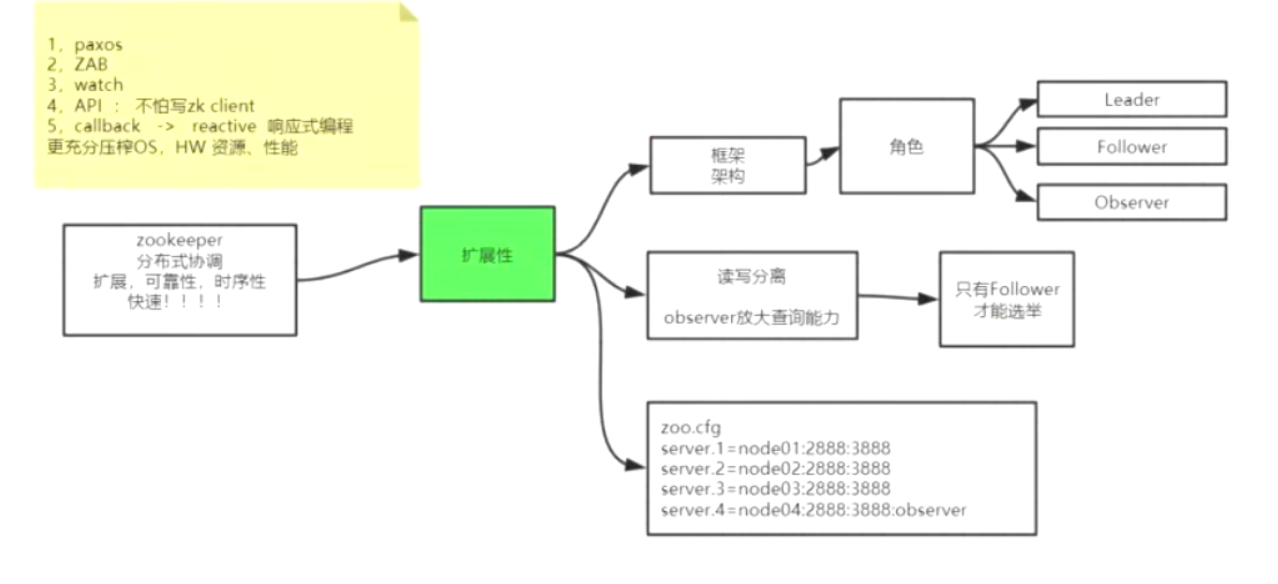
1. 可扩展性

由于只有follower可以选主，因此follower的个数会影响选主效率，个数越少效率越高，因此可以提供Observer去控制follower的个数，并且observer可以提高查询能力。

