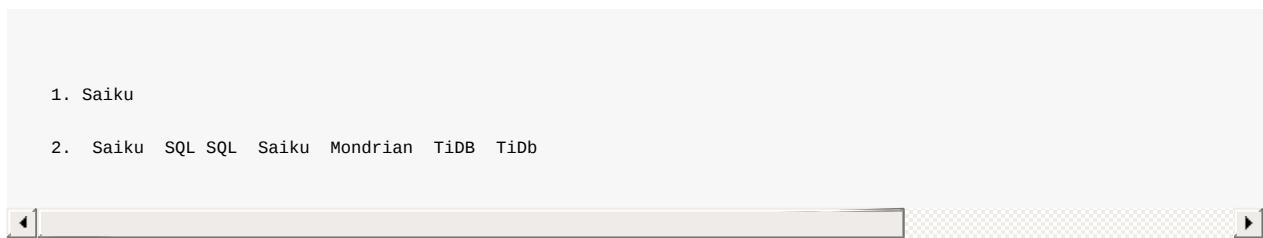
3

2.3

Saiku ,,



4



7

TiDB 7.1 oncall-map 7.27.37.4 3

- 7.1 Oncall
- 7.2
- 7.3 TiKV is busy
- 7.4 TiDB OOM
- 7.5 TiKV

7.1 Oncall

1. TiDB

TiDB TiDB FAQ askTUG PingCAP

PingCAP diagnose-map 7

2.

<https://github.com/pingcap/tidb-map/blob/master/maps/diagnose-map.png> TiDB png markdown
<https://github.com/pingcap/tidb-map/blob/master/maps/diagnose-map.md>

3.

TiDB TiKV OOM PD latency region is unavailable 1.1.1 4

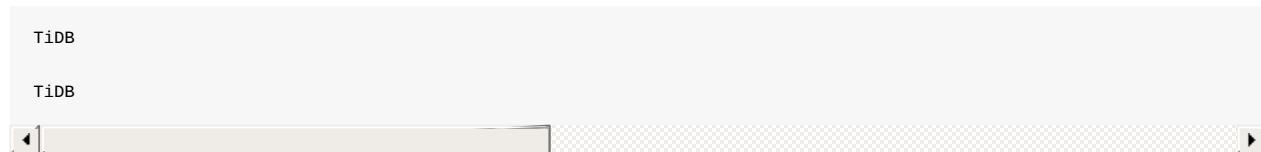
ONCALL-958 <https://github.com/pingcap/tidb-map/tree/master/maps/diagnose-case-study>

4.

TiDB askTUG

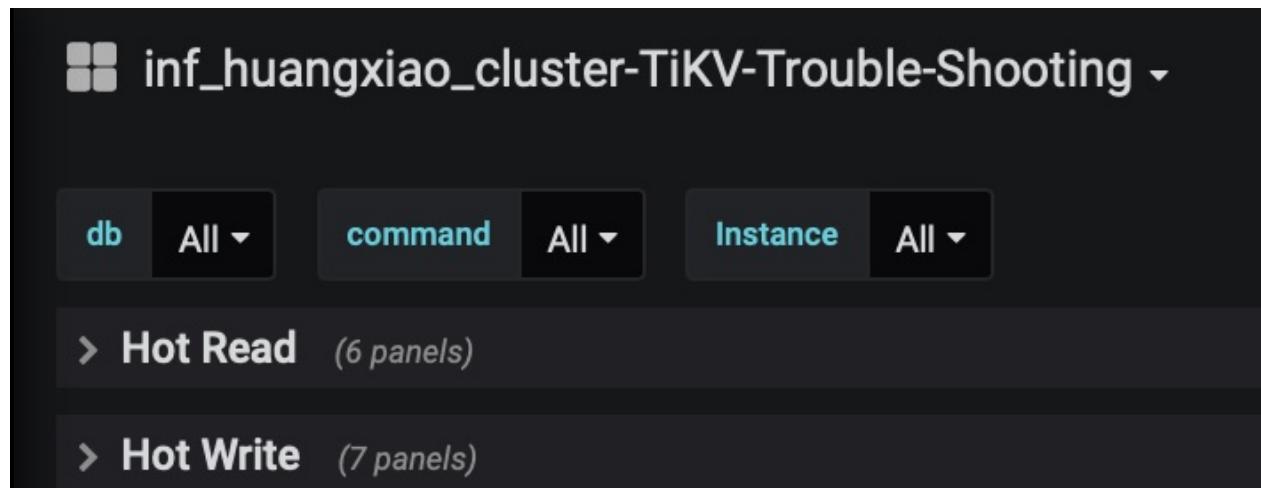
TiDB <https://github.com/pingcap/tidb-map> issue

7.2



7.2.1

TiDB 3.0 Grafana TiKV-Trouble-Shooting Dashboard Hot Read Hot Write



1. Hot Read

- CPU TiKV CPU
- QPS TiKV QPS
- Storage ReadPool CPUReadpool CPU
- Coprocessor CPUCoprocessor CPU
- gRPC poll CPUgRPC CPU 80%
- IO utilization TiKV IO

2. Hot Write

- CPU TiKV CPU
- QPS TiKV QPS
- gRPC poll CPUgRPC CPU 80%
- IO utilization TiKV IO
- Raft store CPURaftstore CPU 80%
- Async apply CPUasync apply CPU 90%
- Scheduler CPUScheduler CPU 80%

Hot Read Hot Write TiKV

7.2.2 /

TiKV TiDB 3.0 SQL information_schema.TIDB_HOT_REGIONS /

```
-- TYPE read write
SQL> select * from information_schema.TIDB_HOT_REGIONS where type = 'read'\G
***** 1. row *****
    TABLE_ID: 21
    INDEX_ID: NULL
    DB_NAME: mysql
    TABLE_NAME: stats_histograms
    INDEX_NAME: NULL
    REGION_ID: 44
    TYPE: read
MAX_HOT_DEGREE: 17
REGION_COUNT: 0
FLOW_BYTES: 248548
1 row in set (0.02 sec)
```

pd-ctl Region

```
$ pd-ctl -u http://{pd}:2379 -d region topread [limit]
$ pd-ctl -u http://{pd}:2379 -d region topwrite [limit]
```

region_id

```
$ pd-ctl -u http://{pd}:2379 -d region topread 1
{
  "as_peer": null,
  "as_leader": {
    "1": {
      "total_flow_bytes": 248535,
      "regions_count": 1,
      "statistics": [
        {
          "region_id": 44,
          "flow_bytes": 248548,
          "hot_degree": 15, -- 1 degree+1
          "last_update_time": "2020-03-07T16:14:51.186168306+08:00",
          "AntiCount": 1,
          "Version": 11,
          "Stats": null
        }
      ]
    }
  }
}
```

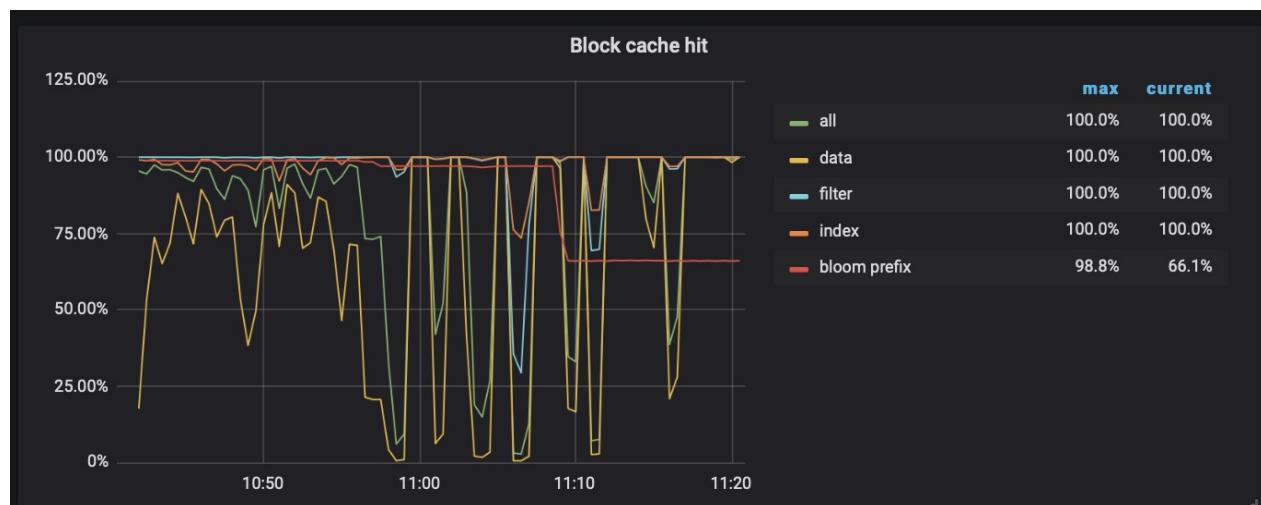
Region ID

```
$ curl http://{TiDBIP}:10080/regions/{RegionId}
{
  "region_id": 44,
  "start_key": "dIAAAAAAAAAB",
  "end_key": "dIAAAAAAAAAX",
  "frames": [
    {
      "db_name": "mysql",
      "table_name": "stats_histograms",
      "table_id": 21,
      "is_record": false,
      "index_name": "tbl",
      "index_id": 1
    },
    {
      "db_name": "mysql",
      "table_name": "stats_histograms",
      "table_id": 21,
      "is_record": true
    }
  ]
}
```

SQL TiDB 4.0 Key Visualizer ""

7.2.3

TiDB TiKV BlockCache TiKV-Details Dashboard RocksDB KV Block cache hit SQL



1. BlockCache BlockCache SQL SQL count(*) SQL "SQL" SQL SQL_NO_CACHE Hints
2. Region Region Region TiKV hash SHOW TABLE REGIONS hash 4

```

SQL> CREATE TABLE t1(
-> id INT NOT NULL,
-> name VARCHAR(30),
-> hired DATE NOT NULL DEFAULT '1970-01-01',
-> separated DATE NOT NULL DEFAULT '9999-12-31',
-> store_id INT
-> )
-> PARTITION BY HASH(store_id)
-> PARTITIONS 4;
Query OK, 0 rows affected (1.29 sec)
SQL> show table t1 regions\G
***** 1. row *****
REGION_ID: 840
START_KEY: t_194_
END_KEY: t_195_
LEADER_ID: 843
LEADER_STORE_ID: 9
PEERS: 841, 842, 843
SCATTERING: 0
WRITTEN_BYTES: 35
READ_BYTES: 0
APPROXIMATE_SIZE(MB): 1
APPROXIMATE_KEYS: 0
***** 2. row *****
REGION_ID: 844
START_KEY: t_195_
END_KEY: t_196_
LEADER_ID: 847
LEADER_STORE_ID: 9
PEERS: 845, 846, 847
SCATTERING: 0
WRITTEN_BYTES: 35
READ_BYTES: 0
APPROXIMATE_SIZE(MB): 1
APPROXIMATE_KEYS: 0
***** 3. row *****
REGION_ID: 848
START_KEY: t_196_
END_KEY: t_197_
LEADER_ID: 851
LEADER_STORE_ID: 9
PEERS: 849, 850, 851
SCATTERING: 0
WRITTEN_BYTES: 35
READ_BYTES: 0
APPROXIMATE_SIZE(MB): 1
APPROXIMATE_KEYS: 0
***** 4. row *****
REGION_ID: 3
START_KEY: t_197_
END_KEY:
LEADER_ID: 475
LEADER_STORE_ID: 9
PEERS: 458, 471, 475
SCATTERING: 0
WRITTEN_BYTES: 217
READ_BYTES: 0
APPROXIMATE_SIZE(MB): 1
APPROXIMATE_KEYS: 0
4 rows in set (0.05 sec)

```

hash TiDB 3.1 Follower Read TiDB 4.0PD Load Base Splitting Region Region QPS Region""

7.2.4

TiDB

1. MySQL TiDB
2. / int /

- 3.
- 4.
- 5.

