

TiDB跨数据中心解决方案

林豪翔
庄天翼
朱贺天
屈鹏



Background

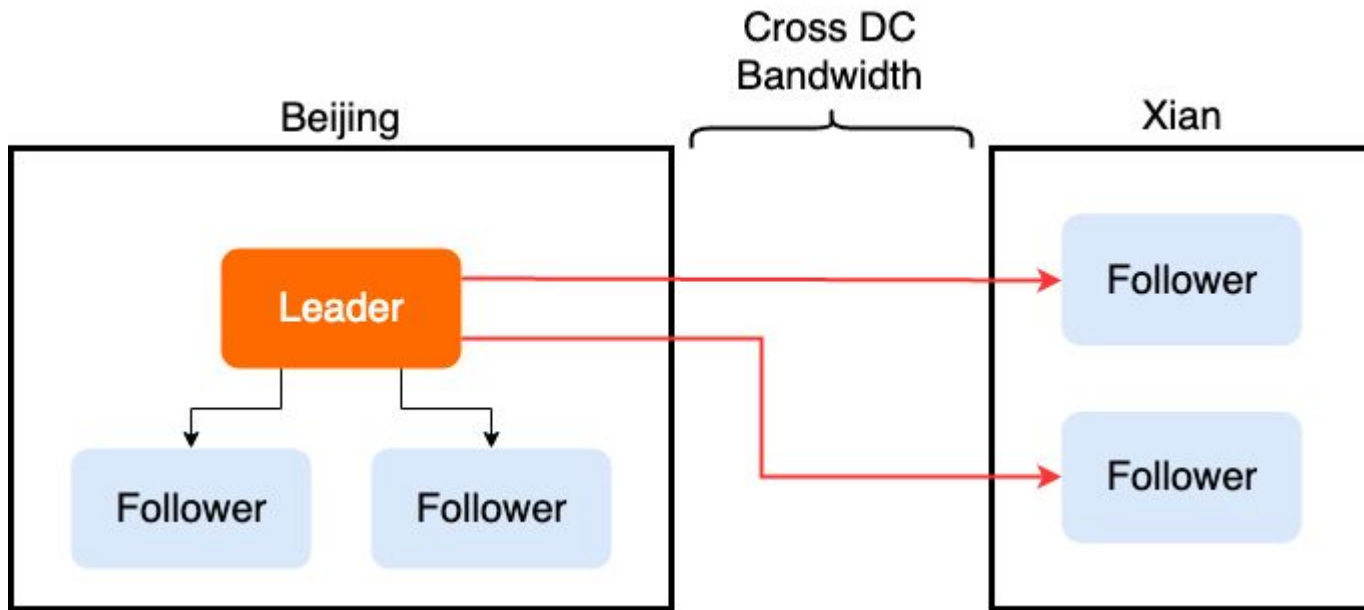
- 两地三中心拓扑在行业内十分常见
- 业务上可能有多个地区的用户需要读取数据
- 跨地区的 RTT 较高
- 跨DC带宽昂贵



