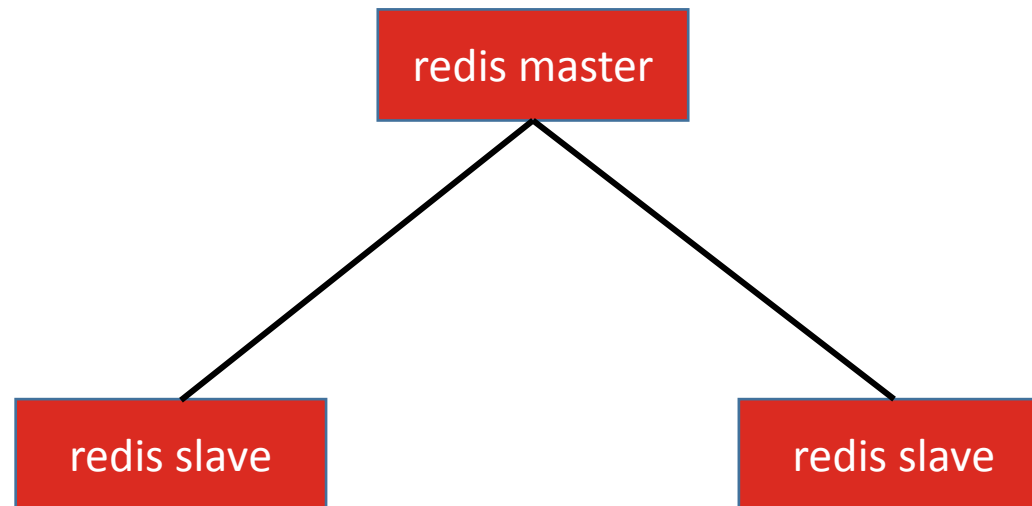


Redis Cluster运维方案

- redis cluster简介
- 运维问题
- 集群标准化
- 其它经验

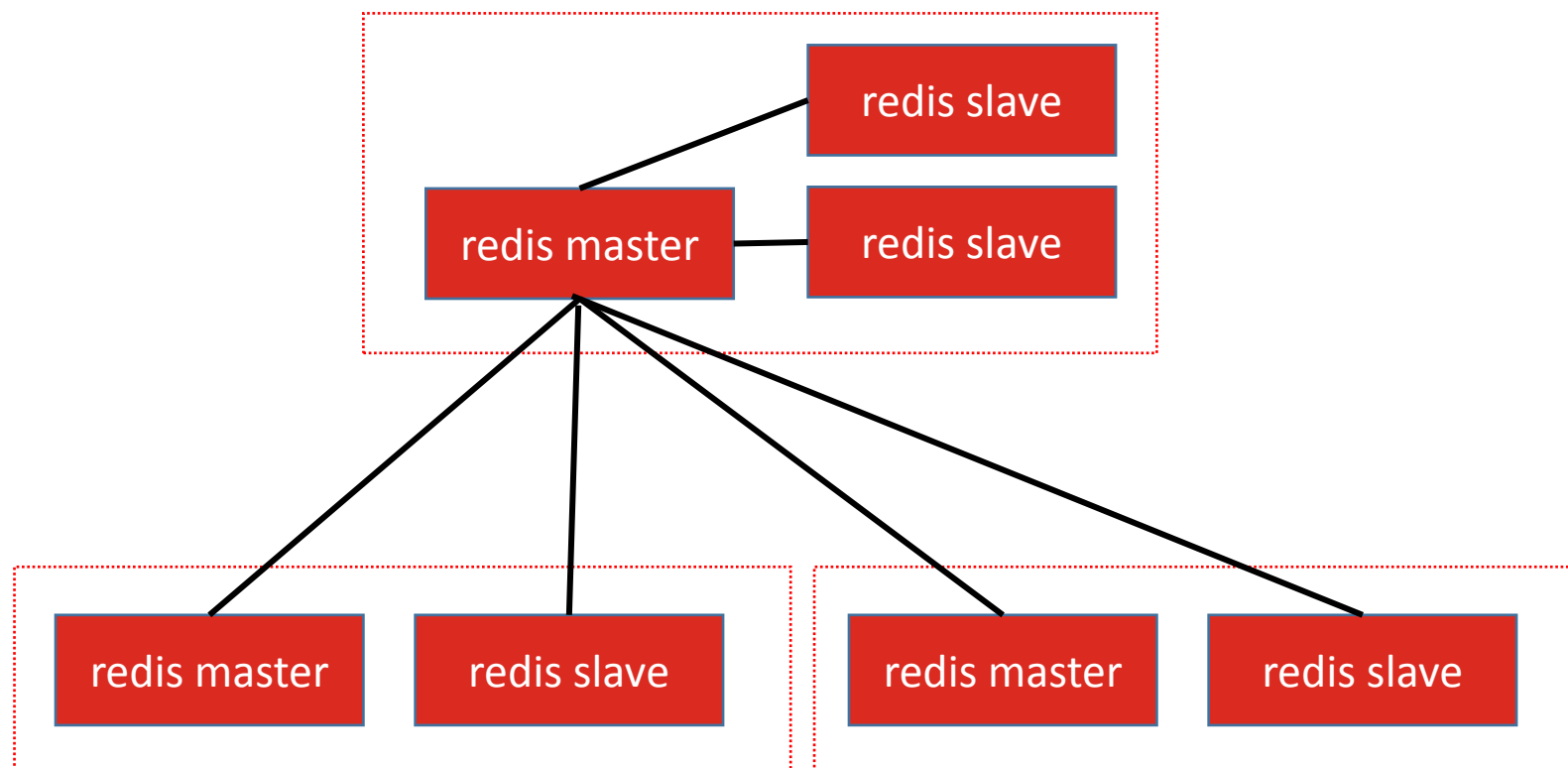
官方集群方案Redis Cluster

- 提供主从



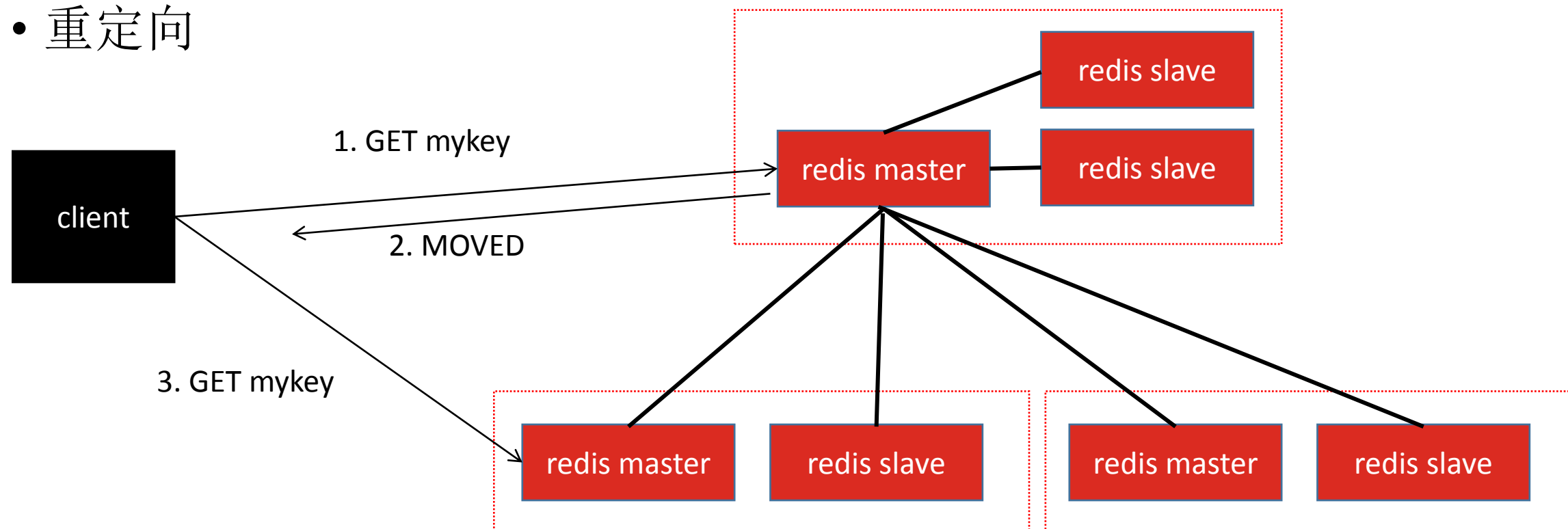
官方集群方案Redis Cluster

- 提供主从和分片



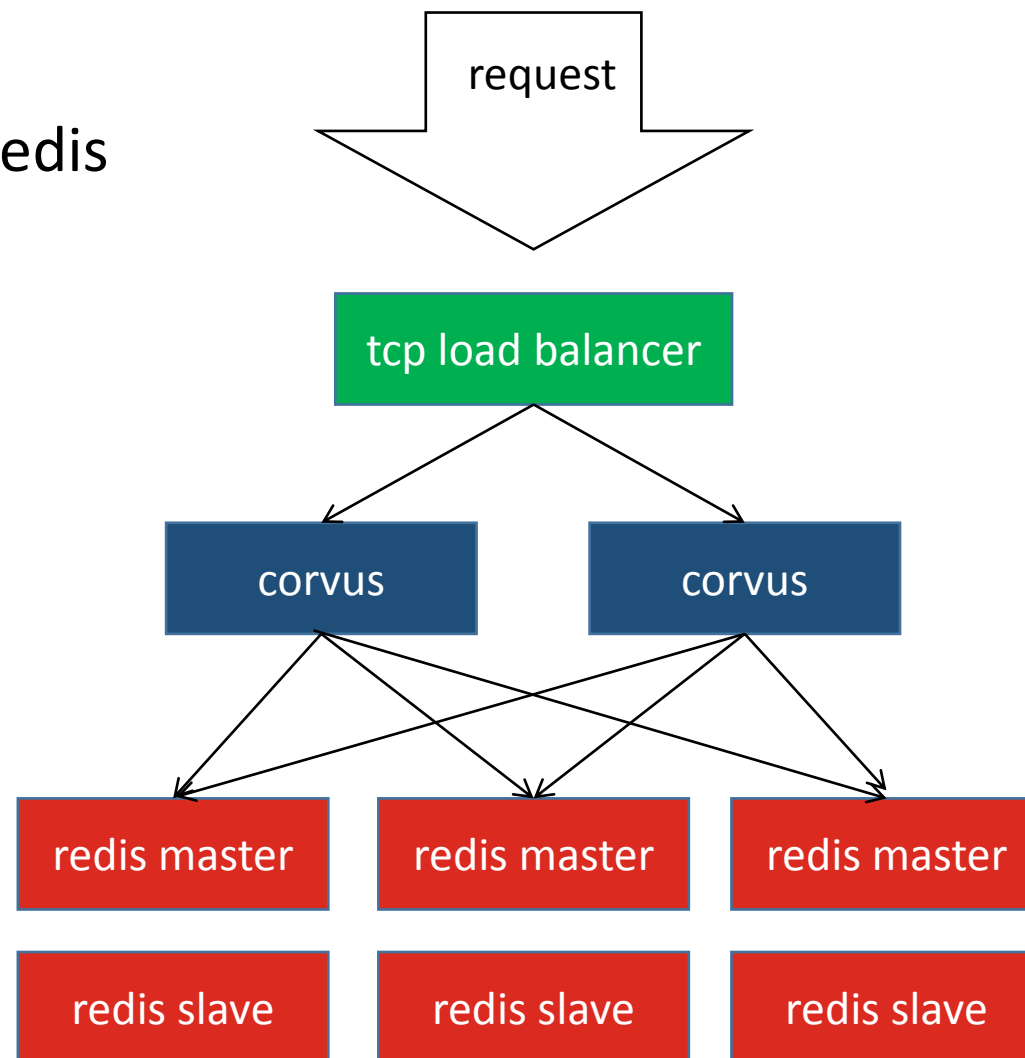
请求流程

- 请求master
- 重定向



redis cluster proxy - corvus

- 屏蔽redis cluster协议，对外暴露单机redis协议



运维方法 - 第一阶段

- CLUSTER命令
- 脚本
- 扩容：
- CLUSTER MEET
- CLUSTER SETSLOT IMPORTING
- CLUSTER SETSLOT MIGRATING
- MIGRATE
- CLUSTER SETSLOT NODE

运维方法 - 第二阶段

- 重造运维工具
- 封装成面板

如果集群没有变更...

host 1	host 2	host 3	host 4
redis master	redis master	redis master	redis master
redis master	redis master	redis master	redis master
redis slave	redis slave	redis slave	redis slave
redis slave	redis slave	redis slave	redis slave

压力不均衡

host 1	host 2	host 3	host 4
redis master	redis master	redis master	redis slave
redis master	redis master	redis master	redis slave
redis master	redis slave	redis slave	redis slave
redis master	redis slave	redis slave	redis slave