7.2.4.1

MySQL ID TiDB Key Region Region TiDB Region TiKV

MySQL TiDB SHARD_ROW_ID_BITS rowid Region RPC CPU

- SHARD_ROW_ID_BITS = 4 16
- SHARD_ROW_ID_BITS = 6 64
- SHARD_ROW_ID_BITS = 0 1

```
CREATE TABLE t (c int) SHARD_ROW_ID_BITS = 4;
ALTER TABLE t SHARD_ROW_ID_BITS = 4;
```

SHARD_ROW_ID_BITS TiDB 4.0 auto_random

7.2.4.2

int TiDB _tidb_rowid ID SHARD_ROW_ID_BITS _tidb_rowid INSERT Region

_tidb_rowid SHARD_ROW_ID_BITS PRE_SPLIT_REGIONS SHARD_ROW_ID_BITS _tidb_rowid ID
pre_split_regions Split region

pre_split_regions shard_row_id_bits

```
create table t2 (a int, b int) shard_row_id_bits = 4 pre_split_regions=2;
```

- SHARD_ROW_ID_BITS = 4 tidb_rowid 16 16=2^4
- pre_split_regions=2 4 (2^2) Region

7.2.4.2

4.0 Region Region

```
$ pd-ctl -u http://{PDIP}:2379 -i
// Region 1 Region
> operator add split-region {hotRegionId} --policy=approximate
// Region 1 Region
> operator add split-region {hotRegionId}1 --policy=scan
```

Region split-merge-interval split-merge-interval Region split merge split Region merge pd-ctl
1h

hash TiDB 4.0 Region

7.2.4.3

ID Region

7.2.4.4

TiDB “”TiDB

7.2.5 PD

PD PD Region Peer Leader Leader Region Leader

7.2.5.1

hot-region-scheduler-limit Store Region Region Region Peer Leader Leader Region Leader

7.2.5.2 PD

TiDB 3.0 region-schedule-limit hot-region-schedule-limit hot-region-schedule-limit limit hot-region-cache-hits-threshold PD

```
// hot-region-schedule-limit
config set hot-region-schedule-limit 8
// limit
config set merge-schedule-limit 2
config set region-schedule-limit 2
```

7.2.5.3 Region

scatter-range-scheduler table Region scatter-range-scheduler

```
curl -X POST http://{TiDBIP}:10080/tables/{db}/{table}/scatter
```

scatter-range-scheduler

```
curl -X POST http://{TiDBIP}:10080/tables/{db}/{table}/stop-scatter
```

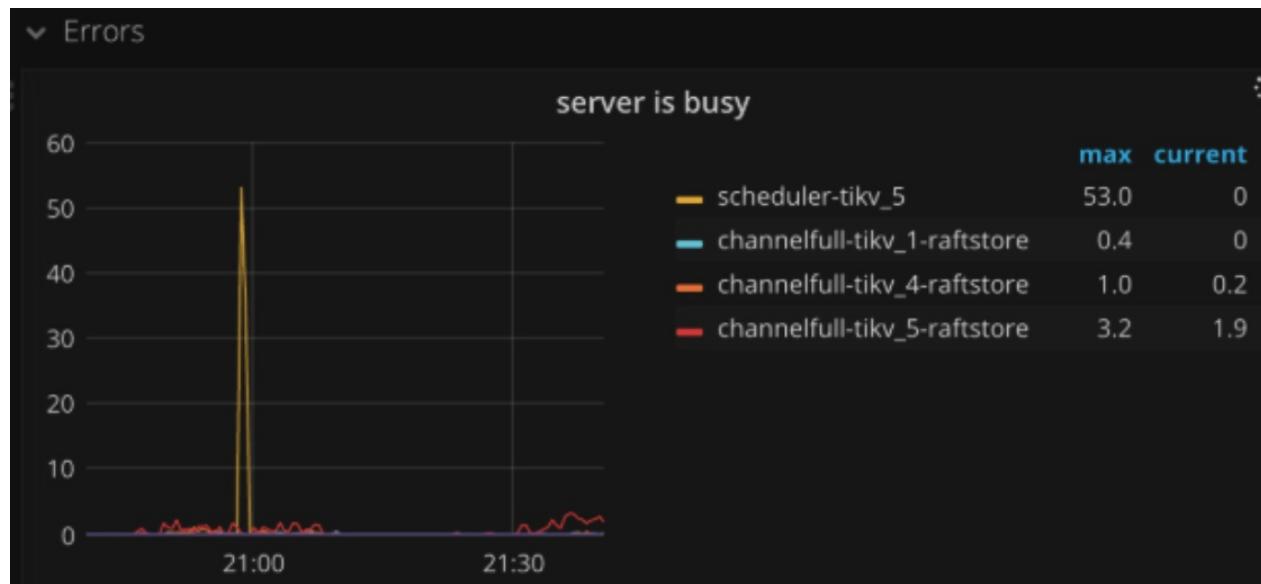
7.2.6

2 A B A B A 1.5K B AB 2

TiDB 2.1 TiKV raft A A bigint 1 0-9 A 10 A

7.3 TiKV is busy

TiDB TiKV Trouble - Shooting - Server Is Busy



TiDB

7.3.1 Server is Busy

Server is Busy tikv-server TiDB

1.

1. TiDB - Query Summary - Duration
2. TIKV TiKV - Server / Thread CPU / Error
3. TiDB server is busy tidb.log Server is Busy
- 4.

2.

1. 5ms -> 5s
2. TimeoutException
3. DB

Server is Busy Server is busy

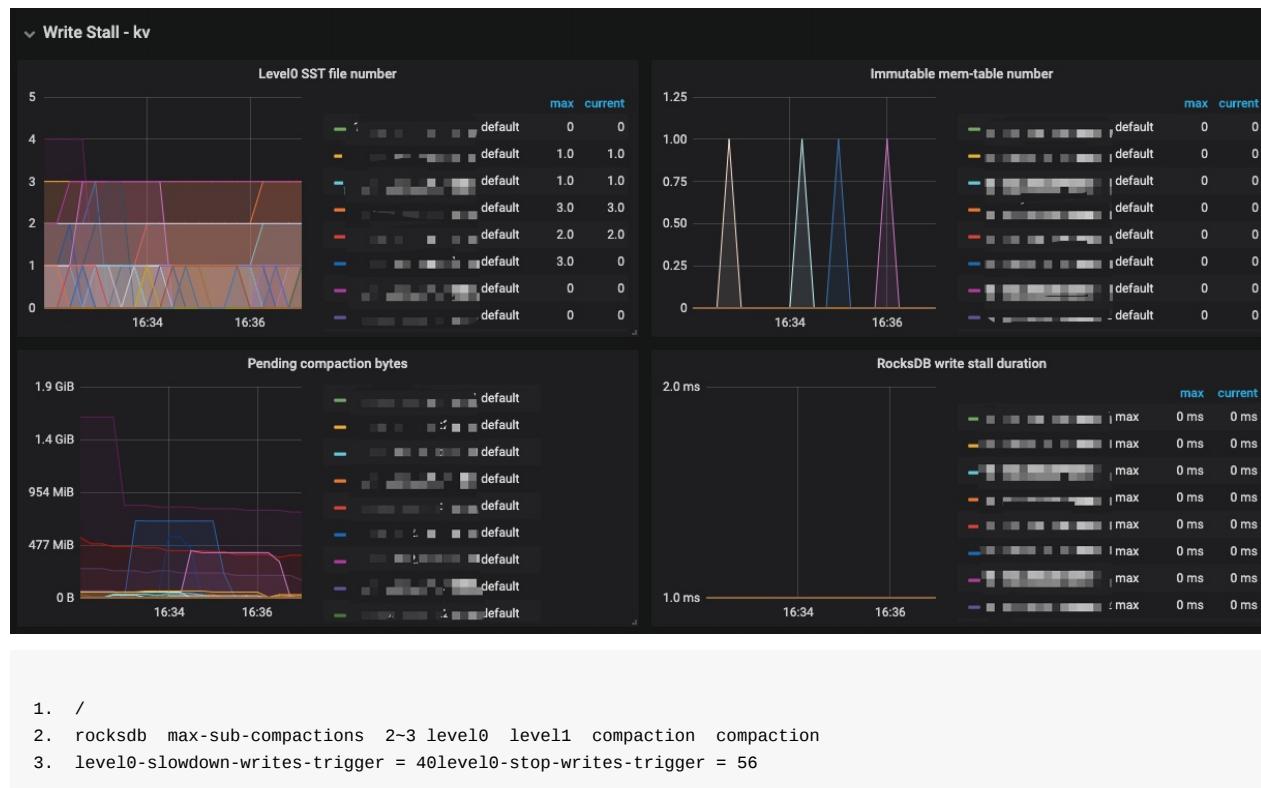
7.3.2 Server is Busy

1. (write stall)

TiKV 2 RocksDB , RocksDB LSM TreeLSM Tree RocksDB RocksDB

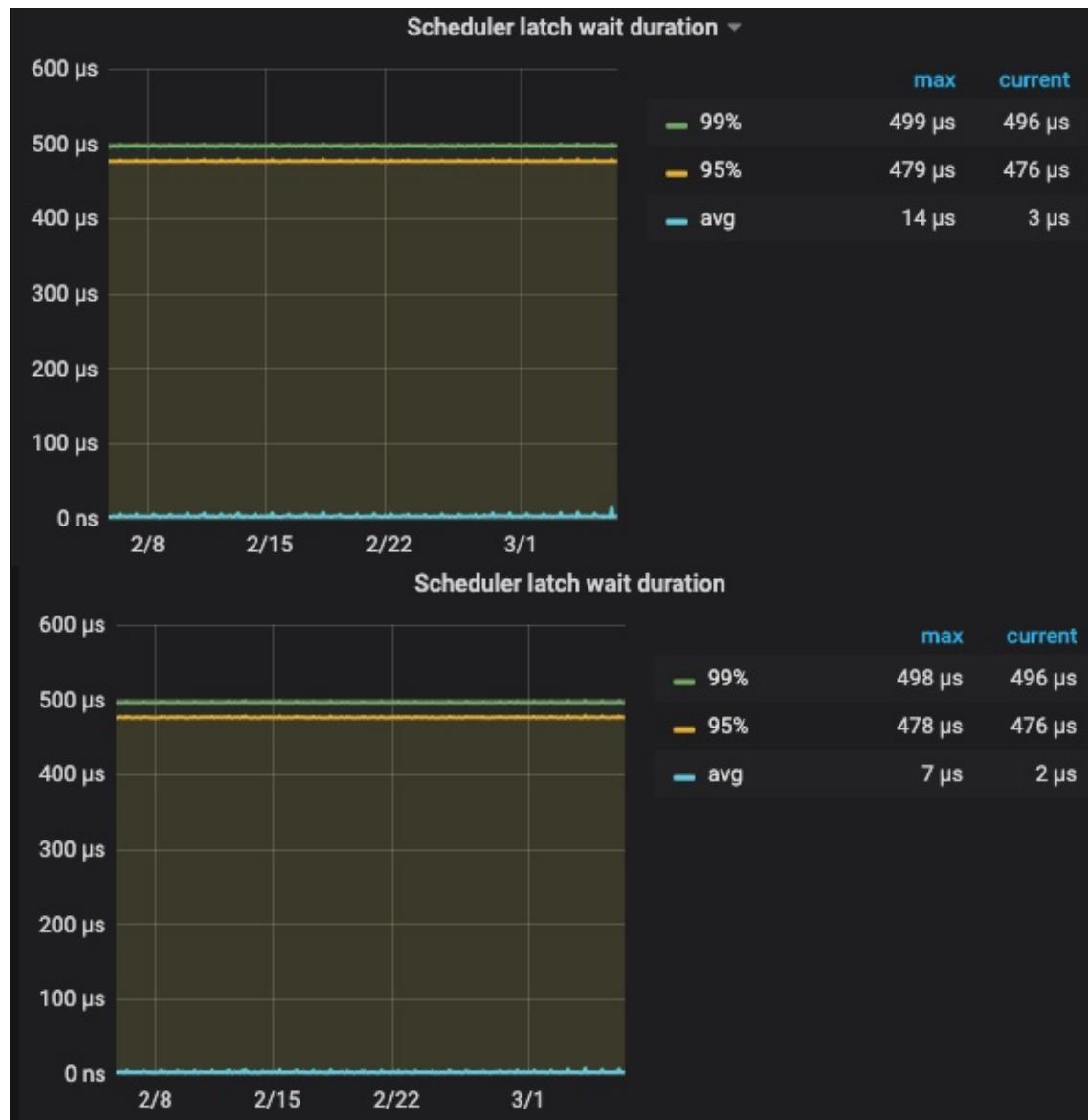
- level0 sst stall
- pending compaction bytes stall IO

- memtable stall



2. scheduler too busy

- latchlatch latch wait duration , TIKV - scheduler prewrite|commit latch wait duration
scheduler prewrite - latch wait duration | scheduler commit latch wait duration



- TiKV [storage] scheduler-pending-write-threshold = "100MB"

- 1.
- 2.
- 3.

