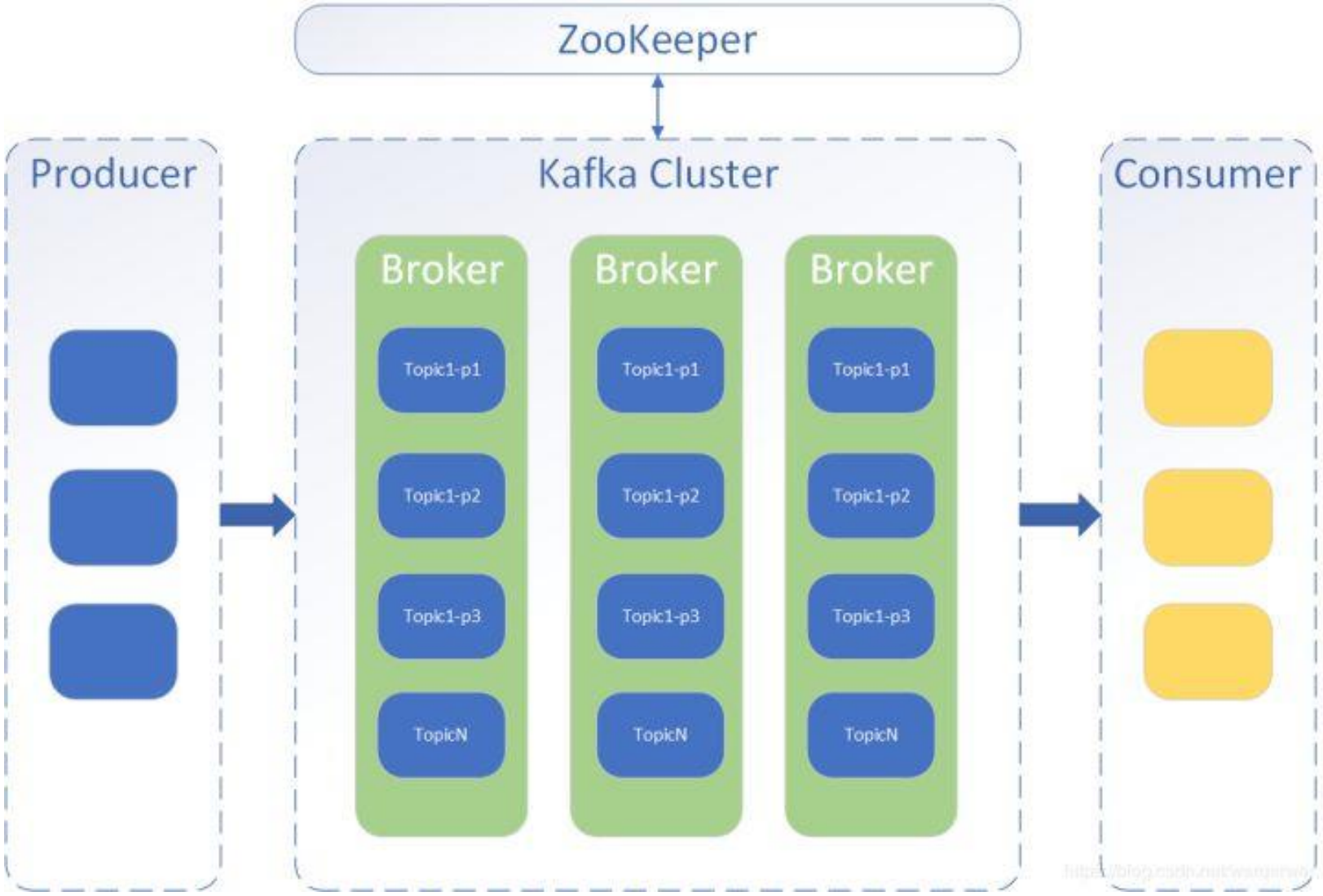
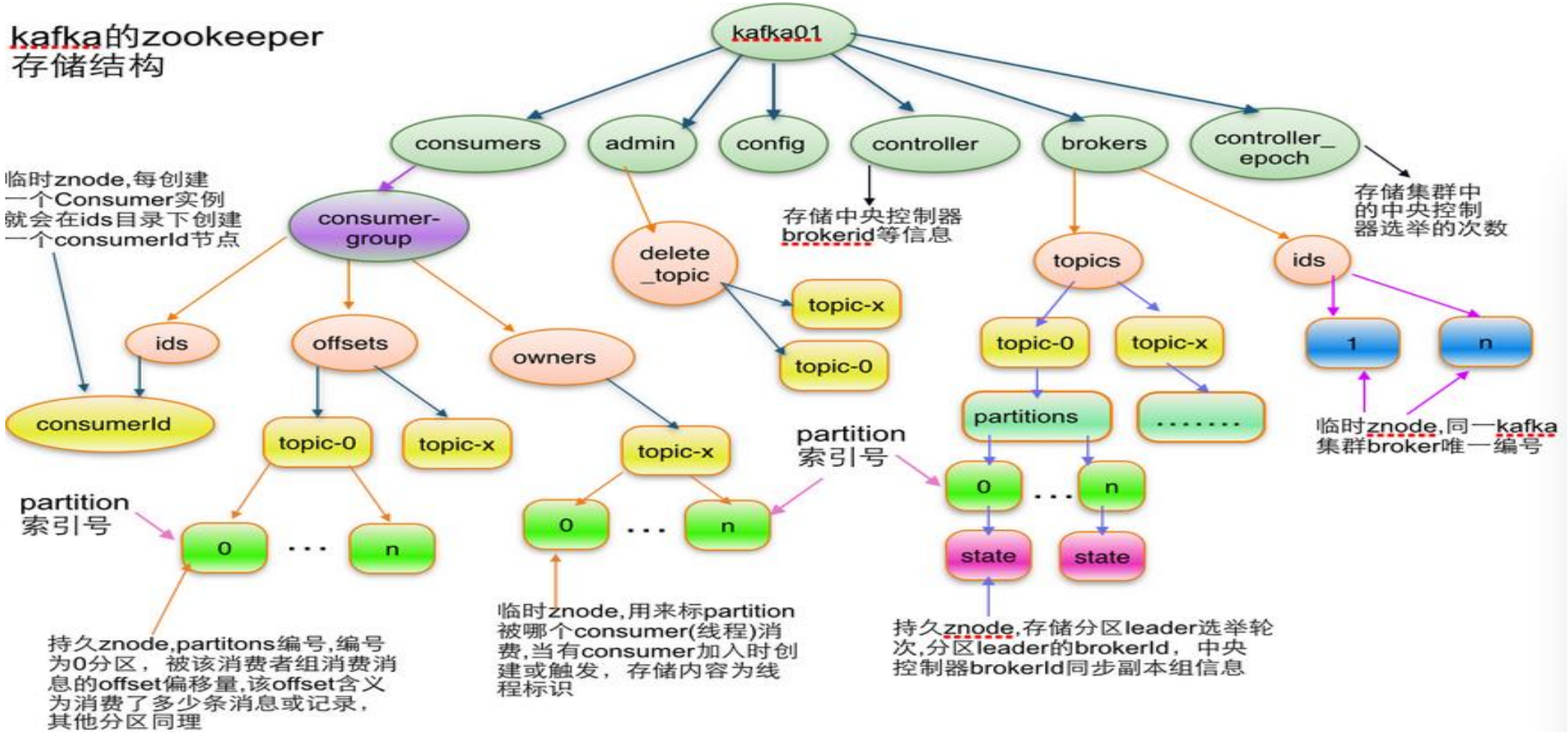
Kafka 课程培训：机制简介

