

# 大型分布式系统案例实战 第12周

DATAGURU专业数据分析社区

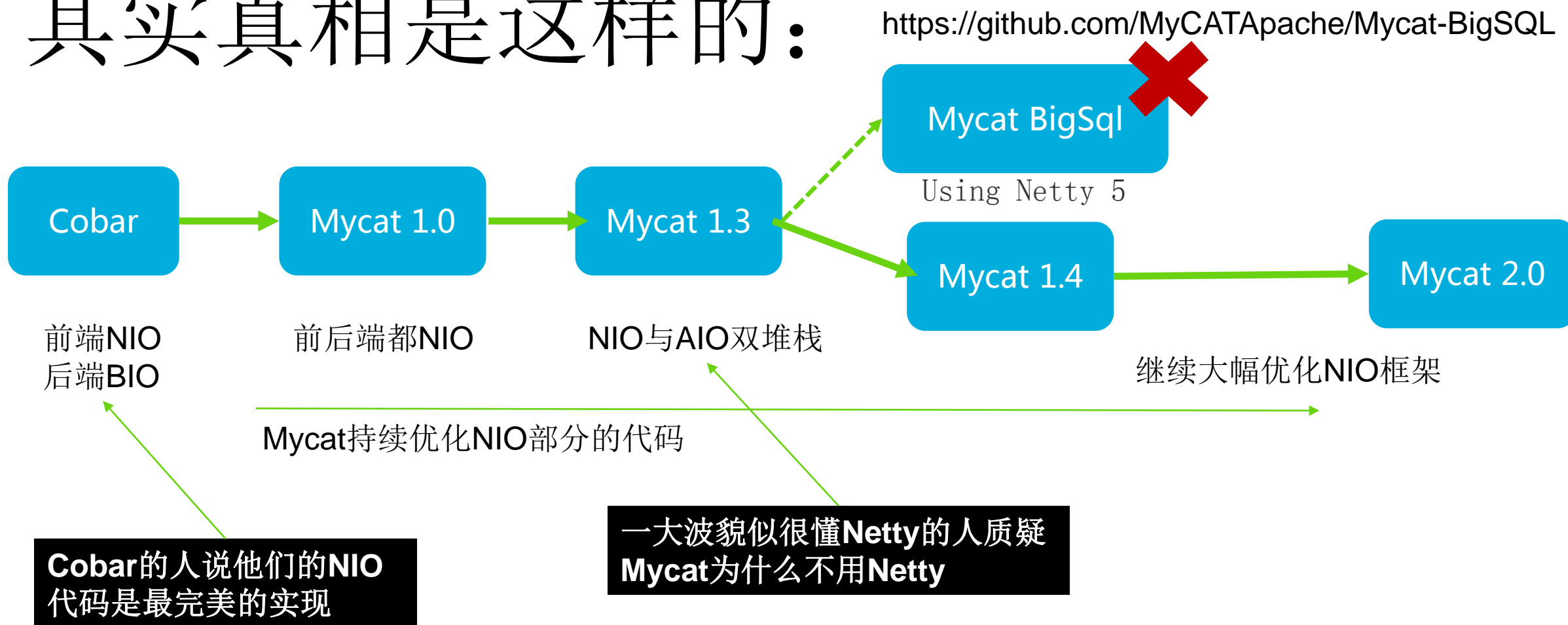
**【声明】** 本视频和幻灯片为炼数成金网络课程的教学资料，所有资料只能在课程内使用，不得在课程以外范围散播，违者将可能被追究法律和经济责任。

课程详情访问炼数成金培训网站

<http://edu.dataguru.cn>

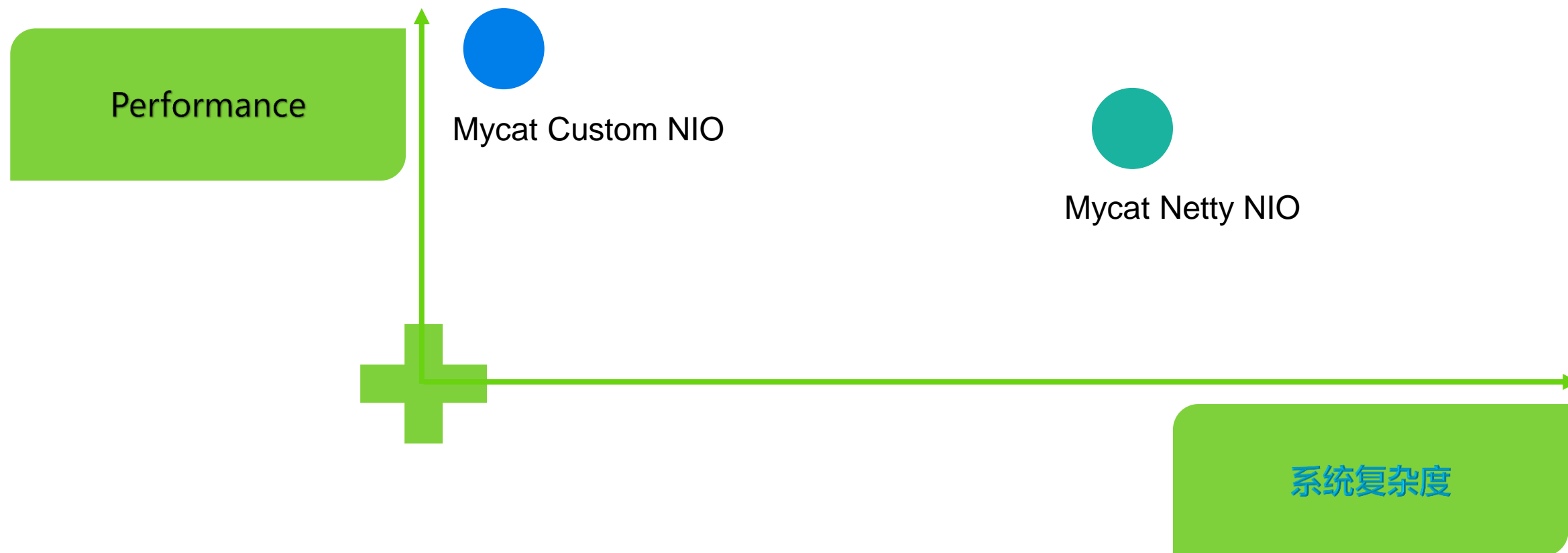
- 从NIO框架说起：为什么没有用Netty
- 从配置文件说起：最终选择Zookeeper的原因
- 从跨分片SQL Join问题说起：为什么这个问题仍然没有现成的答案
- 从HA问题说起：为什么要另外开发一个Mycat-Proxy