3.

Server Is Busy

(1) tikv.yml

```
3.0 2      readpool.storage readpool.coprocessor TiDB
```

```
// , cpu core * 80%, 1
high-concurrency
// , cpu core * 80%, 1
normal-concurrency
// , cpu core * 80%, 1
low-concurrency
//
// 2000, 2000
max-tasks-per-worker-high
//
max-tasks-per-worker-normal
//
max-tasks-per-worker-low
```

```
max-tasks-per-worker-high * high-concurrency = 2000 * 4 = 8000
```

(2) TiDB

- (TiKV) CPU
 - 48 Core 1 tikv-server 48
 - **48**
 - 36
- (TIKV)
 - 48 Core 3 tikv-server 16
 - **16**
 - 12

(3) 4.0

4.0 readpool.storage readpool.coprocessor unified read pool 3

```
[readpool]
# unify-read-pool = true

[readpool.unified]
# min-thread-count = 1
# max-thread-count = 8

## Size of the stack for each thread in the thread pool.
# stack-size = "10MB"

## Max running tasks of each worker, reject if exceeded.
# max-tasks-per-worker = 2000
```

```
1.
2. TIKV
```

4. raftstore is busy

- append log stall 2
 - [tikv-detail]->[RocksDB - raft]->[Write stall duration]
 - [tikv-detail]->[RocksDB - kv]->[Write stall duration]
 - [tikv-detail]->[Errors]->[Server is busy]
- append log duration [tikv-detail]->[Raft IO]->[append log duration]
 - append log IO [tikv-detail]->[Raft IO]->[Write duration]
 - [raftstore] store-pool-size [1,5] [tikv-detail]->[Thread CPU]->[Raft store CPU] [Raft store CPU] [store-pool-size] * 70% store-pool-size

- tikv-server region 5 region raftstore cpu hibernate region

```
[raftstore]
hibernate-regions = true
```

```
**** ****store-pool-size
```

TiKV Server is busy TiDB Server is Busy SQL

7.3.3 Server is Busy

1. SQL

1. ,
- 2.
3. tikv

```
SQL SQL SQL 4.0 unified thread pool
```

2.

tikv is busy

```
**** TiDB v3.0.8 v3.0.8
```

3. region

TiDB 2.1 TiKV raft region region server is busy

1. 3.0 store-pool-size raftstore
2. hibernate-regions region

7.3.4

1. region id region tikv-server server is busy
2. tikv server is busy

```
region shard_row_id_bitspre_split_regions
```

7.3.5

Server is Busy 99% 6ms , 3s ?

Server is Busy Query Duration tidb.log wait exec

coprocessor 2000 2001 2001

7.3.6

1. SQL SQL 4.0 unified thread pool
- 2.
3. region shard_row_id_bitspre_split_regions
4. tikv

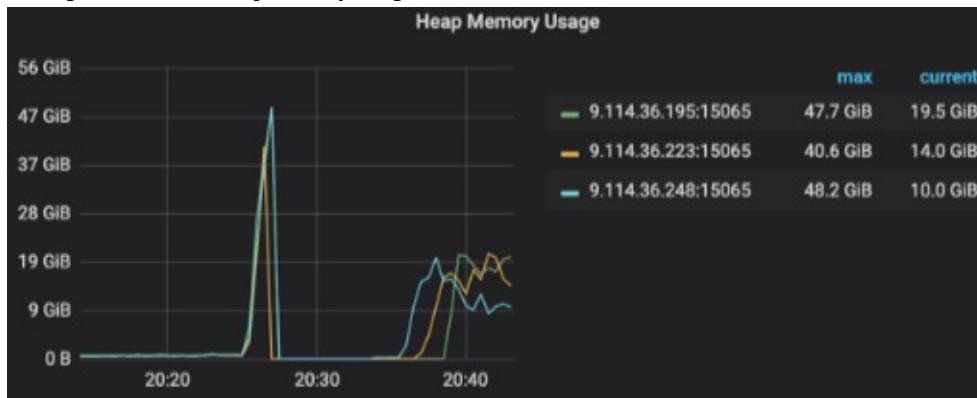
7.4 TiDB OOM

TiDB OOM TiDB-Server TiKV-Server

7.4.1 TiDB-Server

1. TiDB-Server OOM

- tidb-server "ERROR 2013 (HY000): Lost connection to MySQL server during query"
- TiDB grafana Server Heap Memory Usage



- TiDB service

```
$ systemctl status tidb-4000.service
● tidb-4000.service - tidb-4000 service
  Loaded: loaded (/etc/systemd/system/tidb-4000.service; disabled; vendor preset: disabled)
  Active: active (running) since Fri 2020-02-21 12:30:04 CST; 5 min ago
```

- /var/log/message&kern “out of memory” "TiDB service was killed"
 - tidb.log grep "Welcome to TiDB" tidb-server
 - tidb_stderr.log grep "fatal error: runtime: out of memory" "cannot allocate memory"

2. TiDB-Server OOM

TiDB oom TiDB TiDB

-

SQL root TiDB-Server hash joinUpdate&Delete

- Prometheus
-
- mydumper

3. TiDB-Server OOM

- "SQL" SQL
- tidb_mem_quota_query Query 32GB

```
global:
mem-quota-query: 209715200 // 200MB
oom-action: "cancel" // SQL
```

- max_execution_time SQL

```
set @@GLOBAL.max_execution_time=10000; // SQL 10s
```

- SWAP OOM

```
IO
```

- 4.0 oom-use-tmp-storage

```
`true` SQL `mem-quota-query`
```

- TiDB

```
TiDB NUMA TiDB NUMA
membind Only allocate memory from nodes. Allocation will fail when there is not enough memory available on these nodes. nodes may be specified as noted above NUMA NODE NUMA NODES
preferred Only allocate memory from nodes. Allocation will fail when there is not enough memory available on these nodes. nodes may be specified as noted above NUMA NODE
TiDB
sudo -s numactl --preferred=node
```

- TiDB

```
1 SQL
    TiDB OOM tidb_slow_query.log Query_time Mem_max SQL
    SQL
    table scan
    hash join join index loop join
    delete tidb_batch_delete
2 TiDB
    SQL TiDB TiDB oom
```

7.4.2 TiKV-Server

TiKV-Server OOM

1. block-cache

TiKV grafana instance RocksDB block cache size TiKV block-cache 45% container TiKV container

```
[storage.block-cache]
capacity = "30GB"
```

- 1.
2. block-cache

2. coprocessor

gRPC coprocessor OOM [tikv-detail]->[Coprocessor Overview]->[Total Response Size] network outbound

- 1.
2. "SQL"
3. gRPC poll CPU

3. Raft apply raft apply log

Apply apply channel oom apply wait duration apply log duration

- [tikv-detail]->[Raft IO]->[apply log duration]
- [tikv-detail]->[Raft propose]->[apply wait duration]
- [tikv-detail]->[Thread CPU]->[async apply CPU]

- 
1. apply cpu [Async apply CPU] [apply-pool-size] * 70% apply-pool-size
 2. io

7.5 TiKV

TiKV TiDB SQL TiKV TiKV

TiKV

- Number files at each levels TiDB 10G

TiKV LSM-Tree RocksDB INSERT UPDATE DELETE max-bytes-for-level-multiplier RocksDB
10GB (512MB + 1GB + 10GB) * 3 RocksDB key block LZ4 ZSTD 33.5GB. (512MB L0 SST)
)

- TiDB GC ?

TiDB MVCC Snapshot Isolation DELETE UPDATE TiDB TiDB TiDB TiKV TiKV
GC

- UPDATE SQL UPDATE

UPDATE TiDB tikv_gc_life_time 10 UPDATE 10 TiKV GC UPDATE TiDB
GC UPDATE TiKV "10 + * 1.12". 1.12

- TiKV 1020

TiKV GC GC speed TiKV TiKV GC raft leader follower TiKV gc.max-write-bytes-per-
sec GC 128KB ~ 512KB 0KB

- DELETE FROM table_xx;

GC TiDB 10 RocksDB RocksDB RocksDB DROP TA

- TiDB

MySQL TiDB TiDB SELECT * from mysql.tidb tikv_gc_life_time TiDB 10 GC

- GC RocksDB

GC compact TiKV CPU RocksDB compact 12%

Dynamic Level

- TiKV level-1 level-2 level-3 level-4

TiKV RocksDB Dynamic Level Bytes max-bytes-for-level-base 512MB level-6 level-6 level-1
max-bytes-for-level-base max-bytes-for-level-base * max-bytes-for-level-multiplier level-6 level-2
level-5 level-1 level-0 110

- TiKV

TiKV snappy zlib gzip lz4 lz4 hc zstd ["no", "no", "lz4", "lz4", "lz4", "zstd", "zstd"] dynamic level
500G RocksDB 4 500G ZSTD defaultcf writecf compression-per-level [{"
"lz4", "zstd", "zstd", "zstd"}, 50G 500G zstd]

8 TiDB

“” TiDB TiDBTiKVPD

- 8.1 TiDB
- 8.2 TiKV
- 8.3

8.1 TiDB

TiDB TiDB

8.1.1 SQL

SQL SQL SQL Kill TiDB SQL

(1)

max_execution_time

- GLOBAL | SESSION | SQL HINT
- 0
- N SQL Hint

SQL Hint:

```
mysql> SELECT /*+ MAX_EXECUTION_TIME(1000) */ * FROM t1 INNER JOIN t2 WHERE ...;
```

(2)

tidb_mem_quota_query

- SESSION
- 32 GB
- TiDB OOMAction oom-action log SQL cancel Kill SQL SQL

```
--10G
mysql> set @@session.tidb_mem_quota_query=10737418240;
Query OK, 0 rows affected (0.00 sec)
mysql> show variables like '%tidb_mem_quota_query%';
+-----+-----+
| Variable_name      | Value   |
+-----+-----+
| tidb_mem_quota_query | 10737418240 |
+-----+-----+
1 row in set (0.00 sec)
```

mem-quota-query

- GLOBAL
- 32GB
- TiDB Global TiDB tidb-ansible conf/tidb.yaml mem-quota-query TiDB OOMAction oom-action log SQL cancel Kill SQL

```
conf/tidb.yaml
---
# default configuration file for TiDB in yaml format

global:
  ...
  # Only print a log when out of memory quota.
  # Valid options: ["log", "cancel"]
  # oom-action: "log"

  # Set the memory quota for a query in bytes. Default: 32GB
  # mem-quota-query: 34359738368
```

8.1.2

TiDB 2PC TiDB

tidb_retry_limit

- SESSION | GLOBAL
- 10
- tidb_retry_limit = 0 SQL

```
--SESSION
mysql> set @@session.tidb_retry_limit=0;
Query OK, 0 rows affected (0.01 sec)

--GLOBAL
mysql> set @@global.tidb_retry_limit=0;
Query OK, 0 rows affected (0.00 sec)

mysql> show variables like '%tidb_retry_limit%';
+-----+-----+
| Variable_name      | Value |
+-----+-----+
| tidb_retry_limit  | 0     |
+-----+-----+
```

tidb_disable_txn_auto_retry

- SESSION | GLOBAL
- on
- on off TiDB TiDB tidb_retry_limit SQL

```
--SESSION
mysql> set @@session.tidb_disable_txn_auto_retry=OFF;
Query OK, 0 rows affected (0.00 sec)

--GLOBAL
mysql> set @@global.tidb_disable_txn_auto_retry=OFF;
Query OK, 0 rows affected (0.00 sec)

mysql> show variables like '%tidb_disable_txn_auto_retry%';
+-----+-----+
| Variable_name          | Value |
+-----+-----+
| tidb_disable_txn_auto_retry | OFF   |
+-----+-----+
1 row in set (0.01 sec)
```

