**Kafka**

* **How to start kafka locally?**

We should have downloaded Kafka and Zoopkeeper locally. Then we can go in kafka/bin/windows folder(where all .bat files available) and use below command to check Kafka is installed correctly.

***Kafka-topics.bat***

If we want that all .bat files work from any folder of any path in Window, then we can add this folder path to our environment variable, later from any path if we run a kafka command, it will check in all paths given as “all paths” in environment variable.

We create a new folder under kafka folder(which have executable files) named as data. Inside data we will create 2 more folders as zookeeper and kafka.

These folders will hold kafka and zookeeper data.

Copy the zookeeper path and edit zookeeper.properties file under config folder and put this path in dataDir.(need to change to backward slash to forward slash).

Now we are good to start zookeeper server so run below command under kafka folder.

***zookeeper-server-start.bat config/zookeeper.properties***

It will start zookeeper at **port 0.0.0.0/0.0.0.0:2181**

Now we have to edit server.properties to start kafka server. (different command prompt)

Change logs.dirs with kafka folder that we have created under data folder.

We can run kafka command from main Kafka folder with below command.

***kafka-server-start.bat config/server.properties***

or go to inside config folder and run

***kafka-server-start.bat server.properties***

We will get below message once kafka has started:

*[KafkaServer id=0] started (kafka.server.KafkaServer)*

* **Create Topic:**

Now as zookeeper and kafka is running perfectly, we are good to create a topic as:

***kafka-topics.bat --zookeeper 0.0.0.0:2181 --topic topic\_name --create --partitions 3 --replication-factor 1***

* **Console Producer:**

It is used to read data from standard input and publish it to Kafka.

Below are the mandatory field for this command:

**--broker-list <String: broker-list>**  REQUIRED: The broker list string in

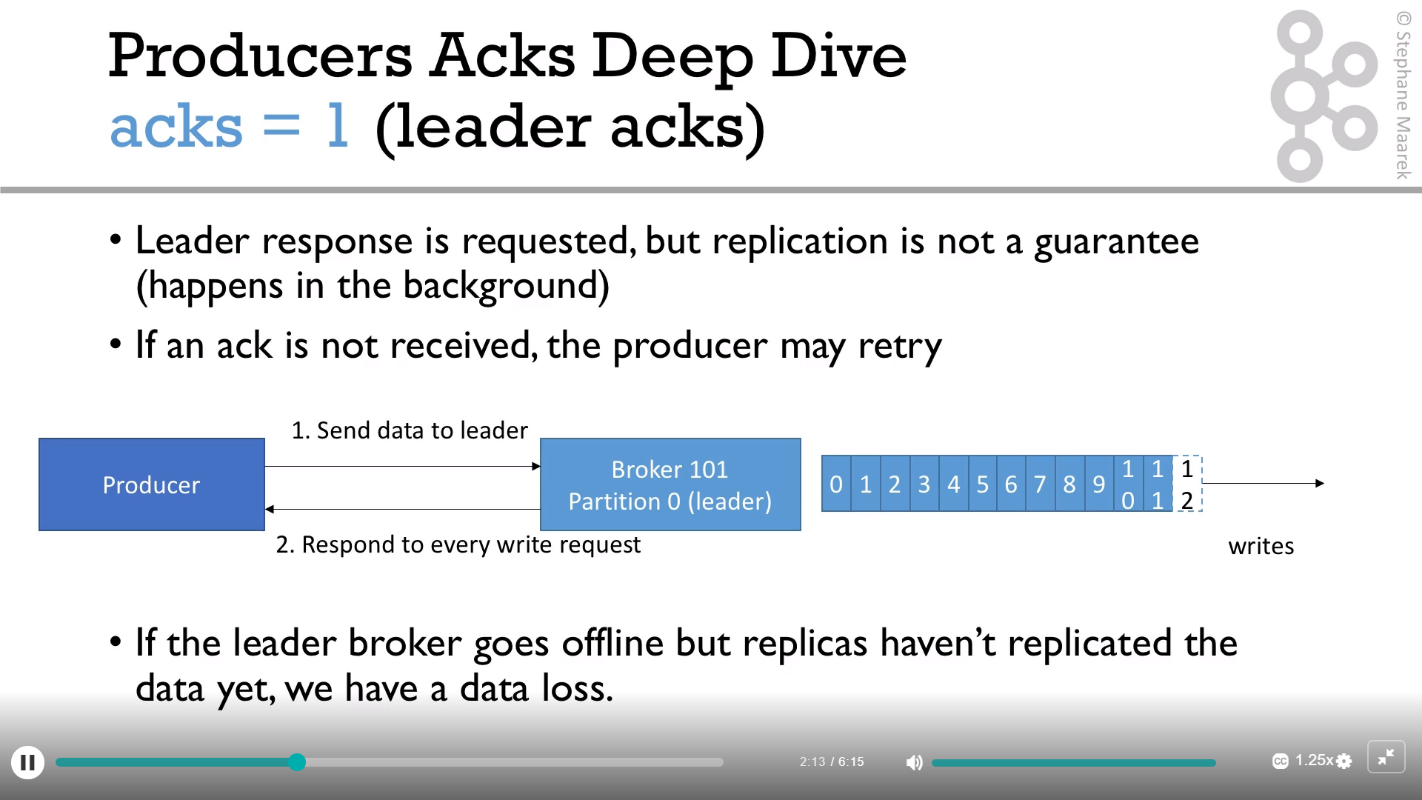
the form HOST1:PORT1, HOST2:PORT2.

