去哪儿网数据库架构变迁

目录

去哪儿MySQL架构

对Galera集群的疑问

Galera架构

Galera复制写集

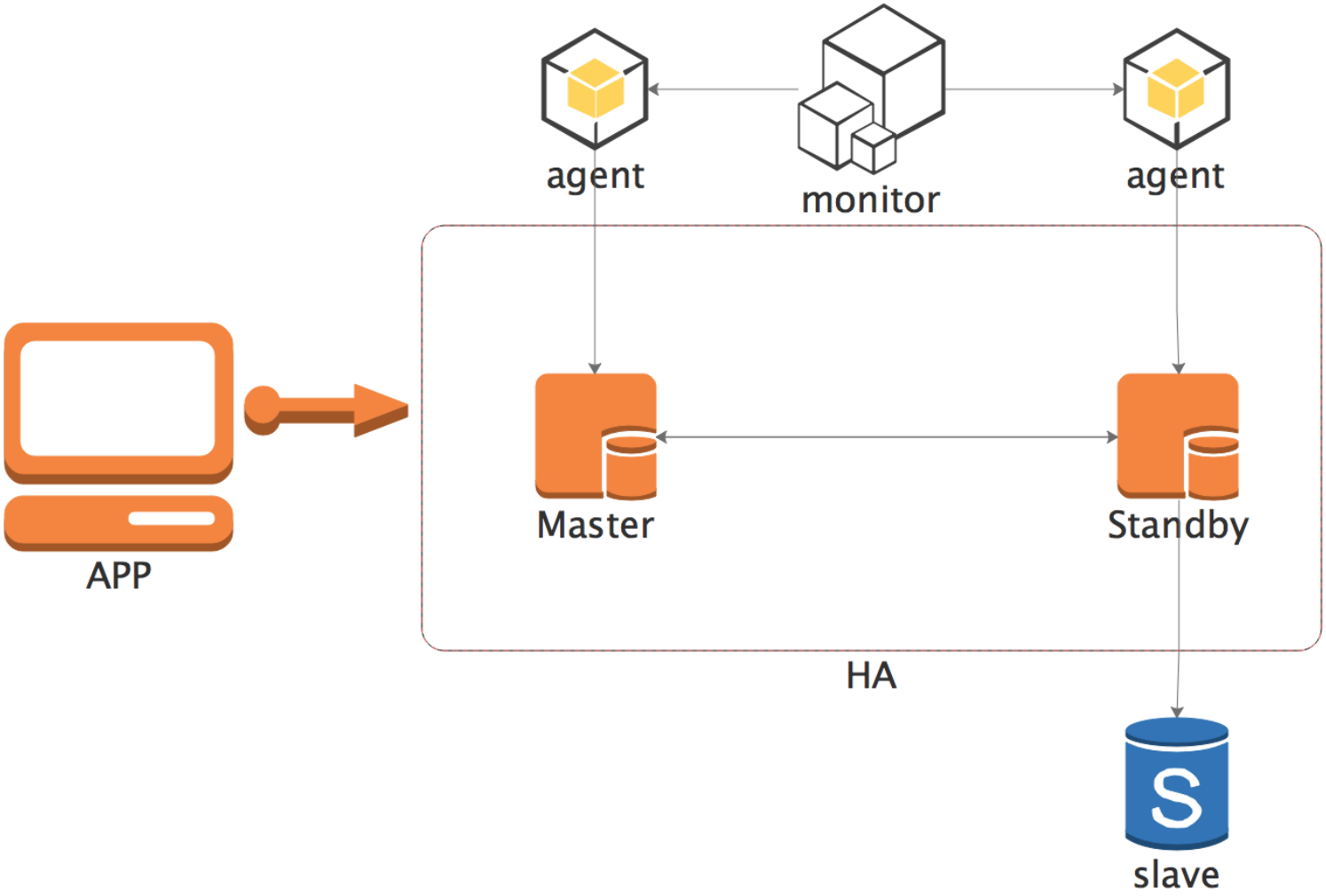
Galera的并发控制

Galera有啥坑?

