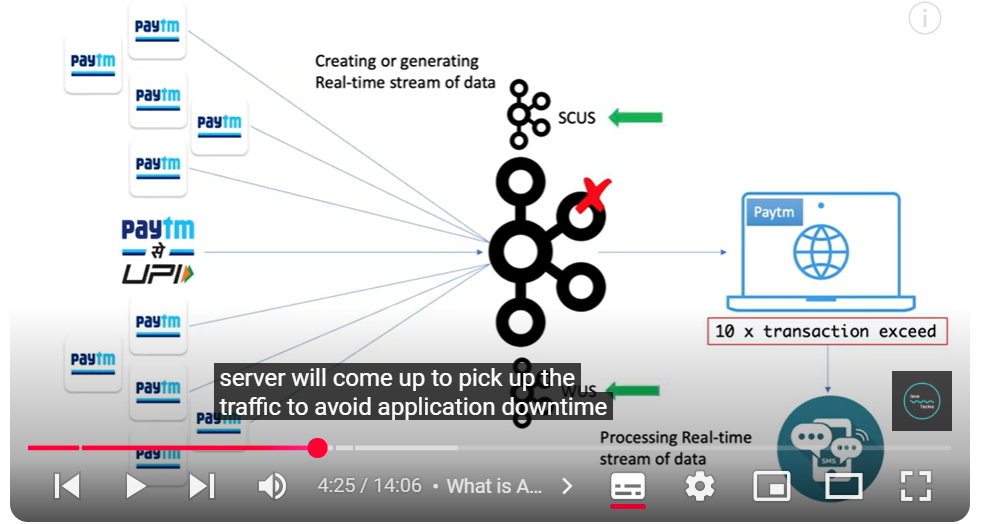
<https://www.youtube.com/watch?v=xGwzuz8F9k0&list=PLVz2XdJiJQxwpWGoNokohsSW2CysI6lDc>

**KAFKA**

Kafka= distributed real time event streaming

Continuously sending messages to the kafka server and reading/process them is called real time event streaming.

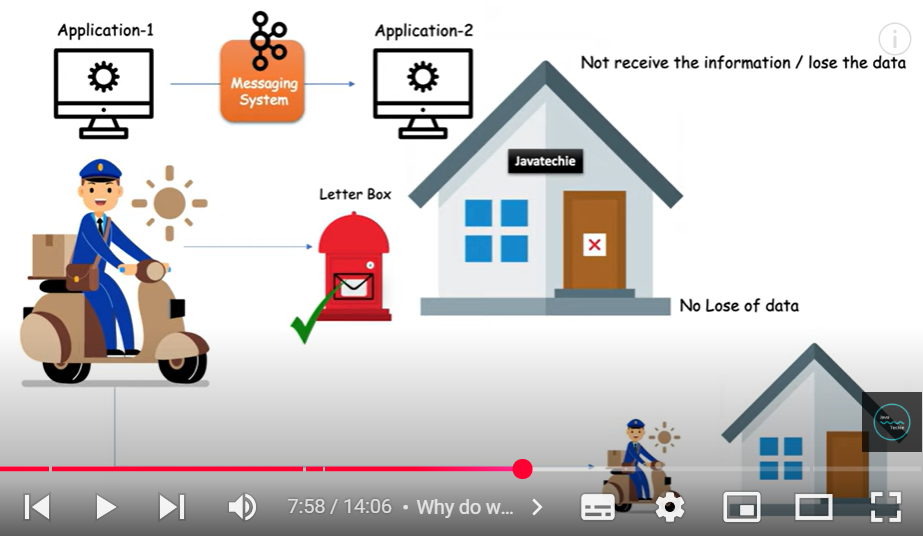
Distributed: Distribute different computers to different nodes to balance the load and minimize the downtime.