- 重新调整主从分布
- 要保存拓扑结构

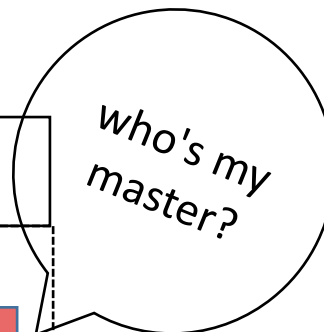
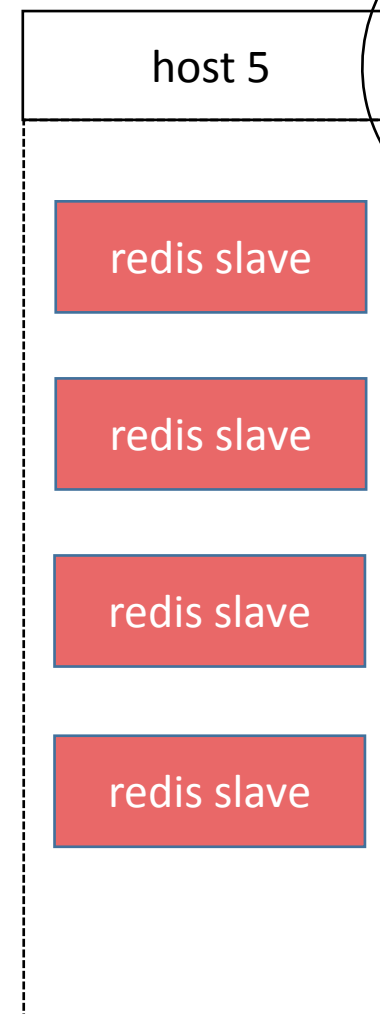
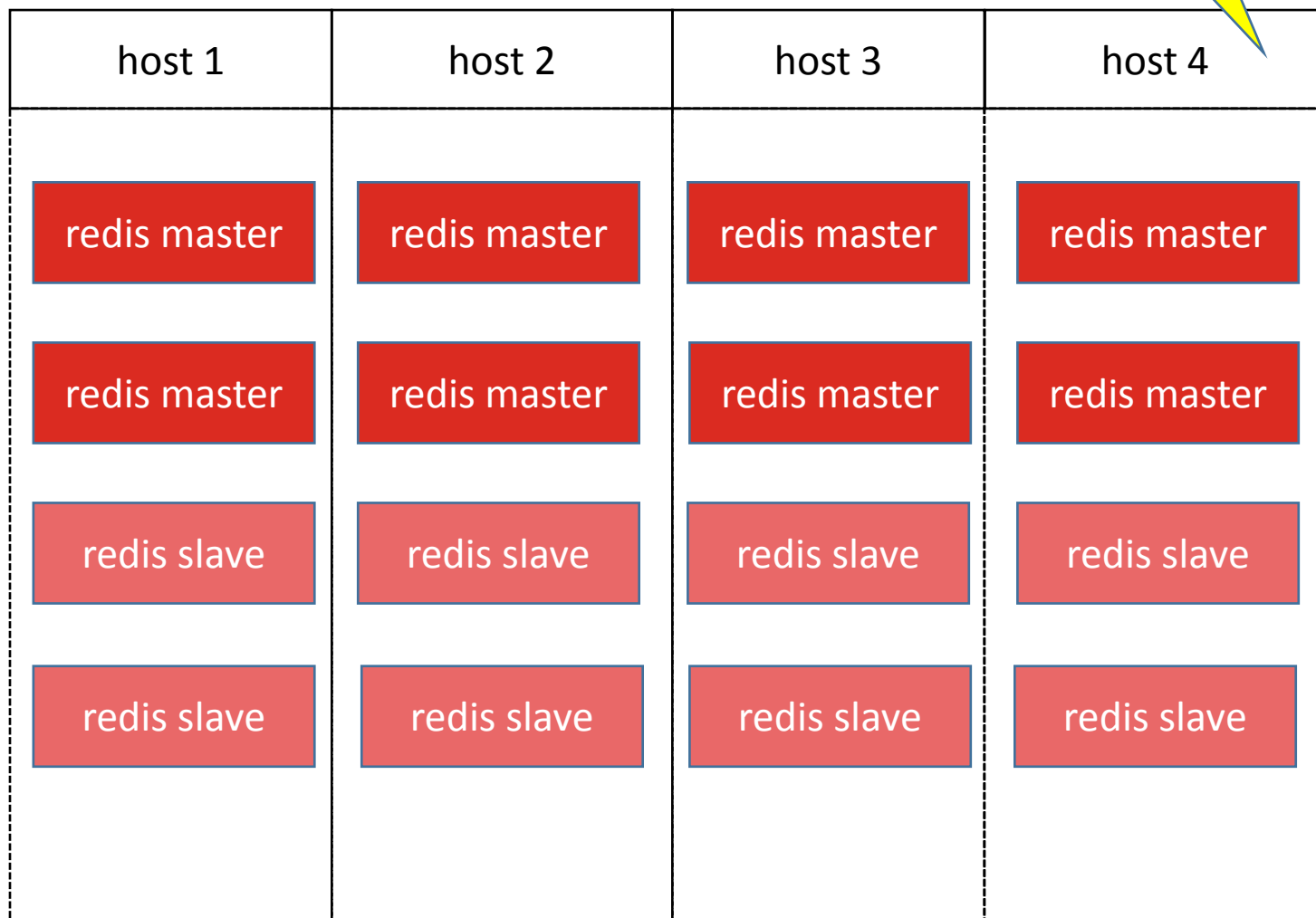
宕机后压力不均衡



host 1	host 2	host 3	host 4
redis master	redis master	redis master	redis master
redis master	redis master	redis master	redis master
redis slave	redis slave	redis slave	redis slave
redis slave	redis slave	redis slave	redis slave

- 分配节点时需要考虑挂机器后的拓扑结构

宕机恢复



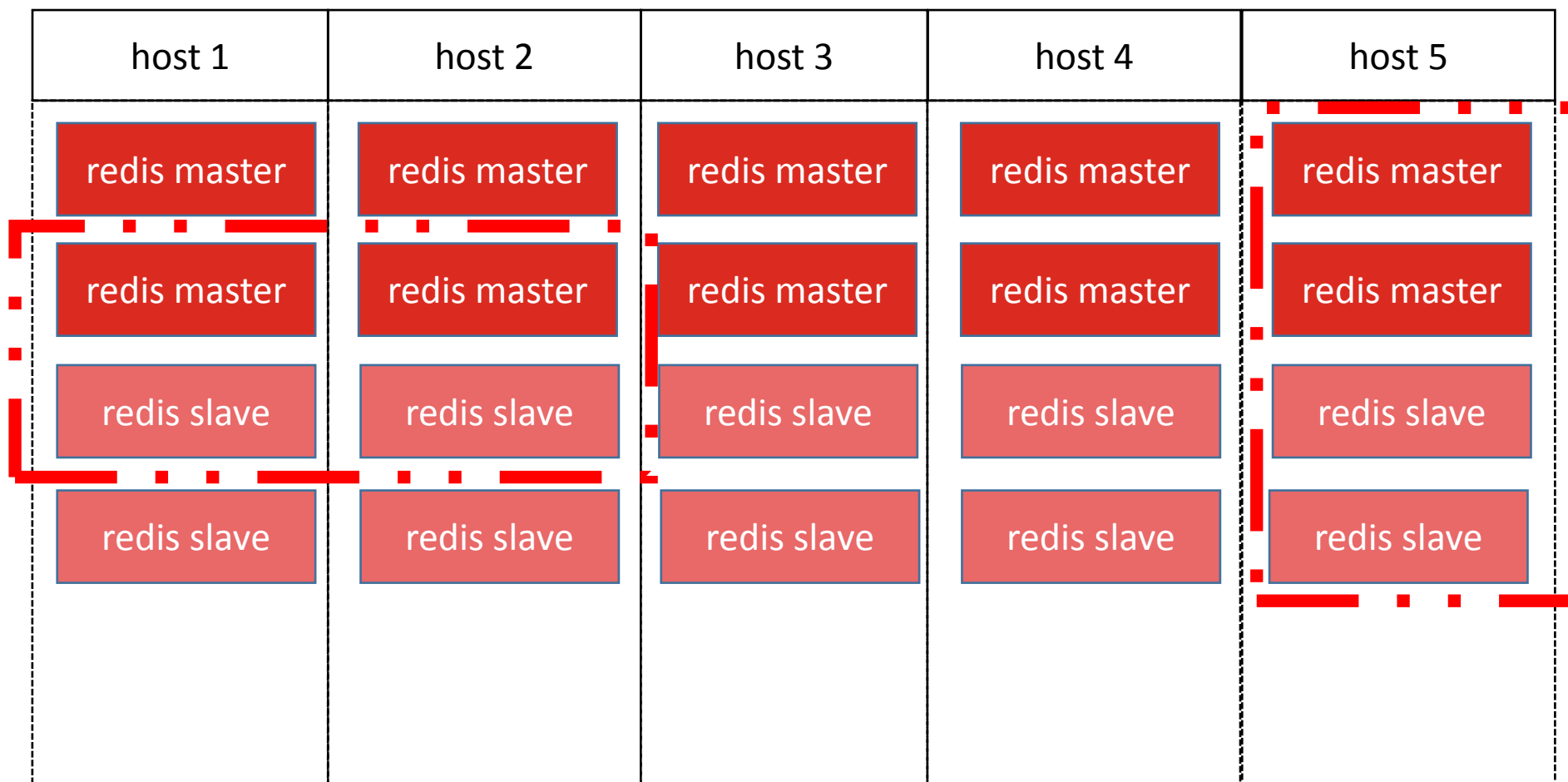
- 要保存拓扑结构

扩容时分配节点

host 1	host 2	host 3	host 4	host 5
redis master	redis master	redis master	redis master	redis master
redis master	redis master	redis slave	redis slave	redis slave
redis slave	redis slave	redis slave	redis slave	new master
new slave	new slave	new master	new master	new slave

