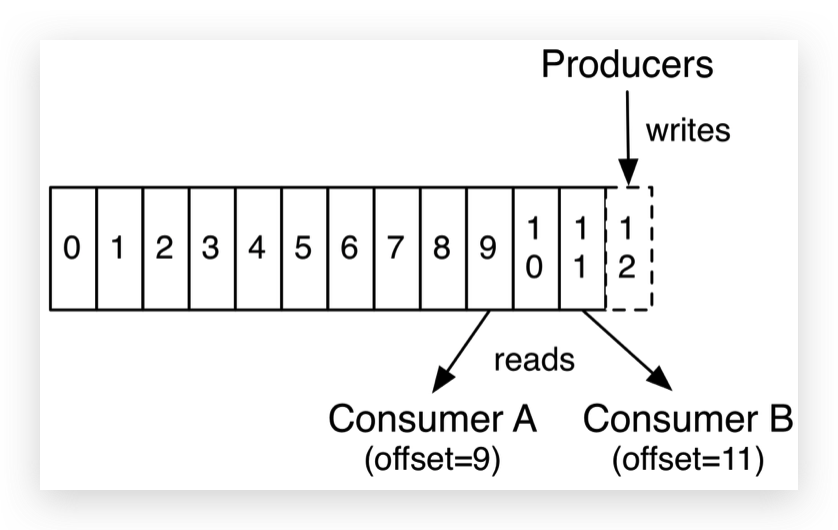
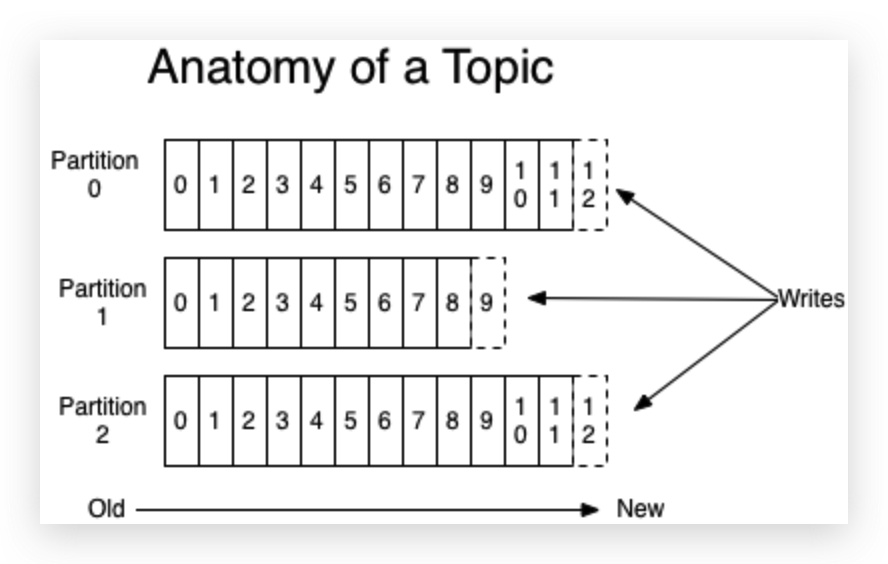
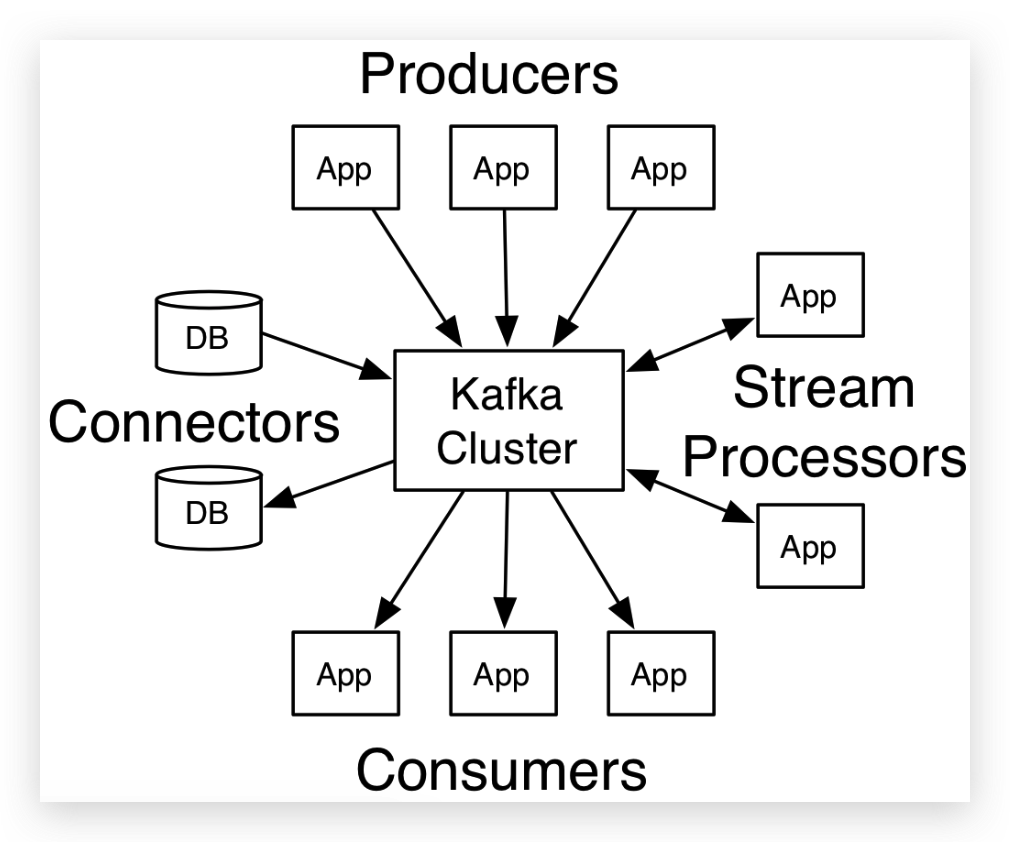
**distributed streaming platform**