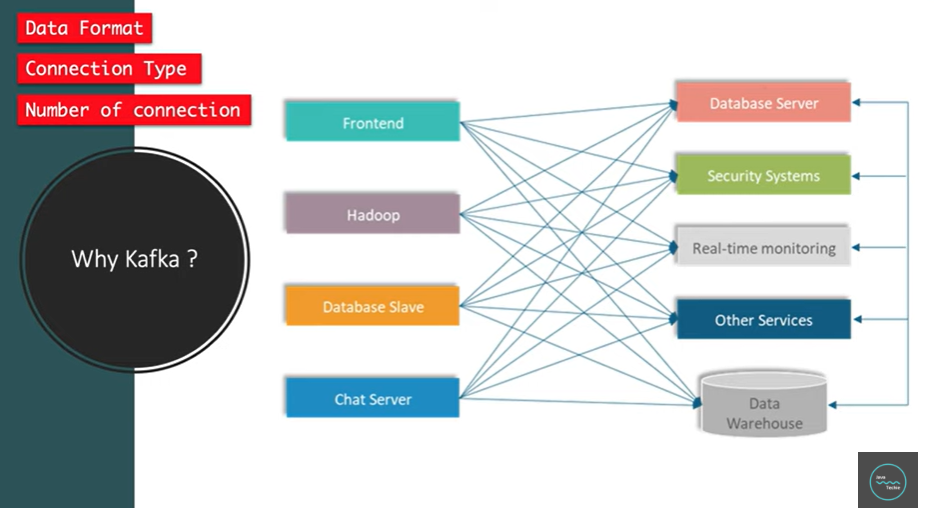




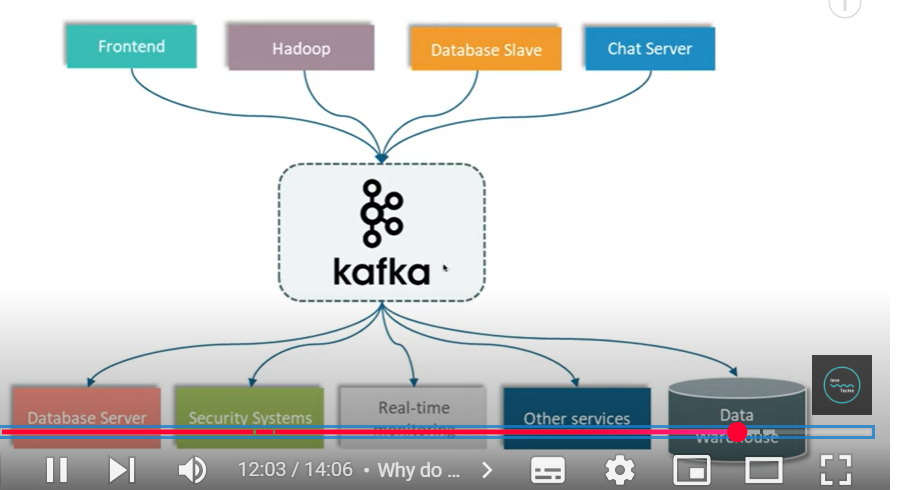
Postbox acts as a middleman between the postman and u. So there is no loss of data even when u r not at home.

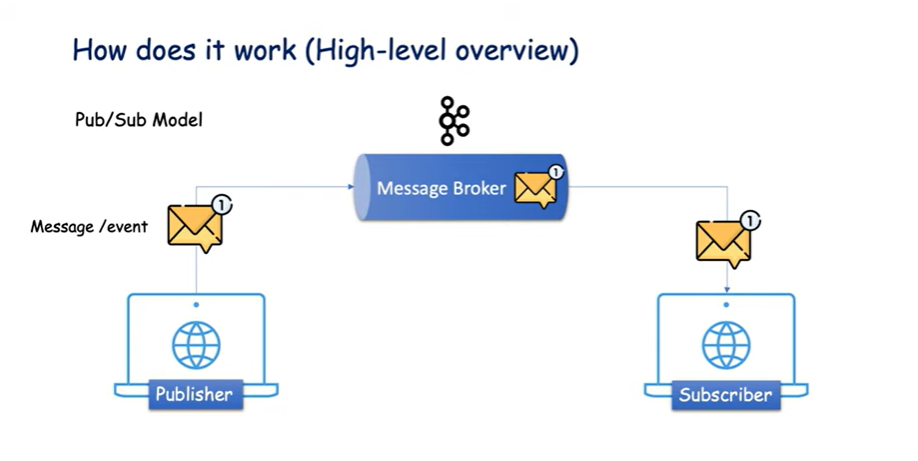
Kafka is like the letter box, it is a messaging system which will keep on collecting the messages like a letter box when the application 2 is down/active.