8.1.3 Join

TiDB SQL Join TP | AP Join SQL

tidb_distsql_scan_concurrency

- SESSION | GLOBAL
- 15
- scan AP TP AP TiKV CPU SQL

```
--SESSION
mysql> set @@session.tidb_distsql_scan_concurrency=30;
Query OK, 0 rows affected (0.00 sec)

--GLOBAL
mysql> set @@global.tidb_distsql_scan_concurrency=30;
Query OK, 0 rows affected (0.00 sec)

mysql> show variables like '%tidb_distsql_scan_concurrency%';
+-----+-----+
| Variable_name | Value |
+-----+-----+
| tidb_distsql_scan_concurrency | 30   |
+-----+-----+
1 row in set (0.01 sec)
```

tidb_index_lookup_size

- SESSION | GLOBAL
- 20000
- index lookup batch AP TP SQL

```
--SESSION
mysql> set @@session.tidb_index_lookup_size=40000;
Query OK, 0 rows affected (0.00 sec)

--GLOBAL
mysql> set @@global.tidb_index_lookup_size=40000;
Query OK, 0 rows affected (0.01 sec)

mysql> show variables like '%tidb_index_lookup_size%';
+-----+-----+
| Variable_name | Value |
+-----+-----+
| tidb_index_lookup_size | 40000 |
+-----+-----+
1 row in set (0.00 sec)
```

tidb_index_lookup_concurrency

- SESSION | GLOBAL
- 4
- index lookup AP TP SQL

```
--SESSION
mysql> set @@session.tidb_index_lookup_concurrency=8;
Query OK, 0 rows affected (0.00 sec)

--GLOBAL
mysql> set @@global.tidb_index_lookup_concurrency=8;
Query OK, 0 rows affected (0.00 sec)

mysql> show variables like '%tidb_index_lookup_concurrency%';
+-----+-----+
| Variable_name | Value |
+-----+-----+
| tidb_index_lookup_concurrency | 8   |
+-----+-----+
1 row in set (0.00 sec)
```

tidb_index_lookup_join_concurrency

- SESSION | GLOBAL
- 4
- index lookup join SQL

```
--SESSION
mysql> set @@session.tidb_index_lookup_join_concurrency=8;
Query OK, 0 rows affected (0.00 sec)
--GLOBAL
mysql> set @@global.tidb_index_lookup_join_concurrency=8;
Query OK, 0 rows affected (0.01 sec)
mysql> show variables like '%tidb_index_lookup_join_concurrency%';
+-----+-----+
| Variable_name | Value |
+-----+-----+
| tidb_index_lookup_join_concurrency | 8    |
+-----+-----+
1 row in set (0.00 sec)
```

tidb_hash_join_concurrency

- SESSION | GLOBAL
- 5
- hash join SQL

```
--SESSION
mysql> set @@session.tidb_hash_join_concurrency=10;
Query OK, 0 rows affected (0.01 sec)

--GLOBAL
mysql> set @@global.tidb_hash_join_concurrency=10;
Query OK, 0 rows affected (0.00 sec)

mysql> show variables like '%tidb_hash_join_concurrency%';
+-----+-----+
| Variable_name | Value |
+-----+-----+
| tidb_hash_join_concurrency | 10   |
+-----+-----+
1 row in set (0.01 sec)
```

tidb_index_serial_scan_concurrency

- SESSION | GLOBAL
- 1
- scan AP TP SQL

```
--SESSION
mysql> set @@session.tidb_index_serial_scan_concurrency=4;
Query OK, 0 rows affected (0.00 sec)

--GLOBAL
mysql> set @@global.tidb_index_serial_scan_concurrency=4;
Query OK, 0 rows affected (0.01 sec)

mysql> show variables like '%tidb_index_serial_scan_concurrency%';
+-----+-----+
| Variable_name | Value |
+-----+-----+
| tidb_index_serial_scan_concurrency | 4    |
+-----+-----+
1 row in set (0.00 sec)
```

8.1.4 Mysql

compatible-kill-query

- false
- Kill TiDB Kill sessionID MySQL TiDB TiDB Kill TiDB sessionID compatible-kill-query true TiDB

```
Ctrl+C MySQL Kill TiDB TiDB Kill Kill compatible-kill-query ,( tidb-ansible conf/tidb.yaml
)
```

```
conf/tidb.yaml
---
# default configuration file for TiDB in yaml format

global:
...
# Make "kill query" behavior compatible with MySQL. It's not recommend to
# turn on this option when TiDB server is behind a proxy.
compatible-kill-query: false
```

```
tidb_constraint_check_in_place
* SESSION | GLOBAL
* 0
* TiDB commit unique key
```

```
-- tidb_constraint_check_in_place
create table t (i int key);
insert into t values (1);
begin;
insert into t values (1);
Query OK, 1 row affected
commit; --commit
ERROR 1062 : Duplicate entry '1' for key 'PRIMARY'
```

```
-- tidb_constraint_check_in_place
set @@tidb_constraint_check_in_place=1;
begin;
insert into t values (1);
ERROR 1062 : Duplicate entry '1' for key 'PRIMARY'
```

8.1.5

prepared-plan-cache txn_local_latches TiDB TiDB tidb-ansible conf/tidb.yaml tidb-server

```
conf/tidb.yaml
--
prepared_plan_cache:
  enabled: true          -- prepare plan cache, false
  capacity: 100           --
  memory-guard-ratio: 0.1 -- performance.max-memory, max-proc * (1 - prepared-plan-cache.memory-guard-ratio) LRU
, 0 1, 0.1

--
txn_local_latches:
  enabletrue            -- , false
  capacity: 2048000      -- Hash slot 2 slot 32 Bytes
```

8.2 TiKV

TiKV TiKV

TiKV RocksDB TiKV RocksDB TiKV RocksDB RocksDB KV Raft RocksDB RaftDB Raft

TiKV RocksDB Column Families (CF)

- RocksDB KV default write lock 3 CF
 - default CF [rocksdb.defaultcf]
 - write CF (MVCC) [rocksdb.writecf]
 - lock CF
- Raft RocksDB Raft log
 - default CF Raft log [raftdb.defaultcf]

CF block cache [storage.block-cache] capacity block cache block cache CF block cache
 [storage.block-cache] shared=false CF block cache [rocksdb.writecf] block-cache-size write CF
 CF write-buffer write-buffer-size

TiKV

block-cache write-buffer

1. page cache
2. TiKV select * from ... TiDB TiKV

TiKV

1. TiKV CPU 8 32GB
- 2.
3. IOPS SSD

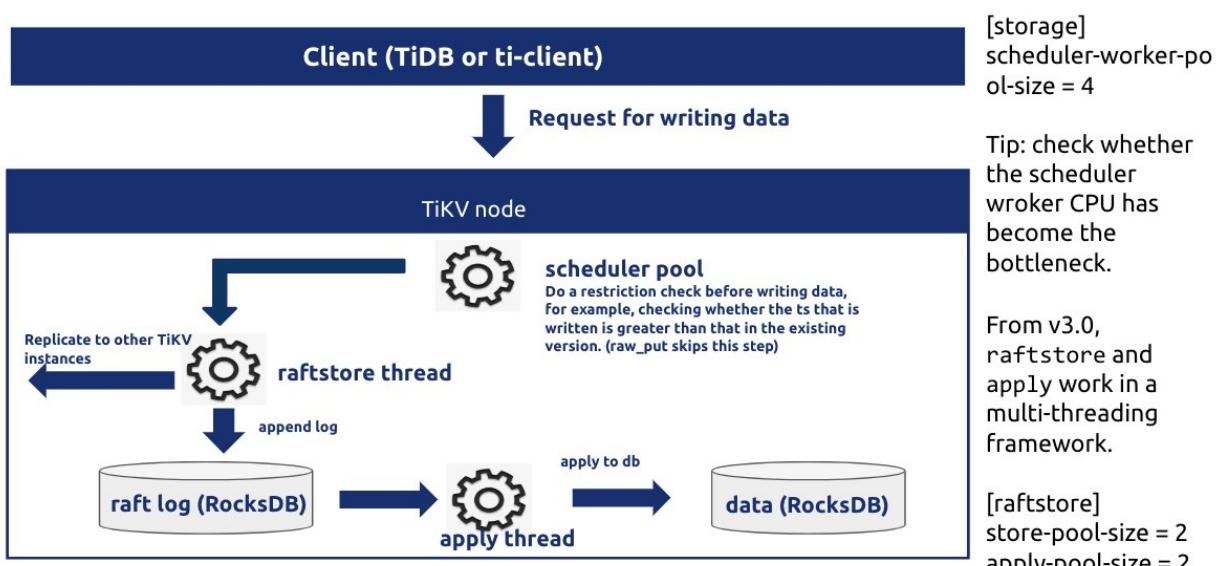
TiKV Region TiKV

8.2.1 TiKV

TiKV 4.0 gRPC Scheduler UnifyReadPool Store Apply RocksDB CPU

- gRPC TiKV
- Scheduler key-value Store Raft
- Raftstore Raft Propose Raft Group Raft Commit Index Apply
- Apply Store key-value RocksDB gRPC
- RocksDB RocksDB Compact Flush RocksDB Compact
- UnifyReadPool TiKV 4.0 Coprocessor Storage Read Pool kv getkv batch getraw kv getcoprocessor

[RocksDB: A Persistent Key-Value Store for Flash and Storage](#)



1. GRPC

gRPC server.grpc-concurrency 4 gRPC IO CPU 8server.grpc-concurrency 2 TiKV Grafana Thread
CPU gRPC poll CPU server.grpc-concurrency 80% server.grpc-concurrency 80%

2. Scheduler

Scheduler	(storage.scheduler-worker-pool-size)	TiKV CPU 16 8 16 4 key-value	sc
Raft raftstore raft scheduler		50% 75% 8 Grafana Thread CPU scheduler worker CPU 400%	

3. Raftstore

Raftstore TiKV raftstore.store-pool-size 2 store fsync RocksDB raftstore.sync-log true raftstore.sync-log false
IOstore 100% CPU RocksDB CPU 40% 60% store

- Raft IO append log duration Store Raft Message RocksDB Raft P99 500ms Node Exporter disk latency
- RocksDB Write Stall RocksDB-raft write stall duration 0 0 write stall

4. UnifyReadPool

UnifyReadPool readpool.unified.max-thread-count CPU 80% CPU 60%90%

5. RocksDB

RocksDB RocksDB Compact Flush CPU rocksdb.max-background-jobs raftdb.max-background-jobs 4.
Stall Grafana RocksDB-kv Write Stall Reason 0 pending compaction bytes rocksdb.max-sub-compactions 2
compaction job memtable count max-write-buffer-number 5 level0 file limit

[Write](#)

```
rocksdb.defaultcf.level0-slowdown-writes-trigger
rocksdb.writecf.level0-slowdown-writes-trigger
rocksdb.lockcf.level0-slowdown-writes-trigger
```

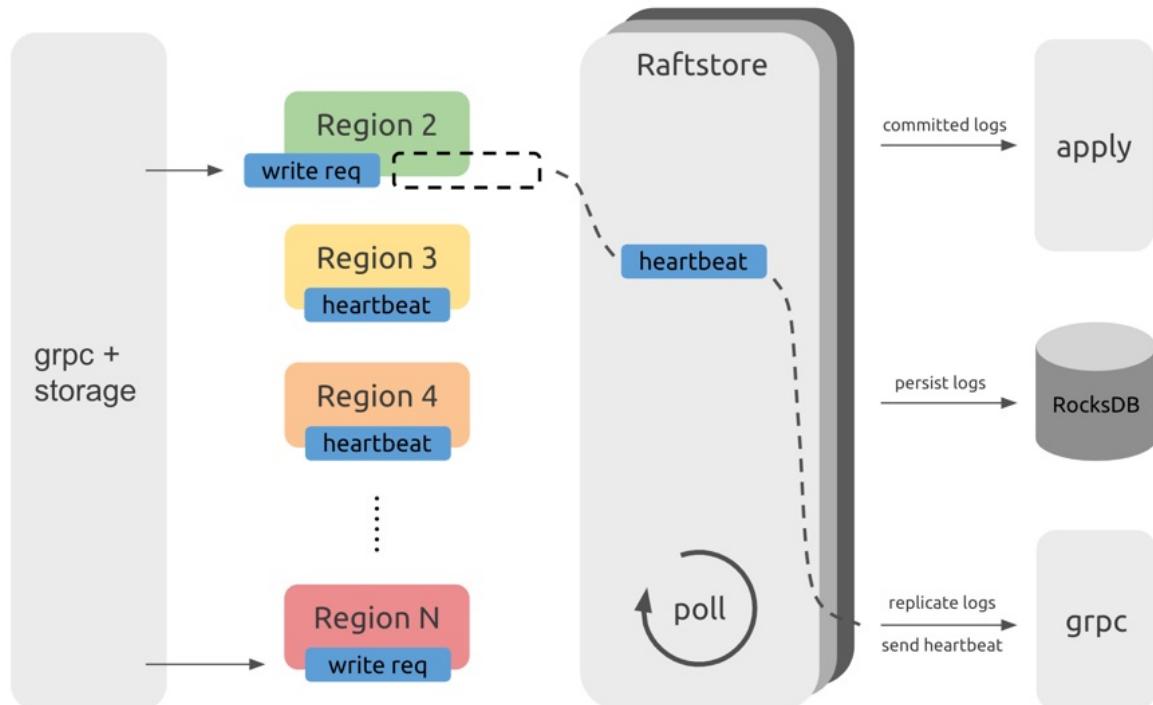
8.2.2 Region

TiDB range key Region TiKV Region TiKV Region

TiKV Raftstore Region

1. Raftstore

TiKV RegionRegion Raftstore Raft Region Raft log Raft Region TiKV Raftstore



Raftstore TiDB gRPC storage KV RegionRaftstore Region Region Raftstore Raft Raft
Region Raft Region

2.