**--topic <String: topic>** REQUIRED: The topic id to produce messages to.

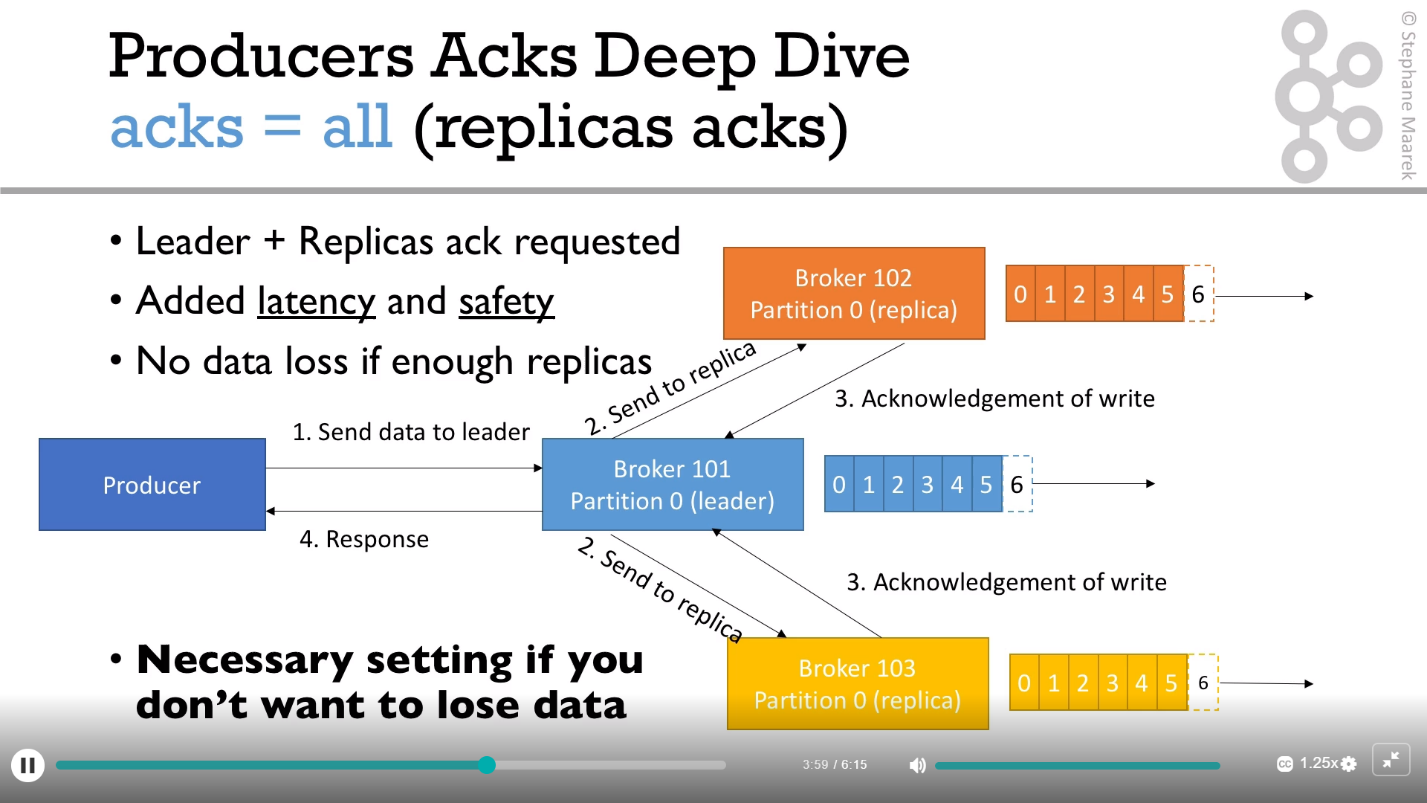
We can use producer’s property acknowledgement value to change it’s default value (acks=1) so the complete command line would be similar to: Acks=0



