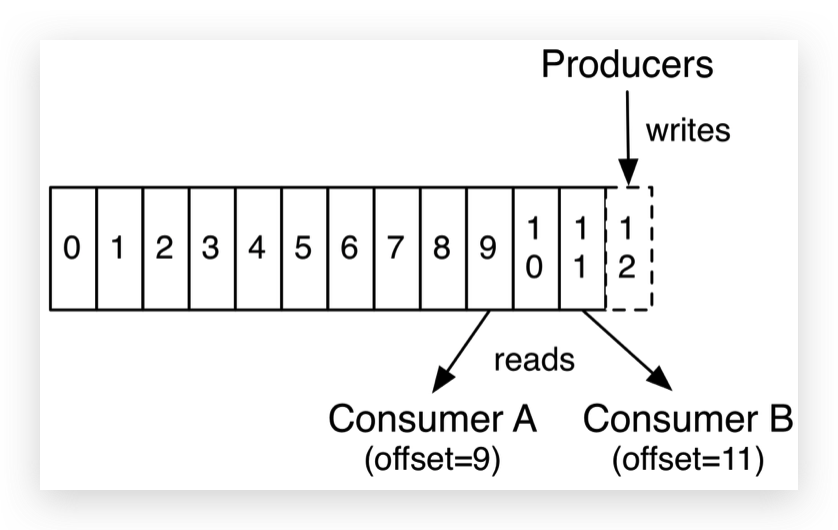
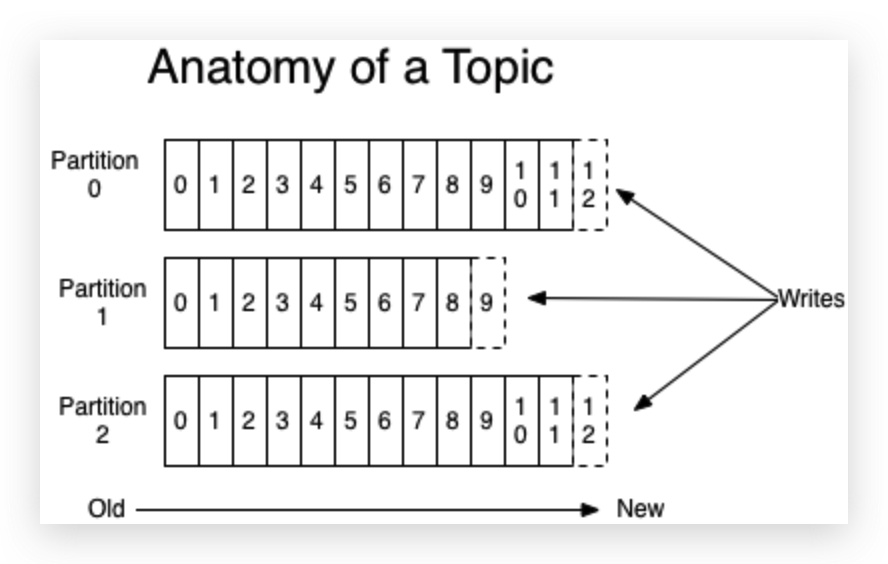