Raftstore Region Region Raftstore Region Raftstore CPU

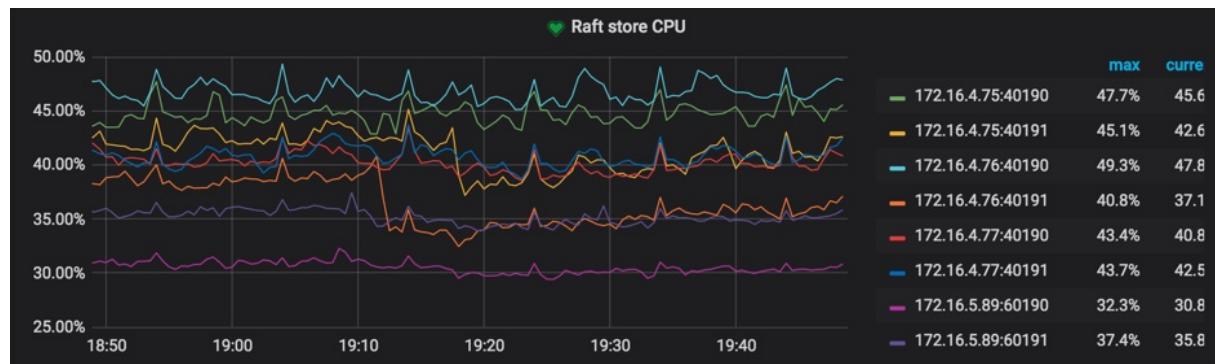
Raftstore CPU 85% Grafana TiKV Raft Propose propose wait duration

Raftstore CPU Raftstore Raftstore I/O CPU 100%

Grafana TiKV metrics

Thread-CPU Raft store CPU

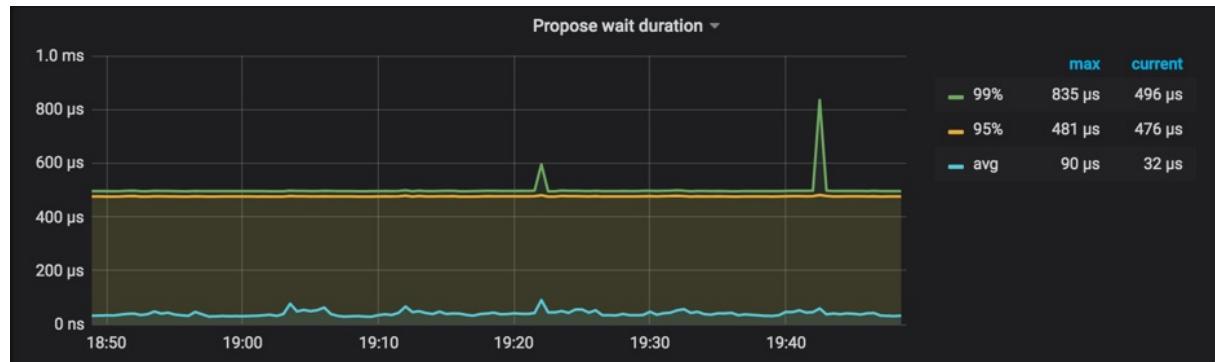
raftstore.store-pool-size * 85%



- Raft Propose Propose wait duration

Propose wait duration Raftstore Raftstore Raftstore append log Raftstore

50-100ms



3.

- TiKV Region
- Region

TiKV

I/O CPU TiKV TiKV Region TiKV

raft-base-tick-interval

Region Region Raftstore TiKV raft-base-tick-interval

```
[raftstore]
raft-base-tick-interval = "2s"
```

raft-base-tick-interval Raftstore Region Raft Raft tick Raftstore tick election timeout heartbeat

```
raft-election-timeout = raft-base-tick-interval * raft-election-timeout-ticks
raft-heartbeat-interval = raft-base-tick-interval * raft-heartbeat-ticks
```

Region Follower raft-election-timeout Leader Leader raft-heartbeat-interval Leader Follower raft-base-tick-interval
Raft Raft Leader

Raftstore

```
TiKV raftstore.store-pool-size 2 Raftstore
```

Hibernate Region

```
Region Region Region Hibernate Region raft-base-tick Region Raft Region Raft Raftstore
```

```
TiDB v4.0 Hibernate Region      TiKV master Hibernate Region      Hibernate Region
```

Region Merge

```
Region Merge TiDB v4.0
```

```
Region Merge Region Region Split Region Merge Region Drop Table/Truncate Table Region Region
```

```
pd-ctl Region Merge Region Merge
```

```
>> pd-ctl config set max-merge-region-size 20  
>> pd-ctl config set max-merge-region-keys 200000  
>> pd-ctl config set merge-schedule-limit 8
```

```
Region Merge PD Region Merge      PD Region Merge
```

8.2.3

1. Block-cache TiKV RocksDB Column Family (CF) KV RocksDB defaultwritelock 3 CF
2. default CF [rocksdb.defaultcf]
3. write CF MVCC [rocksdb.writecf]
4. lock CF
5. Raft RocksDB Raft logdefault CF Raft log [raftdb.defaultcf]
6. TiDB 3.0 CF Block-cache RocksDB TiDB 2.1 block-cache-size CF Block-cache Block-cache
7. CF Write-buffer write-buffer-size

TiDB 3.0 TiKV conf/tikv.yml block-cache-size capacity

```
storage:  
  block-cache:  
    capacity: "1GB"
```

TiKV TiKV capacity = MEM_TOTAL * 0.5 / TiKV

Sync-log

Raft TiKV 50% ACK sync-log sync-log 30%

conf/tikv.yml raftstore sync-log

```
[raftstore]  
sync-log = true
```

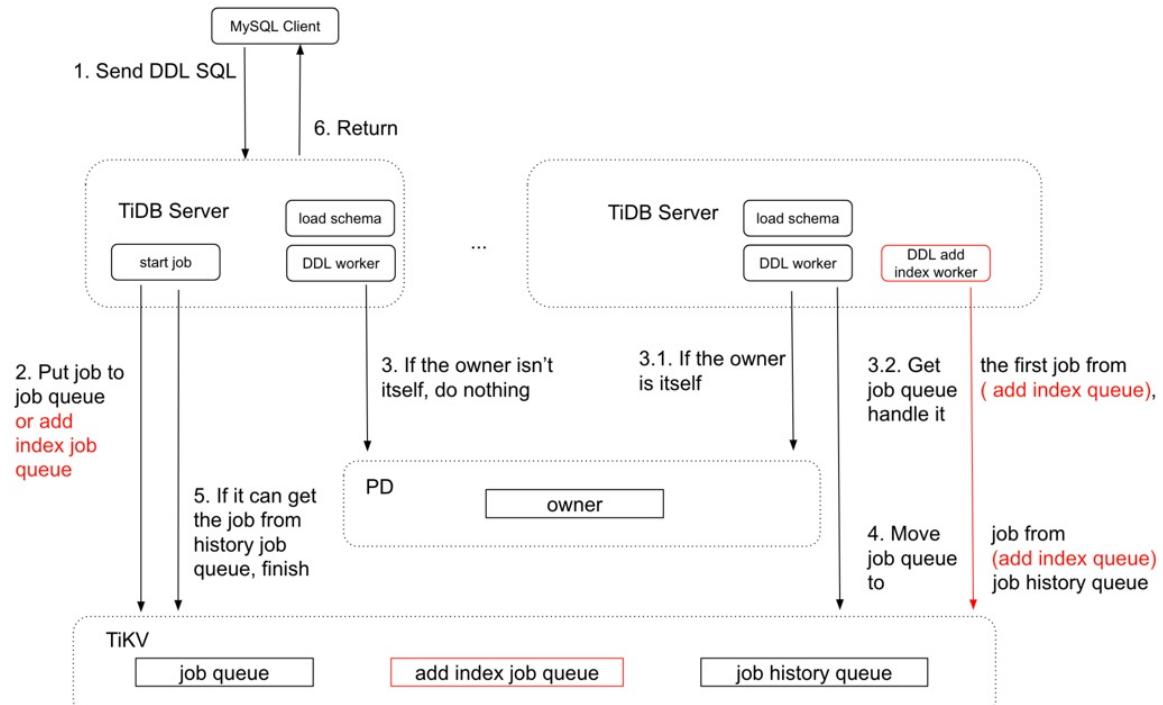
8.3

TiDB Online DDL TiDB DDL DML

8.3.1 TiDB

TiDB DDL

Google F1 schema schema TiDB 2.1 DDL (add index job)



:

- 1.
- 2.

:

1. TiDB TiDB
2. Job (add index job queue)
3. TiDB Worker Job
4. Worker PD region 256 Job region
5. Job Job (history ddl job)

8.3.2

tidb_ddl_reorg_worker_cnt	4	
tidb_ddl_reorg_batch_size	256	
tidb_ddl_reorg_priority	PRIORITY_LOW	PRIORITY_LOW/PRIORITY_NORMAL/PRIORITY_HIGH
tidb_ddl_error_count_limit	512	

```
tidb_ddl_reorg_worker_cnt    tidb_ddl_reorg_batch_size
4 256          tidb_ddl_reorg_worker_cnt    tidb_ddl_reorg_batch_size
tidb_ddl_reorg_priority  PRIORITY_HIGH OLTP
```

1. admin show ddl RowCount START_TIME DDL r1 t1
2. show stats_meta RowCount r0
3. t1/(r1/r2) - t1

- 1.
- 2.