Acks=1



Acks=all



***kafka-console-producer.bat --broker-list 127.0.0.1:9092 --topic first\_topic --producer-property acks=all***

If there is no topic is available and we try to write to that topic from producer, then producer will create a topic with given name with below configuration:

Partition :1, RF: 1, Leader:0

We can change this default property of partition in server.propertgies file

* **Console consumer:**

This tool helps to read data from Kafka topics and outputs it to standard output.

It reads data from producer on real time on a topic. We can get the data from beginning as well using additional command.

Order of message might be different because order is guaranteed over partition not over topic.

***kafka-console-consumer.bat --bootstrap-server localhost:9092 --topic first\_topic --from-beginning***

we can have consumer group where all consumer of that group will get data from producer for a topic randomly so the command would be:

***kafka-console-consumer.bat --bootstrap-server 127.0.0.1:9092 --topic first\_topic –group group\_name***

* **Consumer group**:

This tool helps to list all consumer groups, describe a consumer group, delete consumer group info, or reset consumer group offsets.

--list: to list out all the consumer groups

--describe: to get description about a consumer group

***kafka-consumer-groups.bat --bootstrap-server localhost:9092 --list***

***kafka-consumer-groups.bat --bootstrap-server localhost:9092 --describe --group first\_consumer\_group***

We can reset cg offset to change its behavior to read any message once it’s generated from producer.

**--reset-offsets** Reset offsets of consumer group. Supports one consumer group at the time, and instances should be inactive Has 2 execution options: --dry-run (the default) to plan which offsets to reset, and --execute to update the offsets.