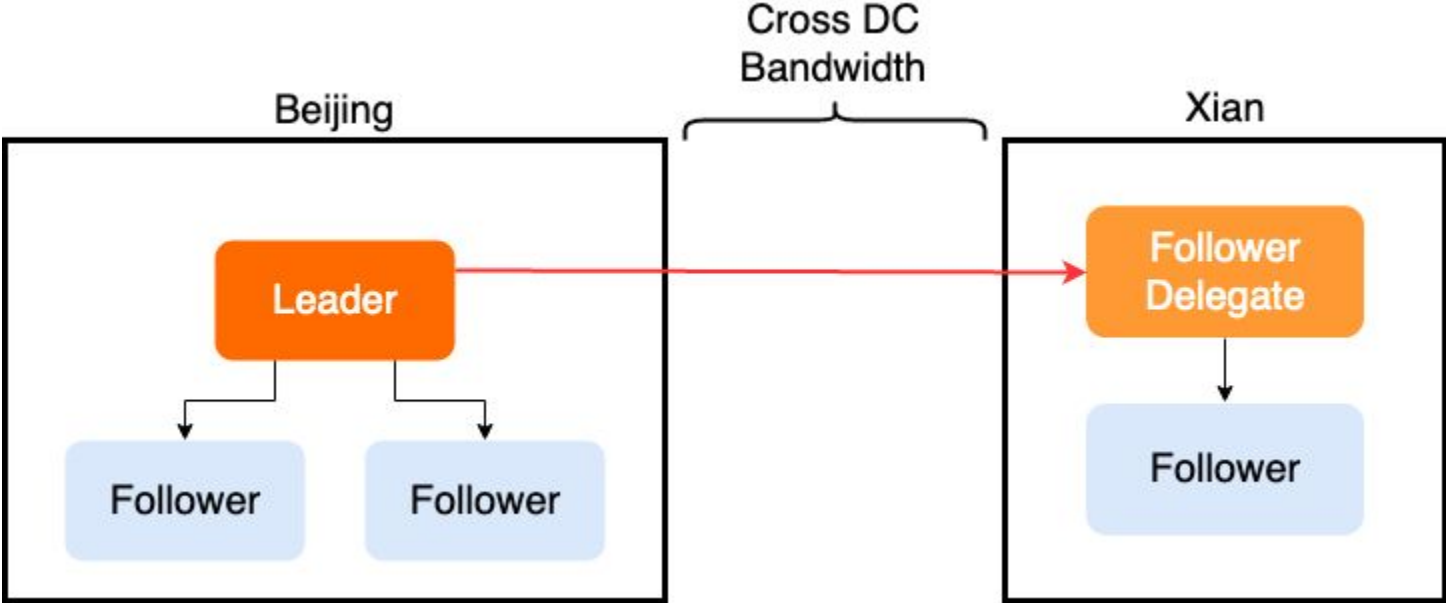
Improve

- **Follower Replication**
- Read improvement

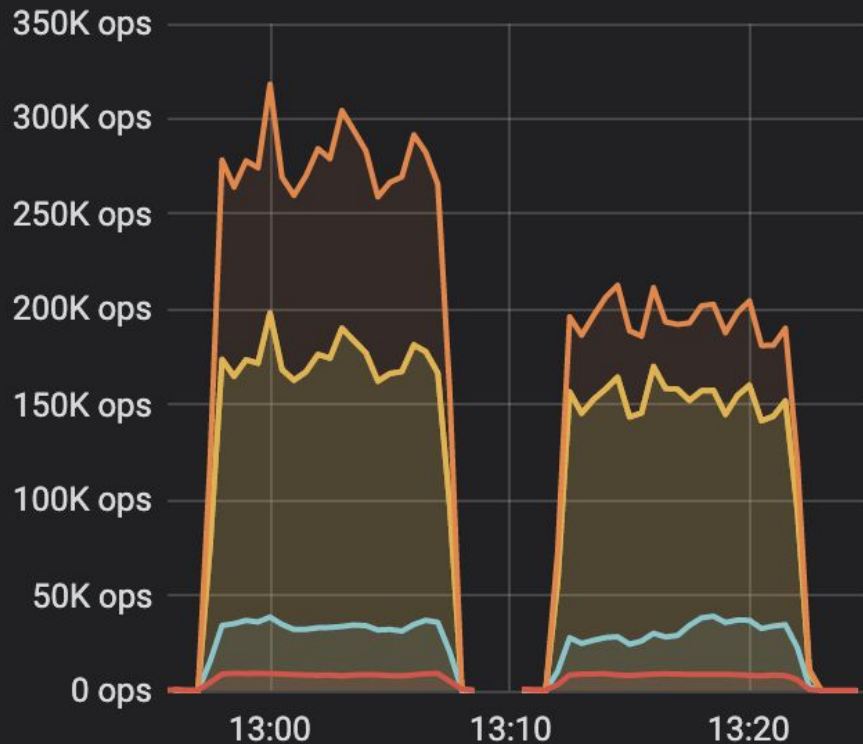
Current



After our improvement

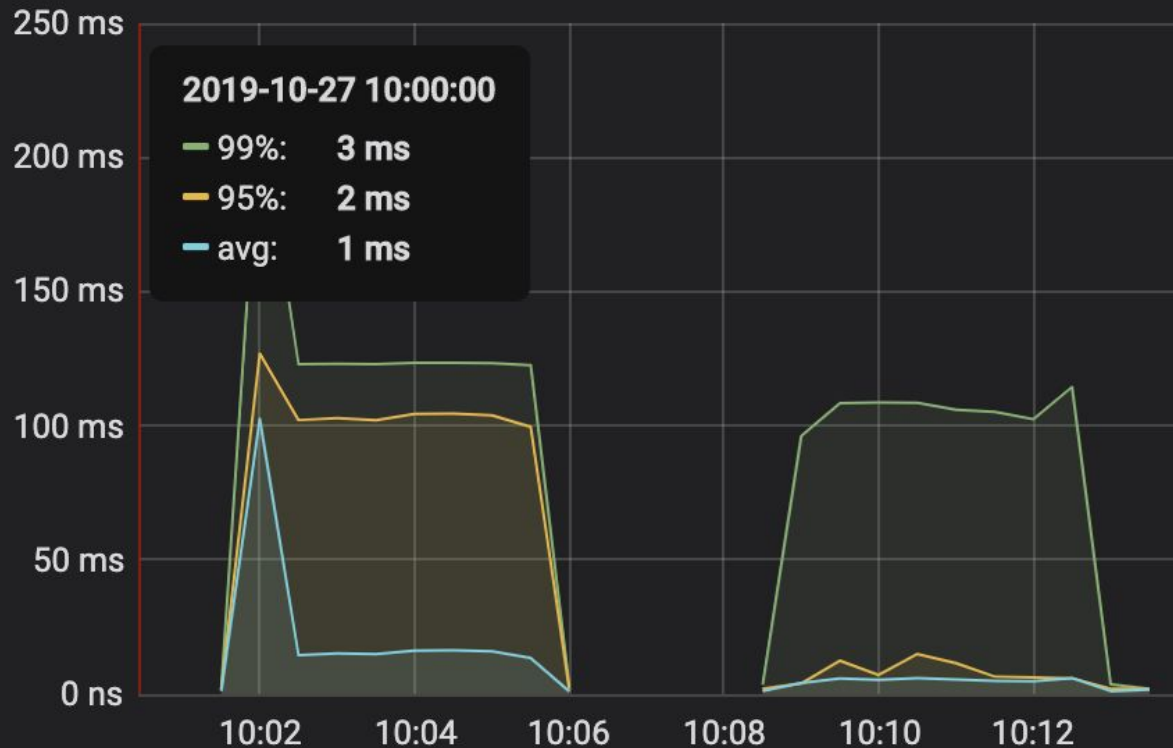


Ready handled ▾



	max	current ▾
count	9.1K ops	87.9 ops
message	318.1K ops	18.6 ops
has_ready_region	39.1K ops	12.2 ops
append	198.1K ops	0.0 ops
commit	198.1K ops	0.0 ops