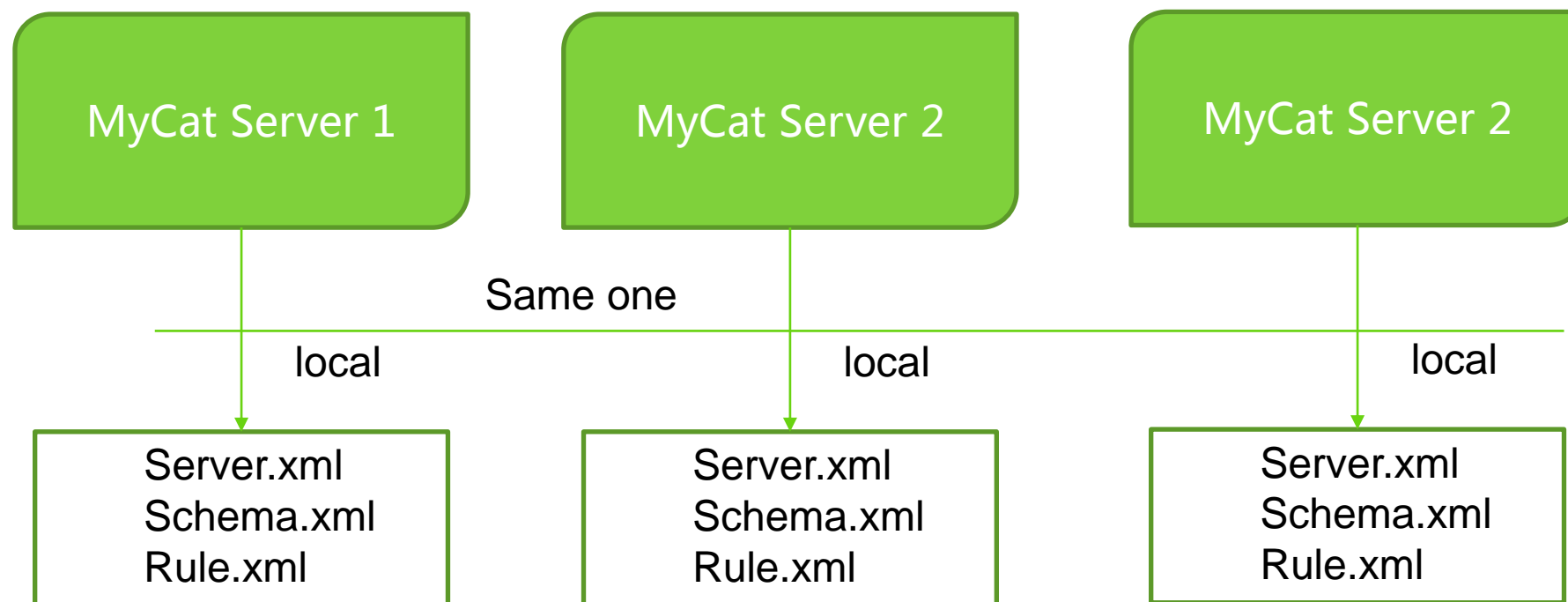
## 其实真相是这样的：

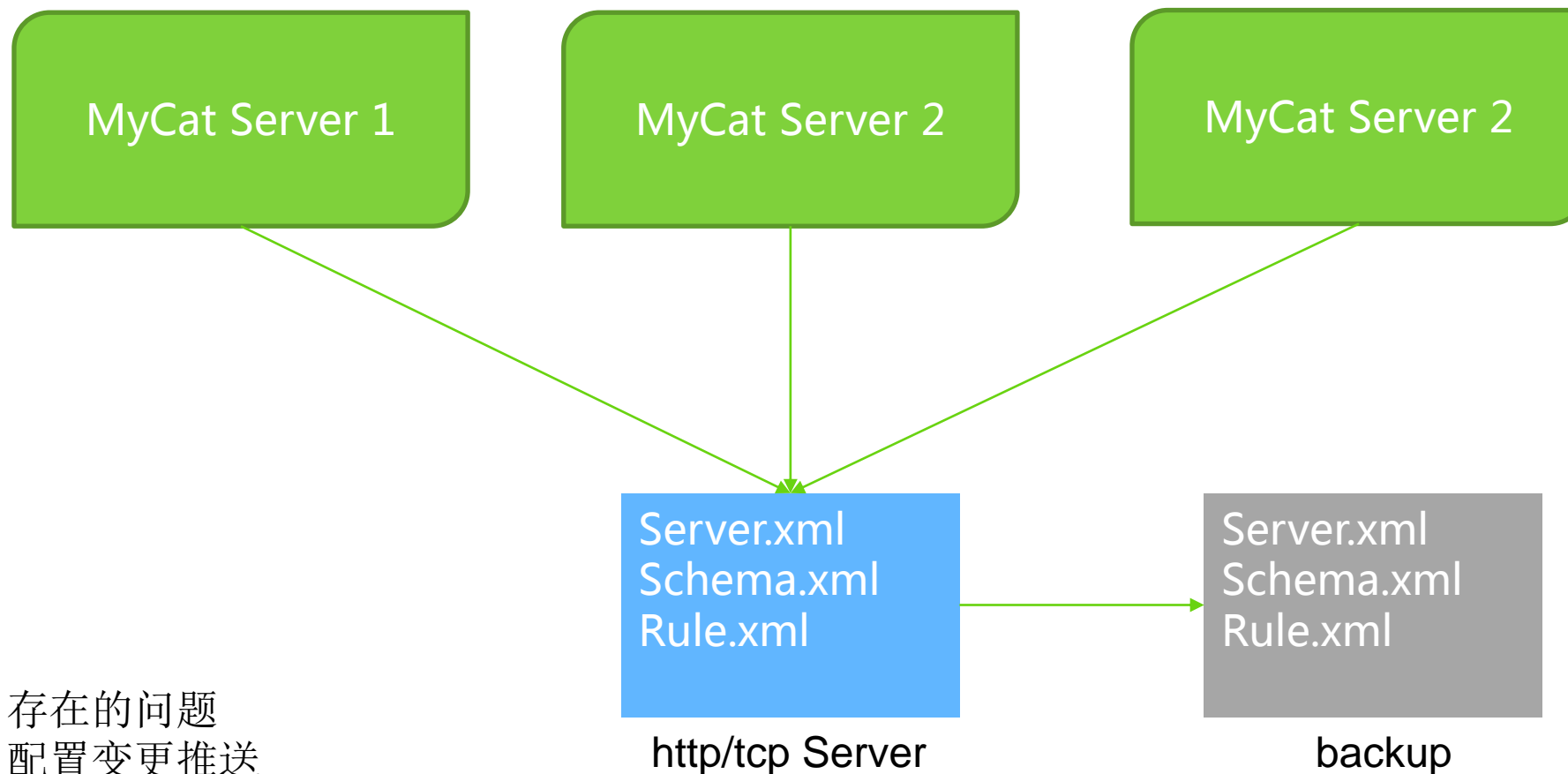