1.1 架构简介：总体架构

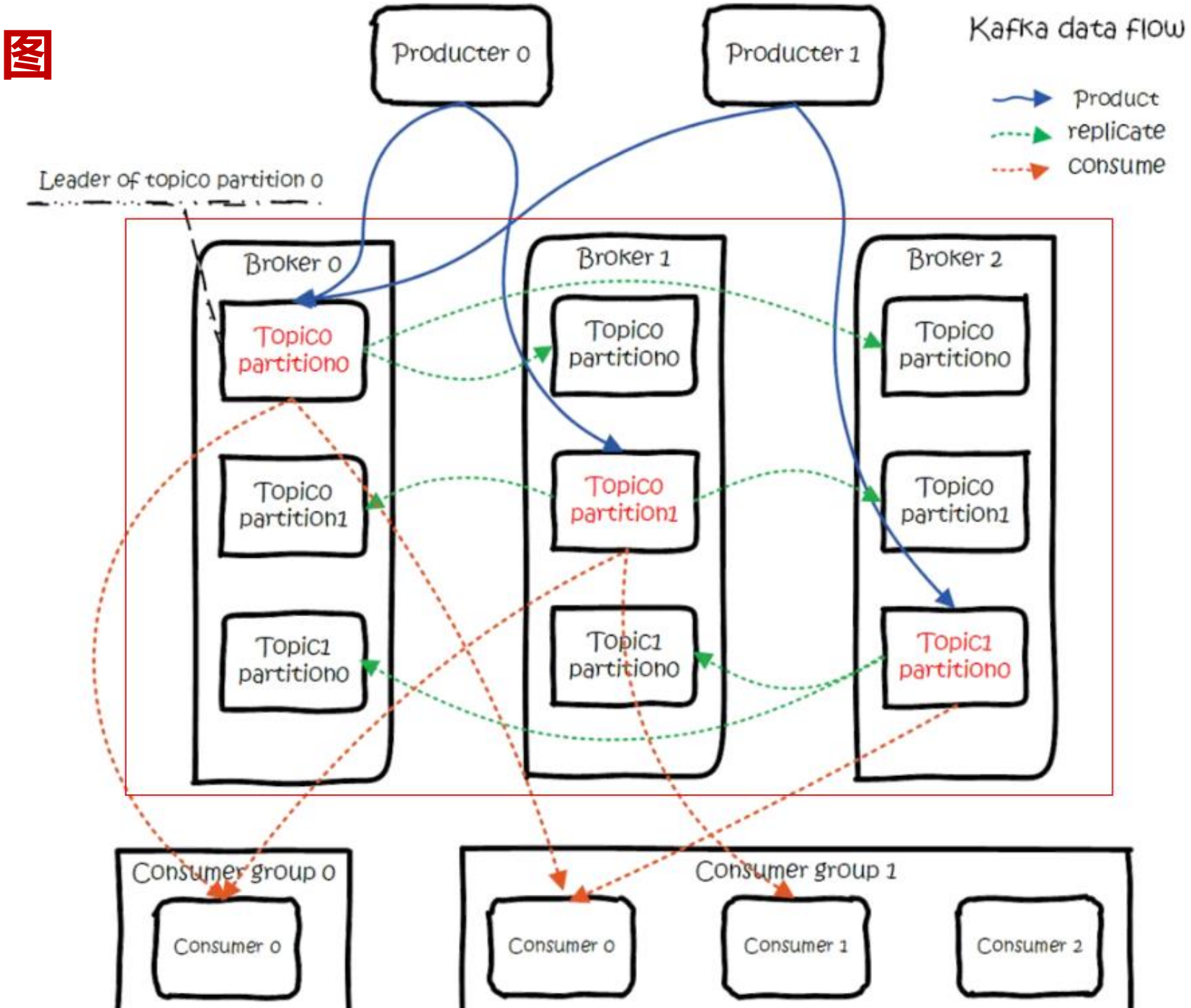
- Zookeeper：存储Kafka元数据
- Broker互为主备
- topic按分区存储
- 副本分部在不同节点



