Imp points

1) While sending a record give partition number- so that message will directly go to that partition if u don’t give that provide at least key so that partition will be calculated based on that key, if u don’t give key , then in round robin fashion it will send messages to all partitions

2) Concept of producer buffer - if producer is sending 1 million records in 2-3 seconds , it won’t send all 1 m records to broker 1 by 1 by hitting 1 million times– we have concept of producer buffer all the records will be in some buffer if buffer reaches max size then it will send in batch, but in real time if we are sending 1 individual message then how come it was sent so fast?? This I am also not sure.

3) Every message in kafka is automatically timestamped , producer api will set the time to the message which is going to send

4) add logic to reject duplicate messages incase of retries

5) Configure enable.idempotence= true for exactly once scenario to avoid duplicates, producer must enable this to avoid duplicates from producer retries, incase even if message already received by broker and if the acknowledgement was lost and if it wasn’t received by producer he may retry which is a duplicate scenario, so to avoid such scenarios keep the flag =true

6) Mandatorily Implement transactions in kafka-- send the messages using transactions – so that in a transaction either if all messages will be sent or no message will be sent, Caution: **- setting a unique transaction id is mandatory for each producer instance,**

And make sure consumer reads only committed data , set isolation levels to read\_comitted

**7) While consuming (ex:- acc transfer request )- make sure consumer should commit the offset frequently**

If he don’t do that- if that consumer went down and came back as he didn’t commit the offset he will read the same message again and debit the money from one of the account

So its danger consumer should always commit the offset

8**)Commit the offsets after consuming** configure commit.interval.ms=100/200ms so that consumer will commit the offsets what ever he has read for each and every 100-200ms

9) After producing each message , register the callback , so that the callback will be fired for each message, once the message successfully received,broker will send the acknowledgement back, producer will get back an acknowledgement when producer got it then callback will be fired

Always think about this worst cases

1. What if the broker went down without committing the offsets

Doubts

1. If producer send 1 million messages , then will we receive 1 million acknowledgements for all messages, if yes but how do we know which acknowledgement is for which message among those million messages

Producer

Producer always sends the data to the leader partition, similarly consumer always consumes from the leader partition

When you are creating a producer its mandatory to tell the application id, then only the broker will come to know from which ap these messages are coming from, so set this ID while configuring the producer

Sample producer code

Purpose of producer id is to track the source of message

Properties props = new Properties();

props.put(ProducerConfig.CLIENT\_ID\_CONFIG, "HelloProducer");// this tell which producer is this

props.put(ProducerConfig.BOOTSTRAP\_SERVERS\_CONFIG, localhost:9092,localhost:9093);

props.put(ProducerConfig.KEY\_SERIALIZER\_CLASS\_CONFIG, IntegerSerializer.class.getName());

props.put(ProducerConfig.VALUE\_SERIALIZER\_CLASS\_CONFIG, StringSerializer.class.getName());

props.put(ProducerConfig.TRANSACTIONAL\_ID\_CONFIG,”diff for each producer instance”);

props.put(ProducerConfig.TRANSACTIONAL\_TIMEOUT\_CONFIG,12000);

KafkaProducer<Integer, String> producer = new KafkaProducer<>(props);

Callback callback=(recMetadata,exception)->{  
 if (exception == null) {  
 int callbackExecutedCount= counter.getAndIncrement();  
 System.*out*.printf("\n %d. msg went to topic --> %s partition --> %d , offset --> %s ",  
 callbackExecutedCount,recMetadata.topic(),recMetadata.partition(),recMetadata.offset());  
 }  
};

logger.info("Start sending messages...");

for (int i = 1; i <= AppConfigs.numEvents; i++) {

producer.send(new ProducerRecord<>(AppConfigs.topicName, i, "Simple Message-" + i), callback);

}

|  |
| --- |
| Producer Record |
| **Topic – Mandatory field because if u don’t know topic name where u will send the message** |
| **Message value - Mandatory field** |
| **Message key –optional (Based on the key partition will be decided)** |
| **Partition – (if u give partition , msg straight away it will goto that partition only)** |
| **Timestamp** |

Partitioner

As per above message for producer record ,

1. if we set partition number message will directly go to that partition,
2. else partition /folder will be decided based on the hash value of the key

new ProducerRecord<>("wellsfargoo",”skey”,"vammov"+e)

here skey is a string , hash code will be calculated on the key and that will decide the partition number