- 扩容计算节点分配相对容易

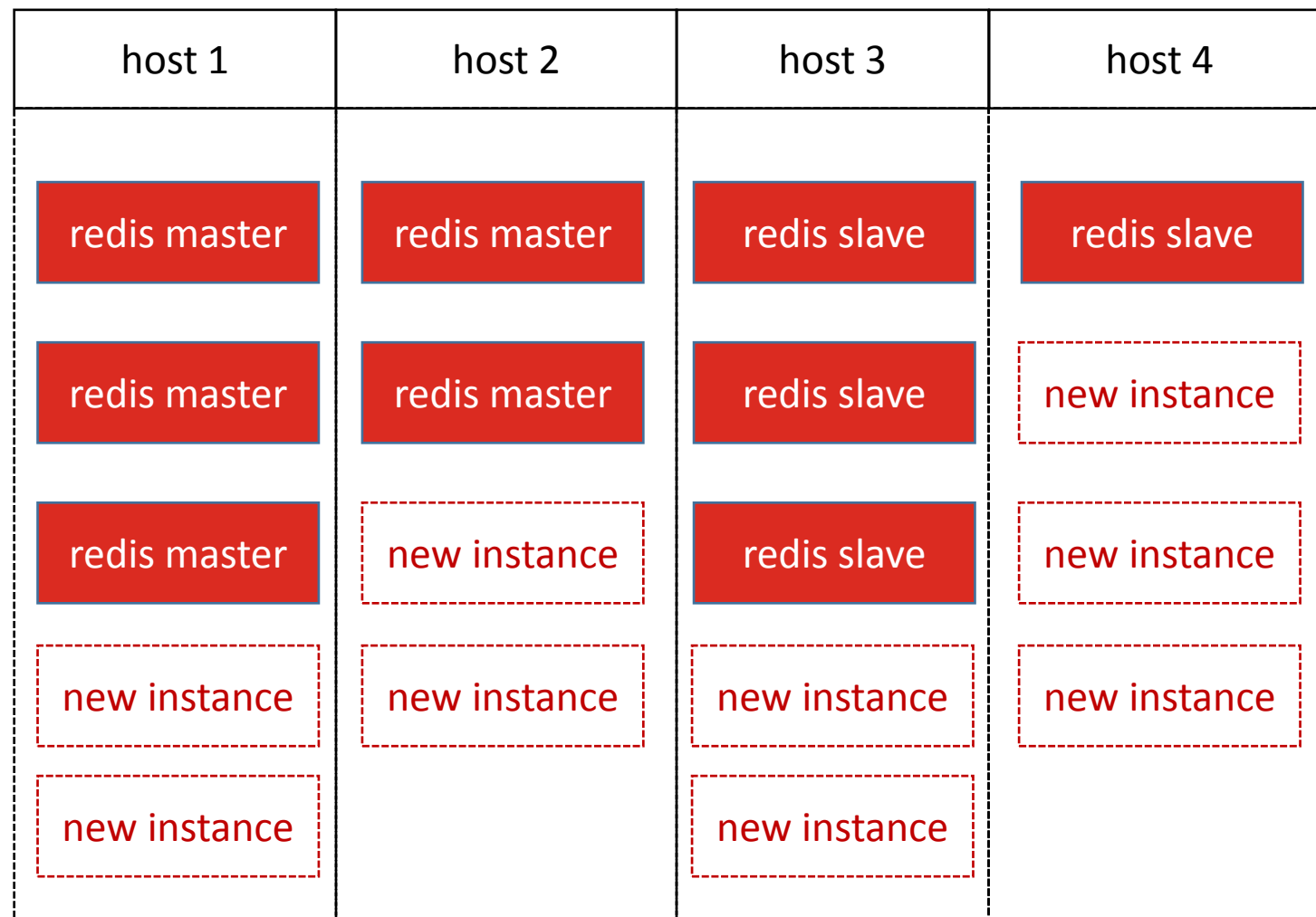
扩容时删除节点



- 扩容保持良好的拓扑结构较难

运维问题

- 节点数量多，管理困难
- 维持集群结构（主从分布,压力均衡）
 - 修复宕机
 - 集群扩容与缩容
 - 主节点分布不均
 - 依赖集群拓扑信息
 - 节点分配要考虑CPU和内存



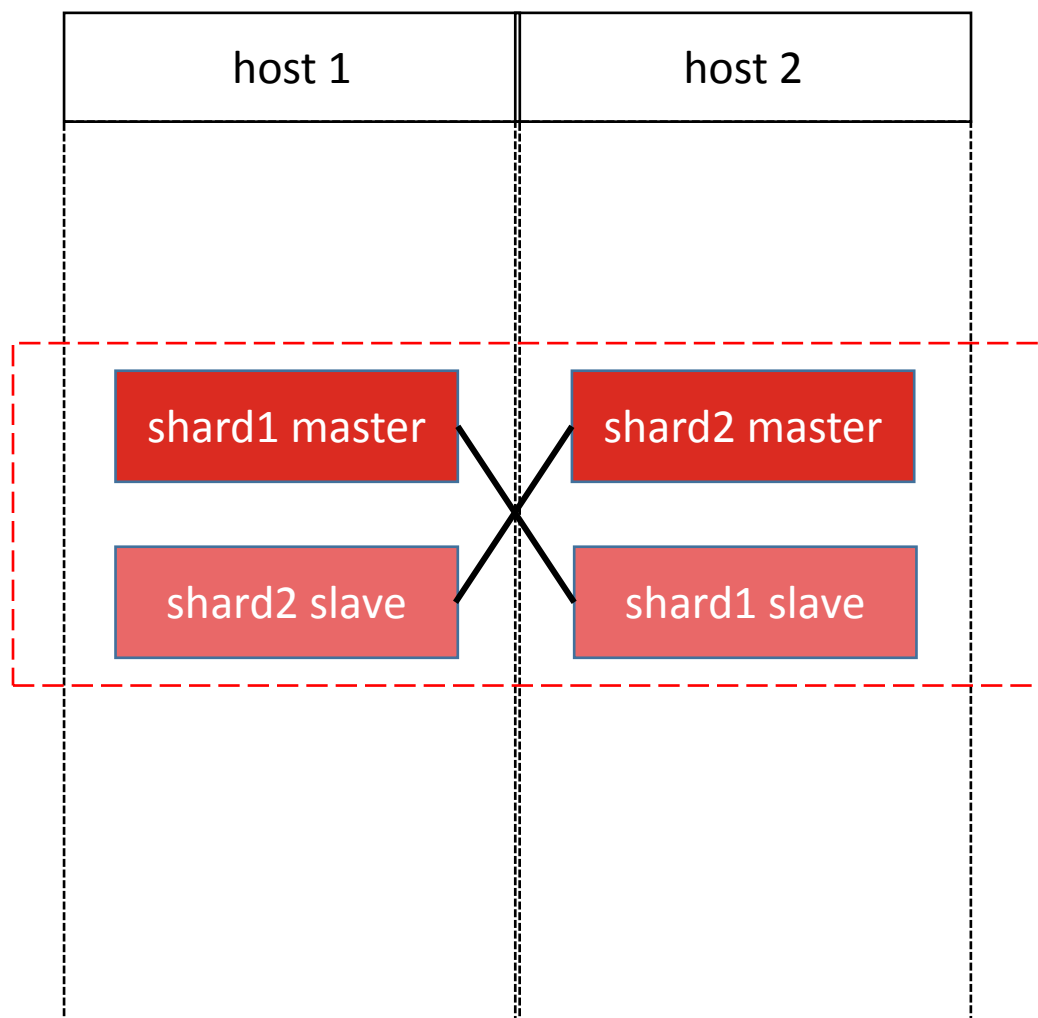
集群标准化 - 第三阶段

- 用规则去约束集群，以方便管理，要求：
 - 集群能在挂一台机器的情况下不影响服务
 - 主节点和从节点不能在一台机器上
 - 不能有过半数（包括半数）主节点在同一台机器上
 - 集群在挂一台机器的情况下, 压力应尽可能平均分流到其它机器
 - 主节点和从节点在各台机器上的分布应当平均