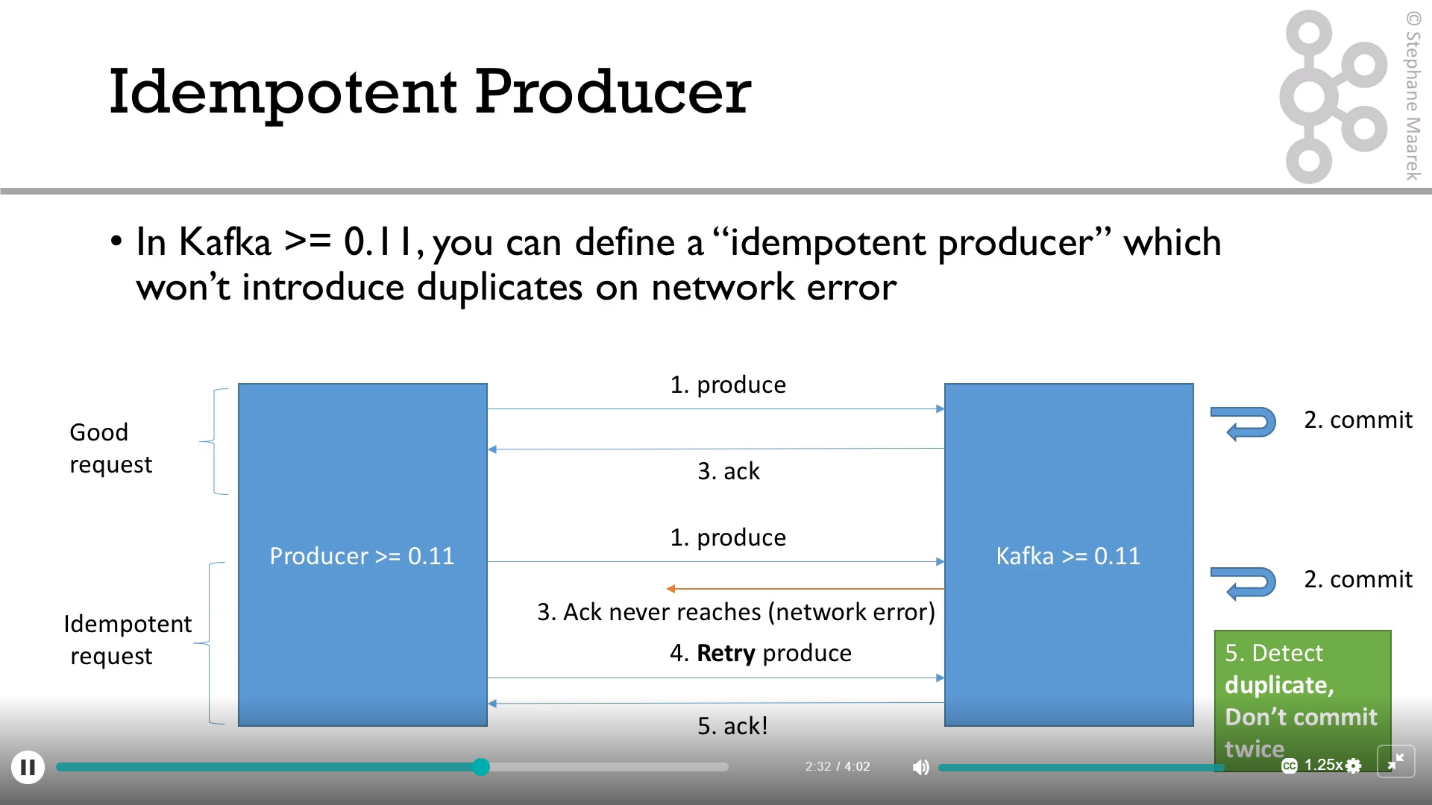
Additionally, the --export option is used to export the results to a CSV format. You must choose one of the following specifications: --to-datetime, --by-period, --to-earliest, --to-latest, --shift-by, --from-file, --to-current.