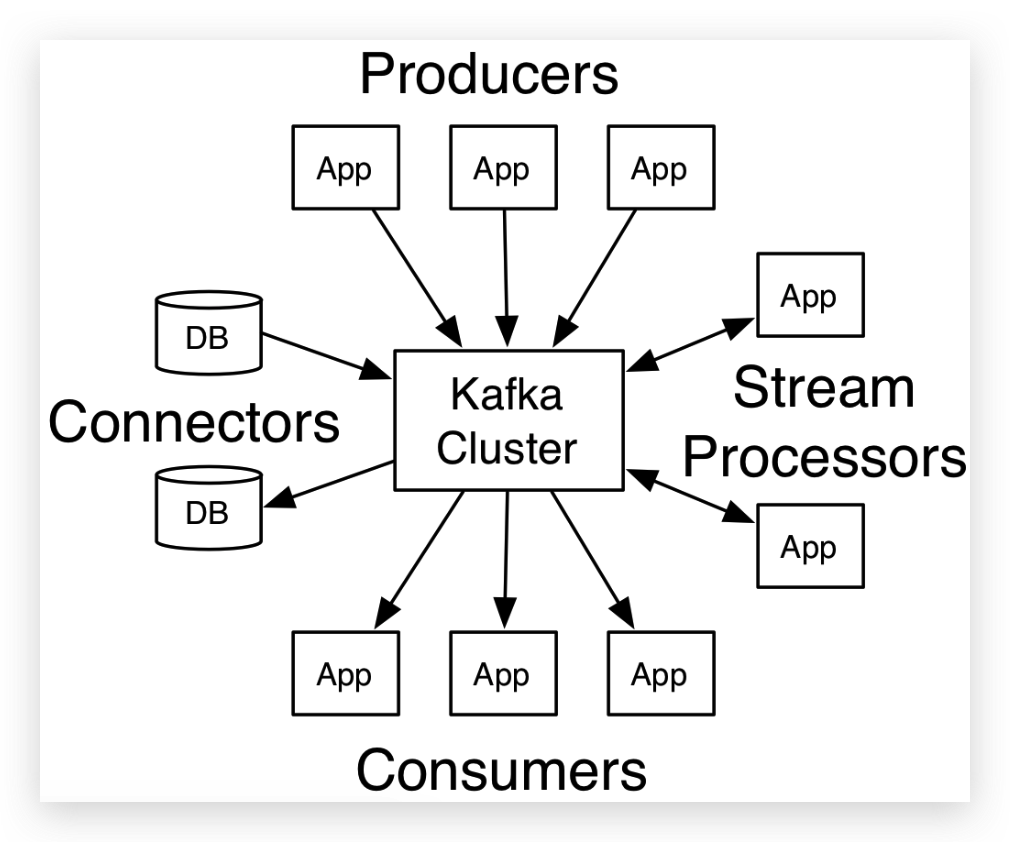
**distributed streaming platform**

