

Basics

What is Kafka Used For? 1. Building real-time **streaming pipelines** that move data between different applications.
2. Building real-time **streaming applications** that are capable of processing streams of data.
3. Building a **fault tolerant storage system** that stores streams of records.

Topic A Kafka topic is a category or feed name under which messages are stored. A Kafka producer publishes messages to a topic, which may be subscribed by zero or more consumers.

Partitions A topic partition is a structured commit log to which the records are continually appended. For each topic, Kafka keeps a minimum of one partition. Each record in the partition is assigned a sequential id called as the offset, which uniquely identifies each of them within the partition. The partitions enable the topic to scale beyond a single server and act as the unit of parallelism.

Benefits of Kafka

Reliability Kafka's distributed design, topic partitioning, and data replication over servers make it reliable.

Scalability Kafka system exists as a cluster of brokers. The number of brokers can grow over time when more data comes. Any failure of an individual broker in a cluster is handled by the system providing uninterrupted service.

Durability Disk-based data retention makes Kafka durable. Messages remain on the disk based on the retention rule configured on a per-topic basis. Even if a consumer falls back due to any reason, the data continue to reside in the Broker till the retention period and is not lost.

High-Performance All the above features make Kafka a High-Performance messaging system.

C

By **Deepak** (deep.inside)
cheatography.com/deep-inside/

Not published yet.
Last updated 22nd February, 2023.
Page 1 of 1.

Sponsored by **Readable.com**
Measure your website readability!
<https://readable.com>