1. If key is also not there ah?? Then default partitioner will send all messages to all partitions(folders) in round robin fashion

properties.setProperty(“partitioner.class”,roundRobinPartioner.class);

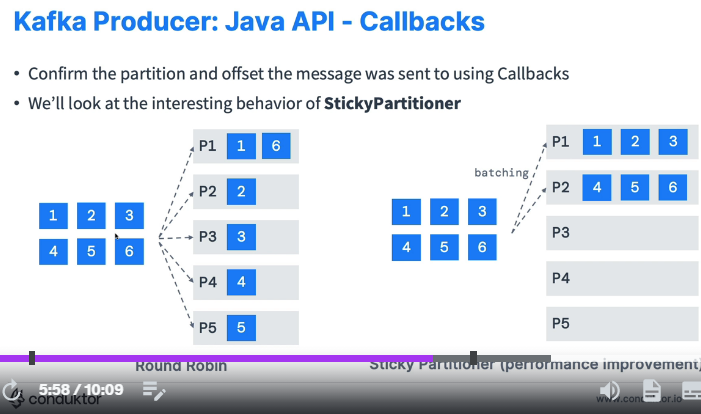
You can also Create a custom partitioner and provide that name to that

Props.put(ProducerConfig.partitioner\_Class\_CONFIG,MyPartitioner.class.getName());

There are 2 default partitioner available

1. **Hash key based partitioning :-** same like hash map internal working , here also for every key hash code will be calculated and based on that partition number will be decided

**Round robin partitioner:-** if we didn’t send the key in the producer record then partition will be round robin fashion , means messages will be sent to all partitions(folders) this is worst performance as here there is no batching extra overhead of each record individually **Sticky partitioner**



In round robin partitioner he will send messages to all partitions equally

Whereas sticky partitioner will send some 10-15 messages to partition-1 and another 10-15 messages to partition-2 like that why? Because this is a batching and it’s a performance improvement it will send all the batch records at once

Kafka default batching size id 16kb ,but if it true why we are getting messages very faslty eventhough 16kb requirement is no met

Serializer

We have 2 types of serializer

1. JSON serializer – heavy load on network, because it embeds lot of schema information while carrying the message
2. Avro serializer – fast and used mostly in big data environment – Fast because AVRO doesn’t embed schema information

Message timestamp

For every message n Kafka time stamp will be there among below 2

Auto time stamping:- As you see, producer record have a field called TimeStamp

1. **Setting Producer time --Create Time** TimeStamp set by producer – its create time like when producer is going to send a message then producer will set this time to the message

Message.timestamp.type=0 this is the default and recommended one

1. **Setting broker time --Log Append Time** Time stamp set by broker on arrival to the broker –In this case time stamp will be set by the broker when he received it

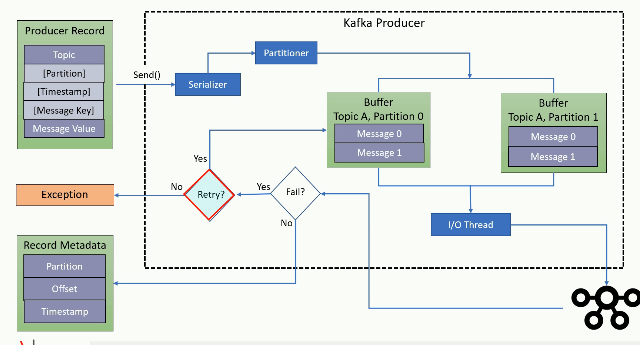
Message.timestamp.type=1

Producer buffer

kafkaTemplate.send(topicName,message) // this will actually send the message to the buffer partition which resides in the producer memory, it will not immediately send the message to the broker, IO threads are real senders when the buffer max capacity is filled then IO threads will send the data from buffer to real kafka broker cluster.