# 从NIO框架说起：为什么没有用Netty

## 关键考虑因素

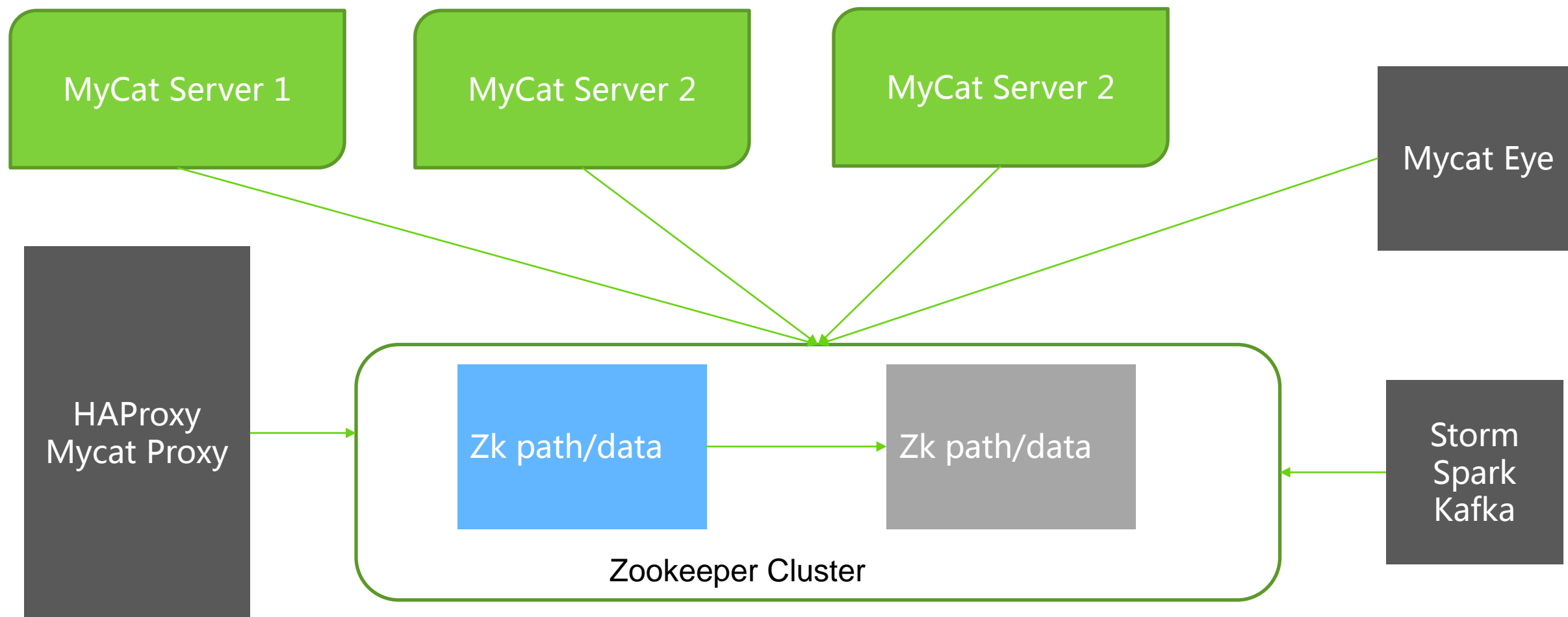