架构升级

终极架构

3m

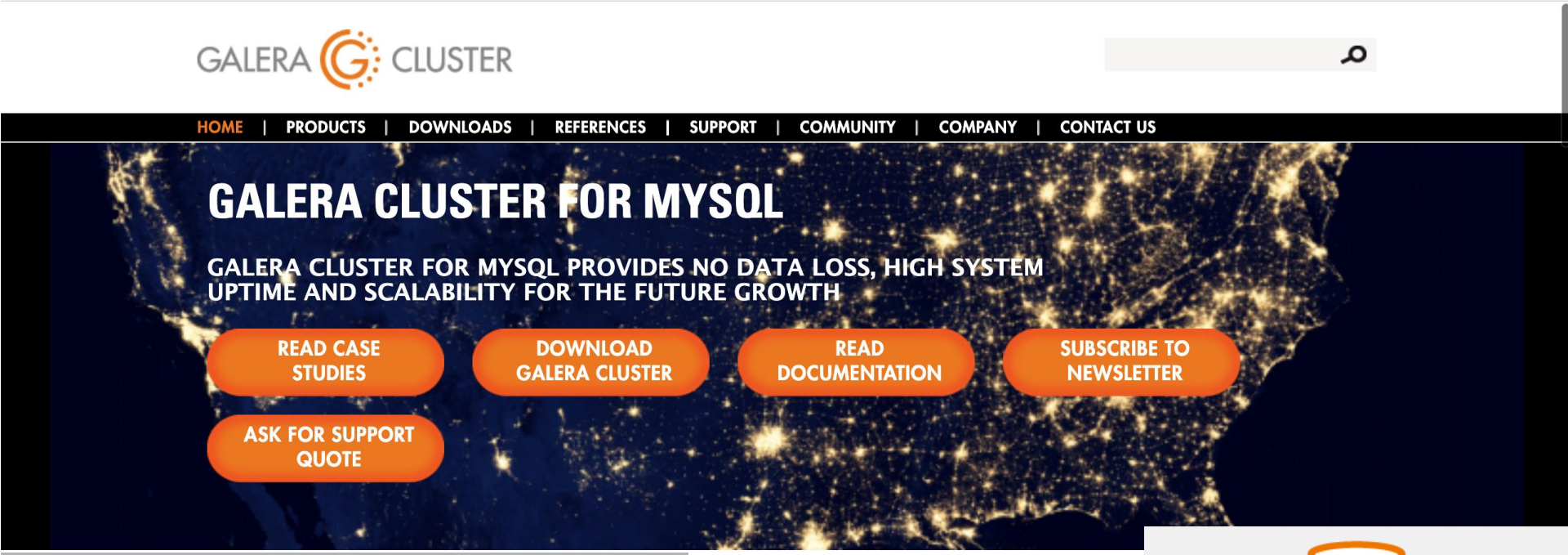


Recommend
火龙果软件

Qmha

Galera

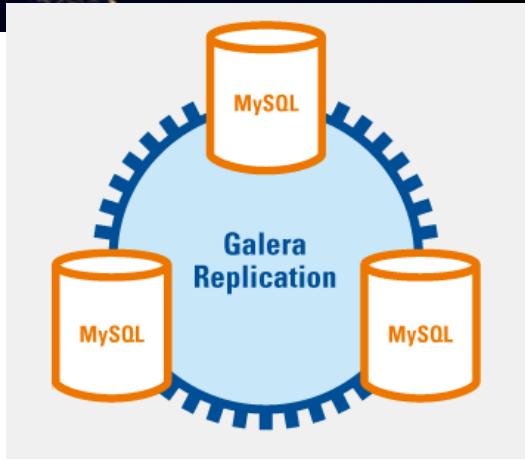
Galera Cluster introduction



集群化大趋势

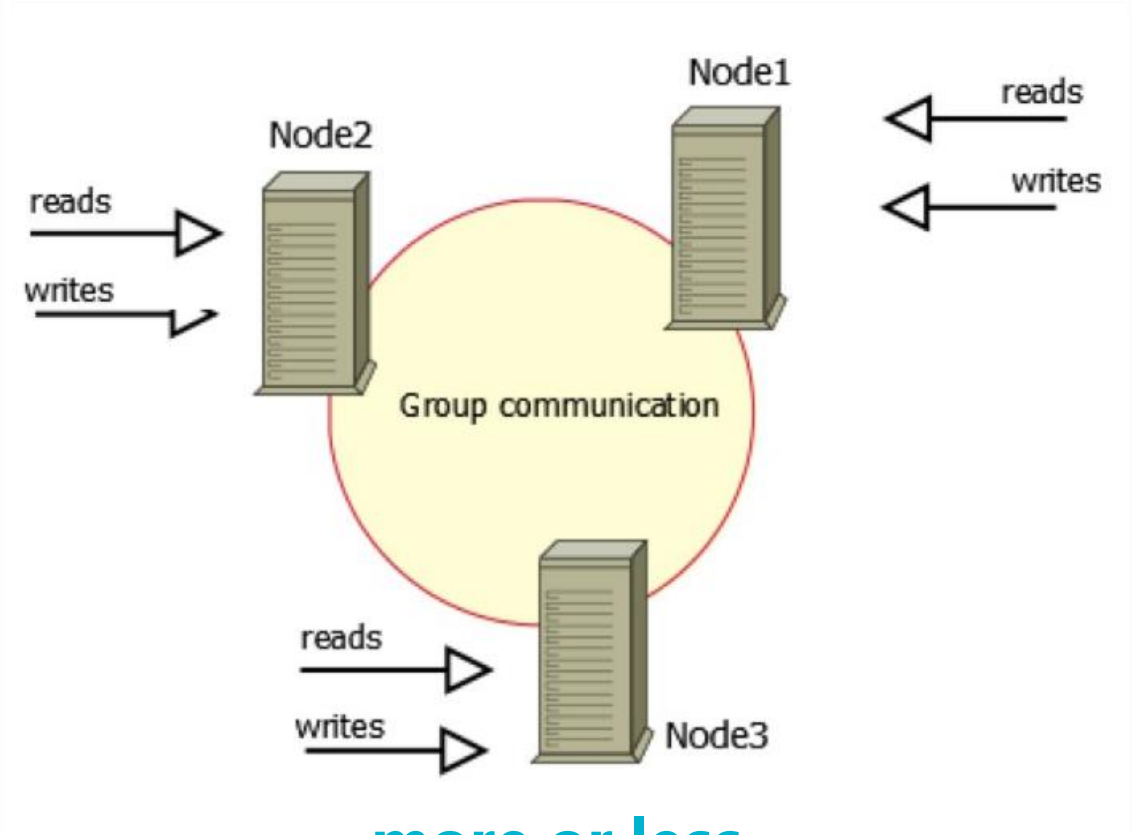
何谓galera cluster?