The default producer memory is 32 MB , if u want you can expand the producer buffer memory

**Buffer.memory for this key** we have to set the value



In case of success we will get the record metadata callback , once we got ack we can find that message was written to which folder(partition) , which file (segment)

Producer memory must be bit high, especially if incoming records are coming in larger volume buffer should be able to hold right. So make sure u have large buffer memory,

Blocking scenario :- Generally producer will send data to buffer and leaves as non blocking , if producer buffer is low and if huge volume of data is coming , producer needs to wait until , until the buffer is cleared means until io threads take data from buffer producer will get blocked, and once buffer is free , producer will push to buffer and leaves

If producer is sending 1 million records at a time producer wont really send them all at once

It will be stored in a producer buffer, once buffer max capacity is filled then only it will send and interesting thing is for every partition we have separate buffer inside the producer itself

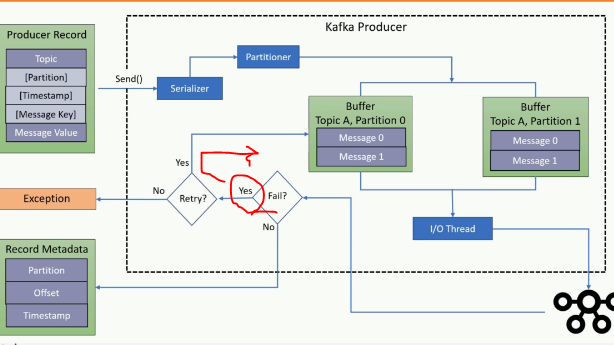
Doubt:- let’s say if buffer max capacity is 1MB if we send 1-2 messages definitely it wont fill , in that case to full the capacity it take hours, but how are we getting instantly and why are we not waiting till buffer is filled, ex:- in our 1DSTR app if u send 1 message having bytes it is immediately coming na

1. Instead of buffer size concept, we should have had buffer count concept, so that when 200 messages came to buffer we will send all of them at once

Producer retries

Generally if we send 1 million messages we should get acknowledgements for each message we received, lets say if we have configured acks=all means leader will send positive acknowledgements only when all followers / all ISR’s received it, lets say leader received the message but followers didn’t received it incase failed acknowledgement will be sent back, and as usual producer will retry to send the same message again now I don’t know what happens then will broker accepts it?

Ok for now understand that it will retry when it gets a negative acknowledgement



Summary

1. Producer.send(key,value);- Seriailization of key and value , decide the partition number

First the key and value will be serialized by the respective serializer

If partition is there it will put that serialized message into that partition buffer ,else it will decide the partition based on the key , if key is also absent then in round robin fashion message will be sent to one of the partition

1. Serialized message will sit into the producer buffer
2. IO thread will send the message to the cluster
3. If the producer received the acknowledgement then no retry , if the producer didn’t received the acknowledgement then retry
4. It will retry for configured number of times, still not working ha then then throw an exception

**Scaling producer**

A single kafka broker is enough to receive 100-1000 messages per second

It says no need to create multiple **KAFKA Producer** objects in single application

Just create multiple threads so that all threads can use same kafka producer object

Refer sample project **“03-multi-threading-completed”**

If you have a trading application where u should send the data at high speed , like 10,000 messages per second then u may need to use multi-threaded producer

Ex:- if u have 2 files which Is having bulk data in each file

Then create 2 threads ,whereas each thread will read parallelly from each file and each thread after reading one line it will send that line to producer buffer internally once that entire producer buffer is reached then all will be sent at once, each producer will have 1 buffer , so as per our schedule lets have 2-3 threads all can happily share the same kafka producer, each thread can use same producer object

Duplicate message –atleast once vs atmost once

Below are the scenarios of coming duplicate messages

1. Generally when we send a message by the producer , broker after receiving the message he should send acknowledge for the received message , incase broker while sending back acknowledgement to the producer if some issue came, then that ack wont reach to the producer in that case producer will retry sending the message which is already received
2. In case of replications – generally replicas are considered as follower brokers , generally these followers will be little behind the leader , like if leader consumed or read 100 messages followers would have read 70-80 messages only , lets say if acks=all means when a message received by the broker 1) leader broker will consume happily 2) when these followers are out of sync and since they didn’t consume properly since ack=all ( means send back ack only when all followers consumed the message) if one of broker has missed to consume it and in this case timeout as we didn’t send the ack back to the producer and producer will think msg not received and he will resend he wont consider that message was already consumed by the leader broker or not , from his side since ack didn’t came he will resend in this case this is duplicate message will be sent to broker even though leader has already consumed the same msg he will get it.