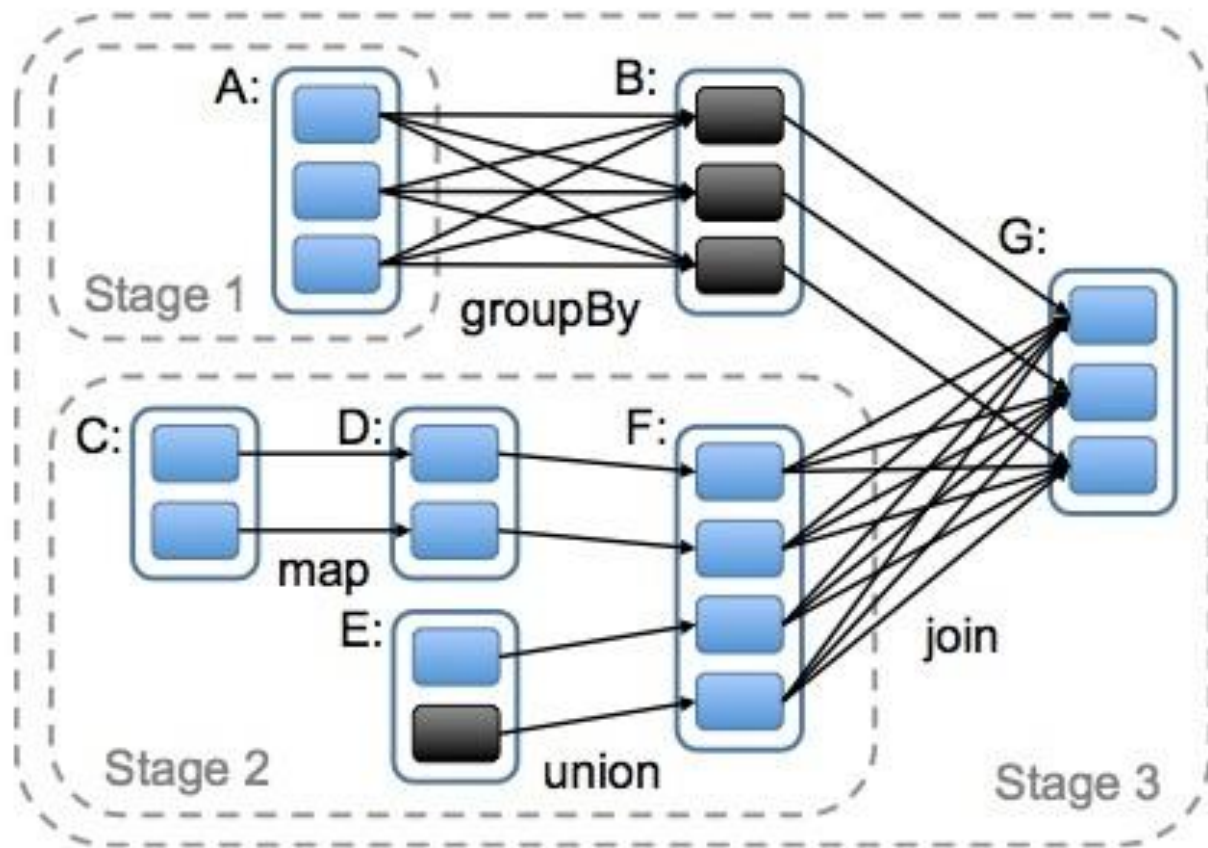
存在的问题  
配置变更推送  
非标准化的实现  
增加了运维和跟其他系统对接的代价