Number of connections reduced to 9 from 20:





**TERMINOLOGIES**

1. **Kafka Broker**

A blue cylinder with a yellow and orange symbol

AI-generated content may be incorrect.

1. **Kafka Cluster**

Cluster: group of computers/servers working for a common purpose.

For publishing huge volume of data, a single kafka server may not be able to handle the load.

A screenshot of a computer

AI-generated content may be incorrect.

1. **Topics**

To categorize different kinds of messages. Topic specifies the category of messages, listeners can just response to those messages.

A screenshot of a video

AI-generated content may be incorrect.

Consumer can subscribe to a specific topic.

1. **Partition**

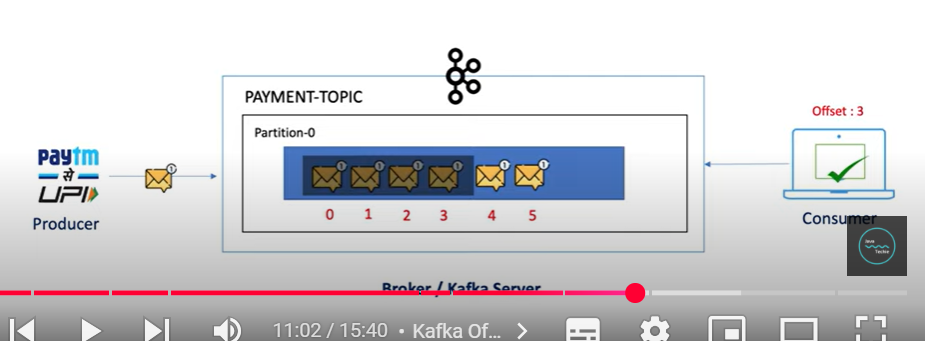
A screenshot of a computer

AI-generated content may be incorrect.

Break the topic into different partitions and each part is distributed into different machines. It gives better performance and high availability.

1. **OFFSET**

The purpose of offset is to find out which messages have already been consumed by the consumer.



If the consumer goes down after reading offset 3, then when it is back again, it knows from which offset to start reading again.

1. Consumer group: instances of consumers.

A diagram of a computer

AI-generated content may be incorrect.

3 consumer instances read through different partitions. Any consumer can talk to any partition. If a consumer gets added It will remain idle.

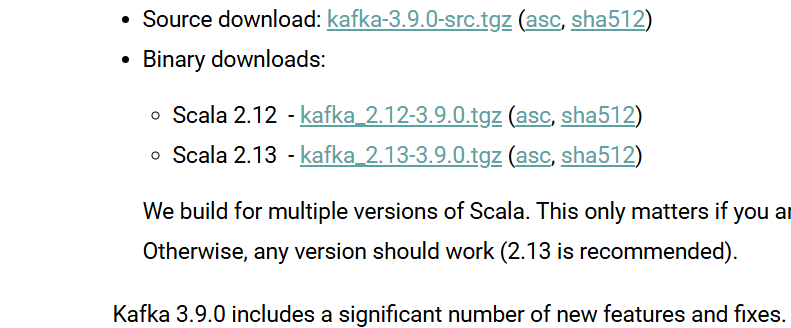
A screenshot of a computer

AI-generated content may be incorrect.

**CLI KAFKA**

<https://kafka.apache.org/downloads>

1. Download the kafka folder: 12 or 13



Extract it.