1 TiDB

1.1 TiDB

1.1.1 TiDB

TiDB 2015 4 5 5 TiDB 1800 PingCAP 2020 3 TiDB 760 Contributor100 30000+
GitHub Stars 1000+ AskTUG 700+ PingCAP University TiDB 1000

2019 TiDB 94 123 38 Infra Meetup 23 Paper Reading 60000+

2020 TiDB 4.0 GA 5 TiDB

1.1.2 TiDB

2015 04	TiDB
2015 09	GitHub Star 2700
2015 12	TiDB 1.0 Alpha Google F1
2016 04	Google Spanner TiKV
2017 10	TiDB 1.0 GA
2018 04	TiDB 2.0 GA, TiDB + TiKV 16000+ Stars, 240+ Contributors
2018 08	TiDB Operator
2018 08	CNCF TiKV CNCF Sandbox
2018 12	Talent Plan
2019 01	TiDB Lightning Toolset & TiDB Data Migration
2019 03	TiDB PingCAP University
2019 05	TiDB Binlog
2019 05	CNCF TiKV CNCF Incubation
2019 06	TiDB User Group
2019 06	TiDB 3.0 GATiDB + TiKV 22000+ Stars, 390+ Contributors
2019 06	TiDB AskTUG
2019 06	TiDB 2.0 Courses
2019 07	TiDB Operator 1.0 GA AWS/GCP/
2019 12	Chaos Mesh
2020 01	TiFlash Beta
2020 03	TiDB Contributor 400
2020 03	TiDB 4.0 RC
2020 03	Talent Plan Courses
2020 05	TiDB 4.0 GA

1.2 TiDB

1.2.1 TiDB

1. [TiDB](#)

TiDB TiDB SQL SQL SQL Load Balancer TiDB
TiDB TiKV TiDB

TiDB MySQL Protocol Packet TiI

2. [Parser](#)

Parser Yacc MySQL Parser SQL SQL AST

3. [TiSpark](#)

TiSpark PingCAP OLAP Spark [TiKV](#) TiDB HTAP (Hybrid Transactional/Analytical Processing) TiSpark
TiKV Placement Driver (PD) Spark

1.2.2 TiKV

1. [TiKV](#)

TiKV K-V [RocksDB](#) Raft Percolator Raft PD K-V ACID TiDB TiKV javac

2. [PD](#)

PD (Placement Driver) TiDB ID TSO PD [TiKV](#) client

3. [grpc-rs](#)

grpc-rs [gRPC Core](#) rust SSLTiKV TiDB

4. [raft-rs](#)

raft-rs Raft rust etcd Raft

5. [rust-rocksdb](#)

rust-rocksdb [Rocksdb](#) rust Rust [Rocksdb](#) TiKV

6. [rust-prometheus](#)

rust-prometheus rust [Prometheus](#) instrumentation Rust [Prometheus](#)

7. [pprof-rs](#)

pprof-rs rust profiling TiKV profiling

1.2.3 Tools

1. [TiDB Lightning](#)

TiDB Lightning TiDB DDL DDL TiDB KV TiKV Mydumper CSV MySQL TiDB

2. Dumpling

Dumpling Mydumper MySQL Dumpling Mydumper Lightning TiDB Dumpling

4. ticdc

ticdc TiDB MySQL, TiDB, Kafka ticdc TiKV KV SQL MySQL TiDB Kafka KV Kafka SQL

5. DM

DM MySQL TiDB DM-master, DM-worker dmctl DM-master DM-worker dmctl DM MySQL binlog
DM-worker TiDB

6. BR

BR TiDB BR SQL BR PD TiKV region leader KV sst TiKV sst Raft

7. tidb-binlog

tidb-binlog TiDB binlog TiDB SQL Pump Drainer Pump binlog Tidb Binlog TiDB, MySQL, Kafka SQL

1.2.4 Cloud

TiDB Operator

TiDB Operator Kubernetes TiDB TiDB TiDB Operator TiDB Kubernetes TiDB Operator
Kubernetes TiDB

1.2.5

1. Chaos Mesh

Chaos Mesh Kubernetes Kubernetes Chaos Mesh Chaos Kubernetes

2. tipocket

tipocket Chaos Mesh TiDB TiDB Operator Chaos Mesh Kubernetes TiDB Chaos

3. TiDB Ansible

TiDB Ansible TiDB

4. docs

docs PingCAP PingCAP <https://pingcap.com/docs/>

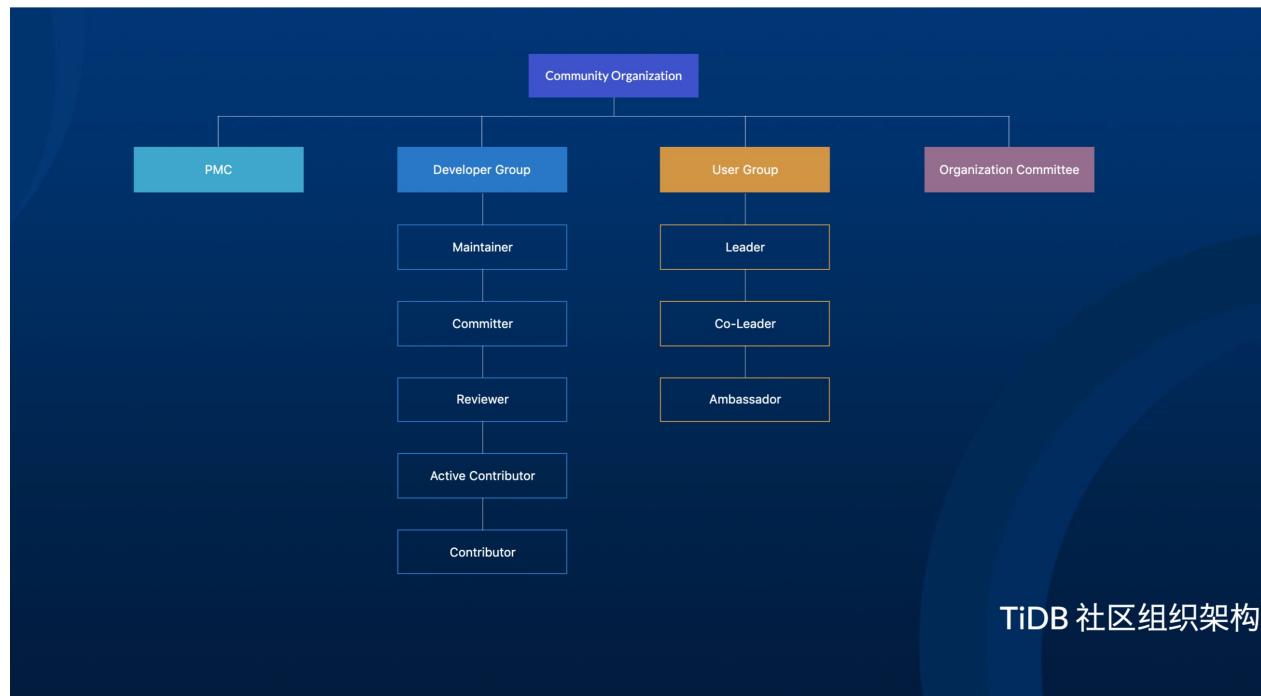
5. docs-cn

docs-cn docs PingCAP <https://pingcap.com/docs-cn/>

1.3 TiDB

TiDB TiDB TiDB community

PMC Developer Group User Group Organization Committee



1.3.1 PMC

TiDB Apache PMC(Project Management Committee) TiDB Roadmap

1.3.2 Developer Group

Developer Group Contributor Committer Maintainer Contributor Committer TiDB Active Contributor Reviewer



TiDB TiDB TiKV TiDB :

- Special Interest Group SIG 70 ARPAnet mailing list TiDB SIG TiDB/TiKV Active Contributor Tech Lead Committer Reviewer Active Contributor Reviewer Committer Maintainer
- Working Group WG SIG SIG

1.3.3 User Group

User Group TiDB User Group TUG TiDB TiDB TUG TUG TiDB

TiDB TiDB TiDB TiDB

1.3.4 Organization Committee

Organization Committee Committee Legal / Quality / Marketing

1.4 TiDB

TiDB TiDB

- ICDE Paper - UniKV: Toward High-Performance and Scalable KV Storage in Mixed Workloads via Unified Indexing
- TiDB Intel Optane DC persistent memory
-