To avoid this w have 1 worst solution that is configure zero retries= atmost Once

So this is bad because even if something went wrong it wont retry leads to data loss

**Best solution is u implement the duplicate record identification** mechanism based on the co relation id

Kafka is by default atleast once- means retry always

Solution for duplicate retry-idempotent producer

Enable.idempotence=true

This solution is only for retry mechanism, when message successfully received by the broker and if broker while sending the acknowledgement if that ack is lost , producer will retry sending same message na eventhough the message successfully received ,producer will send same message again this is the duplicate scenario

How does broker solves this problem??and how it works internally

If you configure this parameter ,Broker while sending the message for each and every message it will append the producer instance id and sequence number

So broker when that message successfully received the message it will store that id+sequence number and if producer sends the same message again with same producer instance id +offset id the broker will reject

This solution won’t work if application itself is wantedly sending the duplicate message

This is the solution only for retry mechanism, if u configure this while retrying ,even if producer send duplicates, broker will reject the duplicates due to producerID+sequence number combo would have already received earlier and sequence also wouldn’t have matched with current sequence number

Enable.idempotence=true

Transactions in Kafka

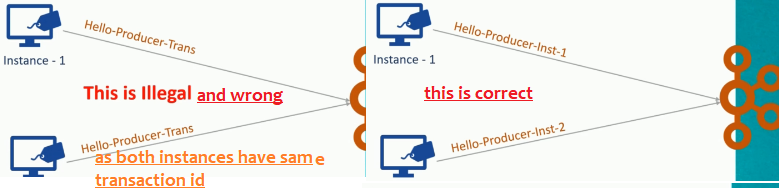
1. Transaction **means – either all or nothing** - It says either all messages within the transaction are committed or nothing is committed
2. Transactions should be committed if all parts are succeeded ,
3. only committed data can be consumed by the consumer, if u have sent to broker and if u didn’t commit the data that data may present in the broker , but the consumer cannot consume it for this u might need to set the isolation level to read\_comitted
4. Transactional id which is set to the producer instance must be unique even in cluster environment- each producer instance in a cluster must have a different transaction id to uniquely identify that producer why?? Because when you are writing the messages with transaction the messages will be written to broker and u will get acks also ,but only thing is those messages will be written with abort flag and each message will have producer instance id, so that if producer aborts the transaction they will be erased , if u don’t give if all producers instances are writing with same id and if all are using transactions whose messages to revert?? So transaction id is mandatory

!!!! Whenever u are using the transactions in Kafka make sure u should have different transaction id for each producer, especially if u scale the application in our prod if we

Have 4 instances of our application then we will have 4 producers then each producer should have different transaction number

Internally it is similar to database commit, like first all instructions will be executed in database once u commit changes are saved, here also in a transaction if we are sending

100 of messages first all messages will go to the kafka and after committing it will persist



You cant run 2 instances of producer with same transaction id-try to know what will happen if we keep same

props.put(ProducerConfig.TRANSACTIONAL\_ID\_CONFIG,” Hello-Producer-Trans”);

if u set as above all producer instances when this appn running on a cluster all instances may take same transaction id

What is the use of transaction id??

Every producer instance should have different transaction id , if same producer was offline and came back after some time and if there are any unfinished transactions they will be aborted and all the messages will be written in abort mode

Assignment:- what will happen when 2 producer instances have transaction id

On say producer.initTransactions(); , producer.beginTransaction()

Whether u start the transaction or not, by default all the message will be written to broker with abort flag, once u commit the flag will be changed to **“c”** for each message and only messages which are marked as committed can be read by the consumer

When u say kproducer.rollback(); as every producer is identified by its transaction id, as messages are already written to the broker, broker finds all the messages written with that transaction id and all those messages will be rollbackedby the broker ,

Now think if 2 producers have same transaction id which is the uniquely identifiable , when 2 producers wrote messages to the broker, if 1 producer issued rollback , as messages are written with same transaction id which producer messages should be rollbacked?

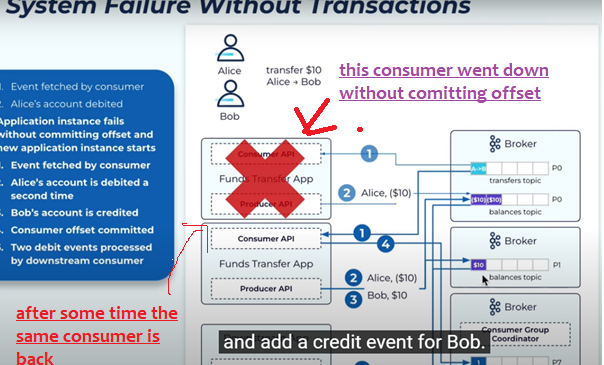
Hence each producer instance should have different transaction id

Ans:- application wont even start if another producer also have same transaction id So in prod also when second instance cant start- because when u say

what is producer.beginTransaction()?