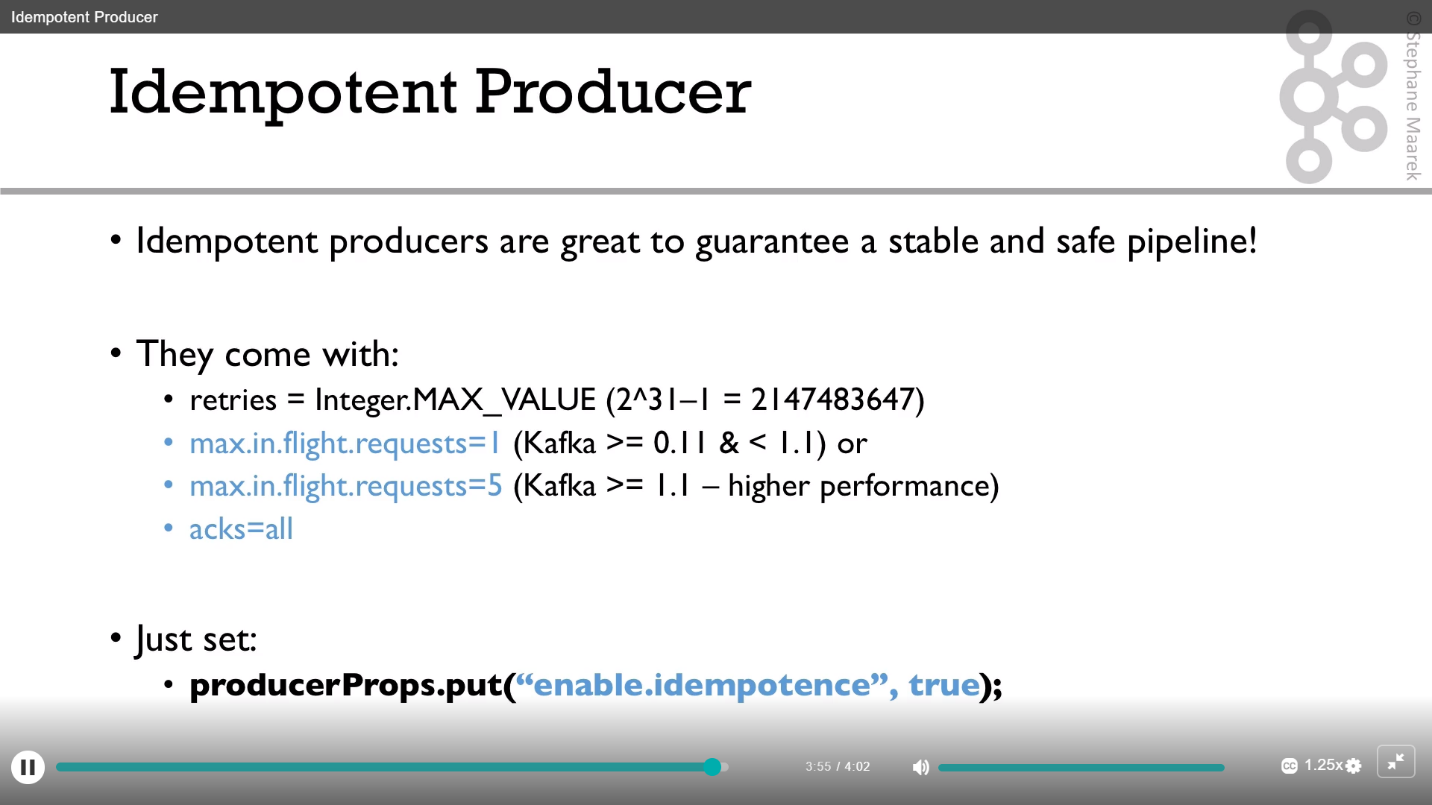
To define the scope use --all-topics or --topic. One scope must be specified unless you use '--from-file'.