TiDB

- TiDB TiDB
- Cloud TiDB scan TiDB 250+ 500+ TB
- TiKV TiKV TiKV KV

TiDB

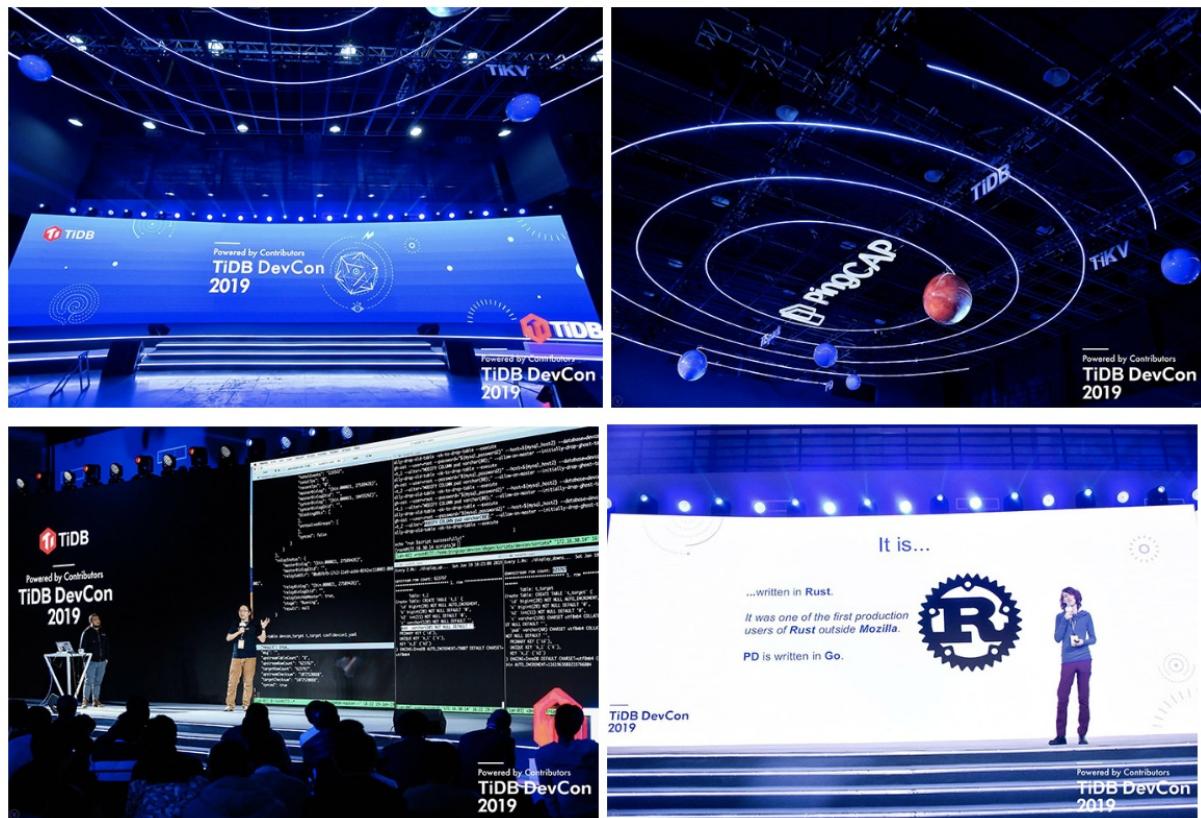
2 TiDB

2.1

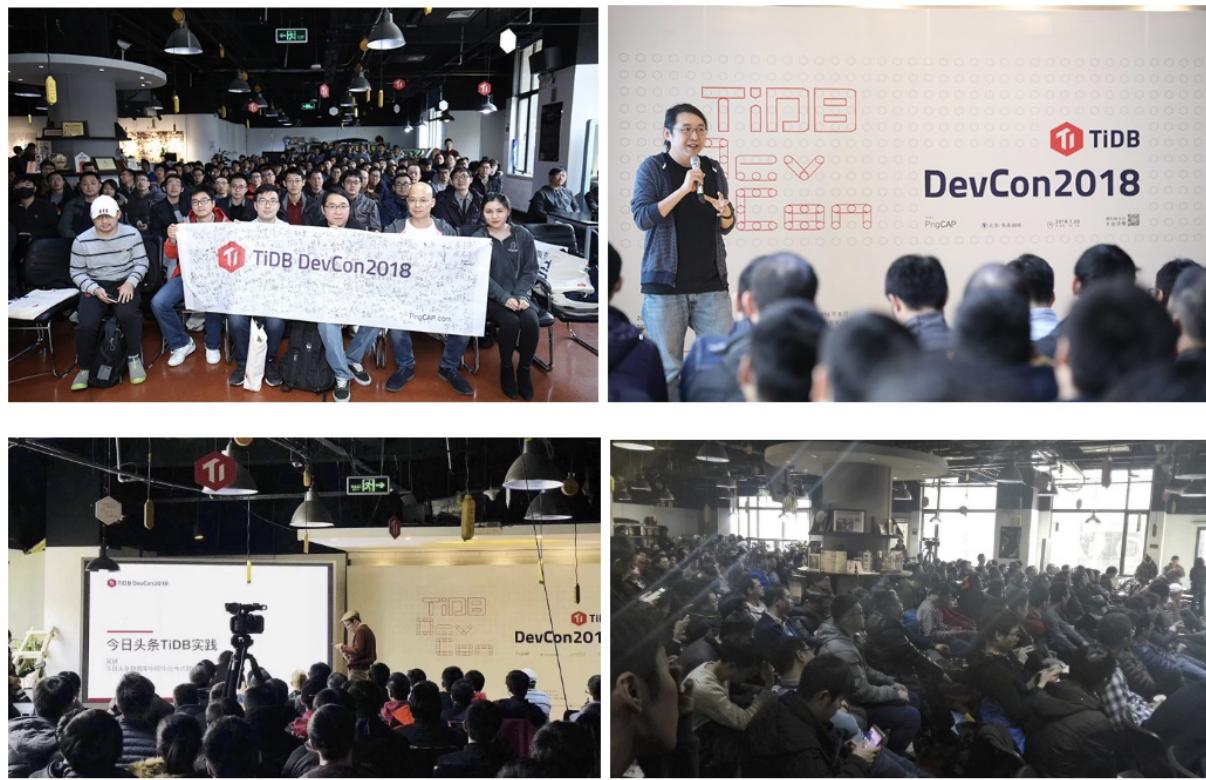
2.1.1 TiDB DevCon

TiDB DevCon PingCAP TiDB 2018 2019 TiDB DevCon TiDB TiDB TiDB DevCon TiDB TiDB

TiDB DevCon 2019 2019 01 19 750+VIPKIDBilibili



TiDB DevCon 2018 2018 01 20 300+Mobike



2.1.2 TiDB TechDay

Get TiDB 2017 PingCAP TiDB TechDay TiDB

TechDay 2017 TiDB Internal, TiSpark, TiDB on Kubernetes

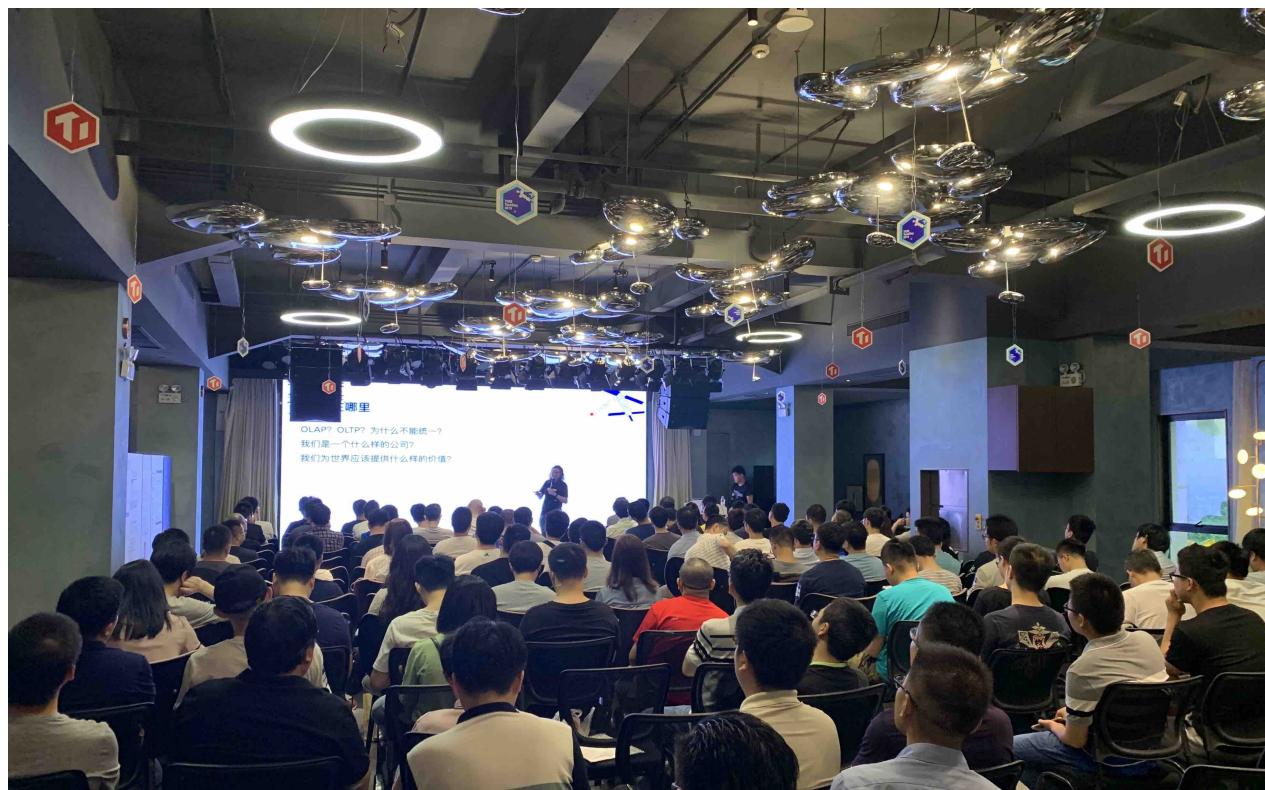


TechDay 2018 TiDB 2.1 Beta Chaos Practice in TiDB



TechDay TiDB TechDay 2019 6 TiDB 3.0 TiDB





2.1.3 Infra Meetup

1.

PingCAP Hackers 2016 3 5 Infra Meetup

2.

2015 TiDB Infra Meetup

2020 3 PingCAP Infra Meetup 125 Infra Meetup 100 TiDB TiDB Meetup Meetup

No.125 Infra Meetup



2.1.4 TiDB Hackathon

1.

Hackathon Hackathon 18 PingCAP TiDB Hackathon TiDB Pizza TiDB 48

2.

PingCAP TiDB TiDB Hackathon

- [TiDB Hackathon](#) 2018 12 1 “TiDB Ecosystem” TiDB 22 6 48 TiDB Batch and Streaming SQL
TiDB CC Region TiEyeAll diagnosis in SQL TiQuery Hackathon



- [TiDB Hackathon](#) 2019 10 26 “Improve” Hacking Time 39 7 Demo Show TiDB
TiDB “” bug TiDB~



3.

TiDB &TiDB Hackathon

Hackathon PingCAPer TiDB Hackathon

Hackathon ~

2.1.5 TUG

TiDB TiDB TiDB Community

[TiDB User Group](#)

TiDB TiDB 11 TiDB TiDB SaaS

" TiDB TiDB "

""

" TiDB "

TUG "TUG "TUG TUG TiDB "" Infra Meetup 2.1.3

1. TUG " " " " Cloud + Database " " " "
2. TUG
3. TUG
4. TUG TUG PingCAP

TUG 2019 8 ShopeeVIPKIDUCLOUD 10+ TUG



2020 TUG TUG TUG TUG

2.2 TUG (TiDB User Group)

2.2.1 TUG

2019 TiDB 1000 TiDB TiDB TUG

TUG 2019 6 TiDB PingCAP 5 Local TUG TUG “”DBA58 Shopee
bilibili 100+ TiDB



TUG TUG TiDB TiDB TUG 2019 TUG 78 78 Talk 5000 TUG TUG

2.2.2 TUG

TUG Management) (Professional)

- TUG Management TUG TUG
 - TUG Governance Committee TUG TUG TUG TUG
 - TUG Infra Team Leader/Co-Leader TUG
 - TUG Local Leader/Co-Leader TUG TUG Leader Co-leader
- TUG Professional TiDB TiDB MVA TUG
 - TiDB MVA (Most Valuable Advocate) TUG TiDB TUG ≥3 TiDB MVA 2019 TUG 29 TiDB MVA

TUG TiDB TUG TUG Ambassador TUG Ambassador TUG TUG Management Professional TUG

2.2.3 TUG

TUG TiDB TUG Talk TUG TUG TUG TUG

2.2.4 TUG

2020 TUG TUG TUG Go Global TiDB TiDB TUG TUG

2.3 Talent Plan: Made for Community, by Community

Talent Plan PingCAP GitHub

3800+ Star Talent Plan TiDB

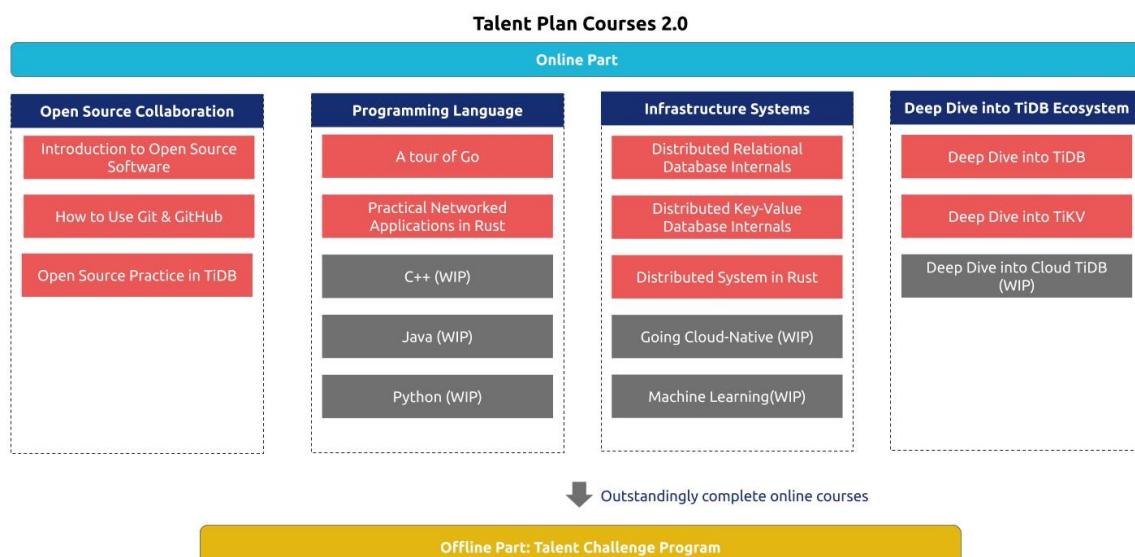
2018.12 Talent Plan 2020.3.7

400 40 4 38 15 Talent Plan PingCAP TiDB Community



2.3.1

Talent Plan Talent Plan Rust TiDB/TiKV TiDB



1.

- Open Source Collaboration
- Programming Language
- Infrastructure Systems
- Deep Dive into TiDB Ecosystem

Series 1: Open Source Collaboration(WIP)

Linux Apache CNCF TiDB

Series 2: Programming Language

Go Rust C++ Python

Rust Brian Anderson Rust —— [Practical Networked Applications in Rust](#) Rust Key-Value

Series 3: Infrastructure Systems

- Go [TinySQL WIP](#)
- Go Key-Value [TinyKV WIP](#)
- Rust [Distributed Systems in Rust](#)

TinySQL SQL

- SQL
- SQL Parser
- SQL DML DDL

TinyKV Distributed Systems in Rust MIT 6.824 TiKV 0 1 KV

- LAB1: KV server
- LAB2: Raft KV server
- LAB3: multi-Raft
- LAB4: percolator

TinyKV Raft Percolator TinyKV TiDB + TiKV + PD TiDB/TiKV/PD

[TinySQL](#) [TinyKV](#)

Series 4: Deep Dive into TiDB Ecosystem

TiDB TiDBTiKVCloud TiDB

2. ——Talent Challenge Program

TiDB 1 [PingCAP Talent Plan](#) [PingCAP //Special Offer](#) [PingCAP/TiDB Meetup](#)

4 41 10 38



2.3.2

1: Distributed Storage Engineer

- Programming Language: "[Practical Networked Applications in Rust](#)"
- Infrastructure Systems: "[Distributed Key-Value Database Internals\(WIP\)](#)" & "[Distributed Systems in Rust](#)"
- "[Deep Dive into TiKV](#)"

2: Distributed Relational Database Engineer

- Programming Language: "[A Tour of Go](#)"
- Infrastructure Systems: "[Distributed Relational Database Internals\(WIP\)](#)"
- "[Deep Dive into TiDB](#)"

3: Cloud TiDB Engineer

- Programming Language: "[A Tour of Go](#)"
- Container & Container Orchestration (DockerK8s ...)
- "[Deep Dive into Cloud TiDB\(WIP\)](#)"

4:

- Open Source Collaboration(WIP): "Introduction to Open Source Software" & "Build a Welcoming Community"
- *[The Art of Community: Building The New Age Of Participation](#), [The Cathedral & the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary](#), [People Powered: How communities can supercharge your business, brand, and teams](#)*

