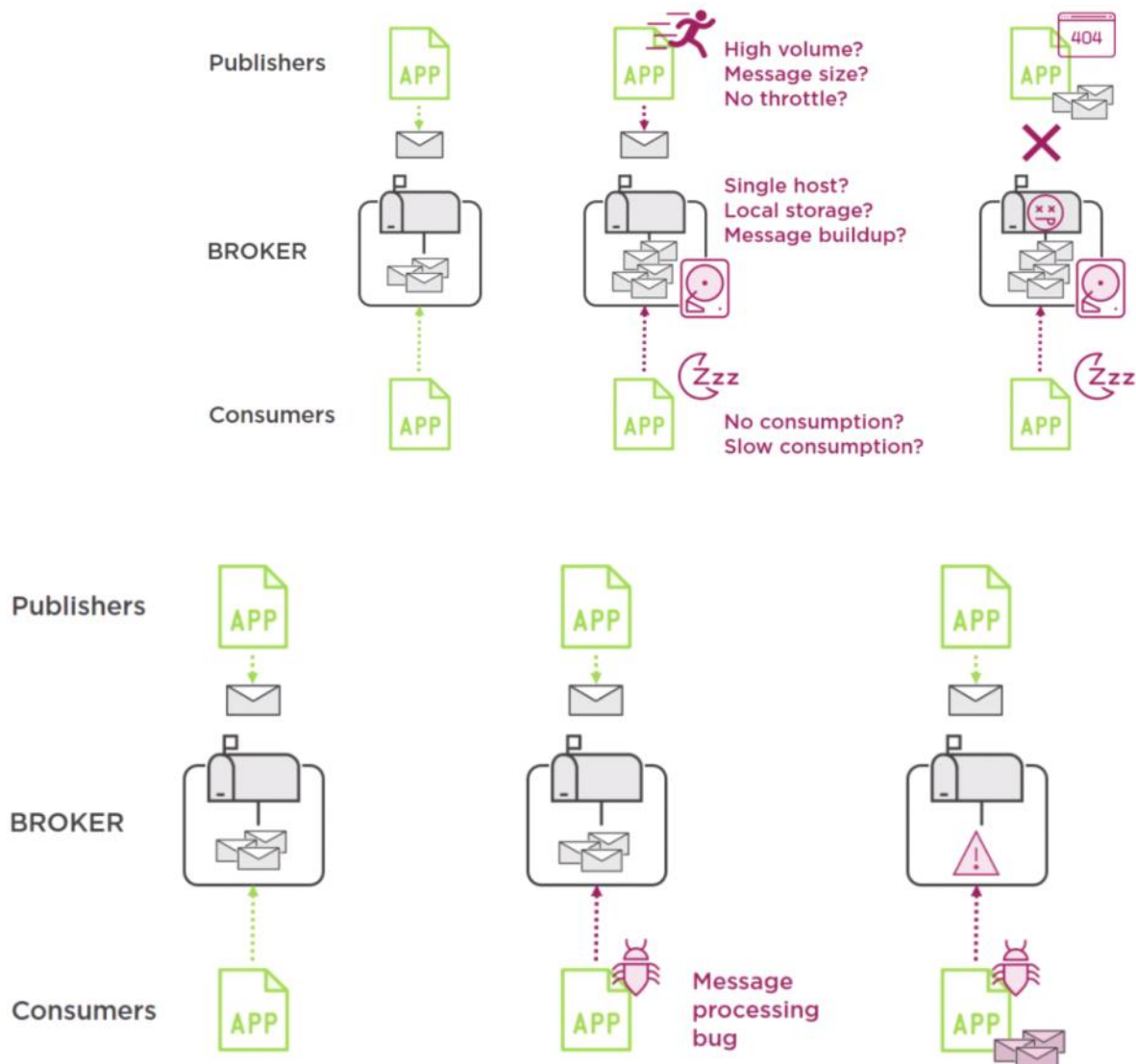


Database replication & Log Shipping - between RDBMS to RDBMS only  
Database - Specific  
Tight Coupling (Schema)  
Performance challenges (log shipping)  
Cumbersome (Subscription)

ETL (Extract, Transform, and Load) - Typically proprietary and costly  
Lot of custom development  
Scalability challenged  
Often times requires multiple instances

Messaging  
Limited scalability  
Smaller messages  
Requires rapid consumption  
Not fault- tolerance (application)



Custom middleware magic - Increasingly complex  
Deceiving  
Consistency concerns  
Potentially expensive

Sunday, December 11, 2016  
7:41 PM

**zookeeper-server-start.bat** C:\Kuiper\Apache\kafka\_2.10-0.10.1.0\config  
**\zookeeper.properties**

# Install

Saturday, December 10, 2016  
7:40 PM

## Installing Apache Kafka 0.8.2.1 on Windows 7

### JDK and JRE environment variable Setup

1. Set the JAVA\_HOME environment variable entry in user variable section

*JAVA\_HOME = C:\Program Files (x86)\Java\jdk1.7.0\_75*

(Modify the path and jdk version according to your usage)

