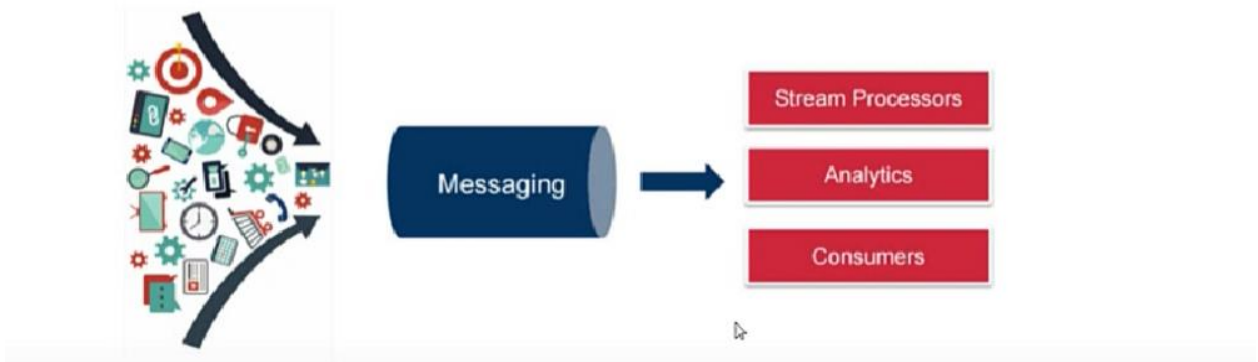


Apache Kafka

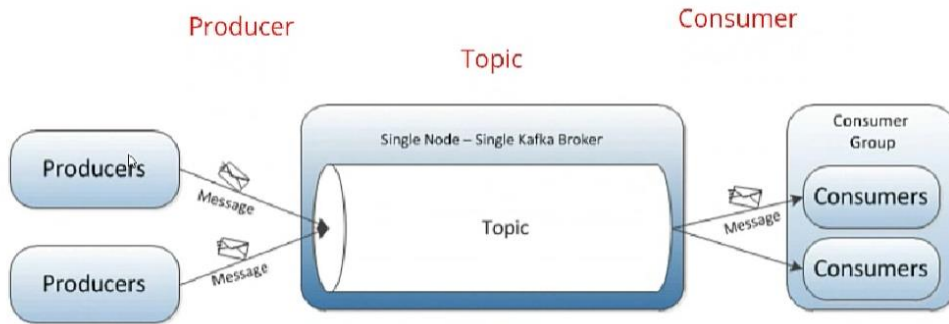
Günümüz uygulamalarında binlerce hatta milyonlarca veri aynı anda gönderilmektedir. Gelen verilerin direkt olarak işlenecek makinalara gönderildiğini varsayalım. Makinada çökme veya bozulma meydana geldiğinde veriler kaybolmuş, işlenemeyecek ve boşuna getirilmiş olacaktır.



Apache Kafka FIFO mantığıyla çalışan bir kuyruk yapısıdır. Yani gelen veriler önce gelen verinin arkasına eklenir ve önce gelen veri kullanılmak üzere gönderilir.

Uygulamadan gelen veriler direkt olarak makinalara göndermek yerine Apache Kafkaya gönderilir ve buradan da makina tarafına iletilir. Makinada bir bozulma meydana geldiğinde veriler kafkanın kuyruk yapısında kalır ve makina düzeldiğinde kaldığı yerden verilerde herhangi bir hata olmadan okumaya devam edebilmektedir.

Kafka çalışma prensibi ve kavramları



Producer: Verilerin gönderildiği kısımdır.

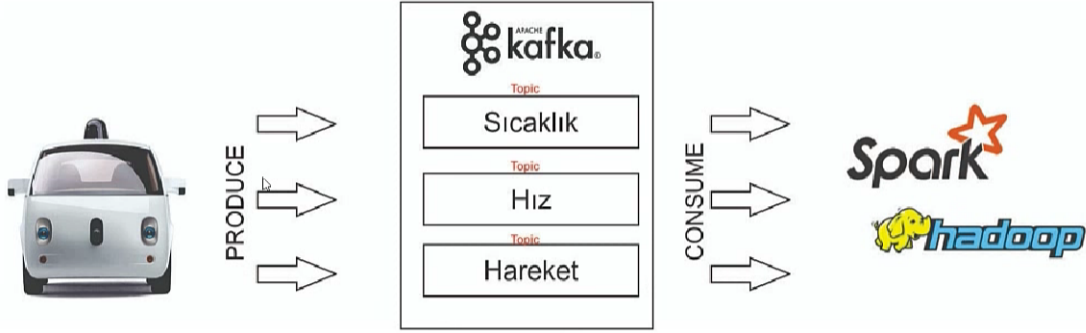
Consumer: Verileri çeken, kullanan yapıdır.

Topic: Konu başlığı, mesela sepet, log bilgileri

Broker: Server, makine

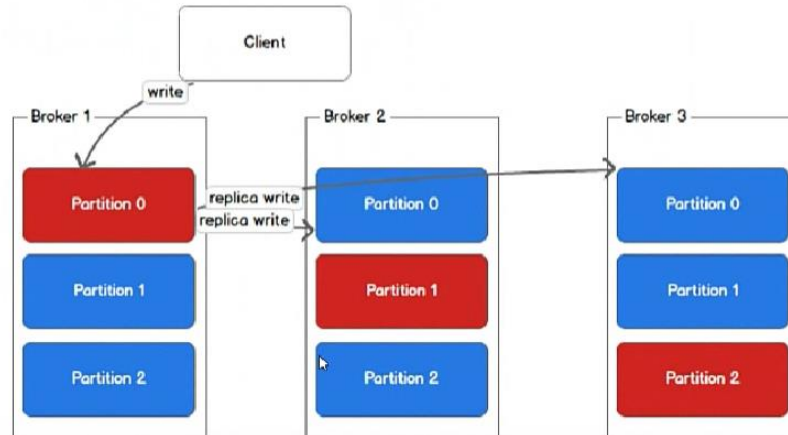
Offset: Kuyruk yapısından verileri okurken nerede kaldığını tutar. Analiz edecek makinalardan biri bozulduğunda offset kaldığı yeri tutacağı için makine düzeldiği zaman kaldığı yerden okumaya devam eder.

Veriler Producer'lar tarafından topic'e gönderilir. Consumerlar tarafından topic'den veriler öncelik sırasına göre hangisi geldiyse çekilir.



Google'ın sürücüsüz aracı (Producer) verileri kafkaya ayrı topiclerde gönderir. Mesela sıcaklık, hız, hareket topiclerini gönderdikten sonra consume edecek spark uygulaması onları ayrı ayrı çekmektedir. Böylece veriler düzenli bir şekilde kaybolmadan iletilir.

Leader (red) and replicas (blue)



Kopyalama işlemi (replication) kafkada yapılmaktadır. Veriler farklı makinalarda kopyaları mevcuttur. Fakat her makinada lider partition dediğimiz bölümler vardır ve brokerın ana görevi lider partition'ı kullanılmaktadır. Diğer makinalardaki bilgisayara bir şey olduğunda kopyası diğer brokerlarda olacağı için veri kaybı olmaz.



Zookeeper dağıtık makinelerdeki kaynak yönetimini kontrol eder. Konfigurasyon için kullanılır.

Zookeeper kafkanın serverler, makinalar yani brokerlar arasındaki etkileşimi sağlar.