These transactions are initiated by the producers, if the same producer instance has initiated the transaction and if he didn’t close that transaction, this producer.beginTransaction() statement will close or it may abort un finished transaction if exists

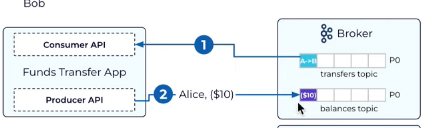
Producer.commitTRansaction(); producer.abortTransaction();

s

In the above image see , as there is no transaction enabled debit request message has been kept 2 times into the queue.

#### Without transaction what is the issue

Let us say a money transfer event, there are no transactions,



1. Step -1 in above message is a message transfer event which the consumer must read
2. consumer after consuming the message it should produce 2 events –debit event, credit event

First consumer will read a message called message transfer event from A --> B and consumer didn’t commit the offset after reading and consumer kept a money debit message from Alice into another topic, and suddenly it went down

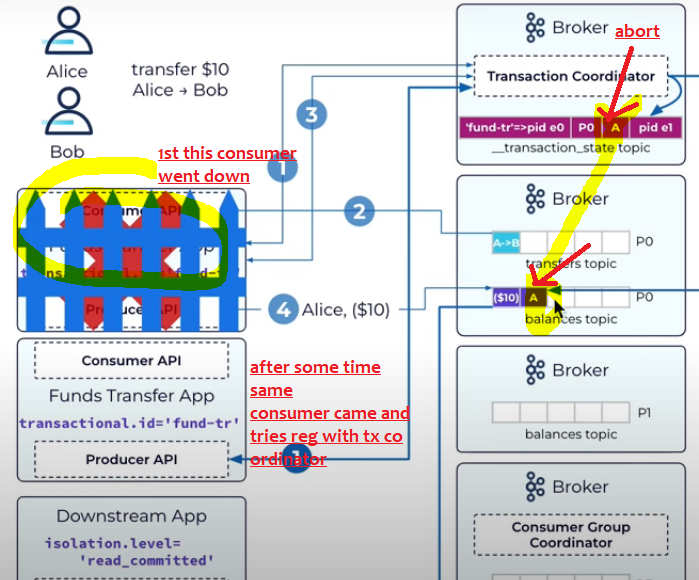
And after sometime it is back and since last time that consumer didn’t commit the offset it will read the same transfer message again **from Alice 🡪 Bob and** puts **Alice money debit request message** into the topic second time which is wrong () and keeps money credit message to Bob in another topic

If first time itself if transaction are there, since credit message to BOB has failed, debit tx wouldn’t have committed issue would have solved, hence transactions are very important, since there is no transaction that’s why this duplicate message came

So when transactions are there even when application fails it will guarantee only single time message will be delivered

#### Exceptions during a transaction

1. first every producer instance should have a transaction id even in a cluster environment also with that this producer it will register to a transaction co-ordinator and get the unique producer instance id



Now see as per above image 1 st consumer read the message and stored in topic and didn’t commit offsets and went offline and after same consumer came and registers with transaction co ordinator and see this transaction id is already registers and *after seeing all that transaction co ordinator tries to abort all pending transactions and those messages which are* ***already written to the broker will be marked as aborted*.**

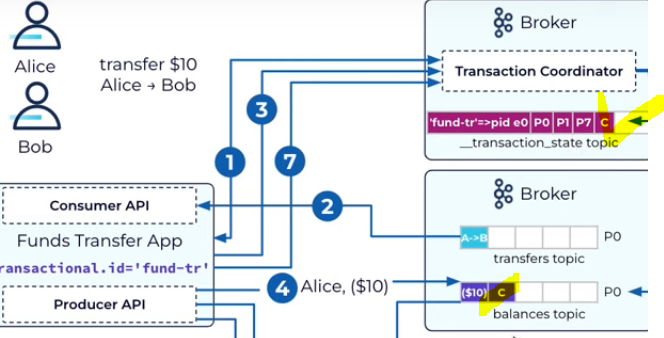
Q) U may get a doubt if these messages are already written what if they are consumed in between ,

A) Simple these messages are not in committed state, consumer reads those messages which are in committed state