2. Append JRE server's bin directory in PATH system variable section

My Path entry looks like following :

*Path = C:\ProgramData\Oracle\Java\javapath;%SystemRoot%\system32;%SystemRoot%;%SystemRoot%\System32\Wbem;%SYSTEMROOT%\System32\WindowsPowerShell\v1.0\;C:\Program Files (x86)\Java\jdk1.7.0\_75\bin;C:\Program Files (x86)\scala\bin;C:\Program Files (x86)\sbt\bin;C:\Program Files (x86)\Java\jre7\bin;*

### Download and Setup of Kafka

1. Download Apache Kafka from <http://kafka.apache.org/downloads.html> (I downloaded kafka\_2.10-0.8.2.1.tgz (asc, md5))

2. Unpack kafka\_2.10-0.8.2.1.tgz using 7-Zip

(You can download 7-zip from <http://www.7-zip.org/download.html>)

3. Copy extracted kafka\_2.10-0.8.2.1 directory to C:\

After copying, your directory structure would look like as follows :

C:\  
kafka\_2.10-0.8.2.1\  
bin\  
config\  
libs\  
LICENSE (file)  
NOTISE (file)

4. Download Kafka080WinBatFiles.zip file from <https://github.com/HCanber/kafka/releases>. Extract the zip and copy the extracted files to C:\kafka\_2.10-0.8.2.1\bin\windows (i.e. overwrite the existing files)

Note : I have followed this blog (<http://blog.canberger.se/2014/01/kafka-080-on-windows.html>) to get the concern configuration files.

5. Open C:\kafka\_2.10-0.8.2.1\config\server.properties and update existing log.dir entry as mentioned below :

*log.dirs=c:/kafka\_2.10-0.8.2.1/kafka-logs*

6. Open C:\kafka\_2.10-0.8.2.1\config\zookeeper.properties and update existing dataDir entry as mentioned below :

*dataDir=c:/kafka\_2.10-0.8.2.1/zookeeper-data*

(Note : Make sure you use forward slash as depicted for step-5 and step-6)

We are done with all the necessary configuration, Let's start the zookeeper and kafka server as mentioned below.

7. Starting Zookeeper and Kafka server

Open a command prompt and start Zookeeper server using following command :

```
cd c:\kafka_2.10-0.8.2.1
.\bin\windows\zookeeper-server-start.bat .\config\zookeeper.properties
```

Note : If you get an error saying "Unrecognized VM option '+UseCompressedOops'  
Could not create the Java virtual machine". Here is the solution.

Open C:\kafka\_2.10-0.8.2.1\bin\windows\kafka-run-class.bat file.  
Find and remove '-XX:+UseCompressedOops' entry

Before removing '-XX:+UseCompressedOops'

```
IF ["%KAFKA_JVM_PERFORMANCE_OPTS%" EQU [""]] (
set KAFKA_JVM_PERFORMANCE_OPTS=-server -XX:+UseCompressedOops -XX:+UseParNewGC -XX:
+UseConcMarkSweepGC -XX:+CMSClassUnloadingEnabled -XX:+CMSScavengeBeforeRemark -XX:
+DisableExplicitGC -Djava.awt.headless=true
)
```

After removing '-XX:+UseCompressedOops'

```
IF ["%KAFKA_JVM_PERFORMANCE_OPTS%" EQU [""]] (
set KAFKA_JVM_PERFORMANCE_OPTS=-server -XX:+UseParNewGC -XX:+UseConcMarkSweepGC -XX:
+CMSClassUnloadingEnabled -XX:+CMSScavengeBeforeRemark -XX:+DisableExplicitGC -Djava.awt.headless=true
)
```

Now open another command prompt and start Kafka server using following command :  
.\bin\windows\kafka-server-start.bat .\config\server.properties

I hope both Zookeeper and Kafka server are executed successfully for you.

8. Now let's create Topic using following command

```
cd C:\kafka_2.10-0.8.2.1
.\bin\windows\kafka-topics.bat --create --zookeeper localhost:2181 --replication-factor 1 --partition 1 --topic test
```

9. List the topics using following command

```
cd C:\kafka_2.10-0.8.2.1
.\bin\windows\kafka-topics.bat --list --zookeeper localhost:2181
```

10. Send message using Kafka console producer

```
cd C:\kafka_2.10-0.8.2.1
.\bin\windows\kafka-console-producer.bat --broker-list localhost:9092 --topic test
```

Note : If you get an error 'Error: Could not find or load main class kafka.producer.ConsoleProducer'. Here is the solution.