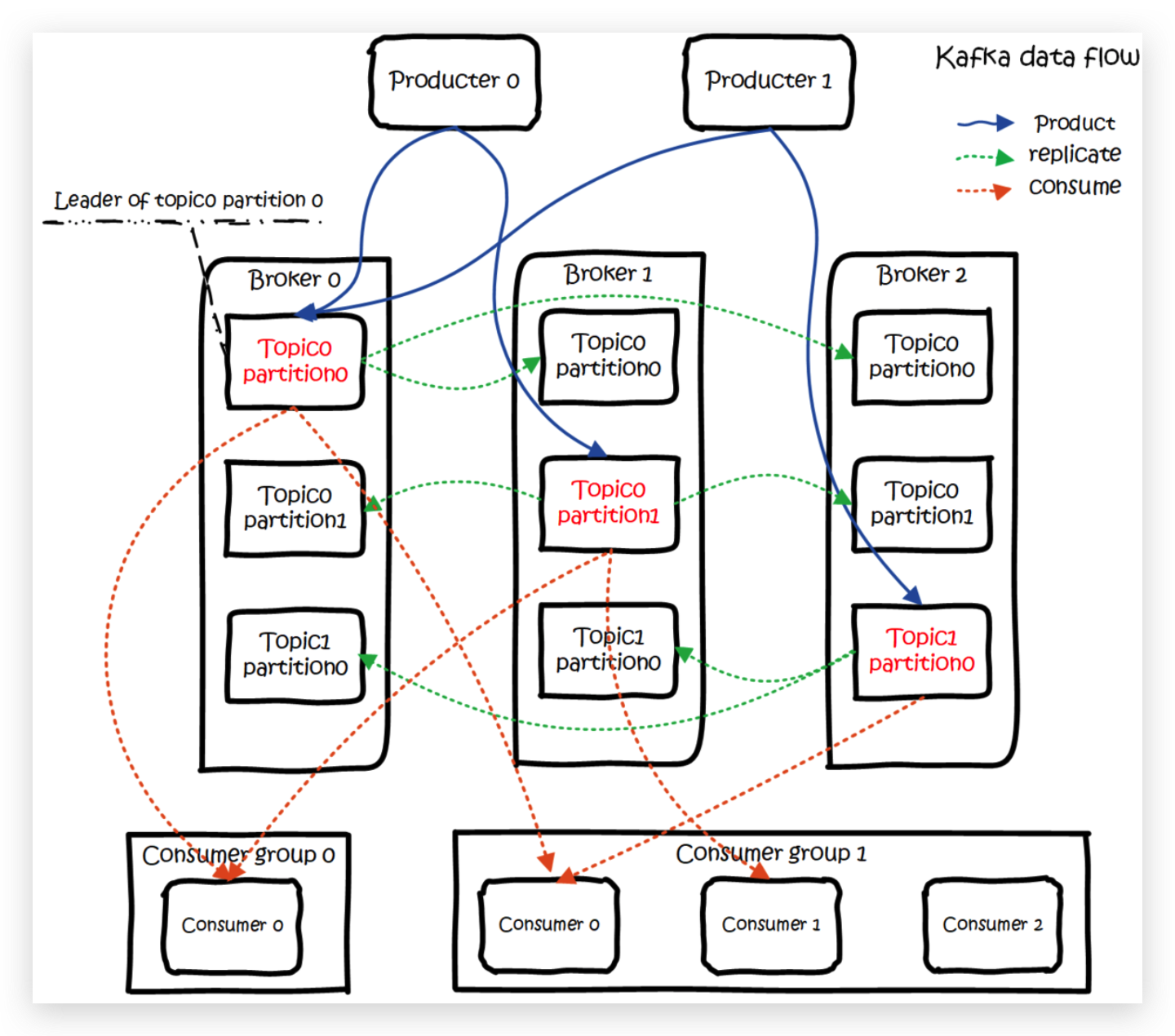
****

**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。

**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.

**tcp协议通信**