规范分片

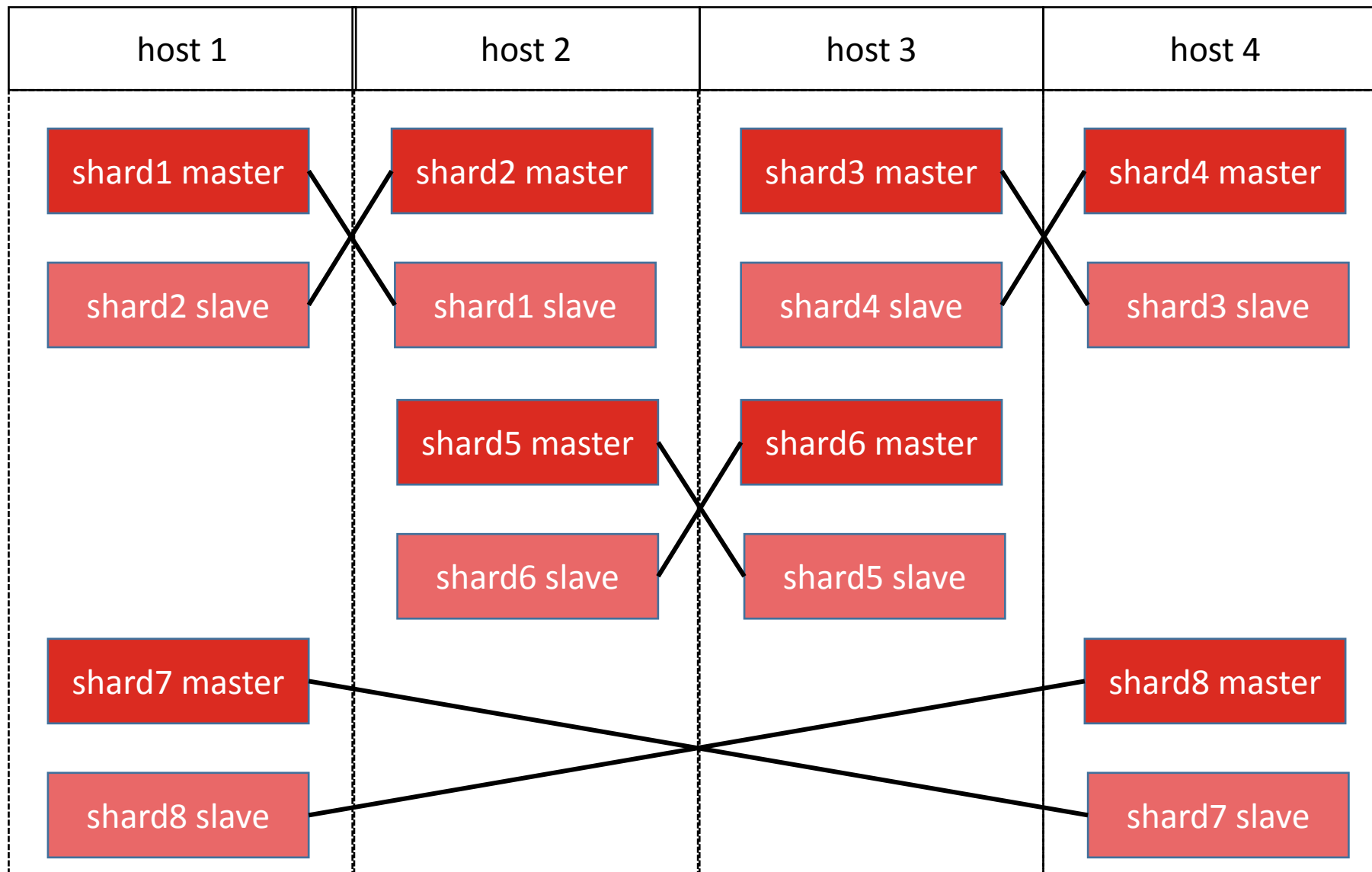
- 每个分片只有一个从节点
- 每个节点的maxmemory一样

分组管理 - chunk

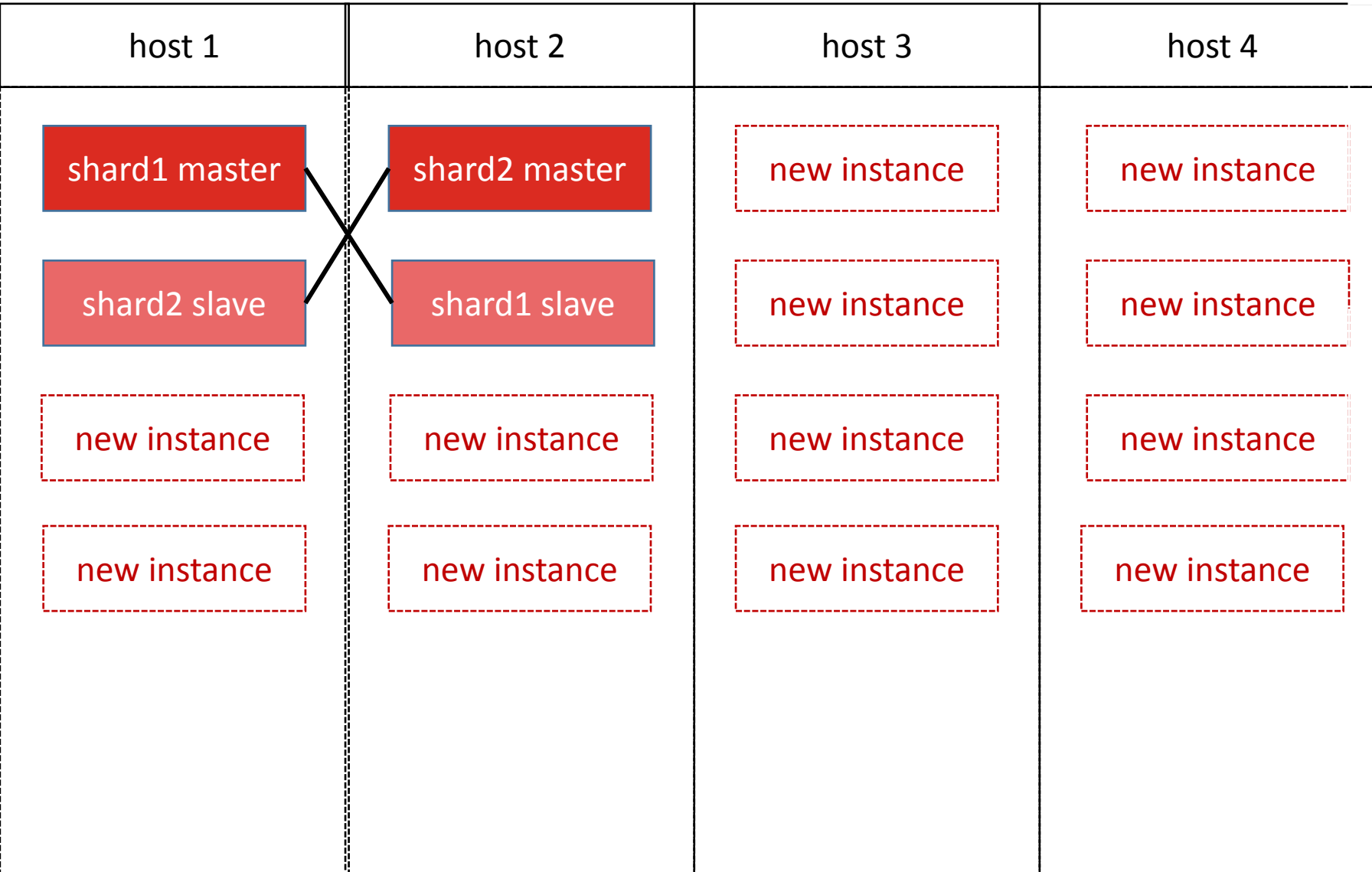


- 保证主从不在同一台机器
- 每台机器的主从比例一样
- 固定了主从角色

分配chunk



分配chunk



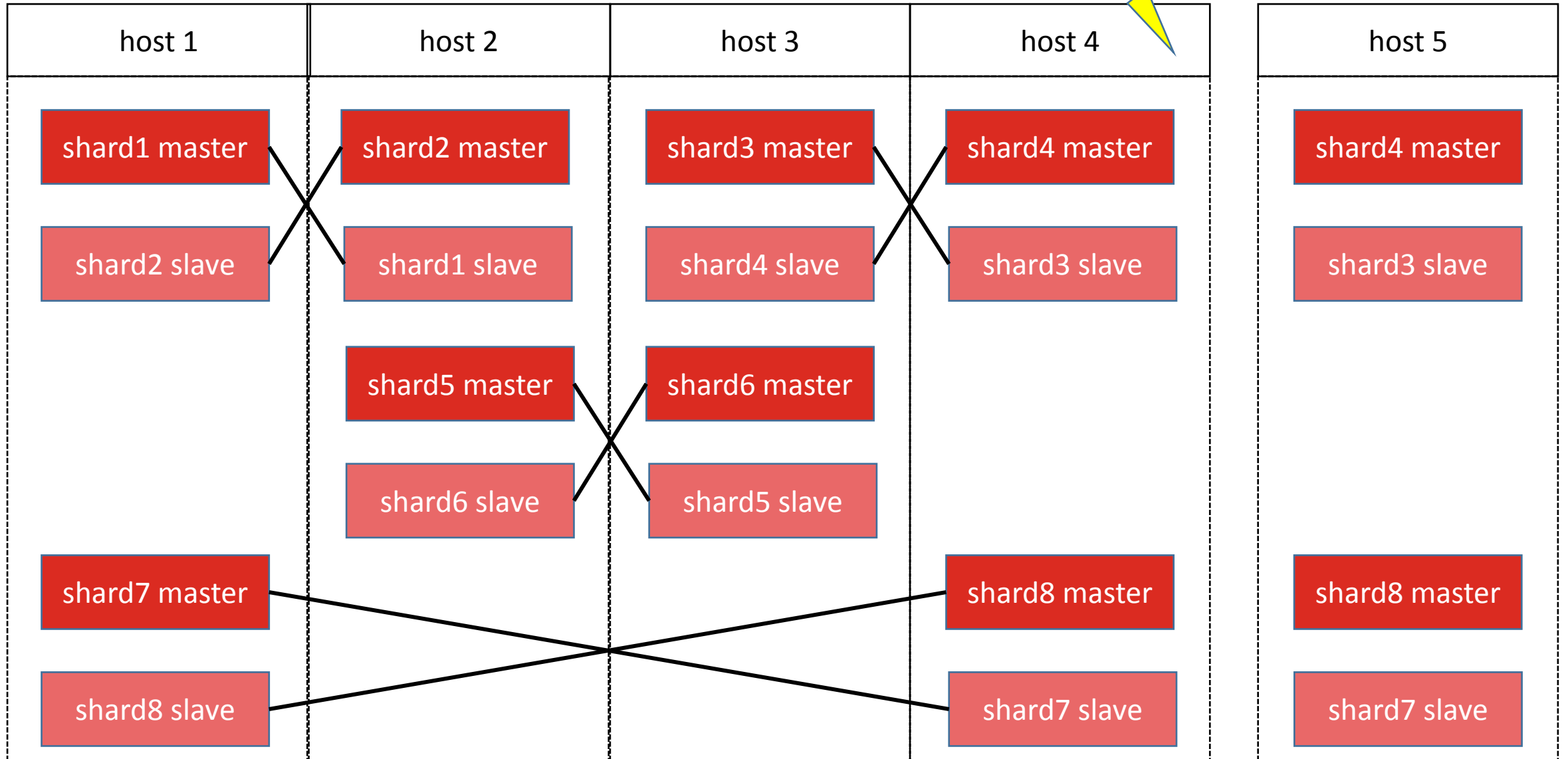
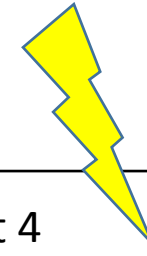