Open C:\kafka\_2.10-0.8.2.1\bin\windows\kafka-console-producer.bat  
Change kafka.producer.ConsoleProducer to kafka.tools.ConsoleProducer

Now you should be able to successfully send the message after fixing above issue. Here is my command prompt output for your reference :

```
C:\kafka_2.10-0.8.2.1>.\bin\windows\kafka-console-producer.bat --broker-list localhost:9092 --topic test
[2015-06-01 14:08:20,196] WARN Property topic is not valid (kafka.utils.VerifiableProperties)
```

1. Hello Kafka from Rashmit
2. Good bye Kafka from Rashmit

11. Consume the above message using Kafka console consumer

```
cd C:\kafka_2.10-0.8.2.1
.\bin\windows\kafka-console-consumer.bat --zookeeper localhost:2181 --topic test --from-beginning
```

Note : If you get an error 'Error: Could not find or load main class kafka.consumer.ConsoleConsumer'. Here is the solution.

Open C:\kafka\_2.10-0.8.2.1\bin\windows\kafka-console-consumer.bat.  
Change kafka.consumer.ConsoleConsumer to kafka.tools.ConsoleConsumer

My command prompt executes the command successfully after fixing above issue :

```
C:\kafka_2.10-0.8.2.1>.\bin\windows\kafka-console-producer.bat --broker-list localhost:9092 --topic test
[2015-06-01 14:08:20,196] WARN Property topic is not valid (kafka.utils.VerifiableProperties)
1. Hello Kafka from Rashmit
2. Good bye Kafka from Rashmit
```

Pasted from <<http://rashmitrathod.blogspot.com/2015/06/installing-apache-kafka-0821-on-windows.html>>

Changes :

1. Kafka-run-class.bat - remove the content in below quotes

"-XX:+UseCompressedOops"

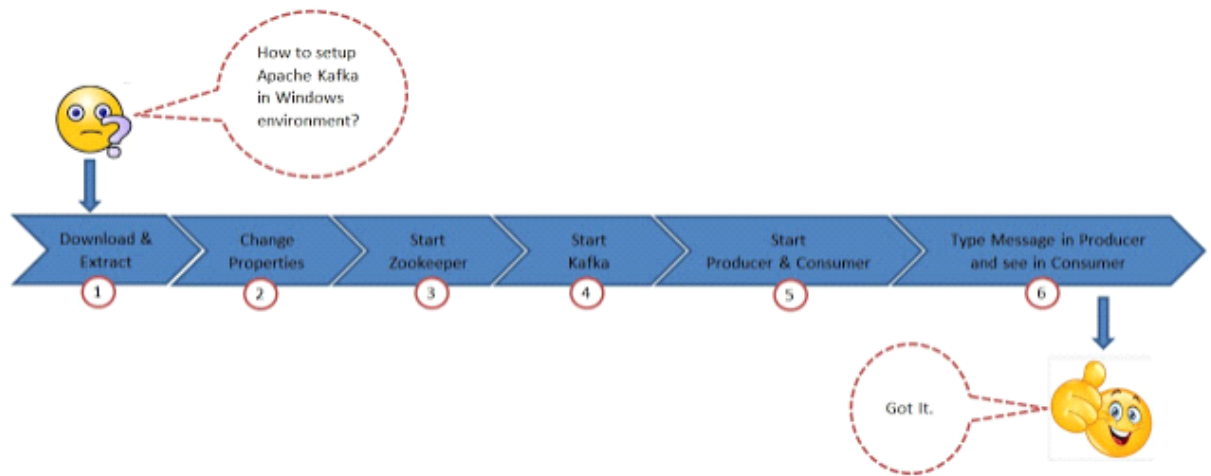
1. If you get an error 'Error: Could not find or load main class kafka.producer.ConsoleProducer'. Here is the solution.

Open C:\kafka\_2.10-0.8.2.1\bin\windows\kafka-console-producer.bat  
Change kafka.producer.ConsoleProducer to kafka.tools.ConsoleProducer

2. If you get an 'Error: Could not find or load main class kafka.admin.CreateTopicCommand'  
Modified : kafka.admin.CreateTopicCommand to kafka.admin.TopicCommand in kafka-create-topic.bat file

**CREATE TOPIC:** *kafka-topics.bat --create --zookeeper localhost:2181 --replication-factor 1 --partition 1 --topic test*

**List :** *kafka-topics.bat --list --zookeeper localhost:2181*



Consumer : `.\bin\windows>kafka-console-consumer.bat --zookeeper localhost:2181 --topic test --from-beginning`

Producer: `.\bin\windows>kafka-console-producer.bat --broker-list localhost:9092 --topic test`

Describe the stat on the topic

`.\bin\windows>kafka-topics.bat --topic test --zookeeper localhost:2181 --describe`

Topic:test PartitionCount:1 ReplicationFactor:1 Configs:  
Topic: test Partition: 0 Leader: 0 Replicas: 0 Isr: 0