Observer作用：减少follower个数；放大查询能力

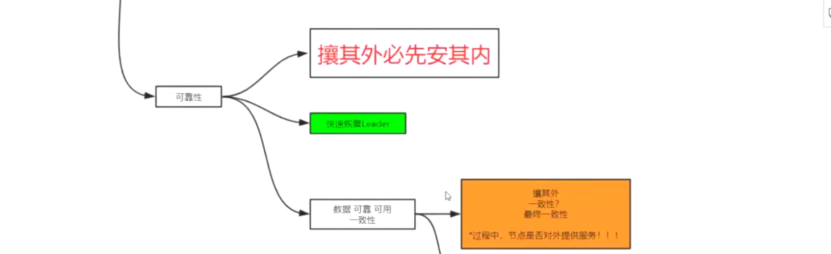


1. 可靠性：对外提供服务（有主情况下，写数据）

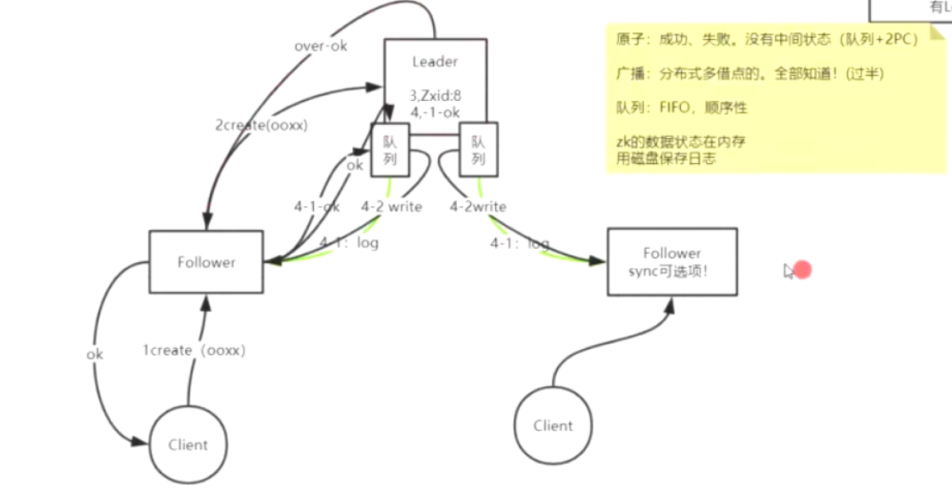
（1）ZAB协议（原子广播协议；数据同步，通过队列实现最终一致性）

paxos 的简化版

两阶段：投票半数通过（通知写log）、投票完成通知写操作（通知写内存）







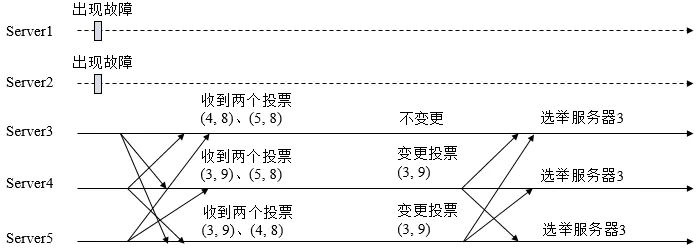
1. 选leader（无主情况下，选leader）

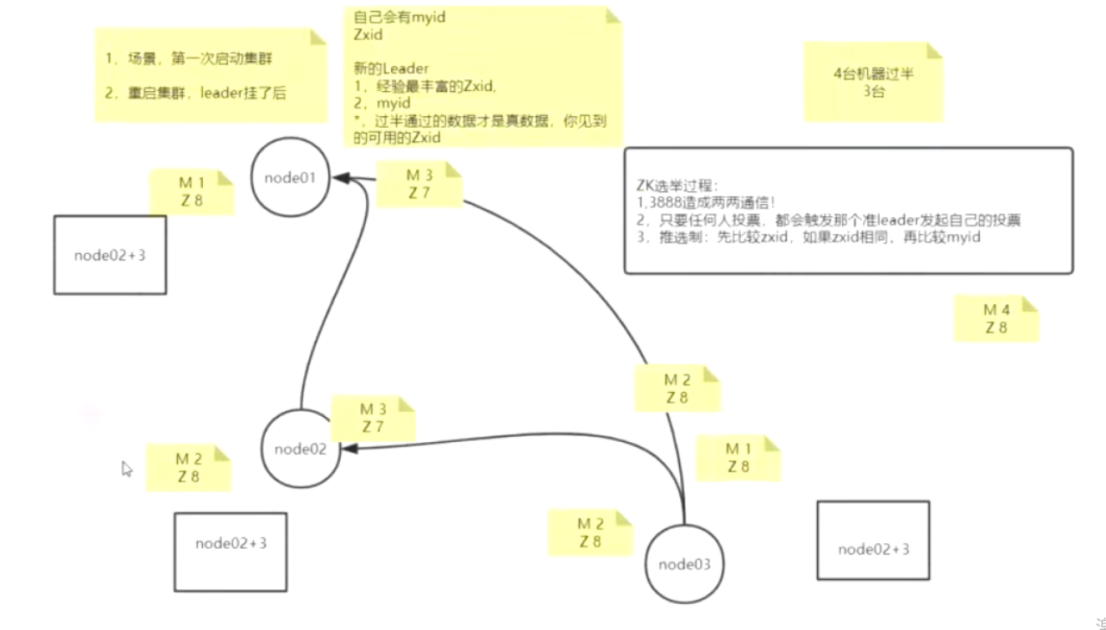
只有过半才可以选择leader