高一致，低延迟，多点写入



Recommend
火龙果软件

Architecture



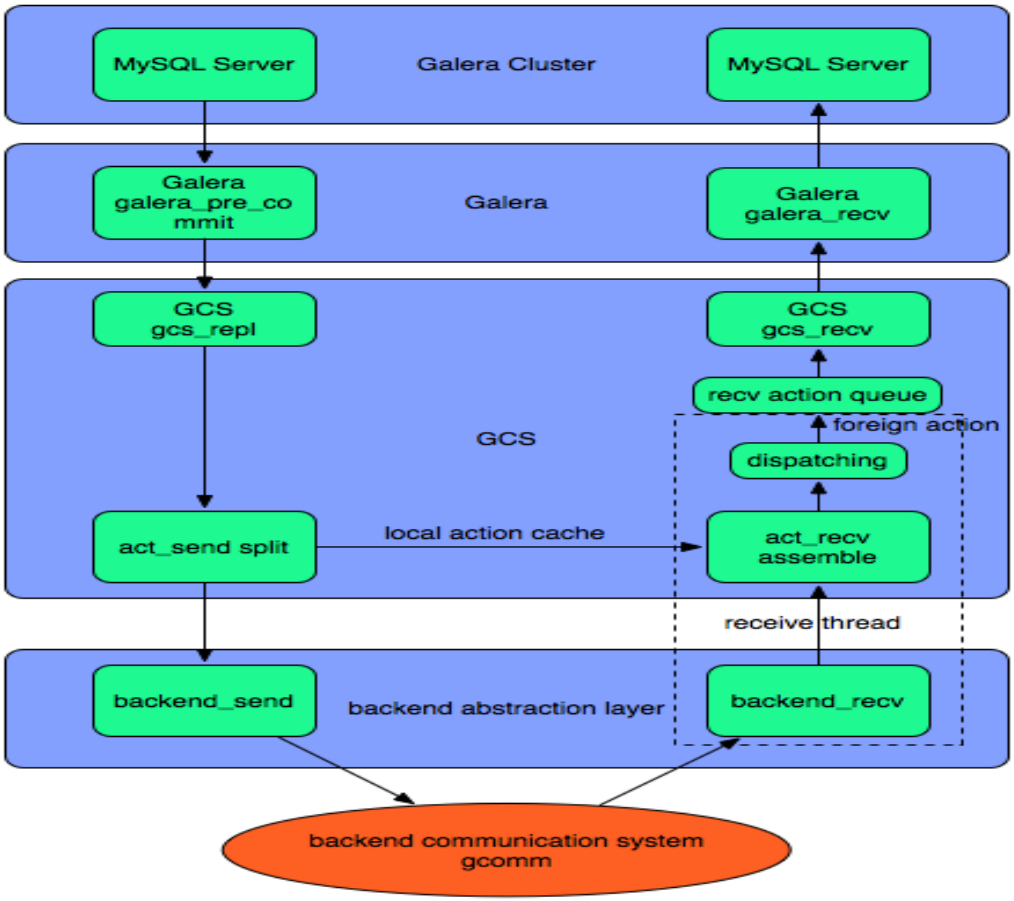
more or less

- multi-master
- synchronous replication
- No replication relay
- concurrency control
- switch over as will
- state transfer automatically
- so-easy for dba