Isr -> In Sync Replic

Monday, December 12, 2016  
1:40 PM

#### Message content

- Timestamp
- Referenceable identifier
- Payload (binary)

#### Message Offset

A placeholder

Last read message position

Maintained by the kafka consumer

Corresponds to the message identifier



# Multi-broker

Monday, December 12, 2016  
2:48 PM

Multi-broker  
Single Partition Topic  
Replication Factor of 3

Look for:

--describe command  
Failure handling  
Continue operation

.\bin\window\kafka-topics.bat --create --topic test --zookeeper localhost:2181 -- replication-factor 1 --partitions 1

Step 1: create server.properties for each instance

Step 2: Run each instance in different windows

Step 3: Create a topic

.\bin\window\kafka-topics.sh --create --topic replicated\_topic --zookeeper localhost:2181 -- replication-factor 2 --partitions 1

Step 4: Describe topic

\kafka\_2.9.1-0.8.2.2>.\bin\windows\kafka-topics.bat --zookeeper localhost:2181 --describe  
Topic:repl\_topic PartitionCount:1 ReplicationFactor:2 Configs:

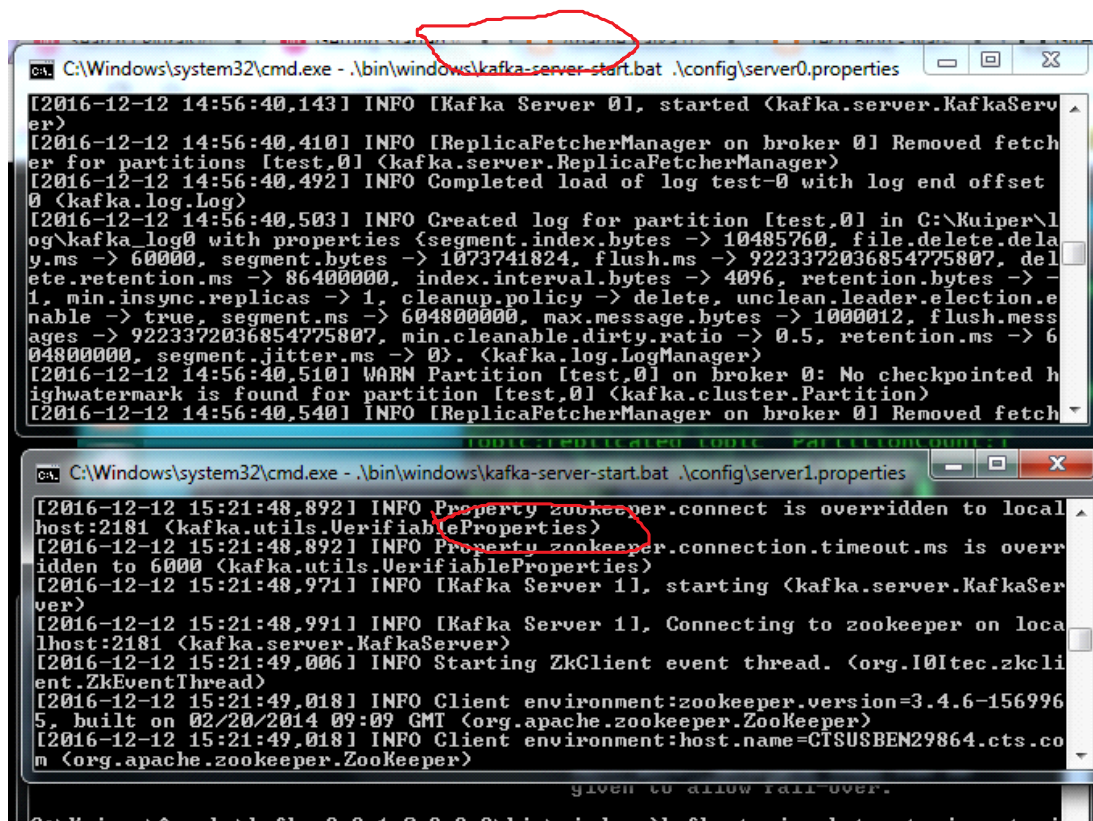
Topic: repl\_topic Partition: 0 Leader: 1 Replicas: 1,0 Isr: 1,0

Topic:test PartitionCount:1 ReplicationFactor:1 Configs:

Topic: test Partition: 0 Leader: 0 Replicas: 0 Isr: 0

Step 5: .\bin\windows\kafka-console-producer.bat -- broker-list localhost:9092 --topic repl\_topic

