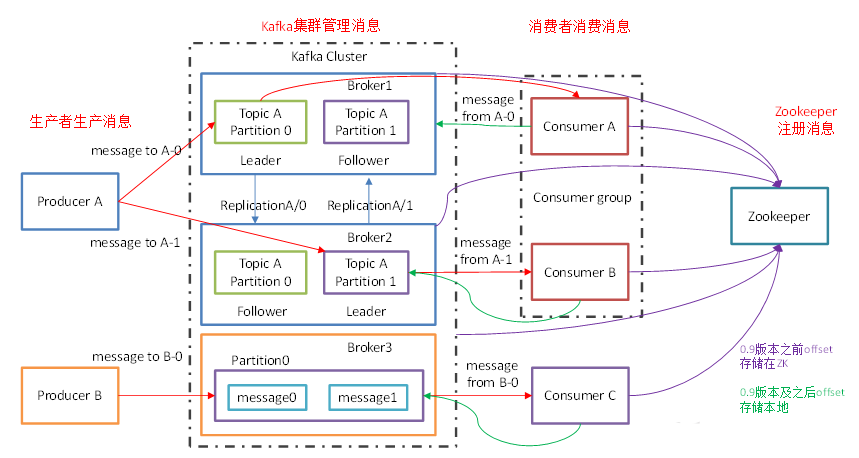
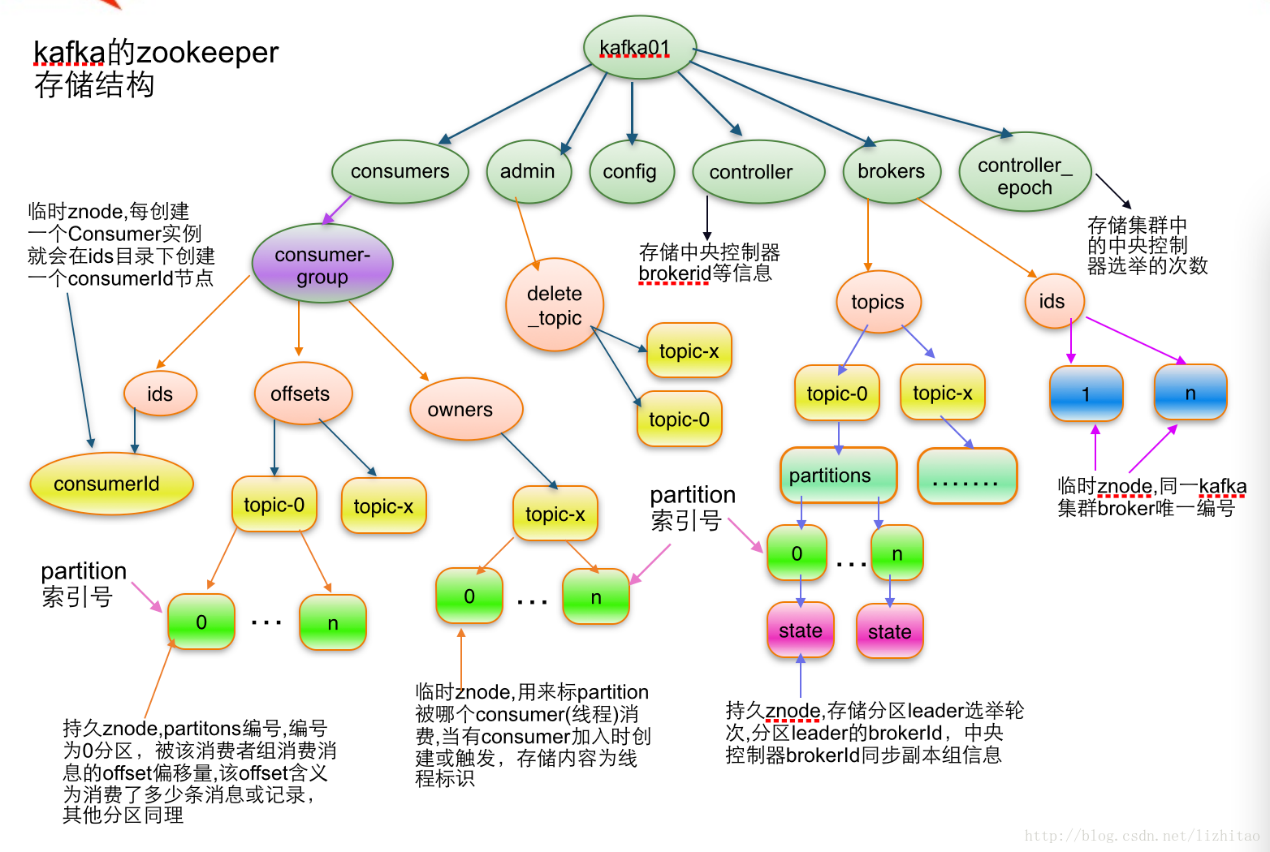
### 1 Kafka架构

生产者、Broker、消费者、ZK；

注意：Zookeeper中保存Broker id和消费者offsets等信息，但是没有生产者信息。





### 2 Kafka的机器数量

Kafka机器数量=2 \*（峰值生产速度\*副本数/100）+ 1

### 3 副本数设定

一般我们设置成2个或3个，很多企业设置为2个。

副本的优势：提高可靠性；副本劣势：增加了网络IO传输

### 4 Kafka压测

Kafka官方自带压力测试脚本（kafka-consumer-perf-test.sh、kafka-producer-perf-test.sh）。Kafka压测时，可以查看到哪个地方出现了瓶颈（CPU，内存，网络IO）。一般都是网络IO达到瓶颈。

### 5 Kafka日志保存时间

默认保存7天；生产环境建议3天

### 6 Kafka监控

公司自己开发的监控器；

开源的监控器：KafkaManager、KafkaMonitor、KafkaEagle

### 7 Kakfa分区数

1）创建一个只有1个分区的topic

2）测试这个topic的producer吞吐量和consumer吞吐量。

3）假设他们的值分别是Tp和Tc，单位可以是MB/s。

4）然后假设总的目标吞吐量是Tt，那么分区数=Tt / min（Tp，Tc）

例如：producer吞吐量=20m/s；consumer吞吐量=50m/s，期望吞吐量100m/s；

分区数=100 / 20 =5分区

<https://blog.csdn.net/weixin_42641909/article/details/89294698>

分区数一般设置为：3-10个

### 8 多少个Topic

通常情况：多少个日志类型就多少个Topic。也有对日志类型进行合并的。