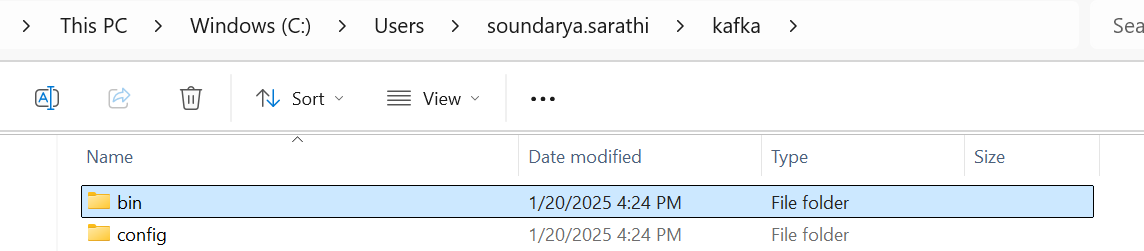
Copy the internal folder to downloads or wherever and rename it to kafka.

Copy and paste, Keep it where the path is short:

A screenshot of a computer

AI-generated content may be incorrect.

There are 2 main folders:



bin & config,

config has all the properties files. For example, zookeeper which keeps track of all the kafka cluster, has to run on port 2151:

A screenshot of a computer screen

AI-generated content may be incorrect.

2, Install Java

1. Install java 8 , after that add them to the path by :

A screenshot of a computer

Description automatically generated

Add both the variables to the path as:

1. To the system variables add:

A screenshot of a computer

Description automatically generated

1. To the user variables add:

JAVA\_HOME

C:\Program Files\Java\jdk1.8.0\_202

A screenshot of a computer program

AI-generated content may be incorrect.

Save the variables and click ok and exit

1. Start Zookeeper as:

navigate to kafka>bin>windows and

A black and white screen with white text

AI-generated content may be incorrect.

.\zookeeper-server-start.bat ..\..\config\zookeeper.properties

Delete and recreate the log directory

rmdir /s /q C:\tmp\kafka-logs

mkdir C:\tmp\kafka-logs

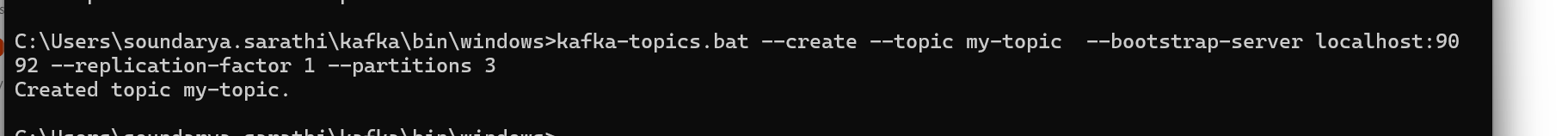
1. Open one more cmd, navigate to kafka>bin>windows and paste the command:

.\kafka-server-start.bat ..\..\config\server.properties

1. Create a topic on another cmd:

navigate to kafka>bin>windows

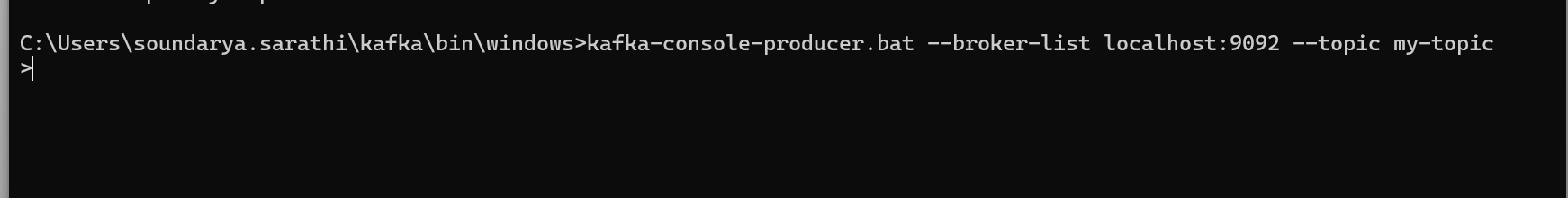
kafka-topics.bat --create --topic my-topic --bootstrap-server localhost:9092 --replication-factor 1 --partitions 3



Bootstrap server is nothing but my kafka server

1. Start a producer

kafka-console-producer.bat --broker-list localhost:9092 --topic my-topic



1. Start a consumer

kafka-console-consumer.bat --bootstrap-server localhost:9092 --topic my-topic --from-beginning

A black screen with white text

AI-generated content may be incorrect.