leader条件：（1）数据最全（Zxid最大） （2）年龄最大（myid最大）

投票过程：

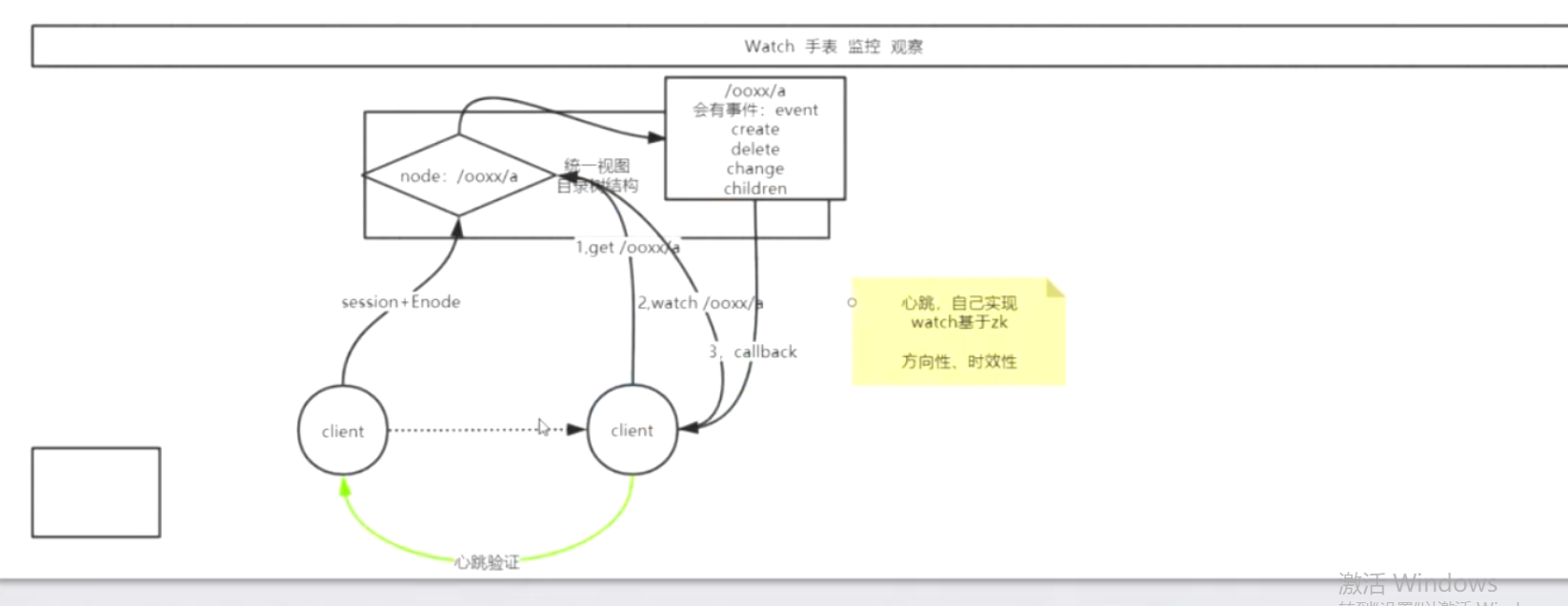
1. 所有节点根据收到的投票，确定leader
2. 所有节点将leader广播出去





1. Watch

Zookeeper是协调别人的，Watch提供监控节点功能。



场景：分布式锁，只有一个客户端可以获得锁

1. 客户端自己实现：

客户端定时查询是否可以获得锁

缺点：时效性不好，可能在查询之前已经可以获得锁了

1. Zookeeper的Watch

当session断开的时候，其对应的临时节点就会被删除，此时Zookeeper会产生事件，并回调通知Watch的客户端。

优点：时点性很好，基本上是实时的。