1.2 Kafka在Zookeeper存储结构



1.3 架构简介：数据流图

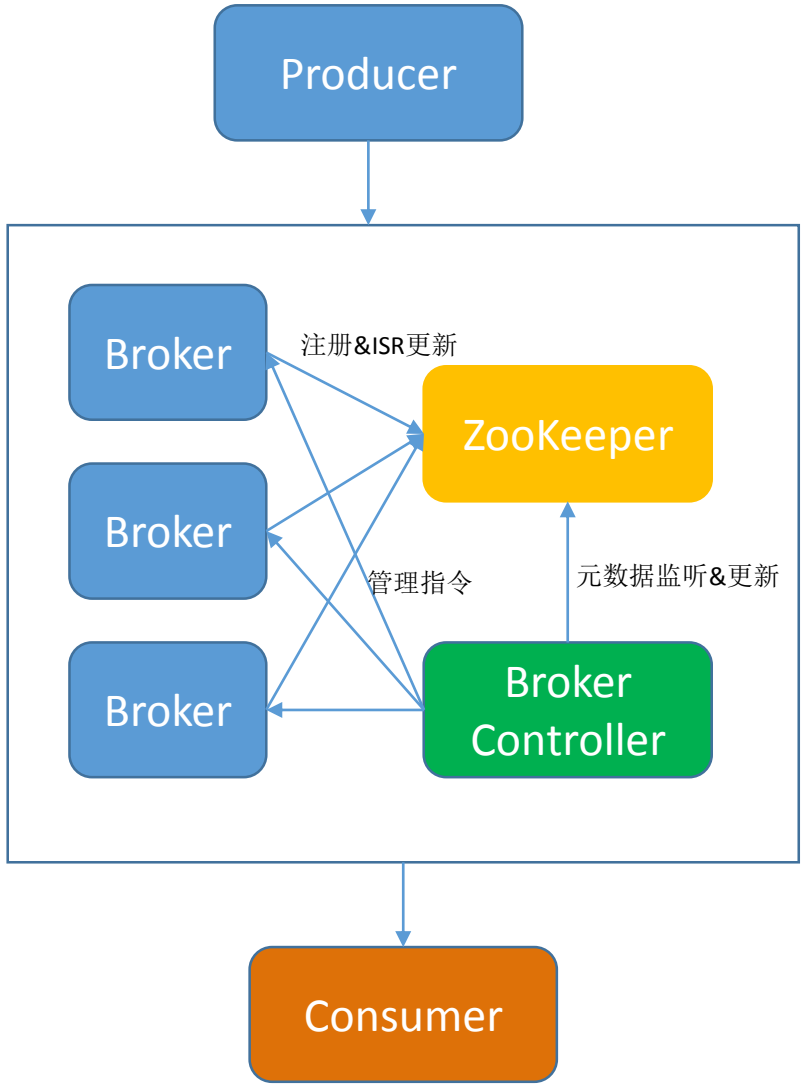

Kafka data flow
Product
replicate
consume
Producter 0
Producter 1
Leader of topico partition 0
Broker 0
Broker 1
Broker 2
Topico partition0
Topico partition1
Topic1 partition0
Consumer group 0
Consumer 0
Consumer group 1
Consumer 0
Consumer 1
Consumer 2