Recommend
火龙果软件

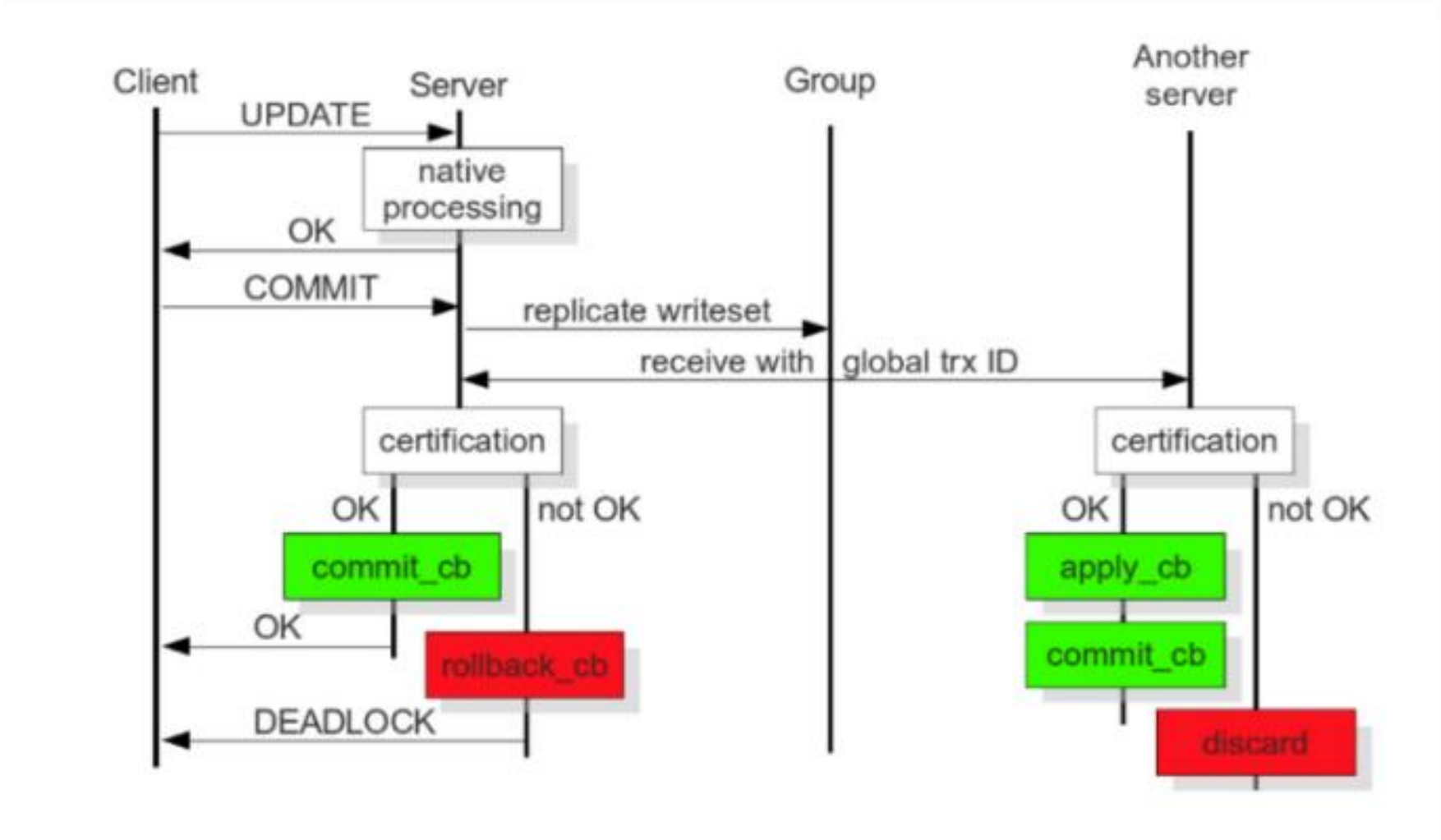
Architecture

- Server
- galera API
- galera logic
- galera message
- network



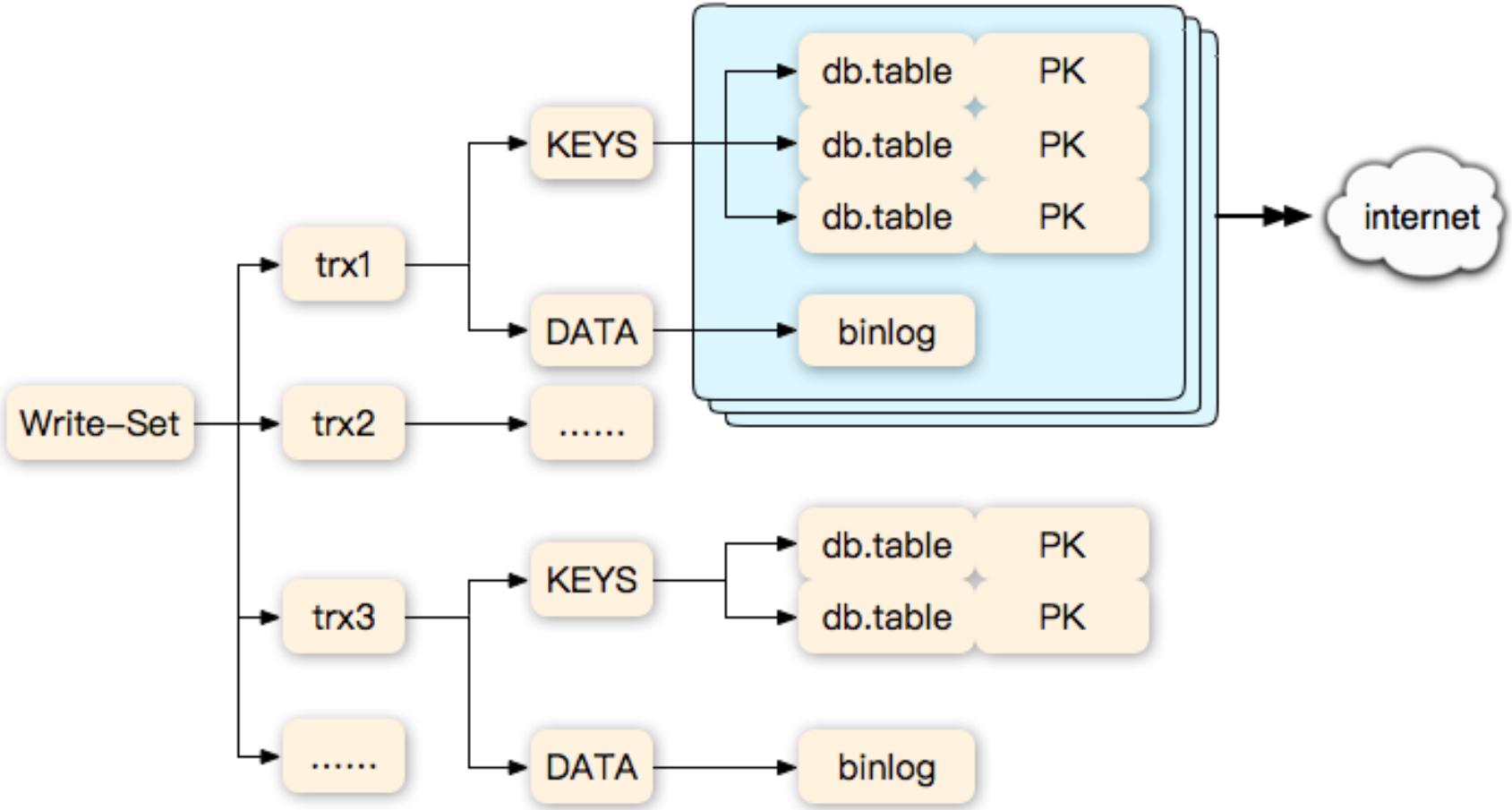
Recommend
火龙果软件

Architecture



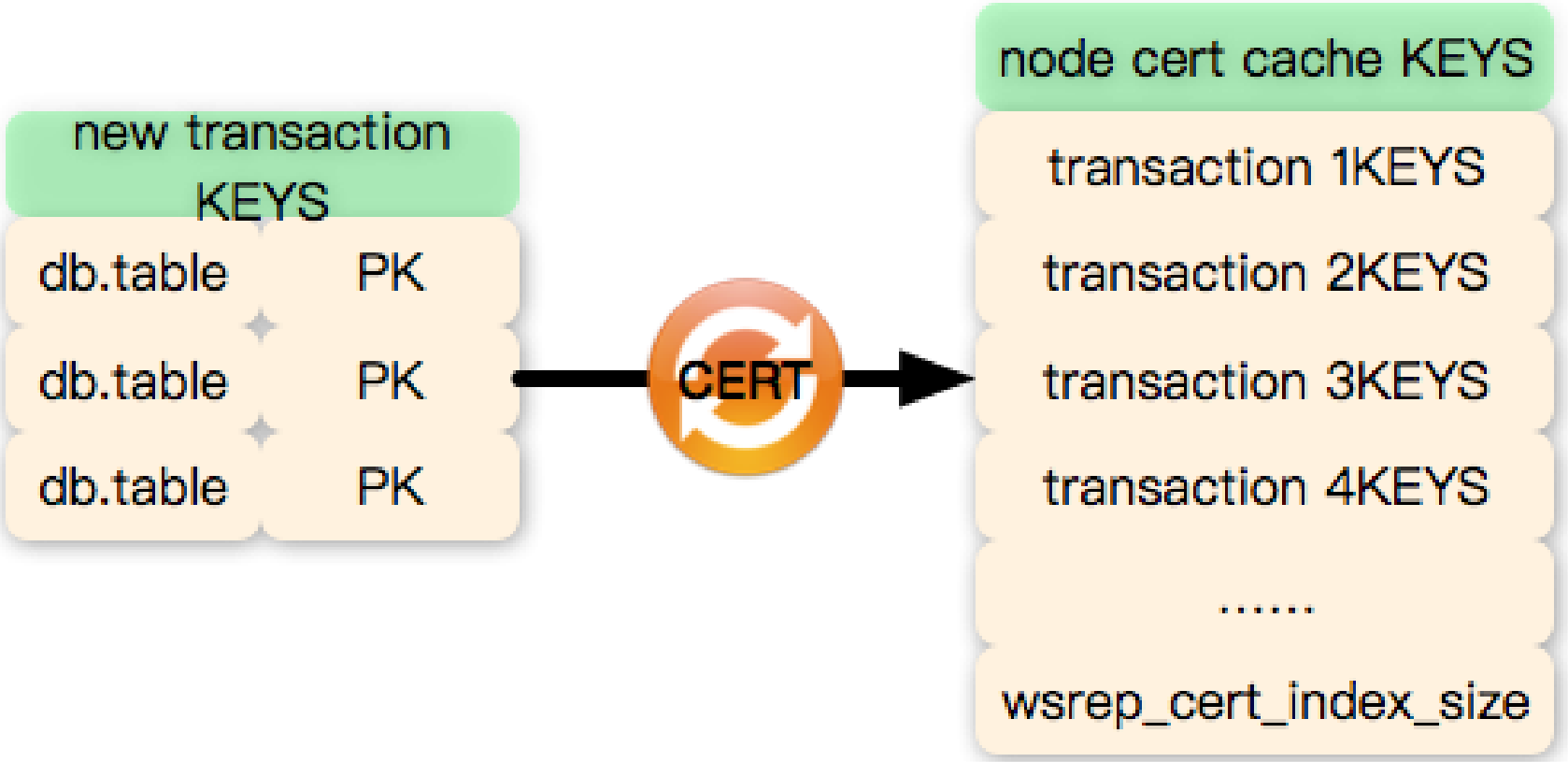
Recommend
火龙果软件

Galera Write-Set



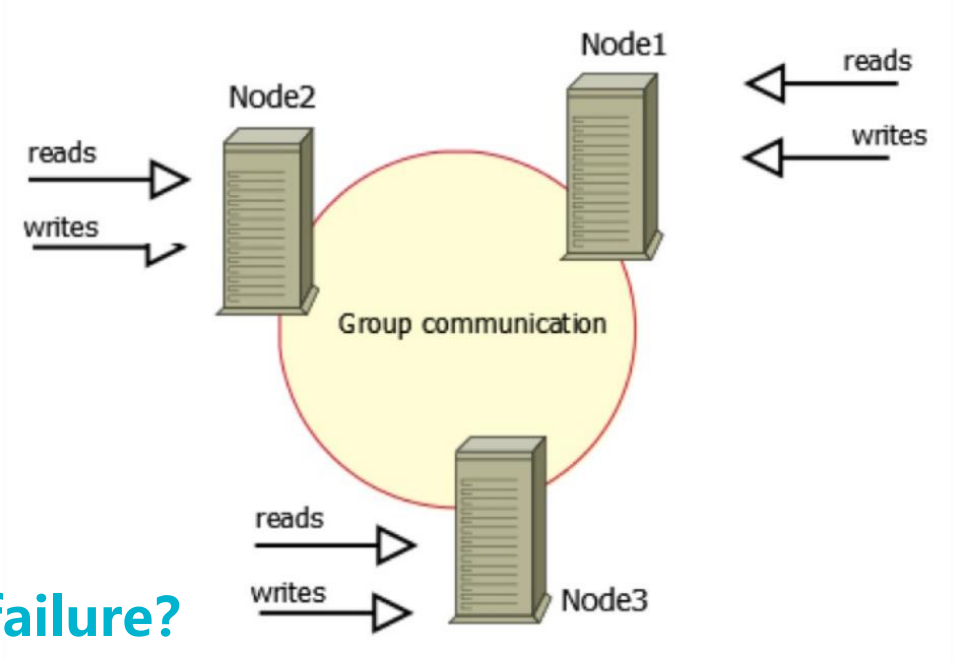