	host1	host2	host3	host4
host1		0	0	0
host2	0		0	0
host3	0	0		0
host4	0	0	0	

	host1	host2	host3	host4
host1		1	0	0
host2	1		0	0
host3	0	0		0
host4	0	0	0	

运维操作

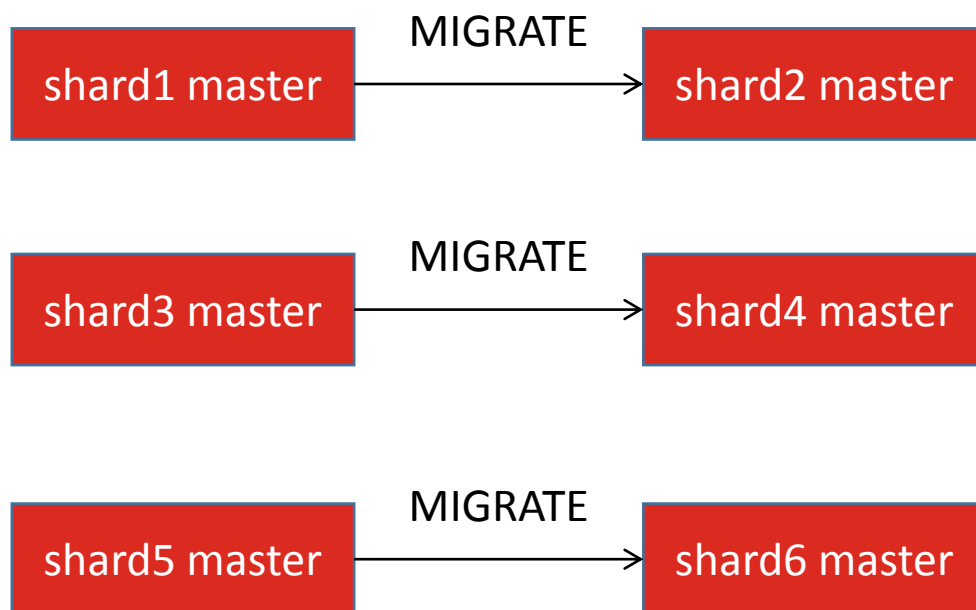
- 创建集群 - 构造 nodes.conf
- 扩容缩容 - 添加删除 chunk
- 处理宕机 - 复刻 nodes.conf