1.4 架构简介：节点角色

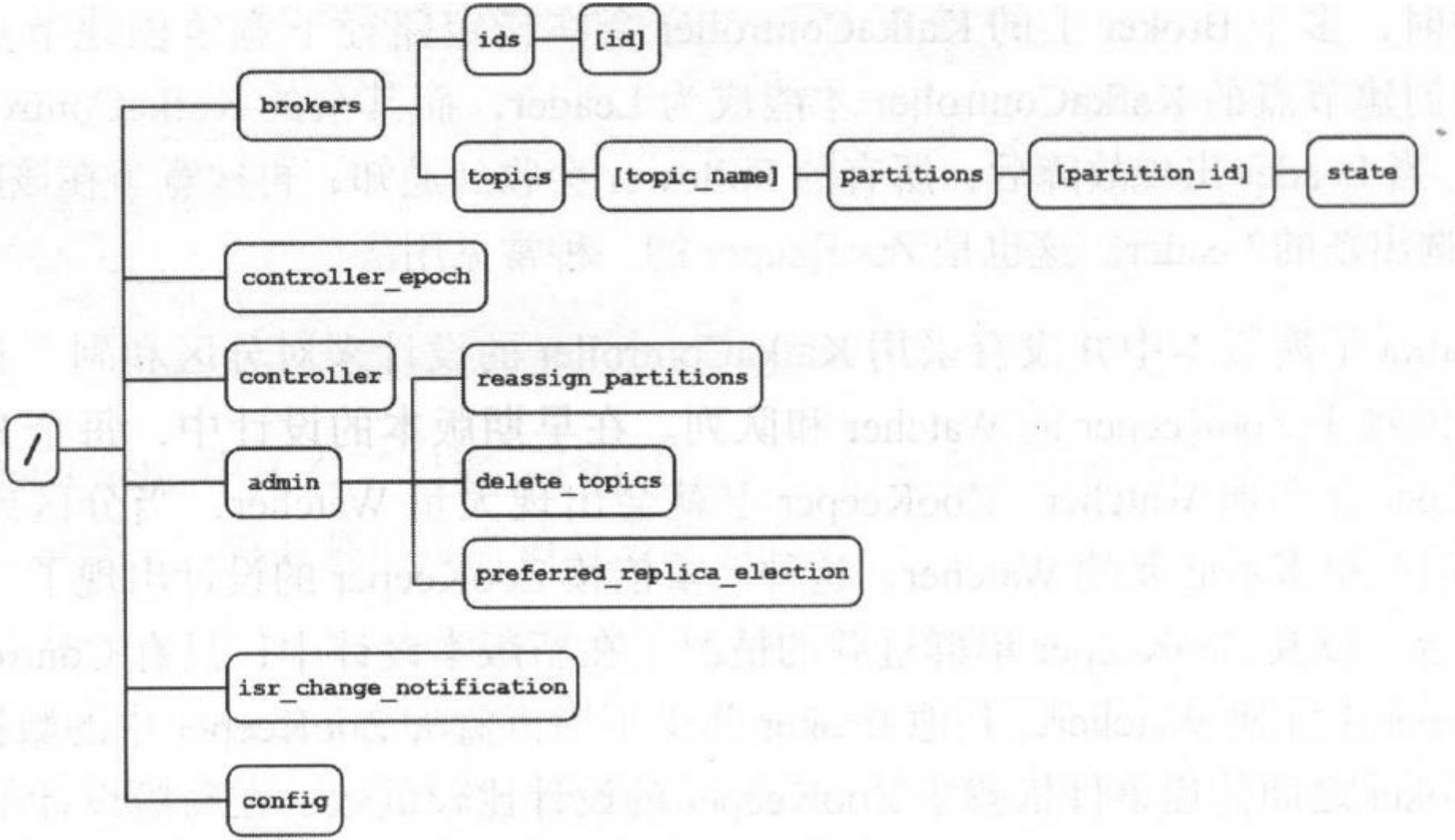
- Controller：主要负责Partition管理和副本状态管理，也会执行类似于重分配Partition之类的管理任务。如果当前的Controller失败，会从其他正常的Broker中重新选举Controller。
- Leader：负责分区消息的读写请求，故障会自动进行切换
- Follower：负责同步leader数据，形成副本，leader故障时，可变成leader，进行故障切换
- Coordinator：负责consumer group的管理，内部消费进度队列的维护等

2.1 Controller

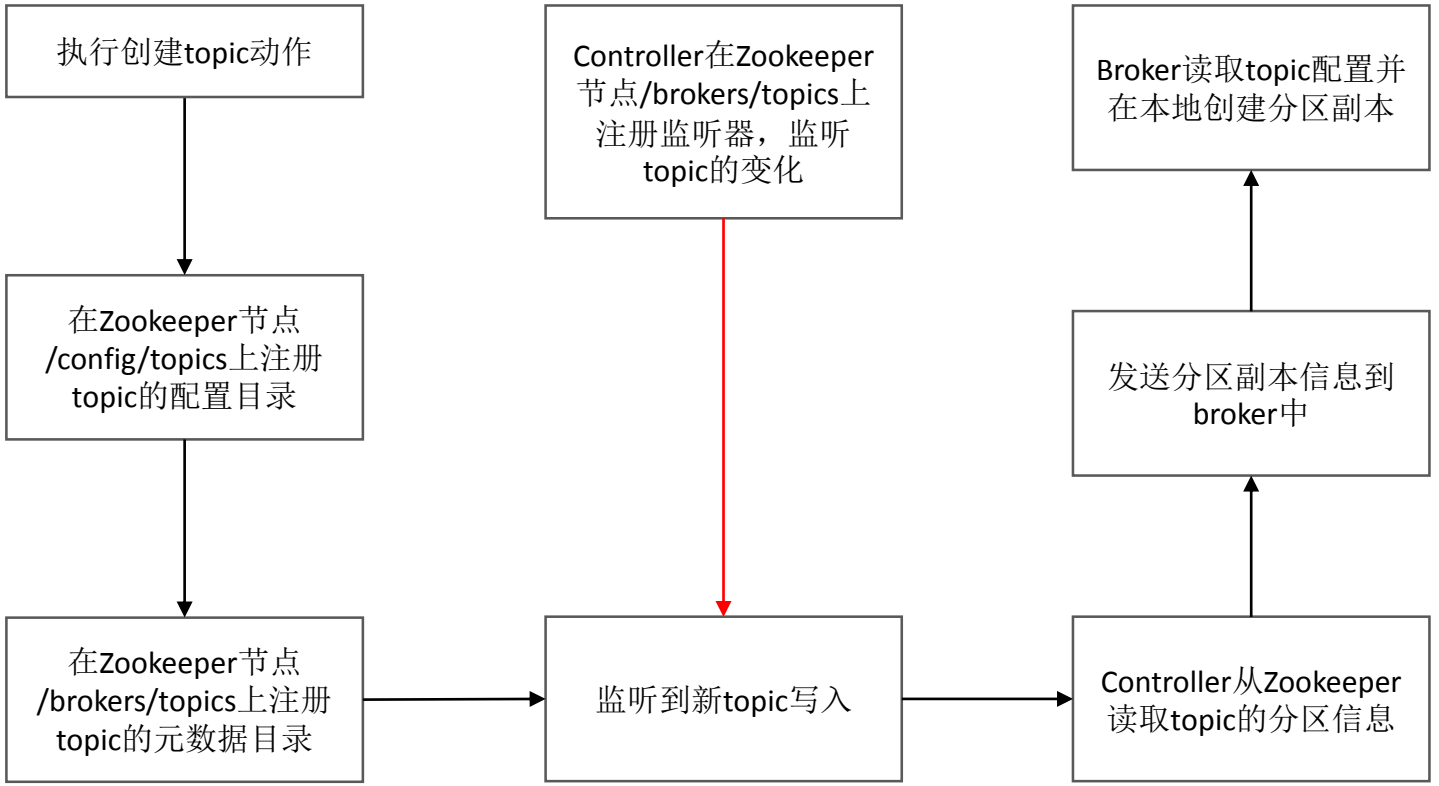
- Broker节点状态管理：新增节点、节点下线、节点故障、节点元数据更新等
- Topic分区状态管理：topic创建删除、分区扩容、分区迁移、leader切换等



