

kafka: 基于发布和订阅的消息系统, 它一般被称为"分布式提交平台"或者"分布式流平台"

消息和批次: Kafka的数据单元被称为消息, 批次就是一组消息

主题和分区: kafka通过主题进行分类, 通过分区实现数据冗余和伸缩性

生产者: 生产者默认情况下把消息均匀的发布到所有分区, 如果指定了消息键和分区器, 分区器会为键生成一个散裂值, 并将其映射到指定的分区上

消费者: 消费者通过检查消息的偏移量来区分已经读过的消息。消费者是消费者群组的一部分, 消费者和分区之间的映射通常被称为消费者对分区的所有权关系

broker和集群: 一个独立的kafka服务器被称为broker。1) broker接受生产者的消息, 为消息设置偏移量, 并且提交消息到磁盘保存。2) broker为消费者提供服务, 对读取分区的请求作出响应, 返回已经提交到磁盘上的消息

集群控制器: 集群中的broker, 有一个会充当集群控制器的角色。控制器负责分区分配给broker以及监控broker。在集群中, 一个分区从属于一个broker, 该broker被称为分区的首领。分区可以被分配给多个broker, 这个时候会发生复制, 也就是为分区提供了消息冗余, 。