# Mycat配置文件的变迁



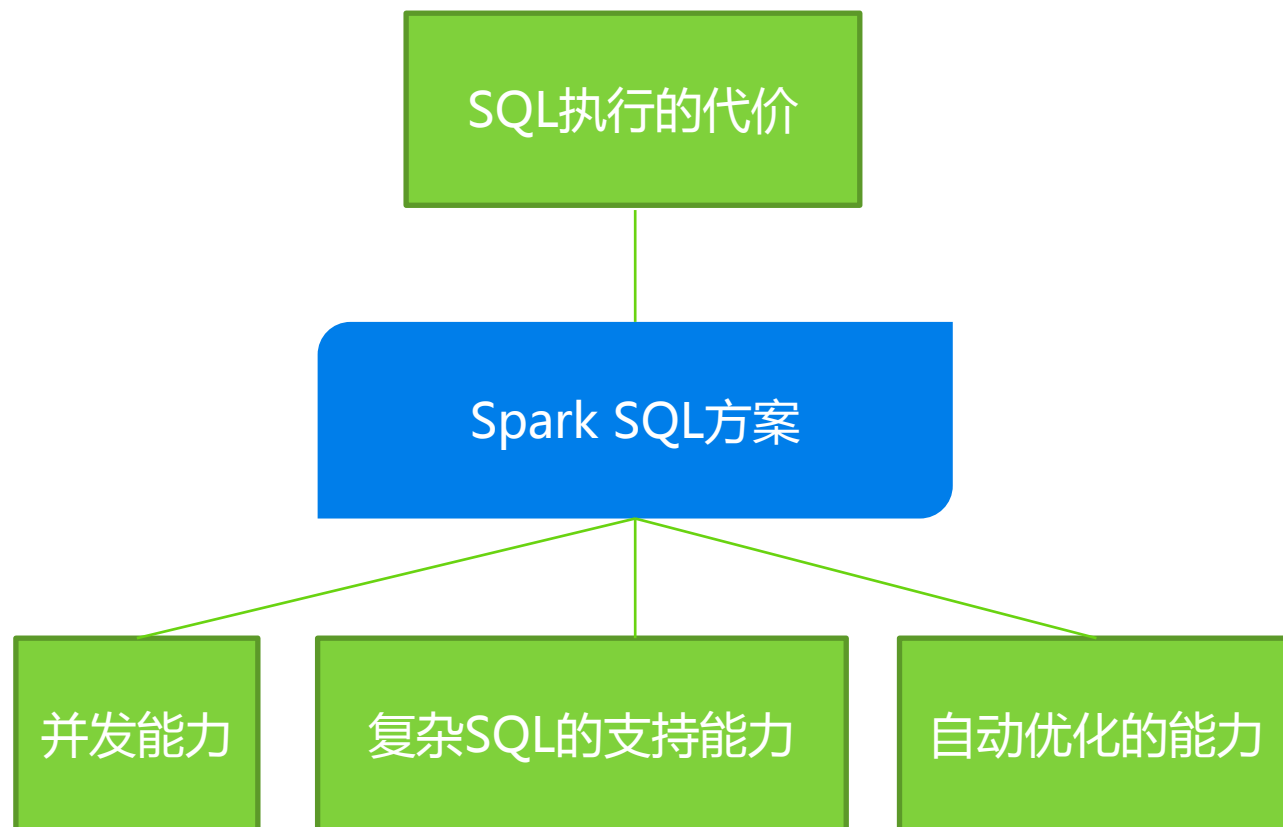


# 跨分片SQL Join的老大难问题



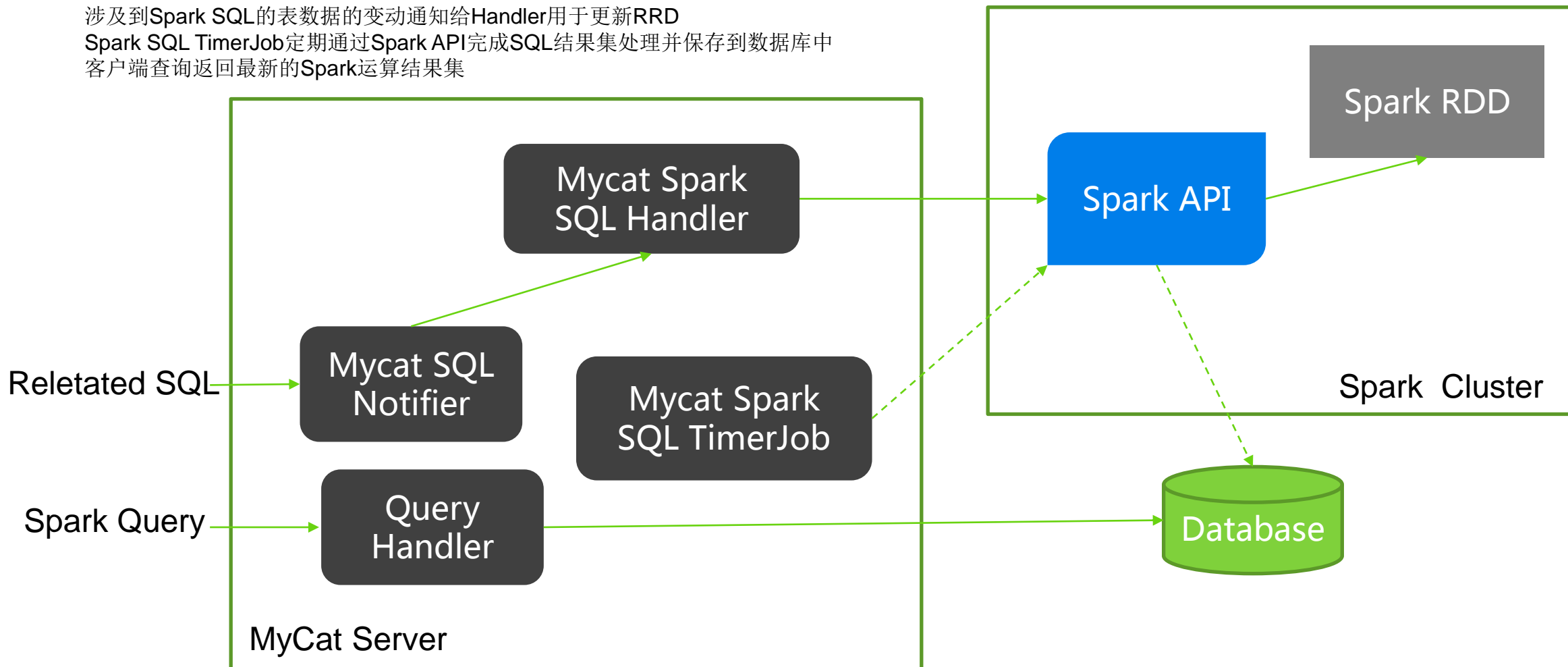
CPU	
一级缓存	1 ns
二级缓存	10 ns
内存	100 ns
一级存储	10,000,000ns (10 ms)
二级存储	20,000,000ns (20 ms)
近线存储	>20,000,000ns (>20 ms)