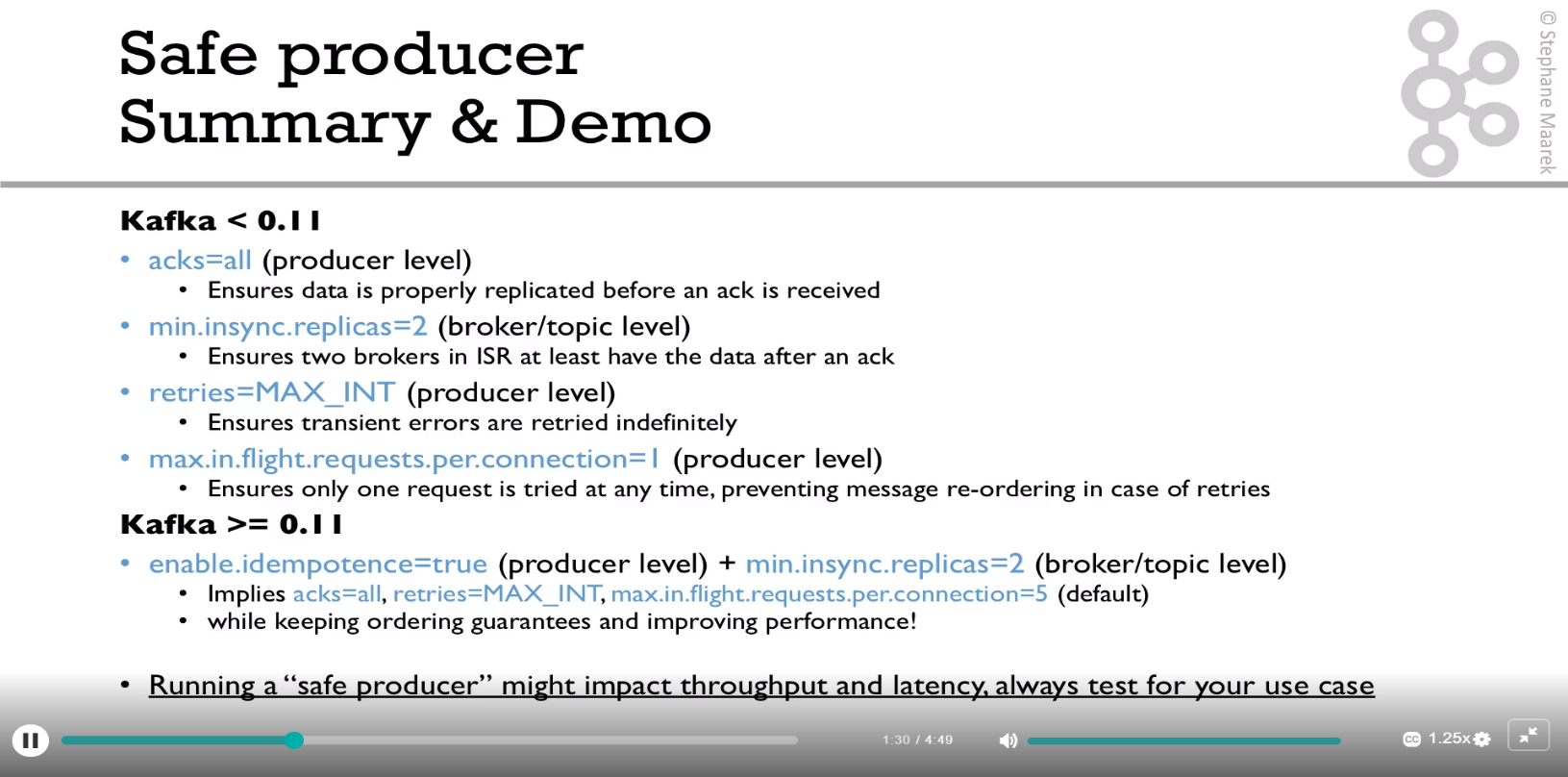
* **Idempotent Producer:**

Sometimes we have a situation when Producer produce the message for Kafka and it went to Kafak but at the time of acknowledgement there is a network issue so Producer didn’t receive ack.

Now Producer will retry to send the same message it will go again and commit in kafka as duplicate message. To stop this duplicate message commit, we can use Idempotent Producer where and id will be generated towards all producer message so for duplicate, Kafka will recognize the message as duplicate and won’t commit it. (will only send acknowledgement)







* **Message Compression:**