Step 6: .\bin\windows\kafka-console-consumer.bat --zookeeper localhost:2181 --topic repl\_topic --from-beginning



```
C:\Windows\system32\cmd.exe - .\bin\windows\kafka-server-start.bat .\config\server0.properties
[2016-12-12 14:56:40,143] INFO [Kafka Server 0], started (kafka.server.KafkaServer)
[2016-12-12 14:56:40,410] INFO [ReplicaFetcherManager on broker 0] Removed fetcher for partitions [test,0] (kafka.server.ReplicaFetcherManager)
[2016-12-12 14:56:40,492] INFO Completed load of log test-0 with log end offset 0 (kafka.log.Log)
[2016-12-12 14:56:40,503] INFO Created log for partition [test,0] in C:\Kuiper\log\kafka_log0 with properties {segment.index.bytes -> 10485760, file.delete.delay.ms -> 60000, segment.bytes -> 1073741824, flush.ms -> 9223372036854775807, delete.retention.ms -> 86400000, index.interval.bytes -> 4096, retention.bytes -> -1, min.insync.replicas -> 1, cleanup.policy -> delete, unclean.leader.election.enable -> true, segment.ms -> 604800000, max.message.bytes -> 1000012, flush.messages -> 9223372036854775807, min.cleanable.dirty.ratio -> 0.5, retention.ms -> 604800000, segment.jitter.ms -> 0}. (kafka.log.LogManager)
[2016-12-12 14:56:40,510] WARN Partition [test,0] on broker 0: No checkpointed highwatermark is found for partition [test,0] (kafka.cluster.Partition)
[2016-12-12 14:56:40,540] INFO [ReplicaFetcherManager on broker 0] Removed fetcher for partitions [test,0] (kafka.server.ReplicaFetcherManager)

C:\Windows\system32\cmd.exe - .\bin\windows\kafka-server-start.bat .\config\server1.properties
[2016-12-12 15:21:48,892] INFO Property zookeeper.connect is overridden to localhost:2181 (kafka.utils.VerifiableProperties)
[2016-12-12 15:21:48,892] INFO Property zookeeper.connection.timeout.ms is overridden to 6000 (kafka.utils.VerifiableProperties)
[2016-12-12 15:21:48,971] INFO [Kafka Server 1], starting (kafka.server.KafkaServer)
[2016-12-12 15:21:48,991] INFO [Kafka Server 1], Connecting to zookeeper on localhost:2181 (kafka.server.KafkaServer)
[2016-12-12 15:21:49,006] INFO Starting ZkClient event thread. (org.I0Itec.zkclient.ZkEventThread)
[2016-12-12 15:21:49,018] INFO Client environment:zookeeper.version=3.4.6-1569965, built on 02/20/2014 09:09 GMT (org.apache.zookeeper.ZooKeeper)
[2016-12-12 15:21:49,018] INFO Client environment:host.name=CTSUSBEN29864.cts.com (org.apache.zookeeper.ZooKeeper)
```

```
[2016-12-12 15:21:49.018] INFO Client environment:host.name=CTSUSBEN29864.cts.co
m <org.apache.zookeeper.ZooKeeper>

given to allow fail-over.

C:\Kuiper\Apache\kafka_2.9.1-0.8.2.2\bin\windows>kafka-topics.bat --topic mytopi
c --zookeeper localhost:2181 --list

C:\Kuiper\Apache\kafka_2.9.1-0.8.2.2\bin\windows>kafka-topics.bat --zookeeper lo
calhost:2181 --list
repl_topic
test

C:\Kuiper\Apache\kafka_2.9.1-0.8.2.2\bin\windows>kafka-topics.bat --zookeeper lo
calhost:2181 --topic repl_topic --describe
Topic:repl_topic      PartitionCount:1      ReplicationFactor:2      Configs:
      Topic: repl_topic      Partition: 0      Leader: 1      Replicas: 1,0
Isr: 1,0

C:\Kuiper\Apache\kafka_2.9.1-0.8.2.2\bin\windows>
```

Stop the server1.properties instance

```
C:\Windows\system32\cmd.exe

afka.server.KafkaServer>
[2016-12-12 15:52:14.538] INFO [Kafka Server 1], Controlled shutdown succeeded <
kafka.server.KafkaServer>
[2016-12-12 15:52:14.548] INFO [Socket Server on Broker 1], Shutting down <kafka
.network.SocketServer>
[2016-12-12 15:52:14.670] INFO [Socket Server on Broker 1], Shutdown completed <
kafka.network.SocketServer>
[2016-12-12 15:52:14.688] INFO [Kafka Request Handler on Broker 1], shutting dow
n <kafka.server.KafkaRequestHandlerPool>
[2016-12-12 15:52:14.692] INFO [Kafka Request Handler on Broker 1], shut down co
mpletely <kafka.server.KafkaRequestHandlerPool>
[2016-12-12 15:52:15.066] INFO [Replica Manager on Broker 1]: Shut down <kafka.s
erver.ReplicaManager>
[2016-12-12 15:52:15.094] INFO [ReplicaFetcherManager on broker 1] shutting down

repl_topic
test

C:\Kuiper\Apache\kafka_2.9.1-0.8.2.2\bin\windows>kafka-topics.bat --zookeeper lo
calhost:2181 --topic repl_topic --describe
Topic:repl_topic      PartitionCount:1      ReplicationFactor:2      Configs:
      Topic: repl_topic      Partition: 0      Leader: 1      Replicas: 1,0
Isr: 1,0

C:\Kuiper\Apache\kafka_2.9.1-0.8.2.2\bin\windows>kafka-topics.bat --zookeeper lo
calhost:2181 --topic repl_topic --describe
Topic:repl_topic      PartitionCount:1      ReplicationFactor:2      Configs:
      Topic: repl_topic      Partition: 0      Leader: 0      Replicas: 1,0
Isr: 0

C:\Kuiper\Apache\kafka_2.9.1-0.8.2.2\bin\windows>
```