Commit log duration



	max	current
99%	217 ms	2 ms
95%	127 ms	2 ms
avg	103 ms	2 ms

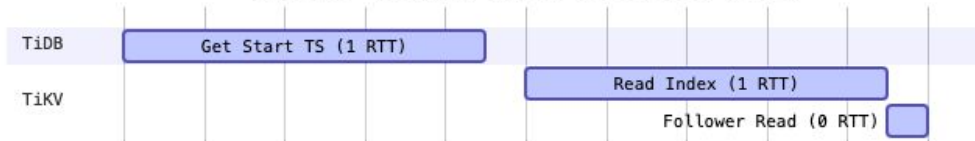
Improve

- Follower Replication
- **Follower Read**

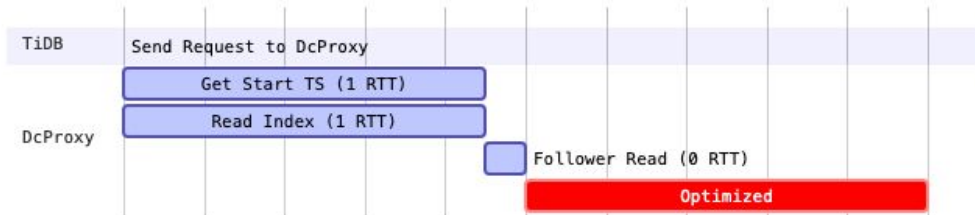
master branch without follower read



master branch with follower read



our branch with concurrent optimize

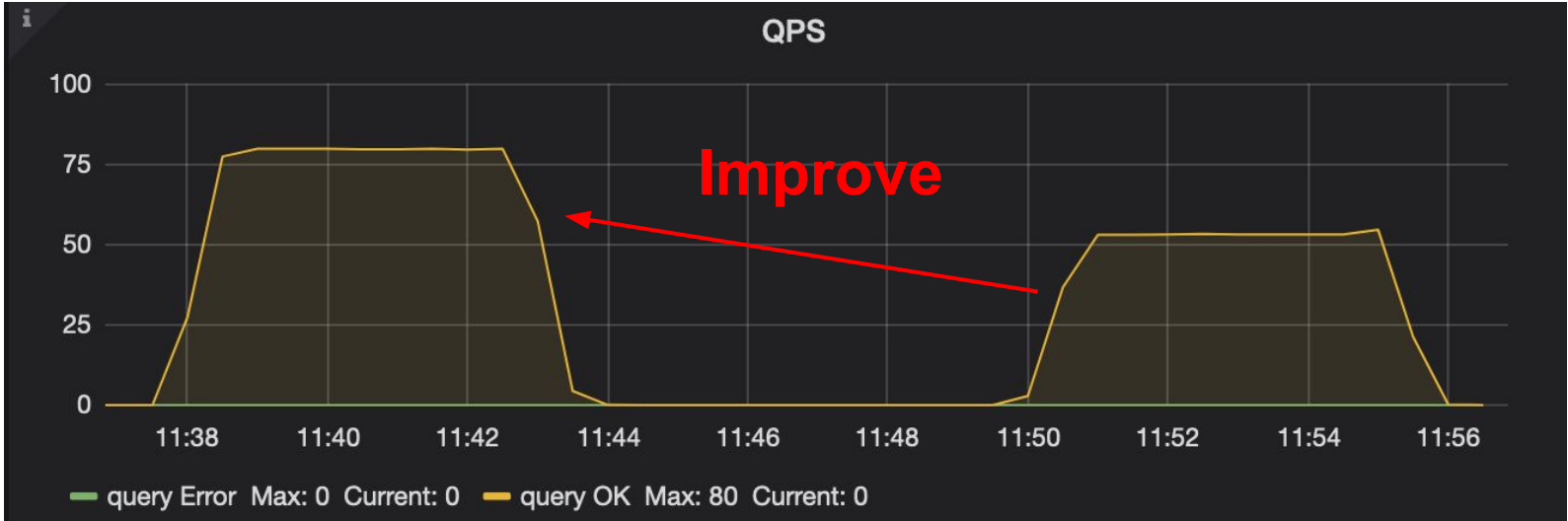


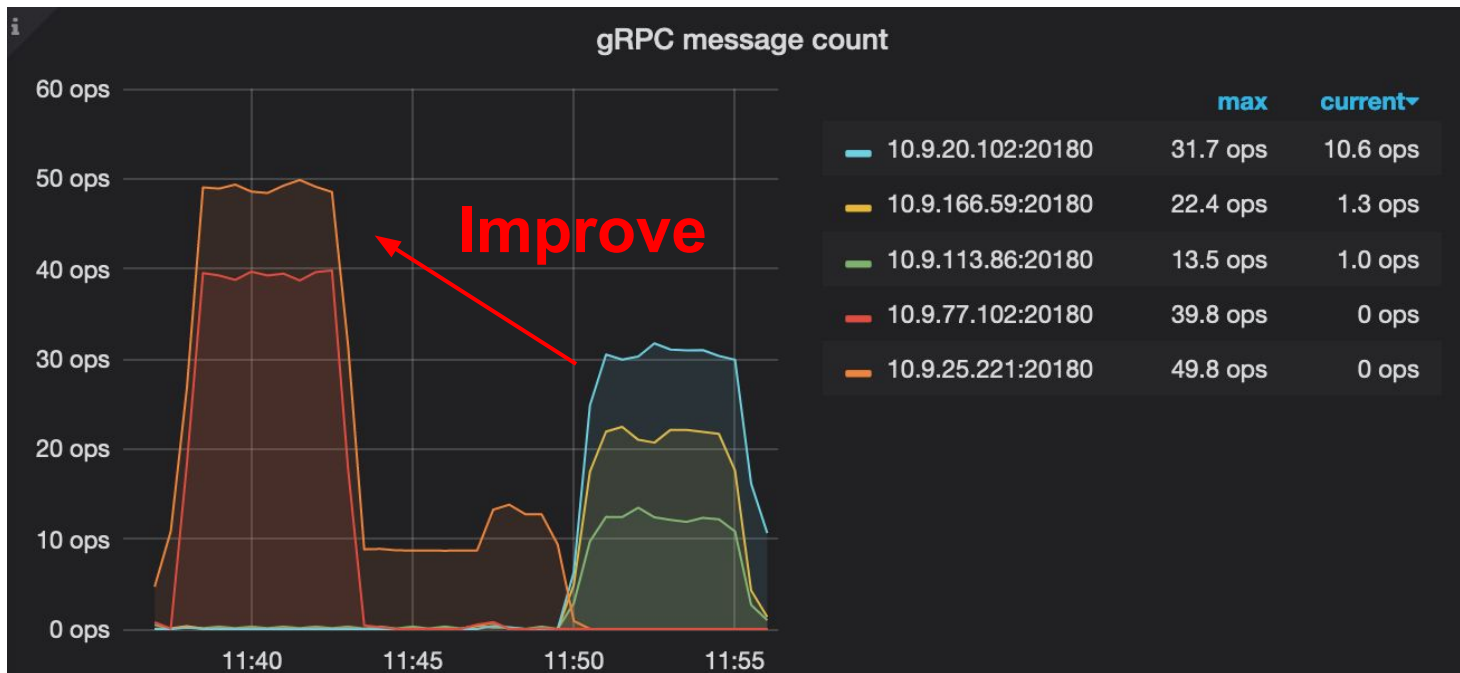
```
MySQL [sbtest]> select count(*) from sbtest1 where id = 2;
+-----+
| count(*) |
+-----+
|         1 |
+-----+
1 row in set (0.20 sec)
```

↓
Improve

```
MySQL [sbtest]> select count(*) from sbtest1 where id = 2;
+-----+
| count(*) |
+-----+
|         1 |
+-----+
1 row in set (0.10 sec)
```







Future Work

- 对交互式事务进行更多优化, 将现在的 n 个 RTT 优化为 1 个 RTT。
- 把 Follower replication 的功能提交到 etcd/raft。
- 利用 Data Center 的功能实现隔离区功能, 部分 DataCenter 的资源用于一些特殊的 workload 的请求。

Q & A