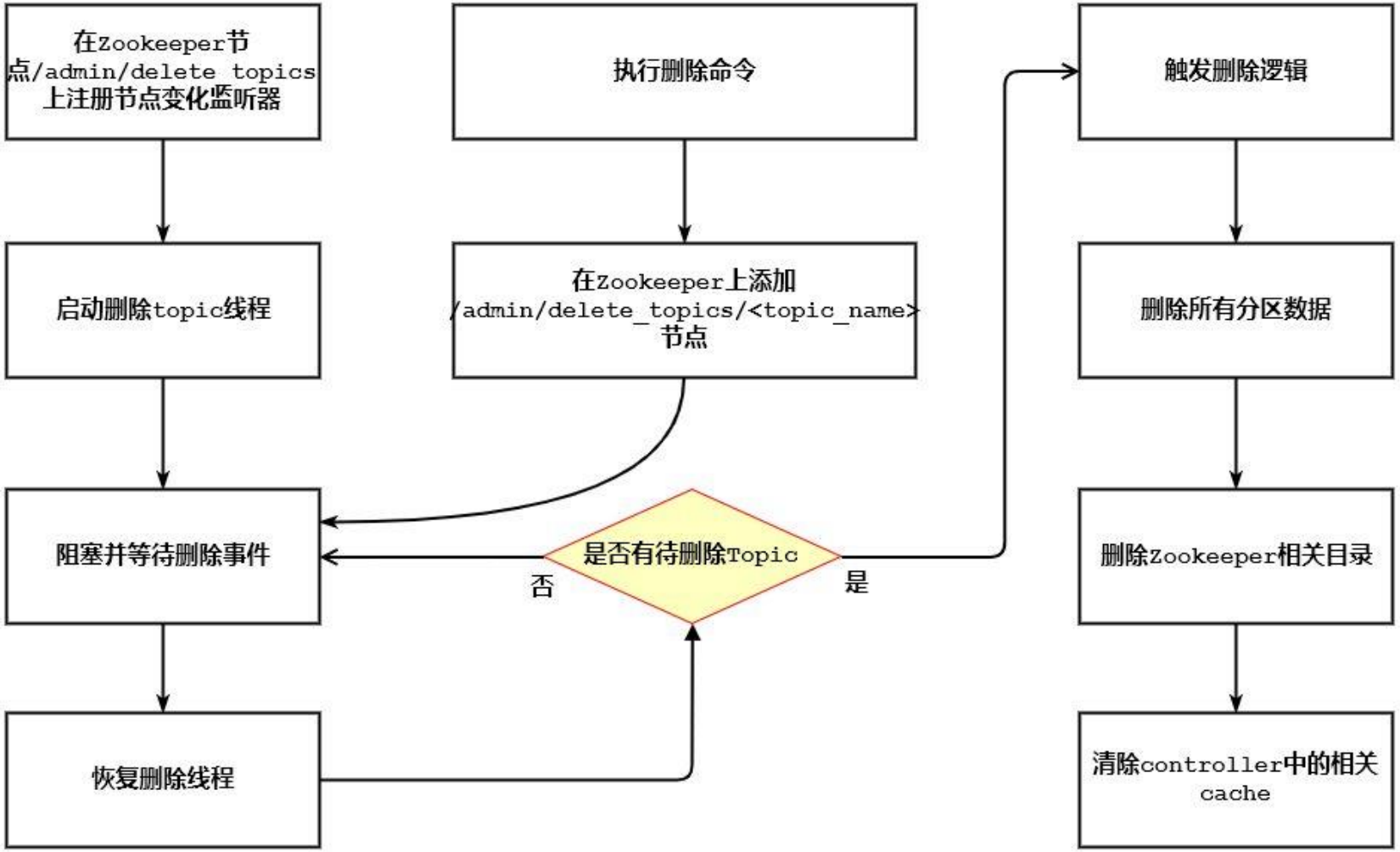
2.2 Controller zookeeper相关节点



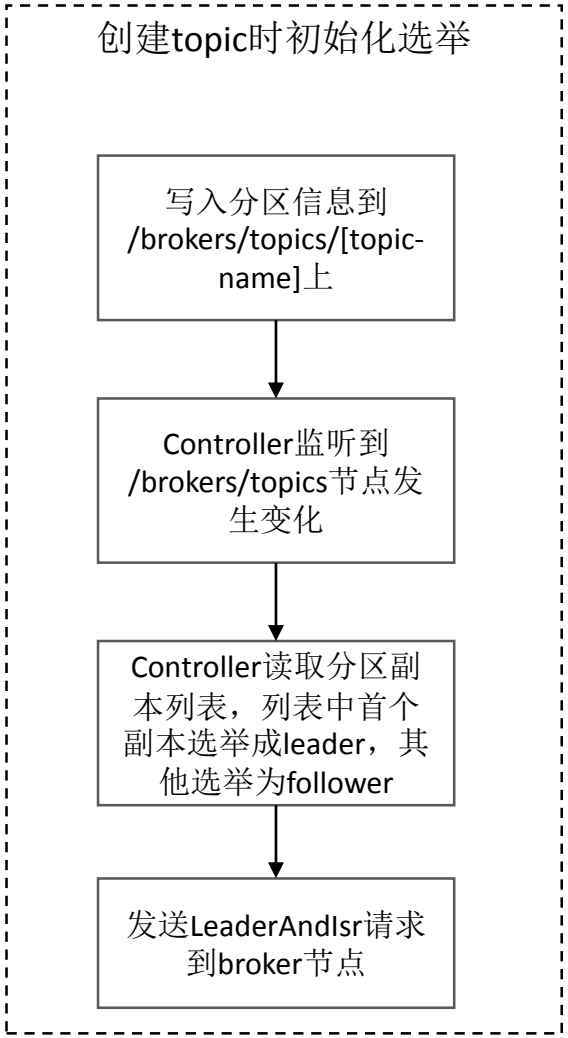
2.3 topic创建流程

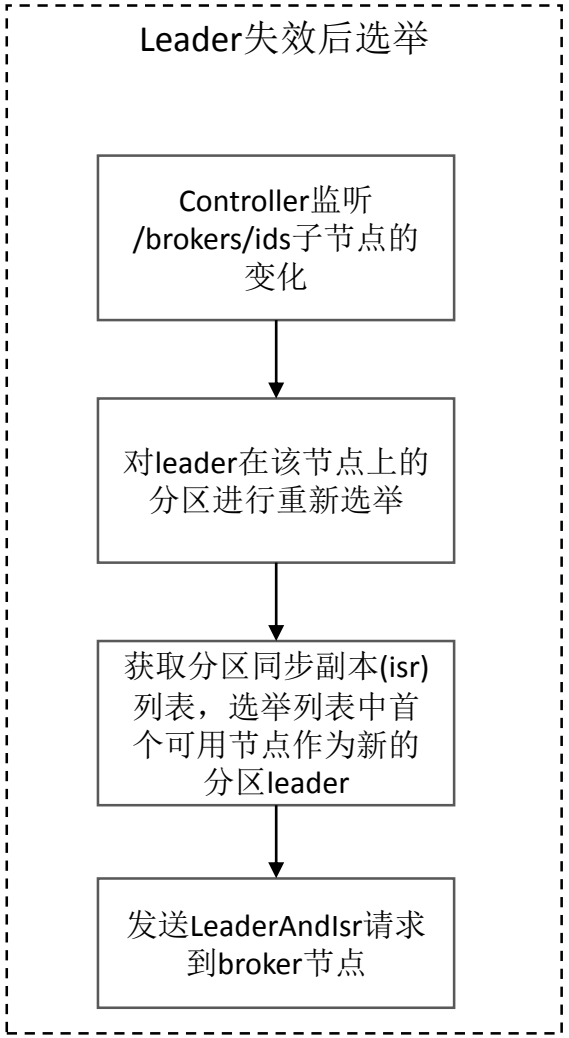


2.4 topic删除流程

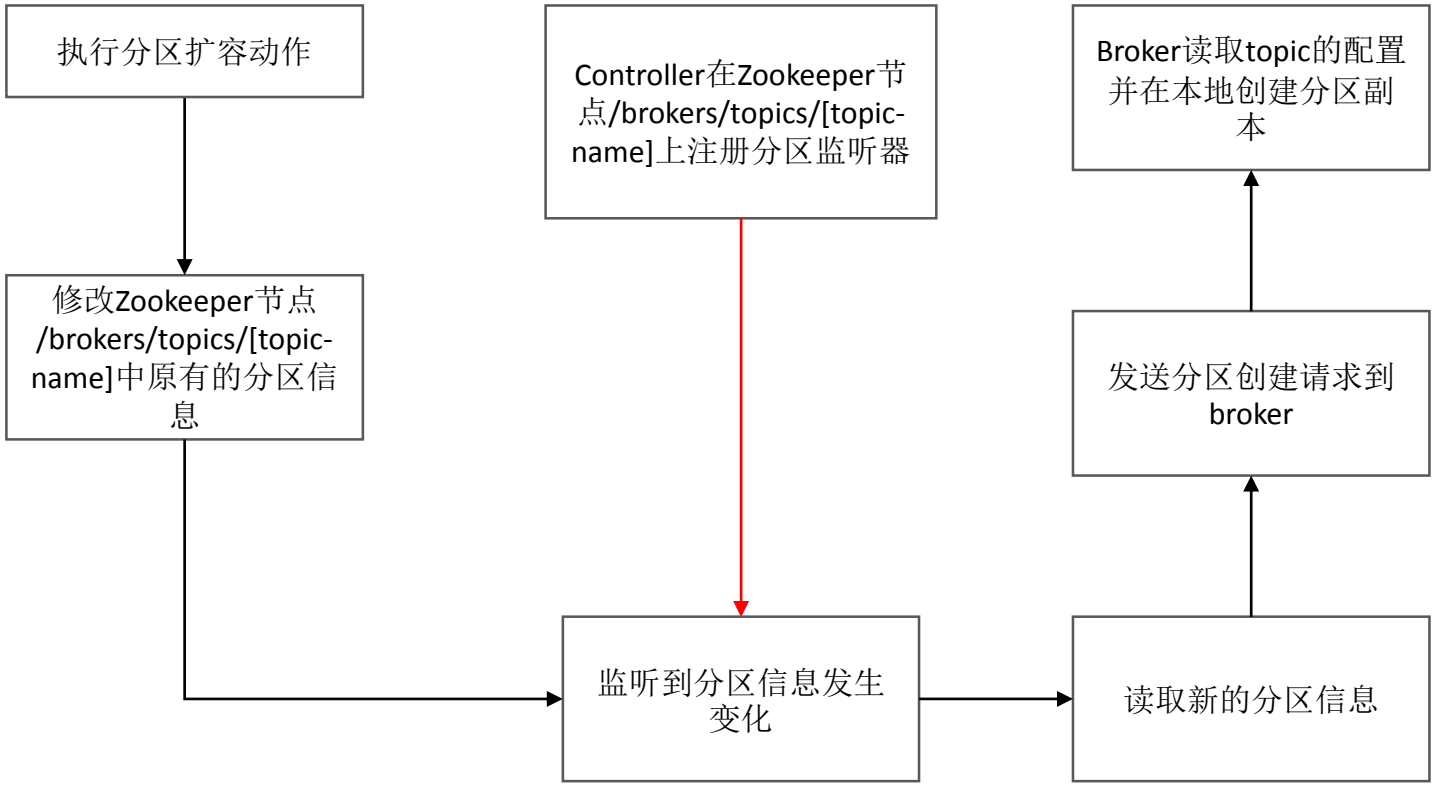


2.5 leader选举流程

