APACHE KAFKA

Apache Kafka 2011 yılında LinkedIn tarafından kendi kullanımları için yüksek verimli bir mesaj aracısı olarak geliştirilmiştir. Daha sonra Java ve Scala diliyle yazılmış olan Apache Kafka LinkedIn tarafından açık kaynak kodlu olarak Apache Software Foundation’a bağışlanmıştır.[1] Kafka, herhangi bir yazılım çözümünün gerçek zamanlı sorunlarına, yani gerçek zamanlı bilgi hacimleriyle başa çıkmak ve onu hızla birden çok tüketiciye yönlendirmek için bir bulunmuş bir çözümdür.Kafka bilgi üreticilerini engellemeden ve tüketicilerin kim olduğunu üreticilere söylemeden üreticilerin ve tüketicilerin arasında sorunsuz bir entegrasyon sağlar.Kafka, mesajları kalıcı depolamada bölümlenmiş bir ileri yazma işlemi günlüğü olarak düzenler ve hem çevrimiçi hizmetler gibi gerçek zamanlı abonelerin hem de Hadoop ve veri ambarı gibi çevrimdışı abonelerin bu mesajları keyfi bir hızda okumasına izin vermek için çekme tabanlı bir mesajlaşma olanağı sağlar.[2]

Açık kaynak kodlu, publish-subscribe mesajlaşma sistemine ve dağıtık bir sisteme sahip olan Apache Kafka aşağıdaki özellikler ile tasarlanmıştır[3];

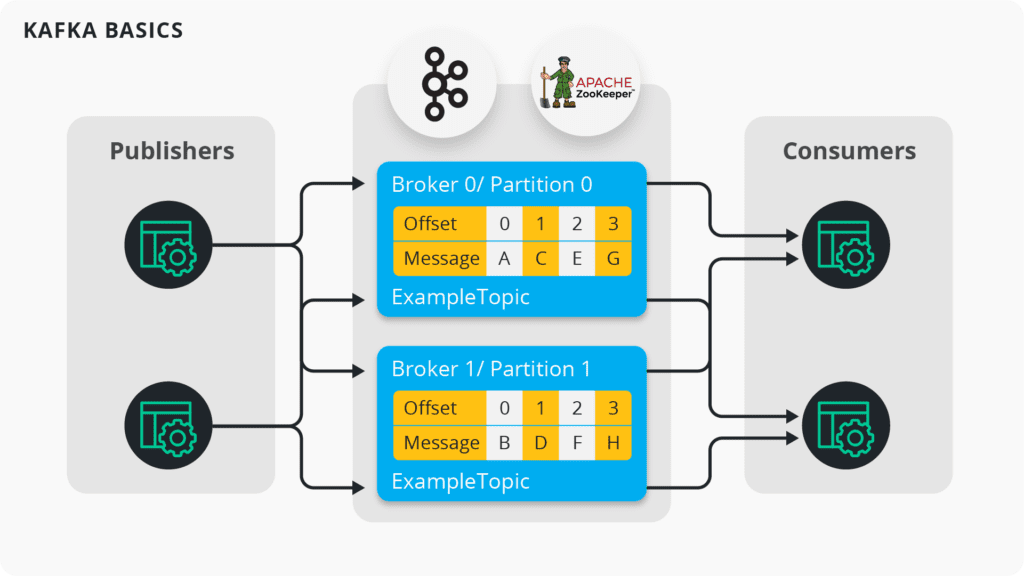
1)Kalıcı Mesajlaşma= Büyük veriden gerçek bir değeri elde etmek için veri kaybı yaşanmaması gerekir.Kafka veri ne kadar büyük olursa olsun sabit zamanlı performans sağlayan disk yapılarıyla tasarlandığından dolayı veri kaybı söz konusu değildir.

2)Yüksek Verim=Büyük verileri göz önünde bulunduran Apache Kafka, ticari donanımlar üzerinde çalışmak ve saniyede milyonlarca veri okuyup yazmak için tasarlanmış bir yapıdır.

3)Dağıtık Sistem= Kafka verileri sunucu üzerinde bölümlendirmeyi ve tüketimi bir tüketici makineleri kümesi üzerinde dağıtmayı açıkça destekler.

4)Çoklu İstemci Desteği=Apache Kafka farklı platformdan istemcilerin entegrasyonunu destekler.(Ex:Java, .NET, PHP, Ruby, and Python)

5)Gerçek Zamanlı=Apache Kafka’nın ana kurulum sebeplerinden birisidir.Üretici tarafından üretilen verinin anlık olarak tüketicide görüntülenme isteğini Kafka karşılar.

Apache Kafka’nın Basit Bir Çalışma Modeli[4]

Kafkanın temel bileşenlerinden bahsedecek olursak;

A. Broker

Broker, Kafka sistemlerinin en temel bileşenlerinden biridir. Kafka kümesinde çalışan her bir servis veya sunucuya verilen isimdir.Brokerlar fiziksel olarak bir sunucuya bağlıdırlar fakat tüm kafka kümelerindeki konulardan ve bölümlerden haberdardır.Kafka Cluster yapısında brokerlerin id’si kendilerine özeldir.

B. Zookeeper

Bir Kafka sisteminin çalıştırılması için ilk ayağı kaldırılması gereken yapıdır. Zookeeper ayrı bir servistir. Zookeeper bulunduğu servisteki tüm işlemleri ve değişiklikleri kontrol eder ve servis koordinasyonunu sağlamaya çalışır.

C. Topic

Topicler verilerin okunmak üzere yazıldığı ve verilerin depolandığı yapıdır.Veriler topic üstüne yazıldığında tüm brokerler verilere erişebilir.Topicler Kafka’nın veritabanı olarak nitelendirilebilir.İstenilen sayı kadar topic oluşturulabilir.

D. Producer

Kafka terminolojisine göre veri yaratan ve bu veriyi ilgili topiclere yazmayı sağlayan yapıdır.Producerler bootstrap server sayesinde tüm broker,topic ve partitionlara erişebilirler bu sayede mesajı nereye yazacağını ve nereye ne kadar yazabileceğini kendisi belirleyebilir.

E. Consumer

Kafka terminolojisine göre Topic’e yazılan verileri okuyan yapıdır.Producer gibi bootstrap server sayesinde tüm kafka cluster yapısına erişebilir.Kendi dengesini sağlayabilecek bir yapıya sahiptir.Consumer groupları belirleyerek Kafka’yı ister queue ister pub-sub modeli ile çalıştırmamıza olanak tanır.

[1]= “What Is Apache Kafka?”, 07 Nisan 2021. https://www.ibm.com/cloud/learn/apache-kafka.

[2]=] Jay Kreps, Neha Narkhede and Jun Rao, ―Kafka: a Distributed Messaging System for Log Processing‖, LinkedIn Corp.

[3]= Garg, Nishant. Apache Kafka. Birmingham: Packt Publishing, 2013.

[4]= Solace. “Solace PubSub+ vs Kafka: The Basics”, 31 Mayıs 2021. https://solace.com/blog/solace-pubsub-vs-kafka-the-basics/.