# 跨分片SQL Join的老大难问题

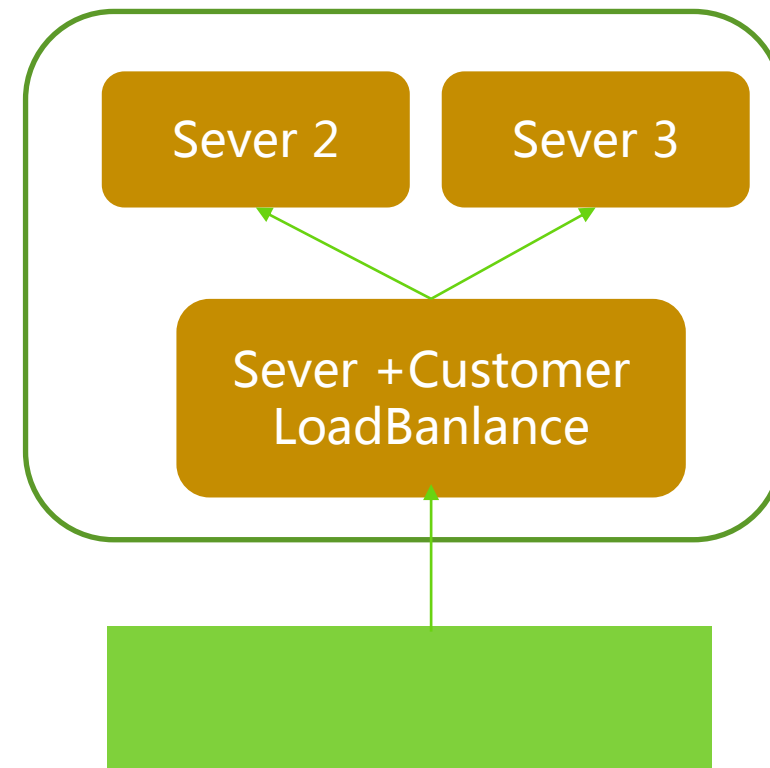
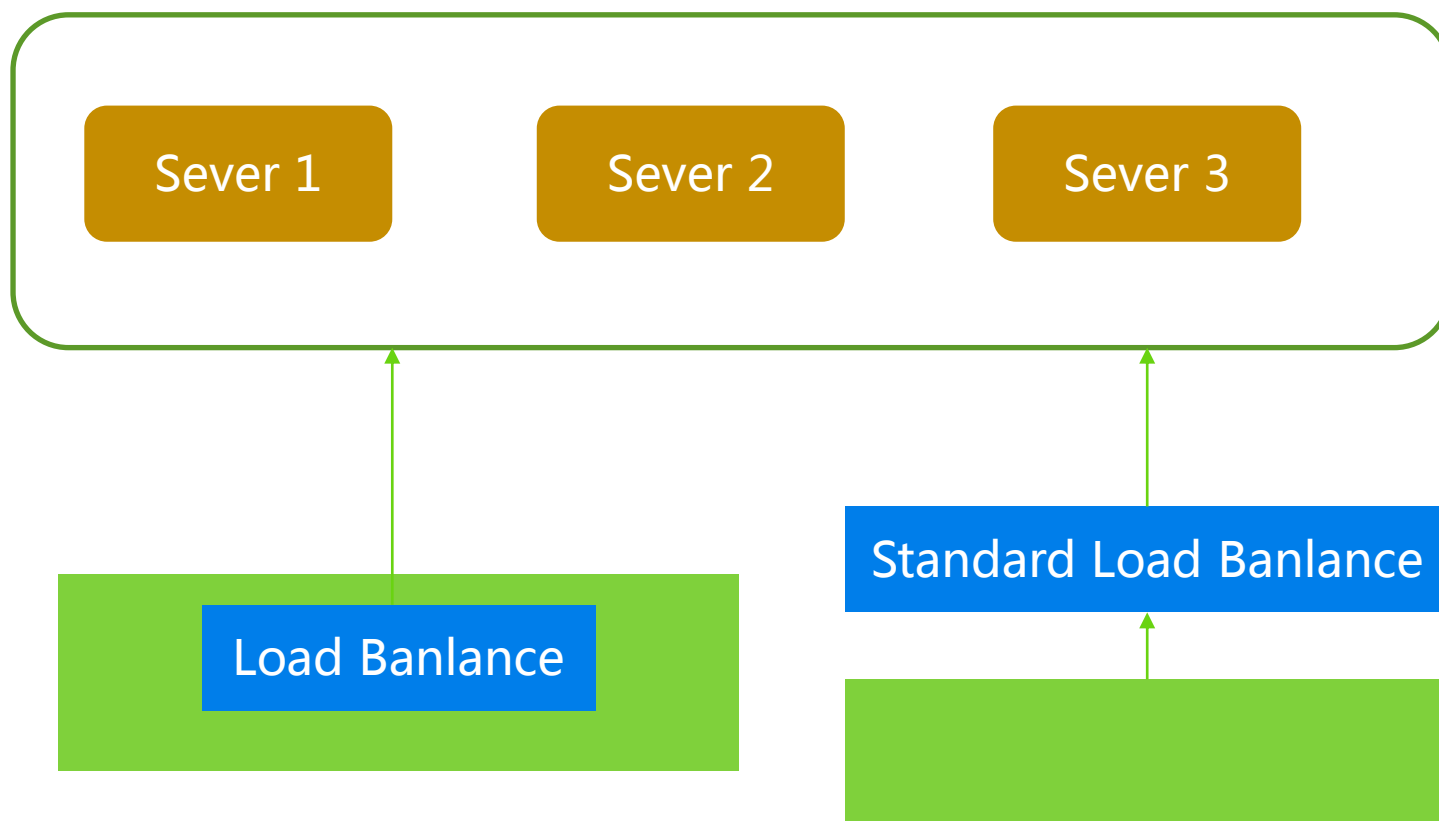