1. List Topics:  
   kafka-topics.sh --list --bootstrap-server localhost:9092
2. Describe Topic:  
   kafka-topics.sh --describe --topic <topic\_name> --bootstrap-server localhost:9092
3. Produce Messages:  
   kafka-console-producer.sh --topic <topic\_name> --bootstrap-server localhost:9092
4. Consume Messages:  
   kafka-console-consumer.sh --topic <topic\_name> --bootstrap-server localhost:9092 [--from-beginning]

For windows use bat instead of sh

Consume Messages with Consumer Group:  
kafka-console-consumer.sh --topic <topic\_name> --bootstrap-server localhost:9092 --group <group\_id>  
List Consumer Groups:  
kafka-consumer-groups.sh --list --bootstrap-server localhost:9092  
Describe Consumer Group:  
kafka-consumer-groups.sh --describe --group cerebra-data-sink-20230903 --bootstrap-server kafka-electric:9092  
Alter Topic Configuration:  
kafka-configs.sh --alter --entity-type topics --entity-name <topic\_name> --add-config <config\_key>=<config\_value> --bootstrap-server localhost:9092  
View Consumer Offsets:  
kafka-run-class.sh kafka.tools.ConsumerOffsetChecker --topic <topic\_name> --group <group\_id> --zkconnect localhost:2181  
Delete a Topic:  
kafka-topics.sh --delete --topic <topic\_name> --bootstrap-server localhost:9092  
Check Broker Health:  
kafka-broker-api-versions.sh --bootstrap-server localhost:9092  
Produce Avro Messages:  
kafka-avro-console-producer.sh --topic <topic\_name> --bootstrap-server localhost:9092 --property value.schema='<avro\_schema>'  
Consume Avro Messages:  
kafka-avro-console-consumer.sh --topic <topic\_name> --bootstrap-server localhost:9092 --property value.schema='<avro\_schema>'

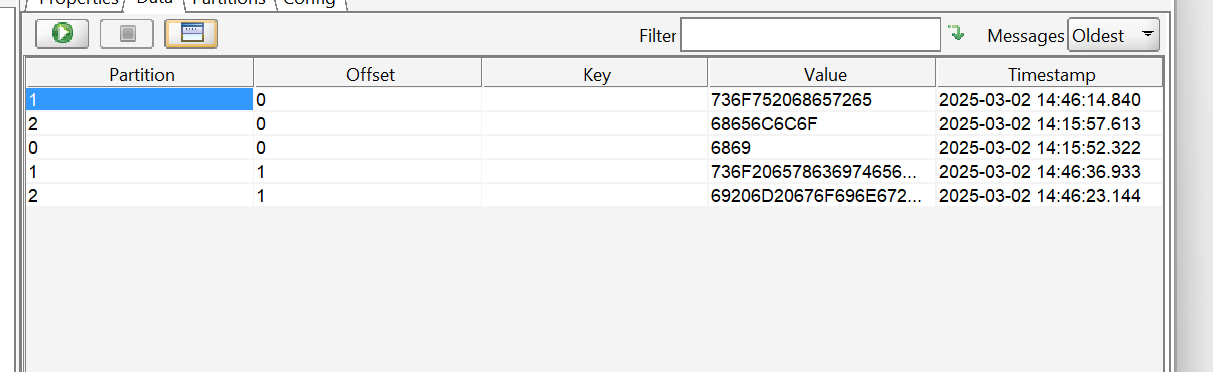
1. OFFSET EXPLORER

Install the offset explorer:

A screenshot of a computer

AI-generated content may be incorrect.

All the messages typed will be visible here along with the partition, the zookeeper decides the partition



1. SENDING BULK CSV FILES

type C:\Users\soundarya.sarathi\Downloads\check\_webid\_data.csv | kafka-console-producer.bat --broker-list localhost:9092 --topic my-topic

A screenshot of a computer screen

AI-generated content may be incorrect.

KAFKA USING SPRINGBOOT