处理宕机

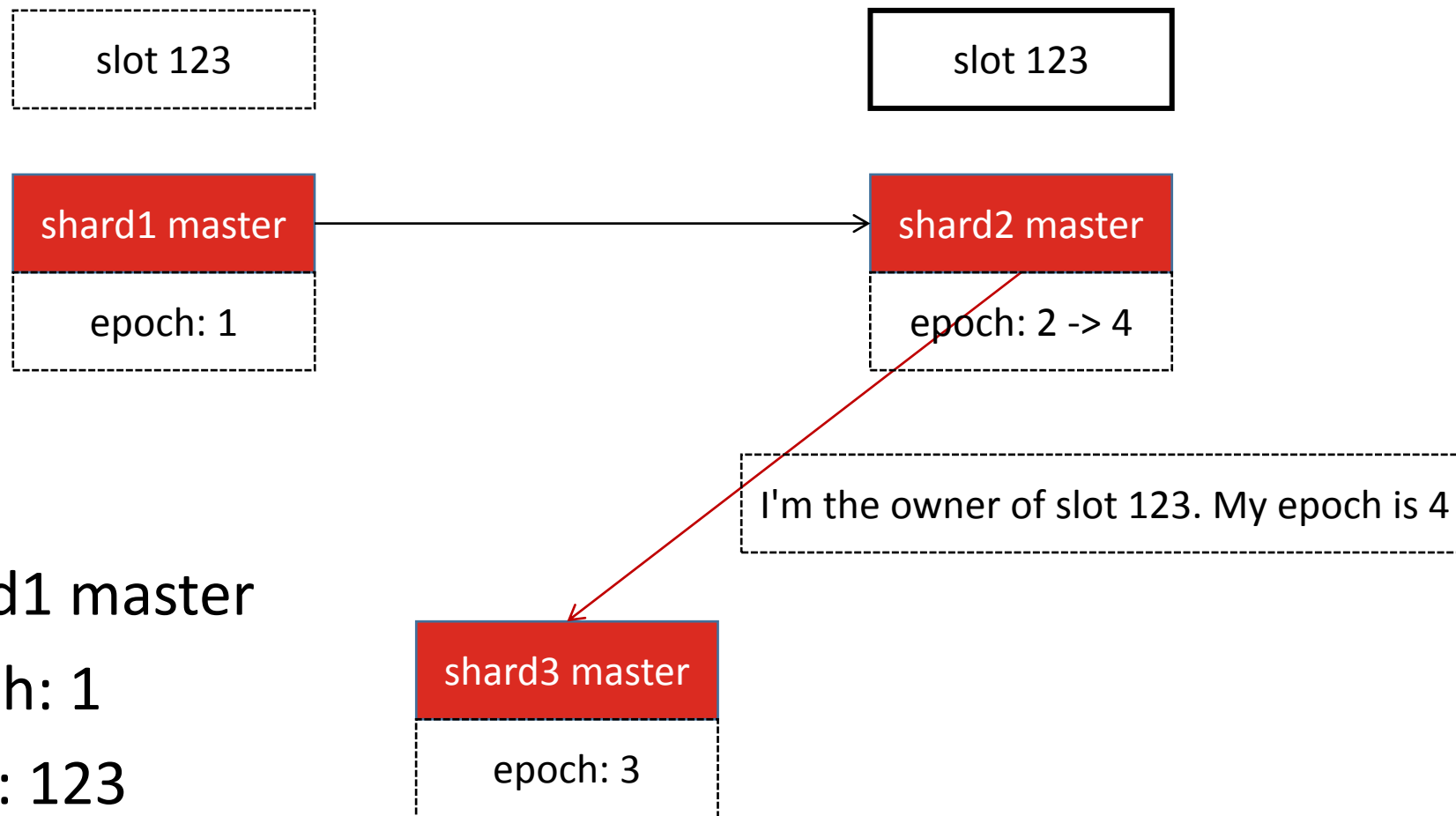


并行迁移slot

- 速度可以提升一个数量级



为什么可以并行？

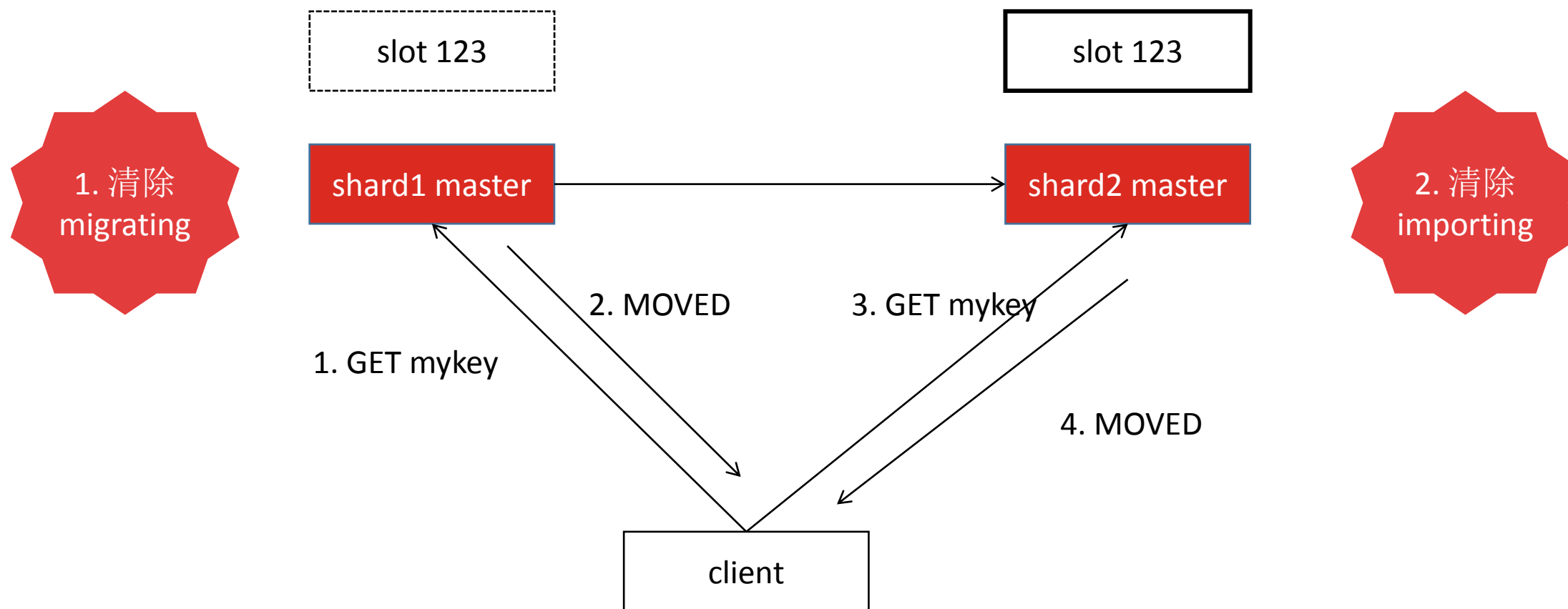


- shard1 master
- epoch: 1
- slots: 123

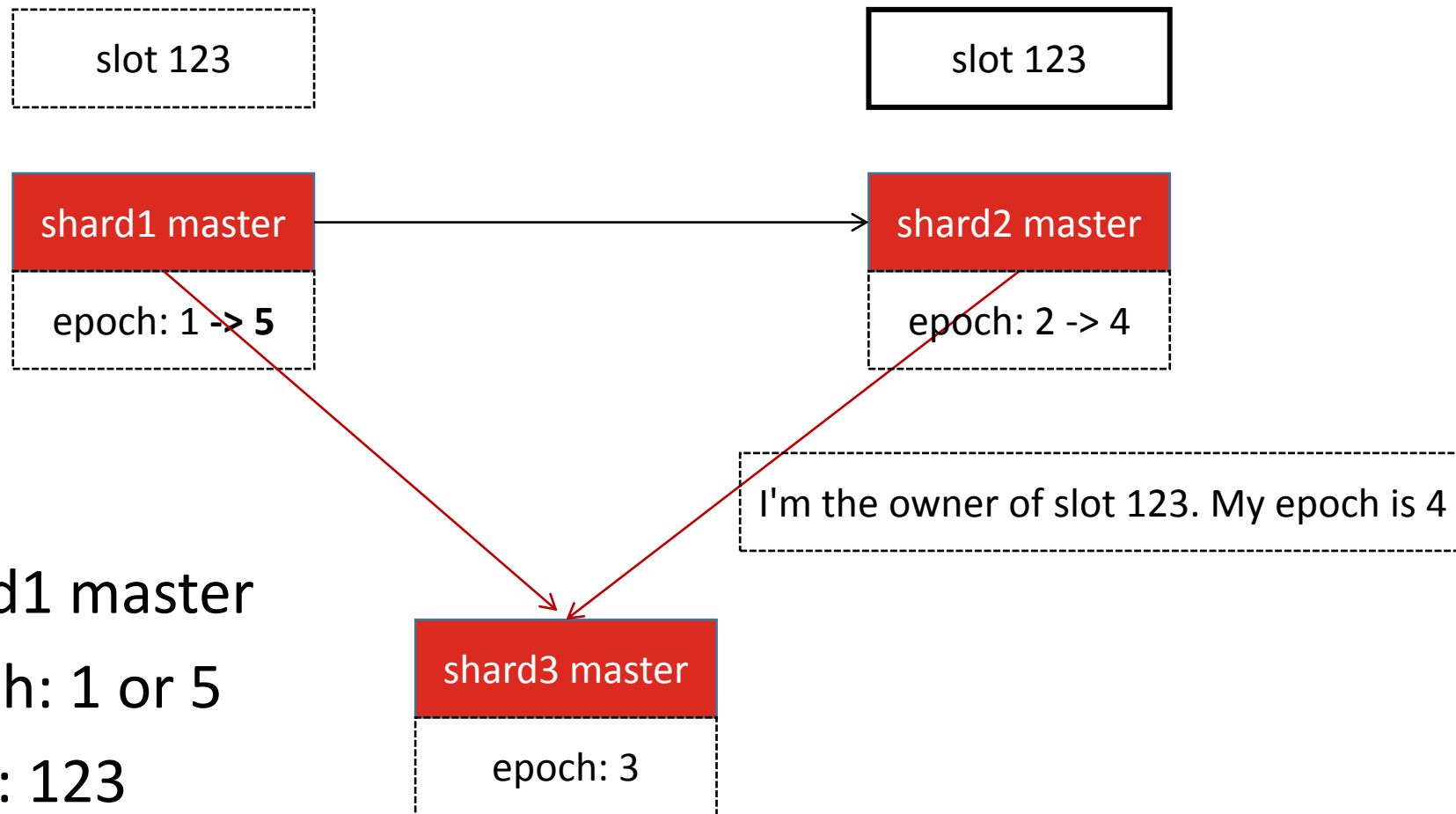