# 跨分片SQL Join的老大难问题

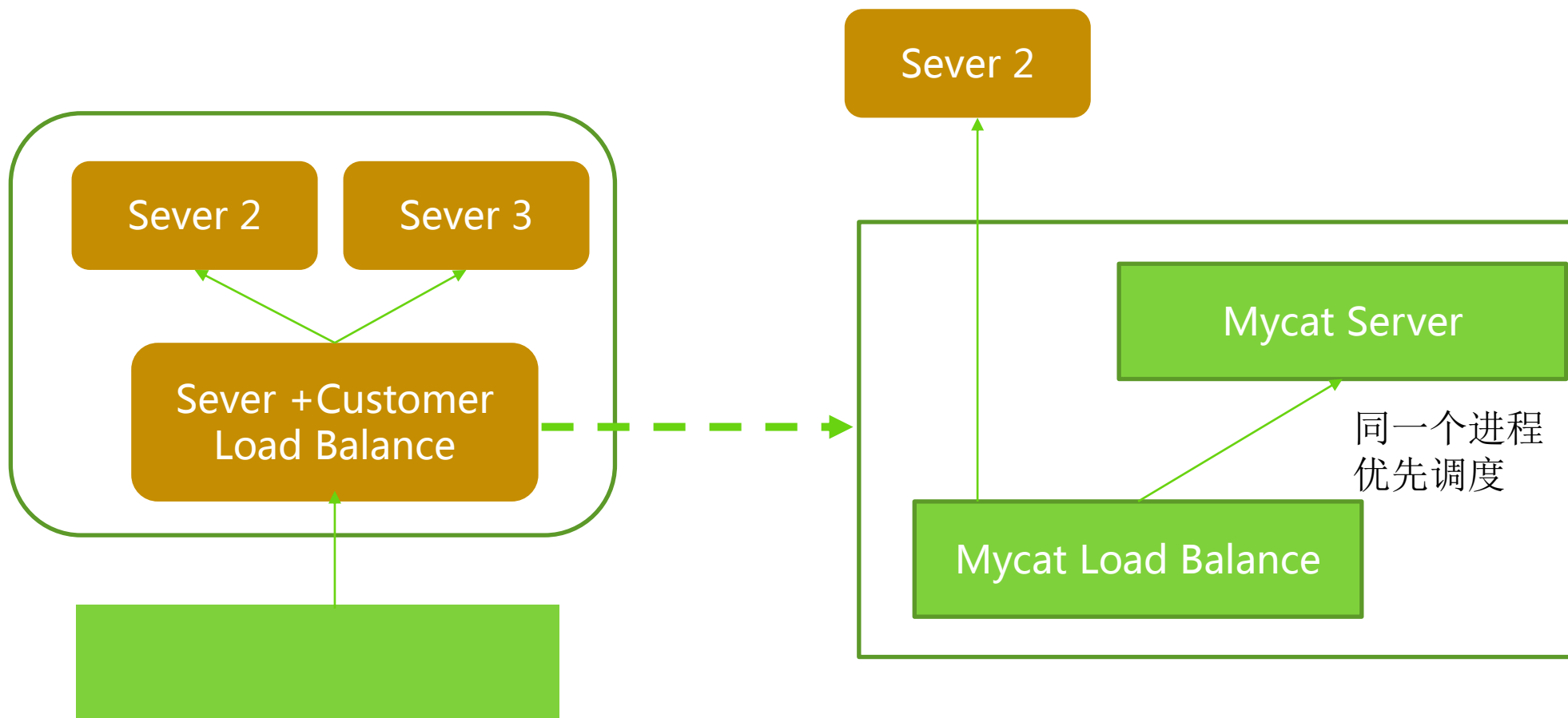
涉及到Spark SQL的表数据的变动通知给Handler用于更新RRD  
Spark SQL TimerJob定期通过Spark API完成SQL结果集处理并保存到数据库中  
客户端查询返回最新的Spark运算结果集