Recommend
火龙果软件

Galera Write-Set



Recommend
火龙果软件

Galera Write-Set

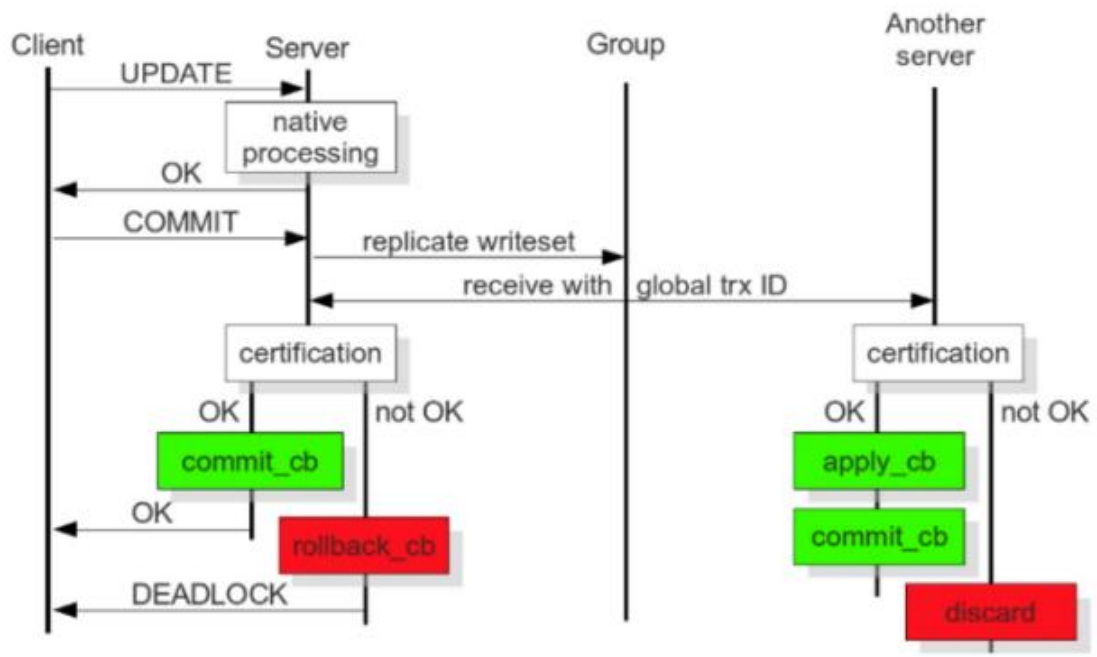


- why to certificate?
- what is certificating failure?
  - two Write-Sets come from different nodes
  - two Write-Sets have the same keys
  - the older GTID is invisible to new transaction