redis cluster的坑

- 迁移slot时redirection过多
- slot信息不一致
- 选举缓慢

迁移slot时redirection过多



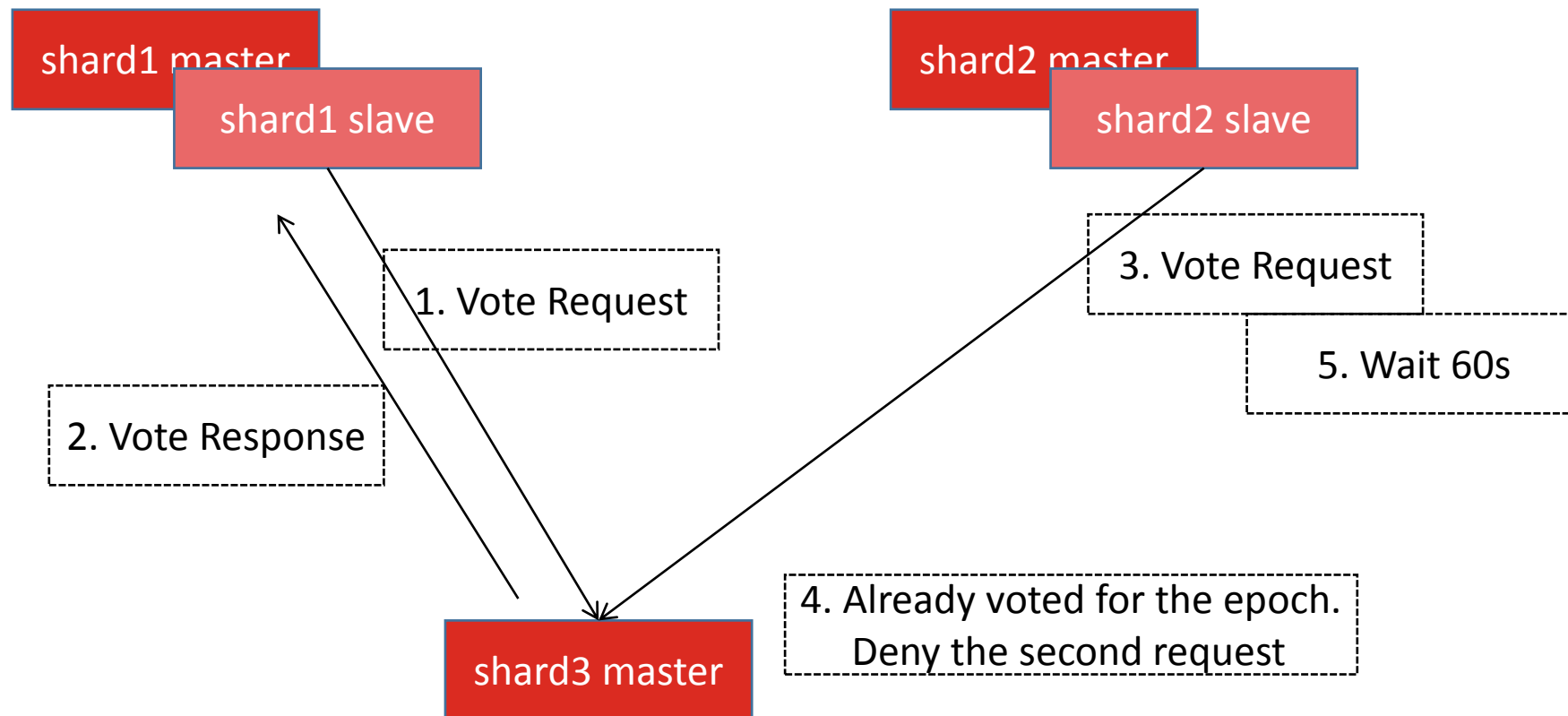
slot信息不一致



- shard1 master
- epoch: 1 or 5
- slots: 123

选举缓慢

- 选举有时候需要花2-3分钟



Question?