2.3.3

- “ PingCAP Demo TiDB TiDB PingCAP Global Meetup ”——
- “ Talent Plan PingCAP PingCAP ” KV | Talent Plan ——
- “ Talent Plan PingCAP Talent Plan ”——
- “ Talent Plan fancy work”——
- “ Talent Plan TiDB PingCAP Cool Geek ”——
- “ Talent Plan TiDB ”——

2.3.4 Talent Plan

Talent Plan “Made for Community, by Community”

- Talent Plan [Talent Plan](#) TiDB TiKV
- Talent Plan [TiDB Community Slack Workspace](#) channel: #wg-talent-plan-courses“ 2.0

2.4 Challenge Program

2.4.1

TiDB TiDB TiDB TiDB

TiDB PingCAP —— TiDB Challenge Program TiDB Issue Issue

2

season	period
Challenge Program season 2	2020.03~2020.05
Challenge Program season 1	2019.11~2020.02

2.4.2



TiDB Performance Challenge 1 2 3 600

- .* Team15050
- niedhui4300 catror3500
- pingyu2600 Renkai2550 js000701800
- ekalinin1450 mmyj1050 AerisNan750 MaiCw4J650 Rustin-Liu650 koushiro650

TiDB SQL in/like TiTan GC GC PD API PD

2.4.3



TiDB Challenge Program [TiDB](#)

[TiDB Usability Challenge](#)

PingCAP AskTUG “TiDB” 2019.12.17 2020.01.12 1 20 2020 2 11 2020 2 20 5

2 2020 3 2 2020 05 30 3

10000 8000 6000 ()*

TiDBTiKVPD [TiDB](#)

2.4.4

TiDB Challenge Program ->->->->->->“” GitHub

2.4.5

PingCAP [TiDB](#) [TiDB](#)

2.5 PingCAP Incubator

PingCAP Incubator CNCF TiDB TiDB TiDB

PingCAP Incubator TiDB TiDB “->->” Feature Project

PingCAP Incubator TiDB TUG TiDB TUG TiDB Community PMC

2020 3 PingCAP Incubator

- TiDB 4.0 TiDB Dashboard(TiDB Key Visualization,)
- TiDB - TiDB In Action48 Book Rush
- TiDB - TiUP
- - Talent Plan Courses(TinySQL/TinyKV)
- AskTUG Discourse PG TiDB
- TiDB - TiDB Wasm

2.6 PingCAP University —— Embrace Unlimited Possibilities

PingCAP University PingCAP

PingCAP University

- TiDB DBA PCTAPingCAP Certificated TiDB Associate— PingCAP TiDB PCTPPingCAP Certified TiDB Professional— PingCAP TiDB
- TiDB Talent Plan

PingCAP PingCAP TiDB DBA TiDB TiDB

2.6.1 TiDB DBA

1.

TiDB DBA 2018 10 2019 1 2020 3 7 2000 11 400 PCTA PCTP TiDB

PingCAP University : Embrace Unlimited Possibilities



2.

- (1) TiDB
- (2) TiDB
- (3) TiDB NewSQL
- (4) PCTA PCTP

2.6.2 TiDB DBA

TiDB PingCAP University TiDB DBA TiDB DBA 2.0 2019 6

- 2.0
- 2.0
- 2.0 TiDB HTAP TiDB

- 2.0 1.0 Demo show

TiDB DBA 2.0

课程类别	培训课程	培训目标	考试认证
基础知识	<p>一、概述 【2课时】</p> <p>二、系统安装部署与管理 【3课时】</p> <p>三、使用管理手册 【2课时】</p> <p>四、生态工具 【6.5课时】</p> <p>五、备份恢复 【1课时】</p> <p>六、TiDB 监控与告警 【2课时】</p> <p>七、业务开发优化 【2课时】</p> <p>总结综述</p>	<p>初级 TiDB DBA</p> <p>需要学习并熟练掌握如下内容：</p> <p>TiDB 概论、架构简介 TiDB 集群管理概述及使用实操 TiDB SQL 与调优 TiDB 对象管理与维护 TiDB 同构和异构平台数据迁移解决方案及实操 TiDB 备份与恢复原理及实操 TiDB 集群监控报警及日志管理 TiDB 常见问题处理 TiDB 业务开发最佳实践及应用案例介绍 TiDB 性能 Benchmark 介绍</p>	PCTA: PingCAP Certified TiDB Associate
高级进阶	<p>八、TiDB 计算内核体系 【4课时】</p> <p>九、TiKV 存储内核体系 【4课时】</p> <p>十、事务 【4课时】</p> <p>十一、系统性能优化 【5课时】</p> <p>十二、SQL 优化 【8.5课时】</p> <p>十三、高可用及容灾 【3课时】</p>	<p>高级 TiDB DBA</p> <p>需要学习并熟练掌握如下内容：</p> <p>TiDB Server 原理 PD 原理 TiKV 原理及架构 深入 Coprocessor 原理 TiDB Server 性能调优 TiKV 性能调优 TiDB 分布式事务原理 TiDB SQL 实现原理 TiDB SQL 性能调优 TiDB 高可用原理、方案及实战</p>	PCTP: PingCAP Certified TiDB Professional

1.

PingCAP University TiDB DBA PCTA TiDB

2.

PCTA PCTP 6*8/=483

11 400 100

- ——
- TiDB——
- —— IT
- ——
- Binlog DM ——Mike
- TiDB——
- TiDB——
- B TiDB——
- PingCAP University ——



2.6.3 TiDB DBA

PCTA PCTP DBA TiDB PingCAP

PCTA PingCAP Certified TiDB Associate PingCAP TiDB PCTA PCTA TiDB

PCTP PingCAP Certified TiDB Professional PingCAP TiDB PCTP PCTP TiDB FeatureSQL Trouble Shooting
PCTP PCTA

PCTA PCTP TiDB 3.0

- [AskTUG](#)
- PingCAP University PCTA 18 20
- PingCAP University PCTA PCTP TiDB
- PCTA PCTP

2.7 AskTUG

2.7.1 AskTUG

“TiDB”TUG TiDB AskTUG 2019 8 TUG “”TiDB TiDB

AskTUG 2019 2 AskTUG 1500+ 1300+ 100+ 100+

The screenshot shows the AskTUG website interface. At the top, there's a navigation bar with links for '问答' (Questions), '博客' (Blog), 'TUG', '活动' (Activities), 'PingCAP University', and a search icon. Below the navigation is a header with '所有分类' (All Categories), '所有标签' (All Tags), and a '分类' (Category) button, which is highlighted in red. Other buttons include '最新' (Latest), '近期 (2)' (Recent 2), and '热门' (Hot). To the right is a '发新主题' (Post New Topic) button.

The main content area is divided into two sections: '分类' (Categories) on the left and '最新' (Latest) posts on the right.

分类 (Categories):

- 数据迁移与同步 (59 / 月, 1 近期): TiDB 生态工具, 包括 DM, Binlog, Lightning 等
- SQL 优化 (16 / 月): 优化相关, 包括系统优化、SQL 优化等
- TiSpark & TiFlash (5 / 月): TiSpark & TiFlash 相关的问题和讨论
- TiDB 系统架构与管理 (98 / 月): 包括: TiDB 架构、原理、部署、升级、系统优化等
- TiDB 技术文章 (11 / 月, 1 近期): TiDB, 数据库, 云原生等相关技术文章和最佳实践分享
- TiDB 活动 (4 / 月): TUG 活动信息、活动资料分享
- AskTUG Weekly (4 / 月): AskTUG 每周精华问答、文章、活动信息汇总
- Cloud TiDB (3 / 月): 包括: TiDB - Operator、K8s 等

最新 (Latest) Posts:

- 为了证明它的速度, 我们一口气对比了 Oracle、MySQL、Greenplum、Apache Spark、MariaDB ColumnStore 和 Oracle..... (1 月前, 25 天前)
- 上游mysql通过DM同步到tidb从库的DDL lock 问题 (1 月前, 1 天前)
- [BUG] max函数和json_extract函数一起用的时候, 取值不稳定 (1 月前, 1 天前)
- tidb后续哪个版本才能支持TiSpark直接写Tidb集群呢? spark thriftserver怎么用JDBC啊 (1 月前, 2 天前)
- tidb在双控场景应用的疑问 (1 月前, 2 天前)
- beeline 连接 tispark thriftserver Set spark.tispark.write.allow_spark_sql=enable 之后还是不能写入数据到tidb (1 月前, 25 天前)

2.7.2 AskTUG

AskTUG 80% UGC

- AskTUG TiDB AskTUG 100% TiDB
- TiDB
- TiDB TUG AskTUG AskTUG
- AskTUG Weekly

TUG AskTUG “”“” TiDB TiDB

2.7.3

- AskTUG TiDB AskTUG AskTUG
- AskTUG AskTUG

2.7.4 AskTUG

AskTUG “”“” 2020 AskTUG

- PingCAP PingCAP “”
- AskTUG POC SOP Case Study FAQ
- AskTUG UI
- AskTUG “”

TUG AskTUG

2.8 Contribution Map

2.8.1 Contribution Map

TiDB 2020.3.7 GitHub 22.7K star 400 contributors TiDB TiDB contributor

[Contribution Map](#)

[Contribution Map](#)

- Contribution Map contributor
- Contribution Map What I Can Contribute contributor
- Contribution Map Learning Materials contributor
- Contribution Map SIGSpecial Interest Group SIG tech-lead mentors SIG TiDB community TiKV community SIG

2.8.2 Contribution Map

Contribution Map TiDB MySQL HTAP TiKV KV

A map that guides what and how contributors can contribute

Table of Contents

- [TiDB is an open-source distributed HTAP database compatible with the MySQL protocol](#)
- [TiKV, distributed transactional key-value database](#)
- [PD Placement driver for TiKV](#)
- [TiKV Clients](#)
- [Libraries depended by TiKV](#)
- [Ecosystem Tools: DM Data Migration Platform](#)
- [Ecosystem Tools - Binlog : A tool used to collect and merge tidb's binlog for real-time data backup and synchronization](#)
- [Ecosystem Tools - Lightning: A high-speed data import tool for TiDB](#)
- [Ecosystem Tools - BR: A command-line tool for distributed backup and restoration of the TiDB cluster data](#)
- [TiDB on K8S/Docker : Creates and manages TiDB clusters running in Kubernetes](#)
- [Deployment Tools - tidb-ansible: A tool to capture data change of TiDB cluster](#)
- [Chaos-Mesh: A Chaos Engineering Platform for Kubernetes](#)
- [Documentation](#)
- [AskTUG\(CN\)](#)
- [PingCAP University\(CN\)](#)
- [SIG - Special Interest Group](#)

TiKV, distributed transactional key-value database

Module	Description	Code Directory	Required Skills	Learning Materials	What I can Contribute	Contributing Tutorials
Util	Utilities like thread-pool, logger, encoding and decoding, etc.	Utilities, Pipeline batch system	Rust	Rust book, Practical networked applications in Rust, Protocol buffers, TiKV source code reading: service layer(CN) , gRPC concepts	Issues want help	Land your first Rust PR in TiKV, Became TiKV Contributor in 30 minutes(CN)
Network	Network layer	Server	Rust, Protobuf, gRPC	Ditto	Ditto	Ditto
Raw KV API		API entrance, Storage struct	Rust	Ditto	Ditto	Ditto
Transaction KV API		API entrance, Implementation of Transaction , GC worker, Pessimistic transaction	Rust, 2PC, Percolator transaction model	Two Phase Commit, Percolator paper , TiKV source code reading : storage(CN) , TiKV source code reading : distributed transaction(CN) , TiKV source code reading MVCC read(CN)		
Multi-raft		Raftstore	Rust, Raft	Rust book, Practical networked applications in Rust, Raft paper , Raft implementation in etcd		
Engine		Engine traits, Engine rocks	Rust, RocksDB		Engine abstraction	
Coprocessor		TiDB query, TiDB query codegen, TiDB query datatype	Rust	TiKV source code reading : Coprocessor Overview(CN) , TiKV source code reading : Coprocessor Executor	Coprocessor Issues	
Backup		Backup source code				

[Contribution Map](#) contributors