Recommend
火龙果软件

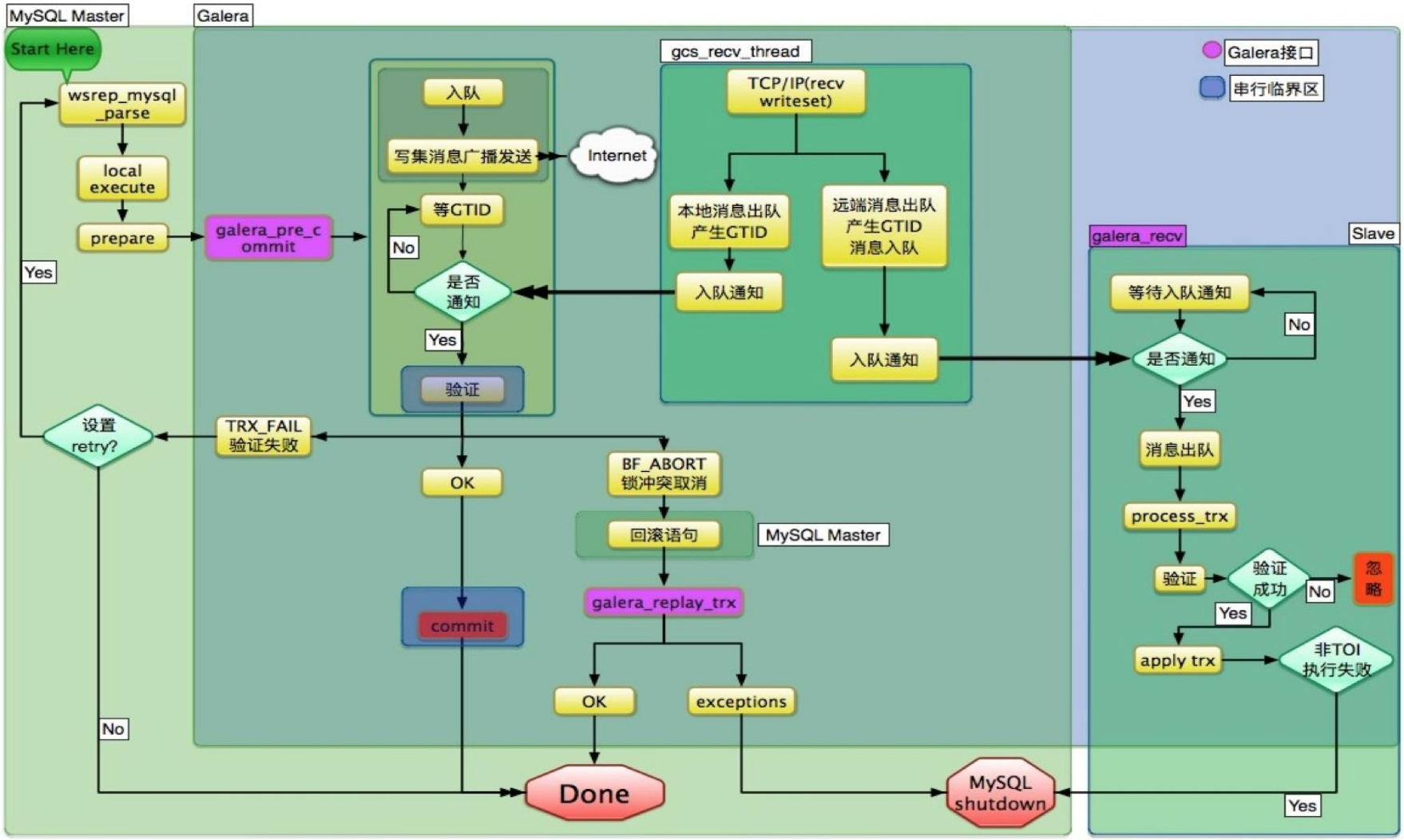
Galera Write-Set

- local certification, replicate, which first? any questions?
- how to deal with certification failure?
- and execution ?



Recommend
火龙果软件

Galera Write-Set



Recommend
火龙果软件

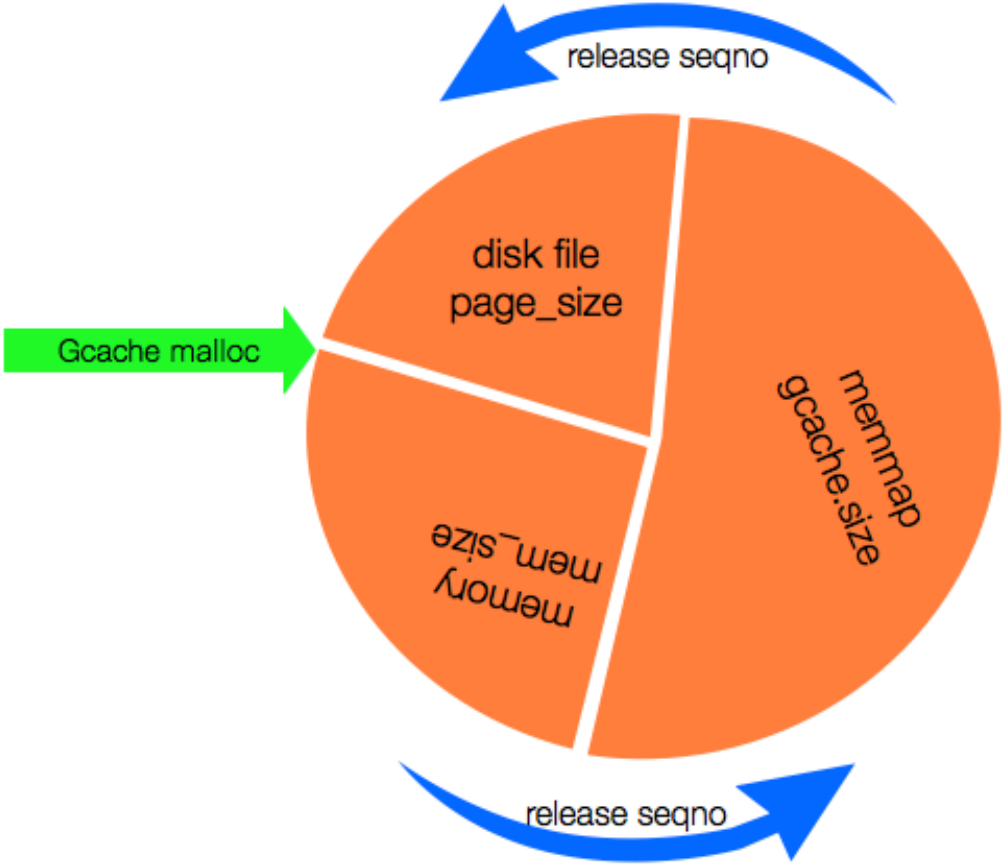
Galera Concurrency Control

- write-set replication
- write-set local certification
- write-set slave apply
- write-set commit

Recommend
火龙果软件

Inside Galera Gcache

- definition
- function
- state transfer
- what size to set
- influence
- theory
- HA