****

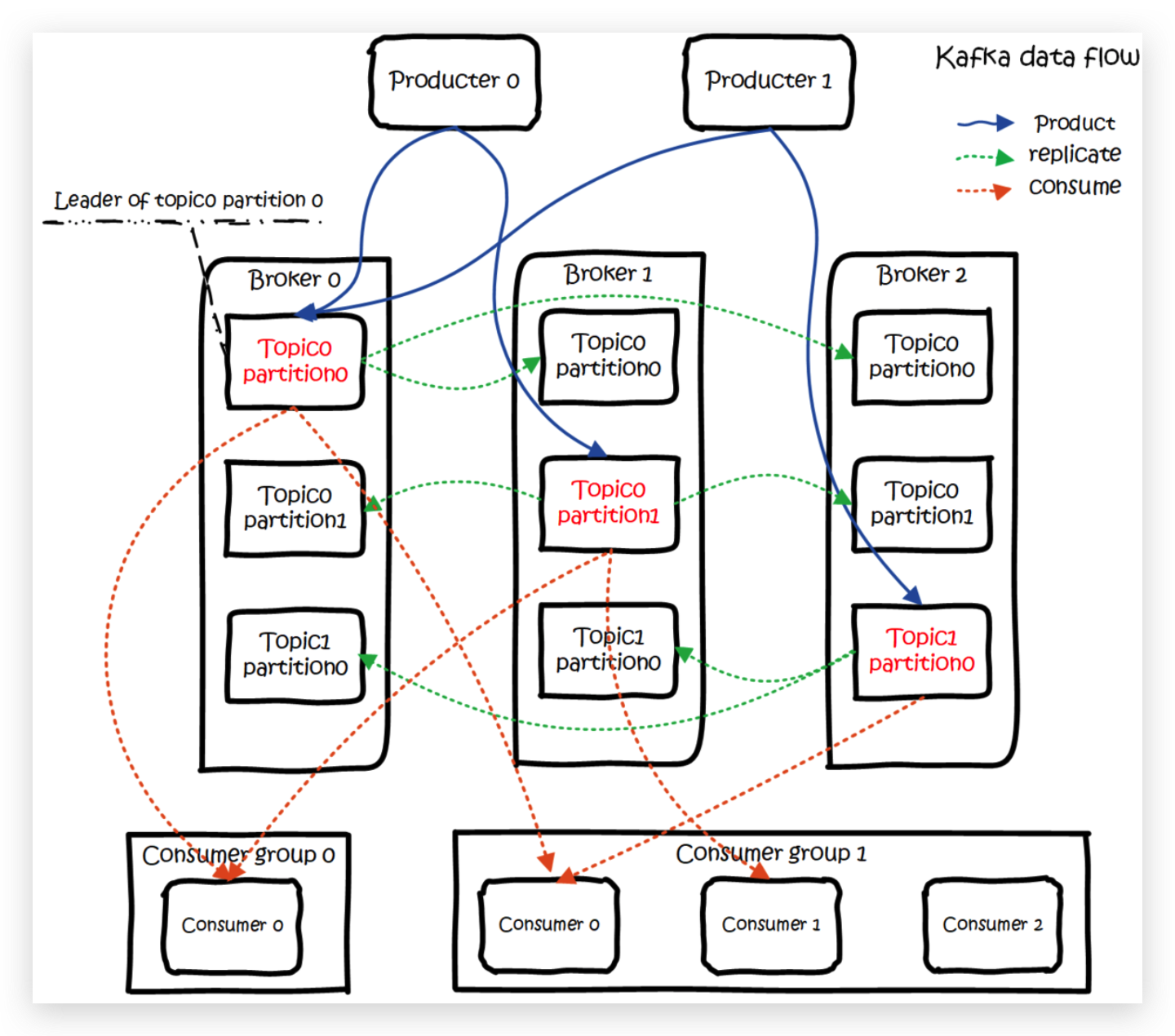
**Broker:** Topic的容器，数据备份

**Topic:** Partition的容器，数据分区

**Partition**: an ordered, immutable sequence of records that is continually appended to—a structured commit log. The records in the partitions are each assigned a sequential id number called the *offset* that uniquely identifies each record within the partition. The Kafka cluster durably persists all published records—whether or not they have been consumed—using a configurable retention period. Each partition is replicated across a configurable number of servers for fault tolerance. 有序的连续不可变记录，以offset作为序列号，需配置记录的保留时间，冗余备份的个数。

**Segment**:

**Record:** key, value, timestamp



**Producer:** publish a stream of records to one or more Kafka topics.

**Consumer:** subscribe to one or more topics and process the stream of records produced to them.

**Connector:** building and running reusable producers or consumers that connect Kafka topics to existing applications or data systems.

**Stream:** effectively transform the input streams to output streams.

**生产**

1. Message要素：(topic，partition，key，value)
2. Send规则：积攒一批消息，再发送
3. Partition分配规则：指定partition -> 按key哈希 -> round-robin
4. Broker冗余规则：为partition选出一个对外服务的leader节点，ISR(in-sync replica)列表作为备用的follower节点。

**消费**

1. Subscribe规则：Consumer Group <--> Topic <--> Producer

(Consumer) (Partition)

1. Offset的位置：groupid|topic|partition -> offset
2. Partition分配规则：partition leader所在的broker即为coordinator

|  |  |
| --- | --- |
| Broker | |
| broker.id | 集群中的进程节点编号 |
| listeners | 监听端口 |
| log.dirs |  |
|  |  |
|  |  |

**运维**

1. 列出所有topic



1. 查看topic详情