Still Producer & Consumer are working with no issues

```
C:\Windows\system32\cmd.exe - .\bin\windows\kafka-console-producer.bat --broker-list localhost:2181,localhost:9092 --topic repl_topic --zookeeper localhost:2181 --describe
Topic: repl_topic PartitionCount:1 ReplicationFactor:2 Configs:
Topic: repl_topic Partition: 0 Leader: 1 Replicas: 1,0
Isr: 1,0
Topic: test PartitionCount:1 ReplicationFactor:1 Configs:
Topic: test Partition: 0 Leader: 0 Replicas: 0 Isr: 0

C:\Kuiper\Apache\kafka_2.9.1-0.8.2.2>.bin\windows\kafka-console-producer.bat --broker-list localhost:9092 --topic repl_topic
[2016-12-12 15:39:35,618] WARN Property topic is not valid (kafka.utils.VerifiableProperties)
My test message 1
My test message 2
My test message 3
Message 4
Hello

C:\Kuiper\Apache\KAFKA_~1.2\bin\windows>kafka-console-consumer.bat --zookeeper localhost:2181 --topic repl_topic --from-beginning
Terminate batch job (Y/N)? y
C:\Kuiper\Apache\KAFKA_~1.2\bin\windows>kafka-console-consumer.bat --zookeeper localhost:2181 --topic repl_topic --from-beginning
My test message 1
My test message 2
My test message 3
Message 4
[2016-12-12 15:52:15,682] WARN [ConsumerFetcherThread-console-consumer-96185-CTS-USBEN29864-1481578840400-9e3de4a0-0-1], Error in fetch Name: FetchRequest; Version: 0; CorrelationId: 6795; ClientId: console-consumer-96185; ReplicaId: -1; MaxWait: 100 ms; MinBytes: 1 bytes; RequestInfo: [repl_topic,0] -> PartitionFetchRequest(4,1048576). Possible cause: java.nio.channels.ClosedChannelException (kafka.consumer.ConsumerFetcherThread)
Hello
```

Bring back the broker1 again

```
C:\Kuiper\Apache\kafka_2.9.1-0.8.2.2\bin\windows>kafka-topics.bat --zookeeper localhost:2181 --topic repl_topic --describe
Topic: repl_topic PartitionCount:1 ReplicationFactor:2 Configs:
Topic: repl_topic Partition: 0 Leader: 0 Replicas: 1,0
Isr: 0,1
```