#### 2- Successful or committed transaction

Success or failure , during a transaction in both cases messages will be written to kafka broker

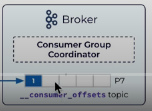
If success and when you **commit the transaction**, then commit marker will be marked on that message see below message, if transaction is aborted then message will be **marked as abort** refer above



Note:- even though both messages are in the broker ,consumer can see only committed messages

Committed offset storage

When the consumer committed the offsets all will be stored in the broker

offsets should be maintained by broker, not consumer, like the team lead should know how far the team member has worked

if consumer stored the offsets , if that consumer went down we don’t know how far the consumer has read, if broker know then he will give the offset number to new consumer to consume from that instance

Callbacks – Acknowledgements

It’s always recommended to attach a callback, for every message we send we will get acknowledgement. When we receive the acknowledgement, our callback method will be fired

In local for debugging purpose , try getting acknowledgements for each message

Just add a callback method

public static void sendAllMessagesToSamePartition(Properties p){  
 KafkaProducer<String,String> producer=new KafkaProducer<String, String>(p);  
 producer.initTransactions();  
 producer.beginTransaction();  
 System.*out*.println("txn began started writing all messages");  
 AtomicInteger counter=new AtomicInteger(0);  
 Callback callback=(recMetadata,exception)->{  
 if (exception == null) {  
 int callbackExecutedCount= counter.getAndIncrement();  
 System.*out*.printf("\n %d. msg went to topic --> %s partition --> %d , offset --> %s ",  
 callbackExecutedCount,recMetadata.topic(),recMetadata.partition(),recMetadata.offset());  
 }  
 };  
 IntStream.*range*(1,6000).forEach(num->{  
 *//Here we are sending all messages to partition number 1, so all msgs will be sent to p1  
 // and key is unique for every message  
 // and as we are already giving partition num,so it will not decide the partition based on the key* ProducerRecord<String,String> record=  
 new ProducerRecord<>("EmployeesInfo","orey naina--"+num);  
 *// this callback method will be executed each and every time for each and every message* producer.send(record,callback);  
 });  
*//even though u didnt committed all messages will be written to broker,  
// once u committed it will change the flag to committed for each and every message* producer.commitTransaction();  
 try {  
 Thread.*sleep*(4000);  
 } catch (InterruptedException e) {  
 throw new RuntimeException(e);  
 }  
 System.*out*.println("sent all 6000 messages and committed too and got acks for "+counter.get());  
 }

As you attached a callback method , if it is a successfully received by the broker producer will get back the acknowledgement

When producer got ack the callback method will be fired automatically

In the above if you see , I have sent 6000 messages continuously to the broker and attached a callback and mostly considering success , 6000 times the callback method will be executed

Sample log:-

sent all 6000 messages and committed too and got acks for 5999

Refer file name

“Send10KMessagesWithAcksCallback.java”

QA’s

1. What is use of producer id-
2. to identify who send this message to the broker   
   b) And while doing a transaction general even though u didn't commit, messages are written to topic in abort mode with producer id , if u didn't commit or if u abort the transaction, these messages will be in abort mode itself and those messages should be deleted right??? So broker will delete all the messages with that producer id + transaction id

if u dont have the producer id , whose messages will be deleted, so its mandatory

and make sure every producer instance should have separate transaction id

Simple theories

1. ProducerRecord<String,String> record=  
    new ProducerRecord<>("EmployeesInfo","orey naina--"+num);

While creating producerRecord if we send the partition number , that record straight away goes to that partition , if u miss it will decide the partition number based on the key’s hash value , if key also not there then value will be sent to all partitions in round robin fashion

Methods in kafka producer

kafkaProducer.flush()

generally if u do kafkaTemplate.send() it will send those records to producer buffer

and IO threads are responsible to take the records from the buffer and send them to broker

flush() if u use flush method , it will immediately send all the records present in buffer

KafkaProducerRecord.send() it will send the data to buffer

KafkaProducer.flush() it will send all the records in buffer to broker immediately

Producer properties

Properties.set(“batch.size”,400); if u set this //I don’t know but try it

I think for that pa

For Every partition, producer will maintain producer local cache and will have individual batch size

Means this is batch size for each partition

properties.setProperty(“partitioner.class”,roundRobinPartioner.class); means all records will be sent to all partitions // this is no recommended in production because all messages will be sent equally to all partitions no batching extra overhead of sending each record individually