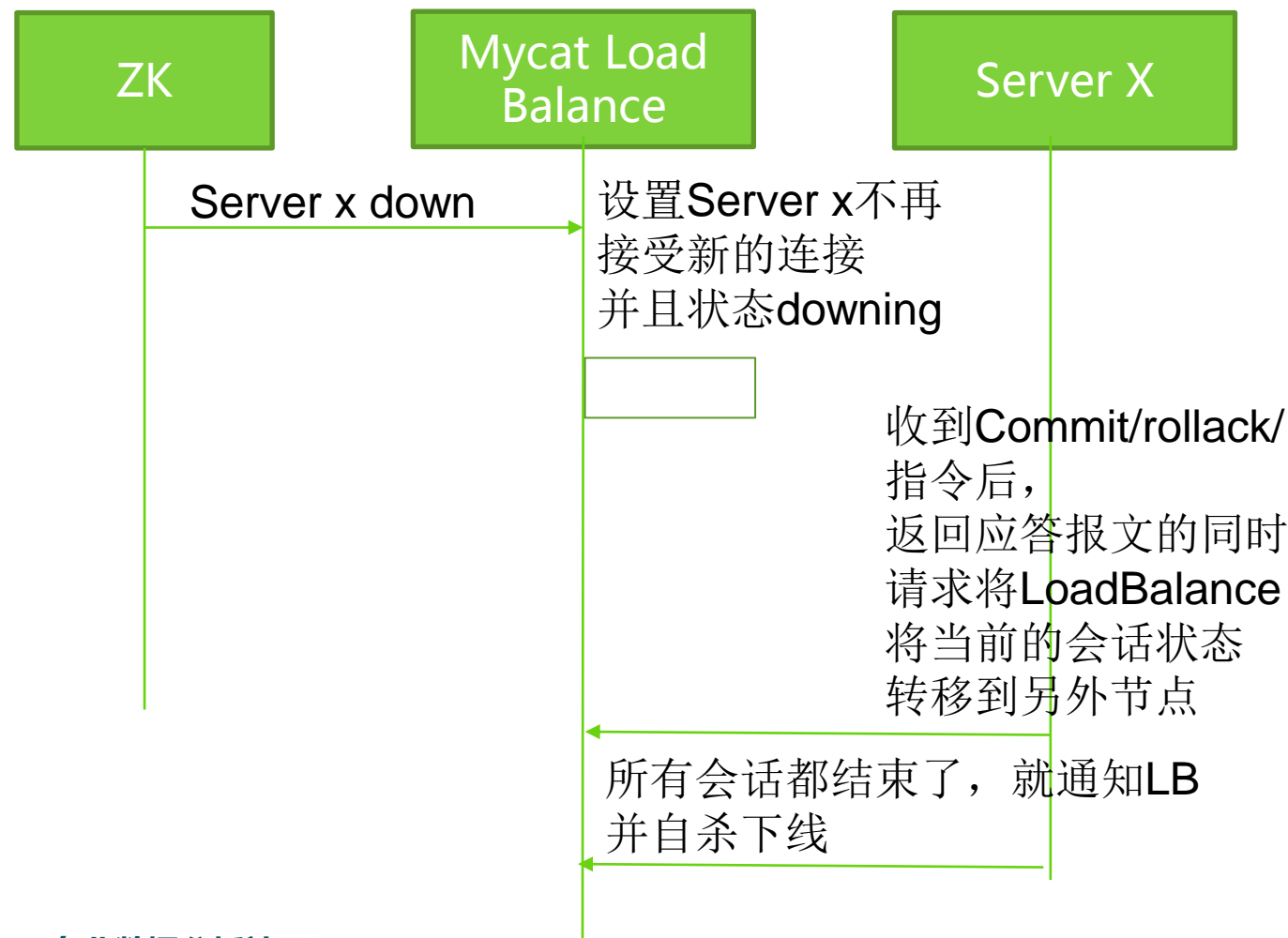
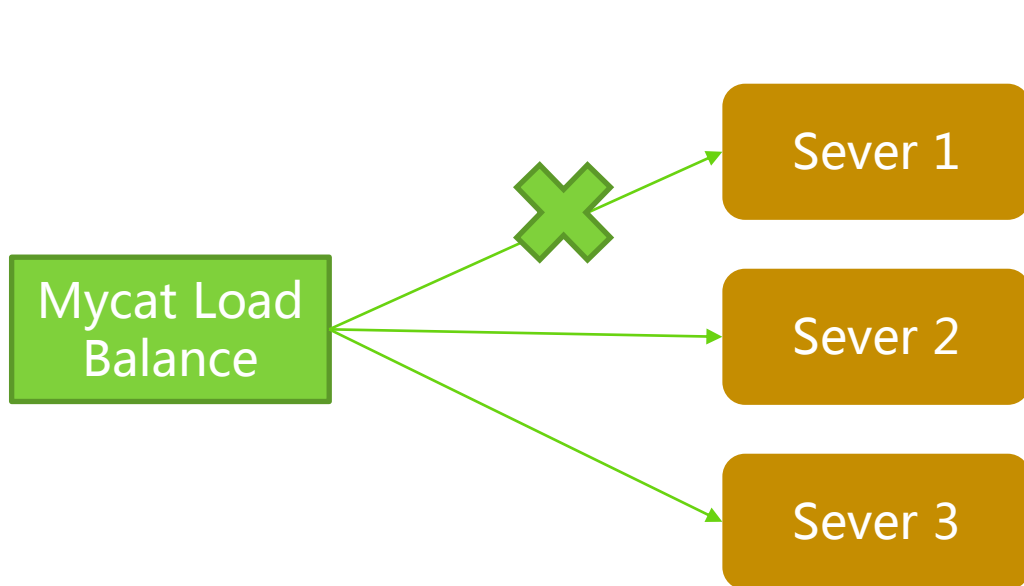
# Mycat的HA问题



# Mycat的HA问题



# Mycat的HA问题



# Thanks

**FAQ时间**