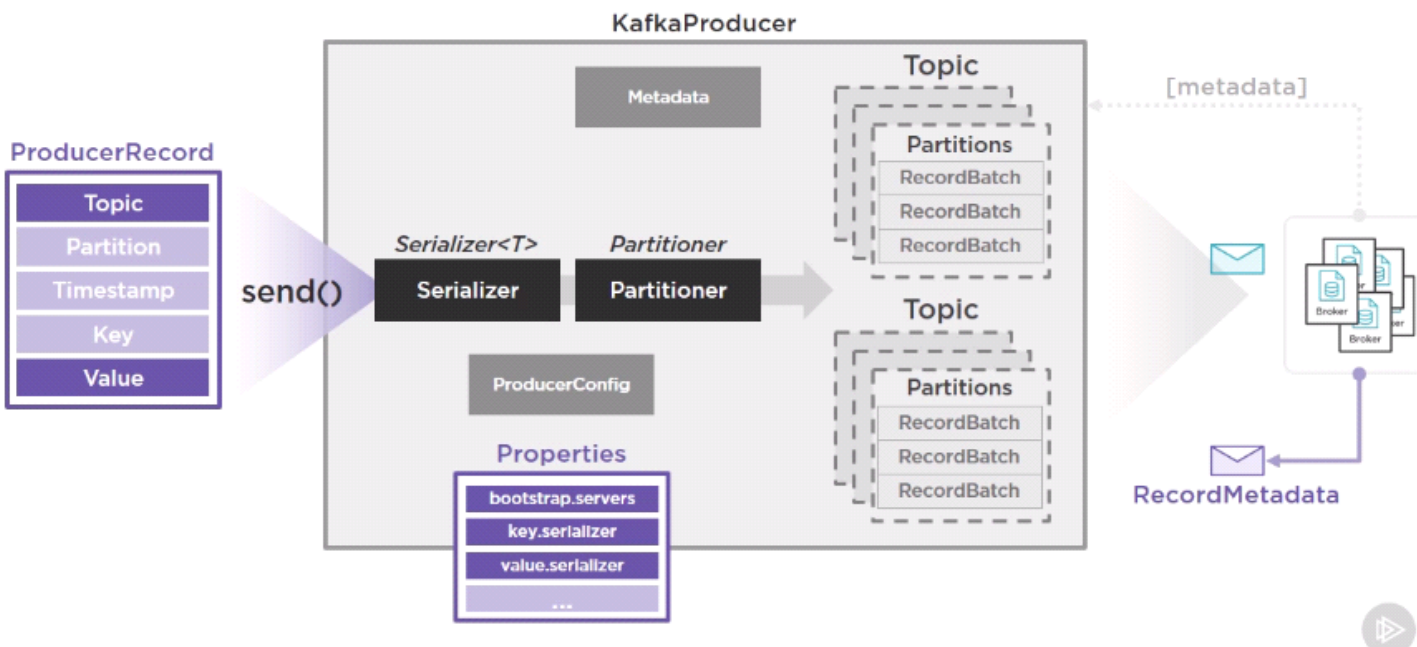
# Producer

Monday, December 12, 2016  
4:07 PM

Maven Project

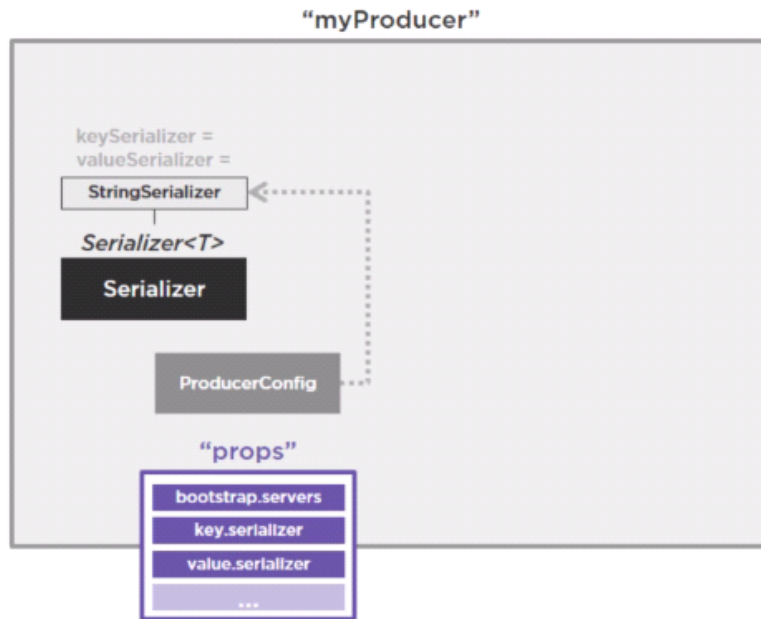
```
<dependencies>  
  <!-- https://mvnrepository.com/artifact/org.apache.kafka/kafka-clients -->  
  <dependency>  
    <groupId>org.apache.kafka</groupId>  
    <artifactId>kafka-clients</artifactId>  
    <version>0.10.1.0</version>  
  </dependency>  
</dependencies>
```