Recommend
火龙果软件

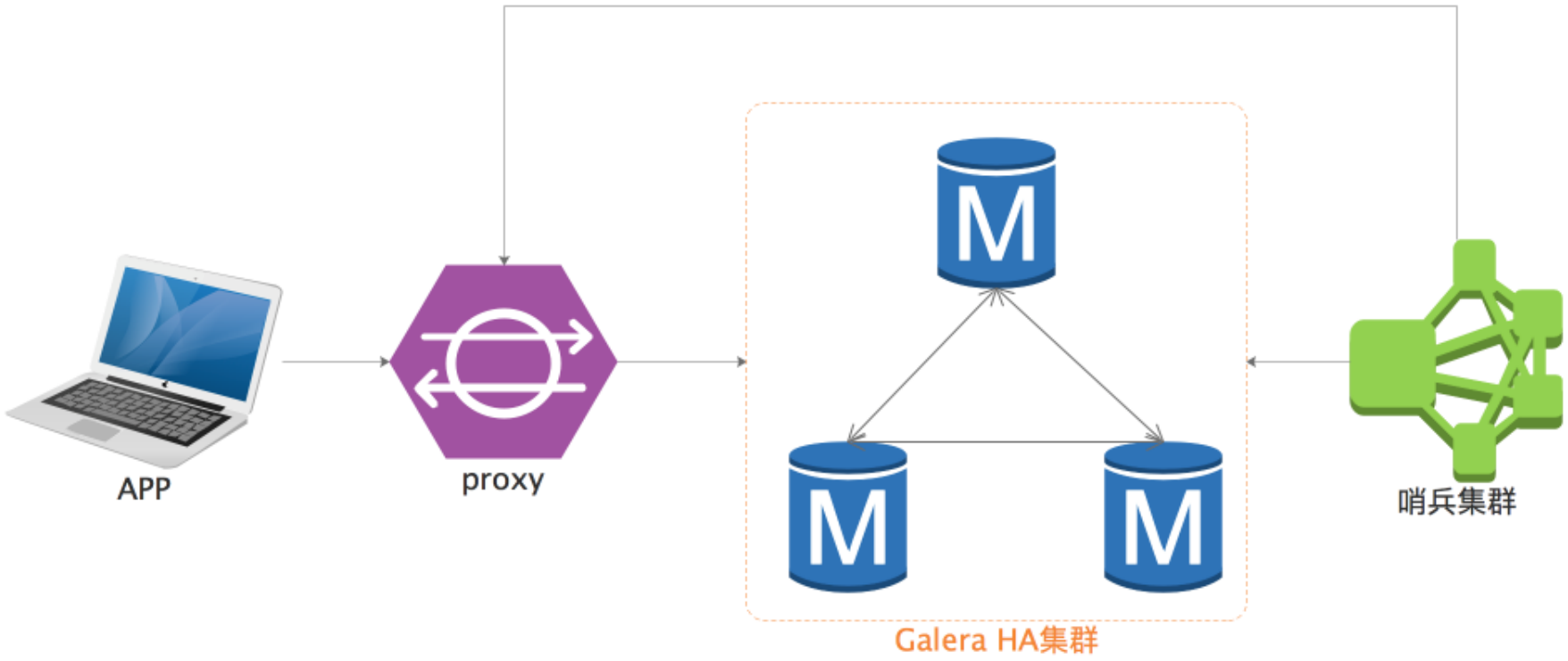
Have many problems?

- DDL result in whole cluster hang ?
- DDL result in deadlock?
- how to cancel DDL execution?
- flow control
- Big Transcation

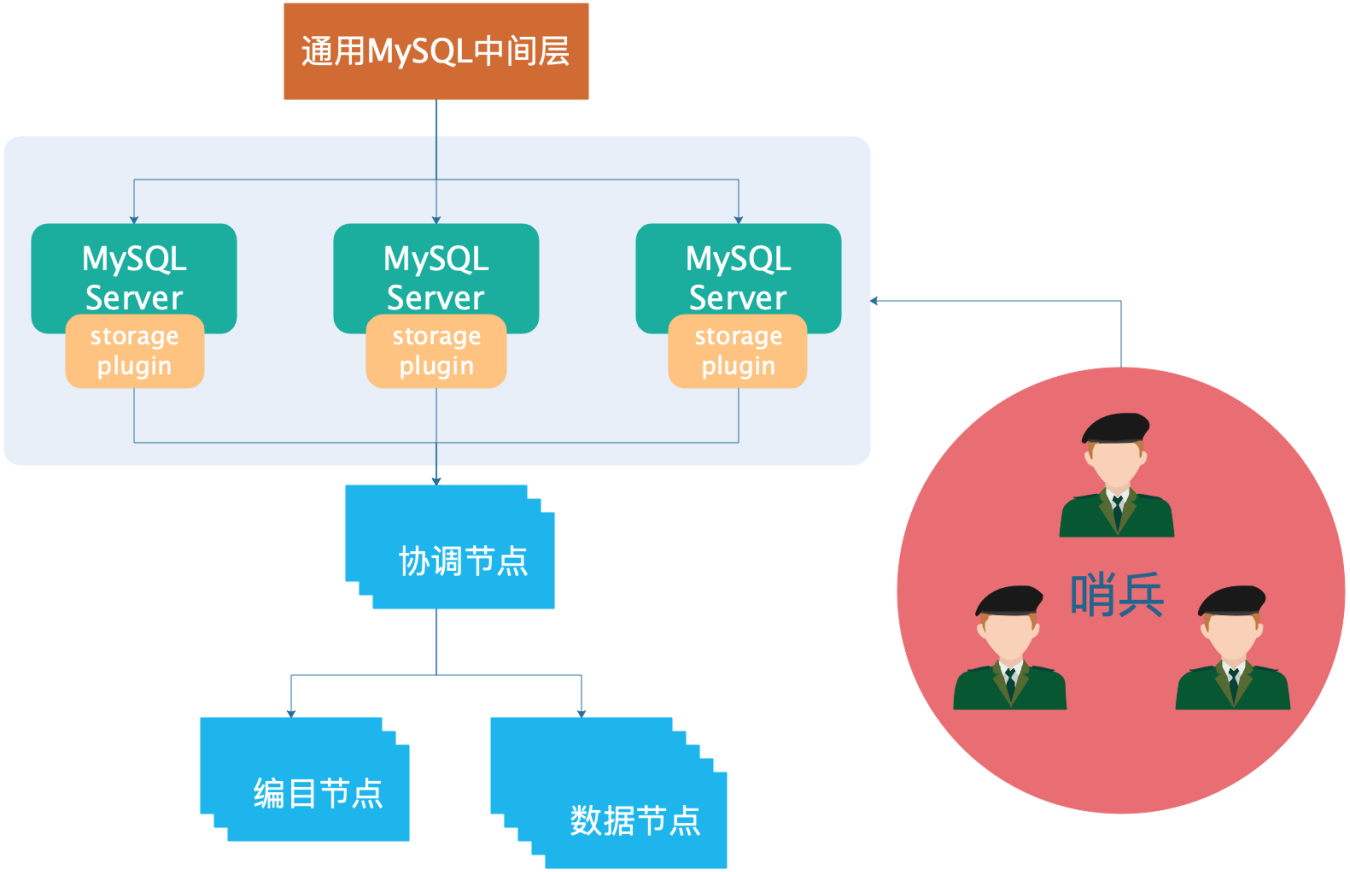
Galera cluster slow? really?

