

高吞吐量：能处理tb（啥是tb我也不知道）级别的海量数据，是他的能力

消息持久化，将消息传入硬盘里，硬盘便宜，空间大，处理海量数据的前提（要先有海量数据）虽在硬盘中，但并不是说在硬盘的读取就一点慢了，kafka的硬盘读取仍然很快，他是通过硬盘的顺序读取，而非随机读取，随机慢，但顺序读取甚至能快过内存，这两点保证了高吞吐量，海量数据，速度快

高可靠：通过分布式保证

高扩展：集群不够了，加一台服务器

Broker：一个服务器成为一个broker

Zookeeper：用于管理集群，单独按一个zookeeper或kafka内置里也有

Topic：主题，消息队列的实现方式有两种：一种是点对点，阻塞队列就是点对点，每个数据被一个消费者消费；发布订阅模式：生产者将消息放到每一个位置，可以有很多消费者同时订阅消息，这个消息可以被多个消费者读到（同时或先后），kafka用的第二种发布订阅消息，生产者将消息发送到的位置就是topic

Partition：分区，对主题的位置分区，增强并发和服务器处理，又将主题分区，可以采用多线程的方式，同时向多个分区写数据，增强并发能力，每个分区按从前往后向里追加写入数据

Offset：消息在分区内的索引

副本，因为kafka是分布式消息引擎，为了让数据更可靠，将数据存放多份

Leader Replica：主副本，能力较强，当你从分区读数据时，主副本可以做出响应

Follower Replica：从副本，只用来存放数据，当主副本挂掉，集群会从 从副本中选一个作为leader