## Kafka Producer Internals

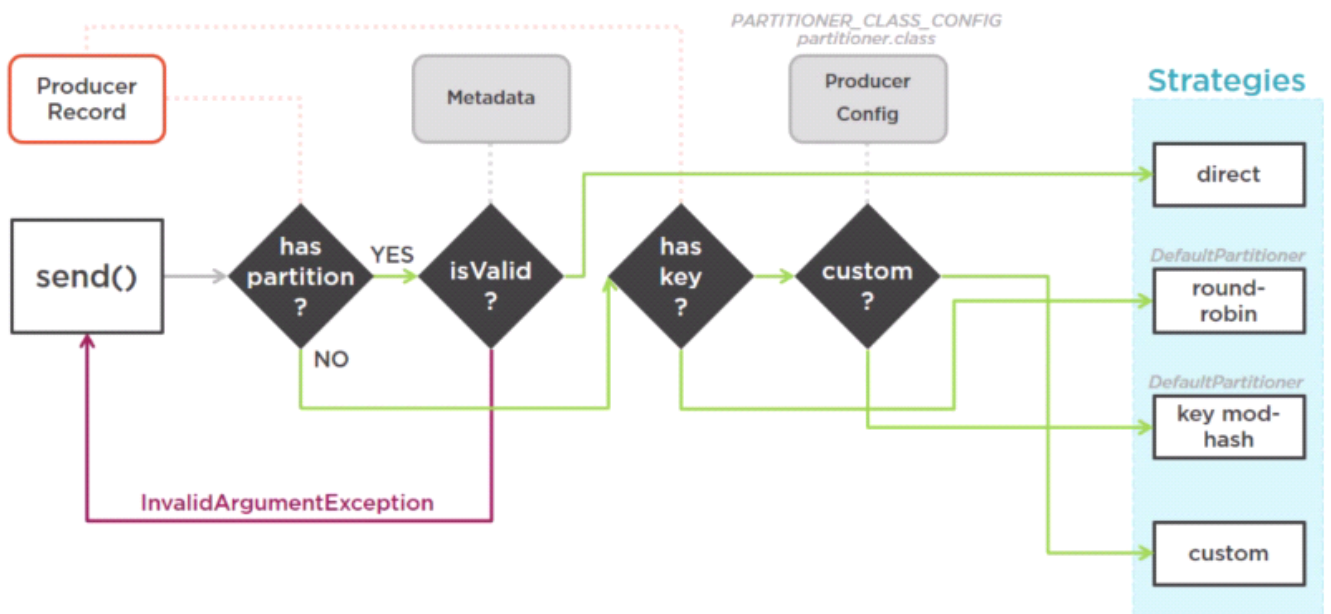


```
Properties props = new Properties();  
props.put("bootstrap.servers", "BROKER-1:9092, BROKER-2:9093");  
props.put("key.serializer", "org.apache.kafka.common.serialization.StringSerializer");  
props.put("value.serializer", "org.apache.kafka.common.serialization.StringSerializer");
```

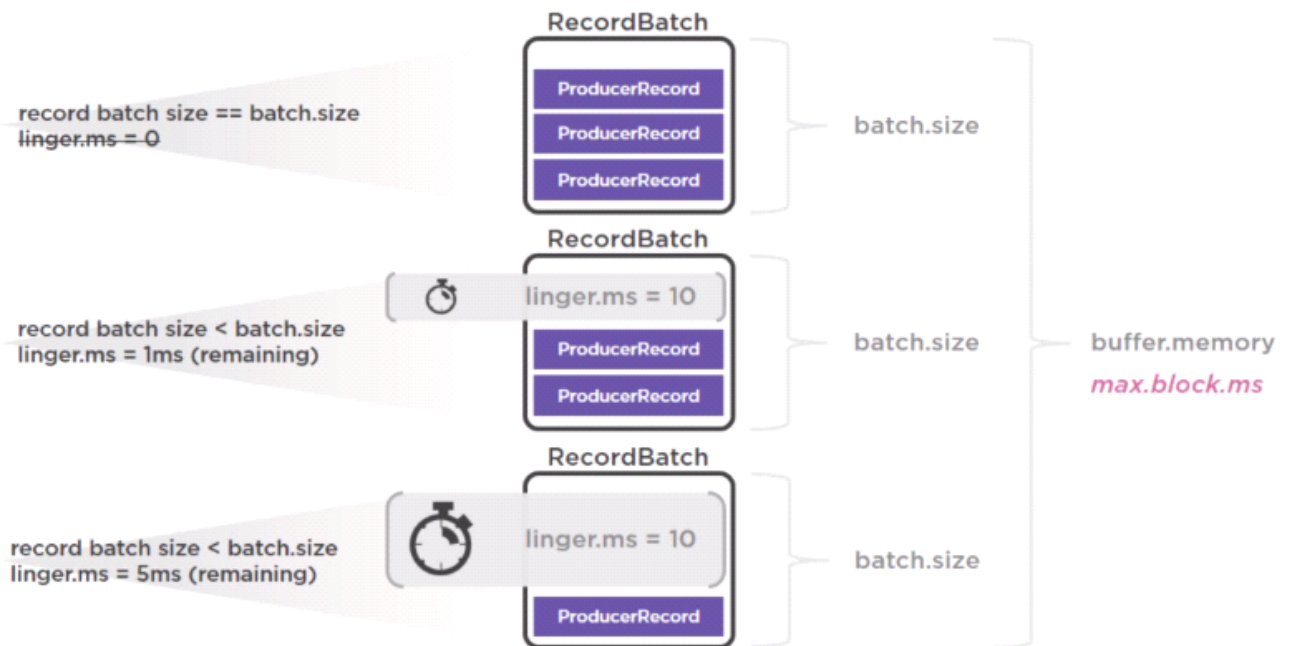
# Basic Kafka Producer



## Kafka Producer Partitioning Strategy



# Message Buffering



## Delivery Guarantees

### Broker acknowledgement (“acks”)

- 0: fire and forget
- 1: leader acknowledged
- 2: replication quorum acknowledged

### Broker responds with error

- “retries”
- “`retry.backoff.ms`”



Tuesday, December 13, 2016  
4:09 PM

Tuesday, December 13, 2016  
4:09 PM

# Kafka Consumer Internals