- galera concurrency control
- local certification
- wait GTID
- synchronous replication
- multi node write

架构升级



架构再升级



三思！大规模MySQL运维陷阱之基于MyCat的伪分布式架构——分享

终极架构理论及需求

通过计算，已经发现，一个只有4层的B+树，可以存储这么巨大的数据量，每层之间都是相差千倍，这棵树要有多矮，多胖啊！简单算了一下，如果把每条记录想象为1mm的正方体的话，把这些记录连起来，可以绕地球旋转6.4圈，太神奇了！

存储资源池

节省资源
(一台机器即可部署)

Redo并行复制

弹性扩缩容

秒级快照备份

分布式高可用

复制低延迟

存储与计算彻底分离
(具有分布式弹性扩容的SSD)

分布式
(现阶段最流行的做法)

高度兼容性
(无缝接入管理平台)

支持在线改表及OSC

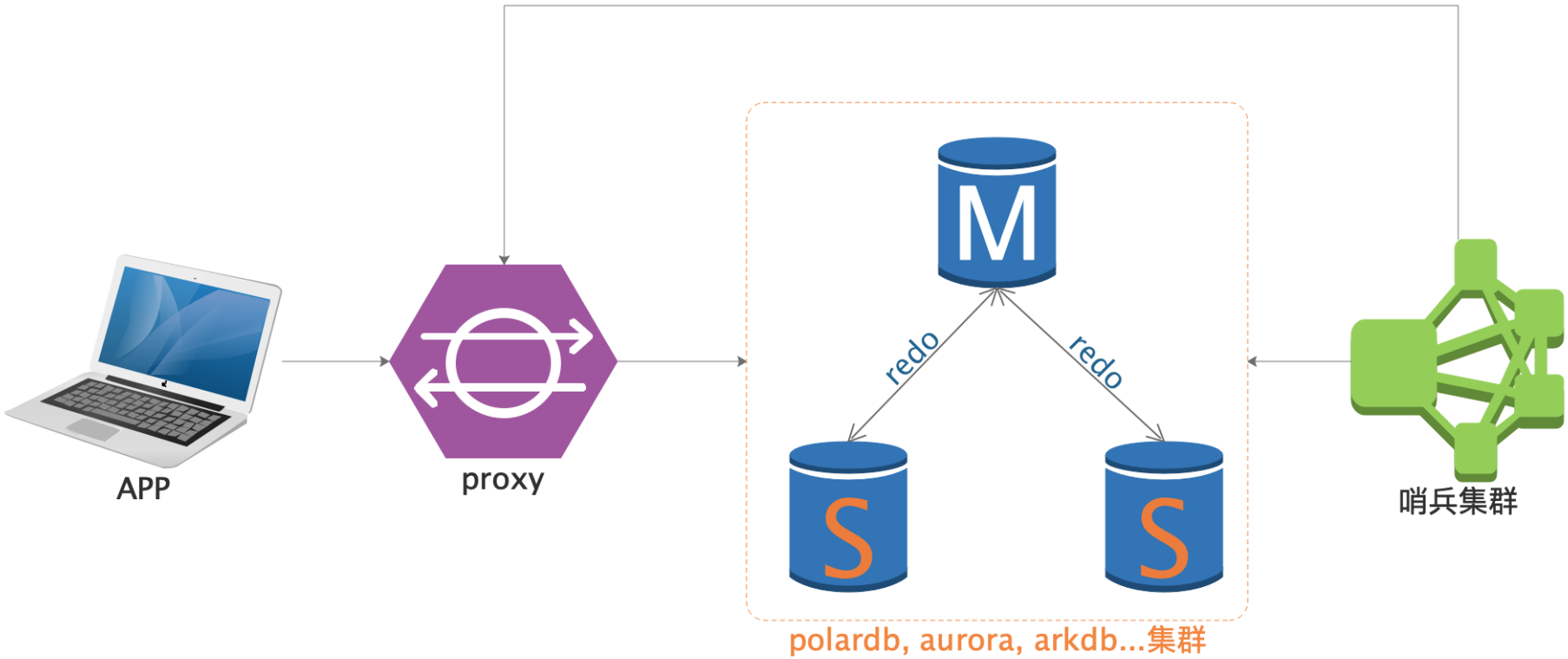
拒绝分库分表
(分库分表就是DBA给自己找罪受)

Recommend
火龙果软件

```
 5s] threads: 80, tps: 0.00, reads: 0.00, writes: 48818.31, response time: 2.67ms (95%), errors: 0.00, re
.00
 10s] threads: 80, tps: 0.00, reads: 0.00, writes: 49686.95, response time: 2.48ms (95%), errors: 0.00, re
.00
 15s] threads: 80, tps: 0.00, reads: 0.00, writes: 50453.83, response time: 2.49ms (95%), errors: 0.00, re
.00
 20s] threads: 80, tps: 0.00, reads: 0.00, writes: 48776.18, response time: 2.55ms (95%), errors: 0.00, re
.00
 25s] threads: 80, tps: 0.00, reads: 0.00, writes: 50965.59, response time: 2.42ms (95%), errors: 0.00, re
.00
 30s] threads: 80, tps: 0.00, reads: 0.00, writes: 51057.42, response time: 2.39ms (95%), errors: 0.00, re
.00
 35s] threads: 80, tps: 0.00, reads: 0.00, writes: 50526.01, response time: 2.37ms (95%), errors: 0.00, re
.00
 40s] threads: 80, tps: 0.00, reads: 0.00, writes: 49244.69, response time: 2.42ms (95%), errors: 0.00, re
.00
 45s] threads: 80, tps: 0.00, reads: 0.00, writes: 50468.29, response time: 2.44ms (95%), errors: 0.00, re
.00
 50s] threads: 80, tps: 0.00, reads: 0.00, writes: 49976.59, response time: 2.48ms (95%), errors: 0.00, re
.00
 55s] threads: 80, tps: 0.00, reads: 0.00, writes: 50484.99, response time: 2.41ms (95%), errors: 0.00, re
.00
 60s] threads: 80, tps: 0.00, reads: 0.00, writes: 51079.63, response time: 2.34ms (95%), errors: 0.00, re
.00
```

Recommend
火龙果软件

终极架构



Recommend
火龙果软件

THANK YOU !

Recommend
火龙果软件