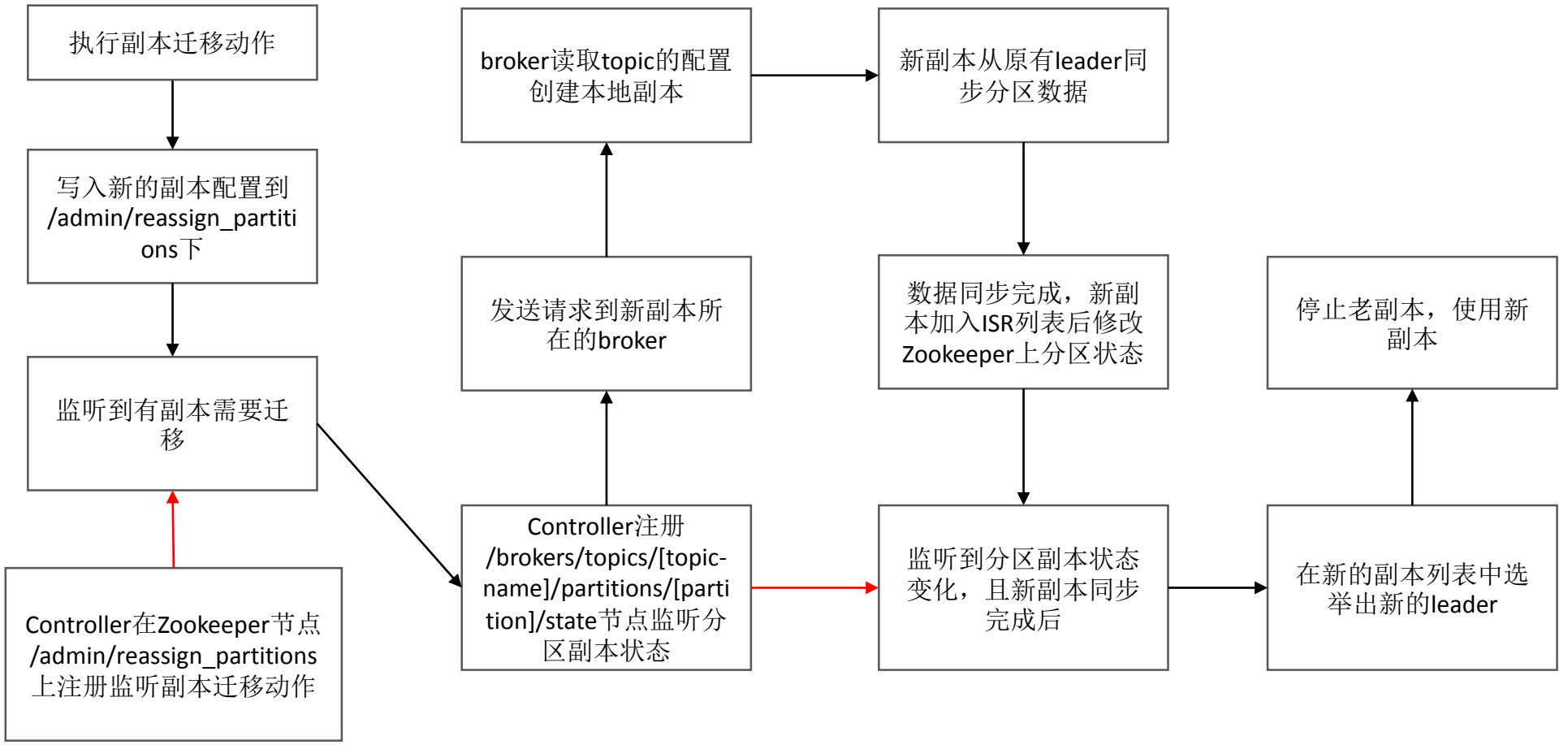




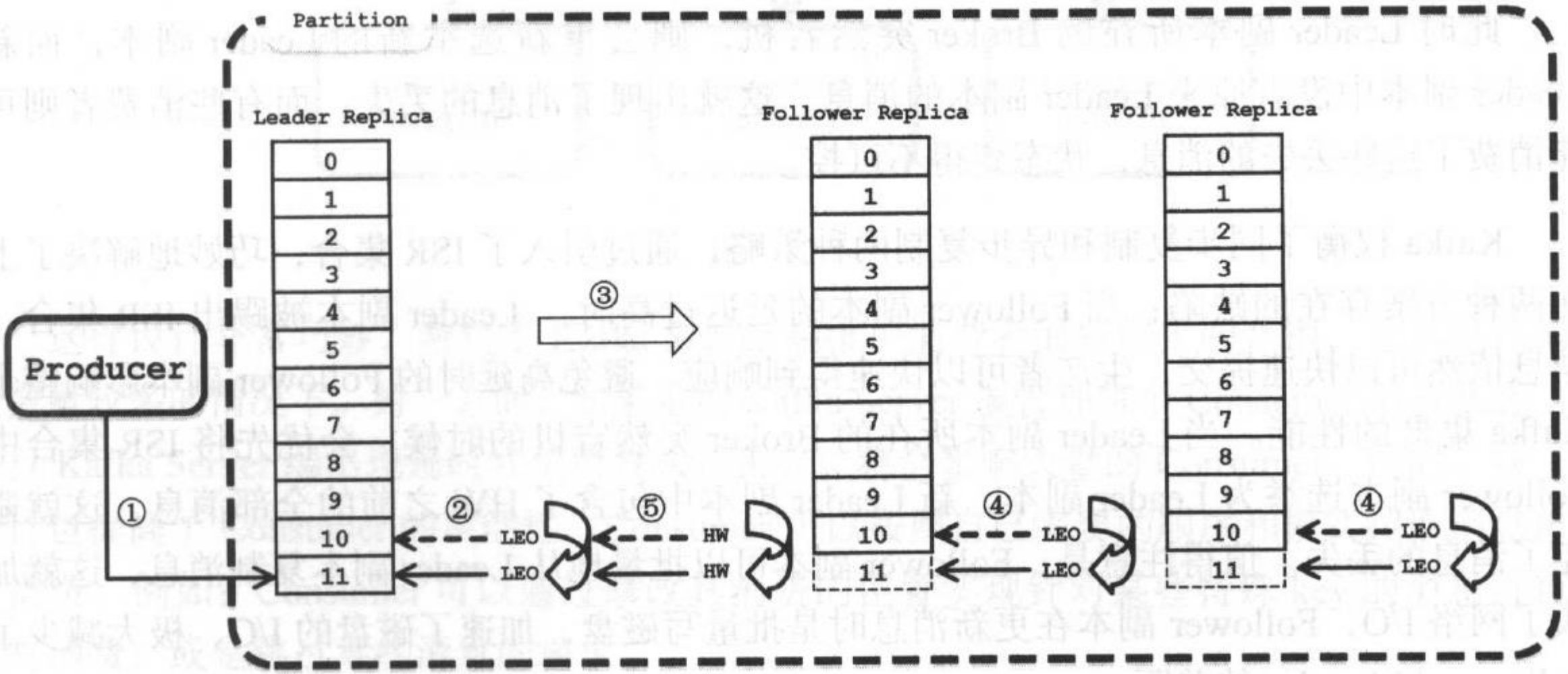
2.6 分区扩容流程


执行分区扩容动作
修改Zookeeper节点/brokers/topics/[topic-name]中原有的分区信息
Controller在Zookeeper节点/brokers/topics/[topic-name]上注册分区监听器
监听到分区信息发生变化
读取新的分区信息
发送分区创建请求到broker
Broker读取topic的配置并在本地创建分区副本

2.7 副本迁移流程


执行副本迁移动作
写入新的副本配置到/admin/reassign_partitions下
监听到有副本需要迁移
Controller在Zookeeper节点/admin/reassign_partitions上注册监听副本迁移动作
Controller注册/brokers/topics/[topic-name]/partitions/[partition]/state节点监听分区副本状态
发送请求到新副本所在的broker
broker读取topic的配置创建本地副本
新副本从原有leader同步分区数据
数据同步完成，新副本加入ISR列表后修改Zookeeper上分区状态
监听到分区副本状态变化，且新副本同步完成后
在新的副本列表中选举出新的leader
停止老副本，使用新副本

2.8 生产请求流程


Partition
Leader Replica
Follower Replica
Follower Replica
Producer
①
②
③
④
⑤
LEO
HW
0
1
2
3
4
5
6
7
8
9
10
11

Thank You