Guía de instalación

Introducción a Kafka

Apache Kafka (broker)

Es un message broker distribuido escrito por dos personas de LinkedIn y recibe su nobre por el escritor Franz Kafka.

Puede procesar sobre 4.5 trillones de mensajes por día (dato del 2018)

Tiene una tecnología para procesamiento de Streams llamado Kafka Streams (Big Data)

¿En qué consiste Apache Kafka?

* Trabaja basado en tópicos identificados con un nombre (se pueden tener tantos como se quiera) y essimilar a una tabla en una base de datos; mientras más particiones tenga un tópico, soportará más paralelismo
* Los tópicos tienen particiones ordenadas y cada uno de los mensajes tienen un ID incremental (offset)

Los Broker son los que permiten a Kafka ser distribuido y tener alta disponibilidad

Un cluster está conformado por n Brokers , identificado cada uno por un identificador.

La teoría indica que debemos tener un número impar de brokers para tener siempre un “líder” y un “respaldo”

**El factor de replicación** indica en cuántos broker está replicado un tópico y siempre debe tener un factor de replicación mayor a 1.

Cuando hablamos de factores de replicación se habla de **líderes** y **seguidores**  y el keeper se encarga de lo demás, así,un sólo broker será líder en una partición.

Palabras clave: tópico, broker, partición

Tipos de ACK:

ACK 0 : entrega y no revisa qué pasa con el mensaje

ACK 1: Se asegura que el líder reciba el mensaje

ACK all: Se asegura que tanto el líder como seguidores reciban el mensaje

Apache zookeper (orquestador)

* Controla los broker, líderes y seguidores
* Envía a kafka cambios en particiones, brokers, tópicos, etc.
* Trabaja siempre en quorum (cluster)
* Kafkano puede trabajar sin zookeper

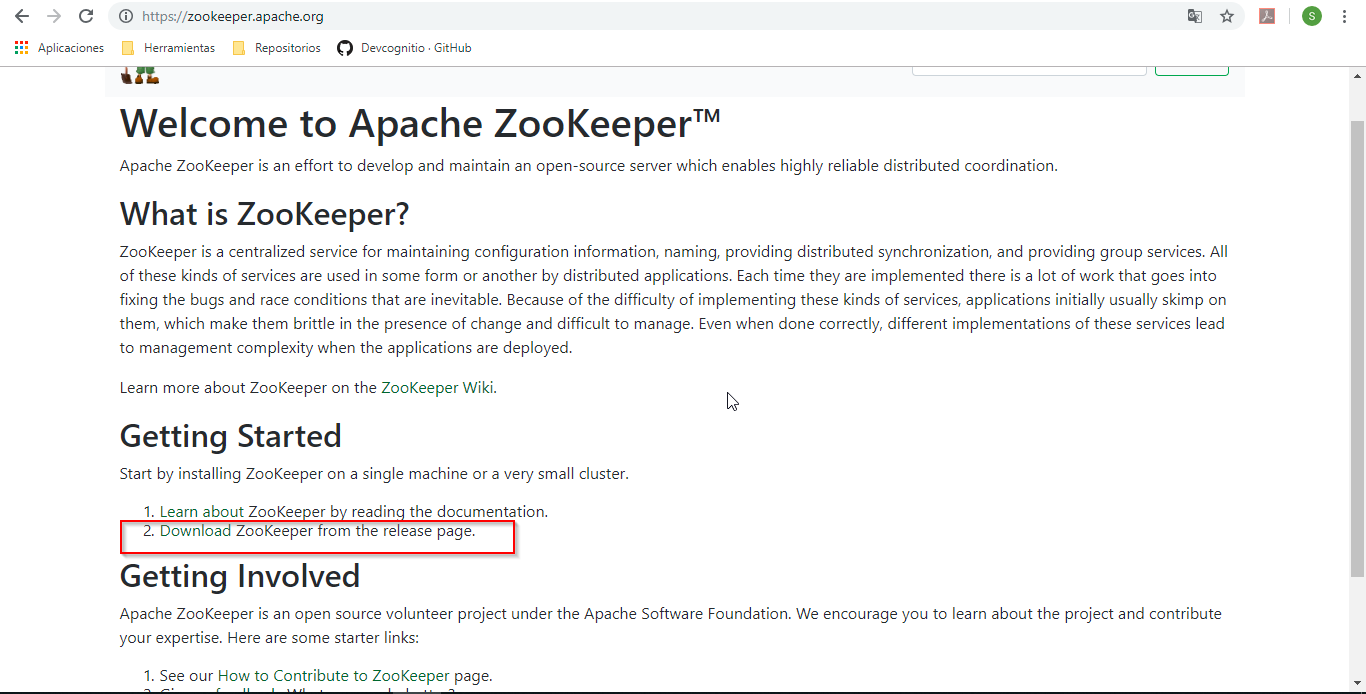
Buenas prácticas

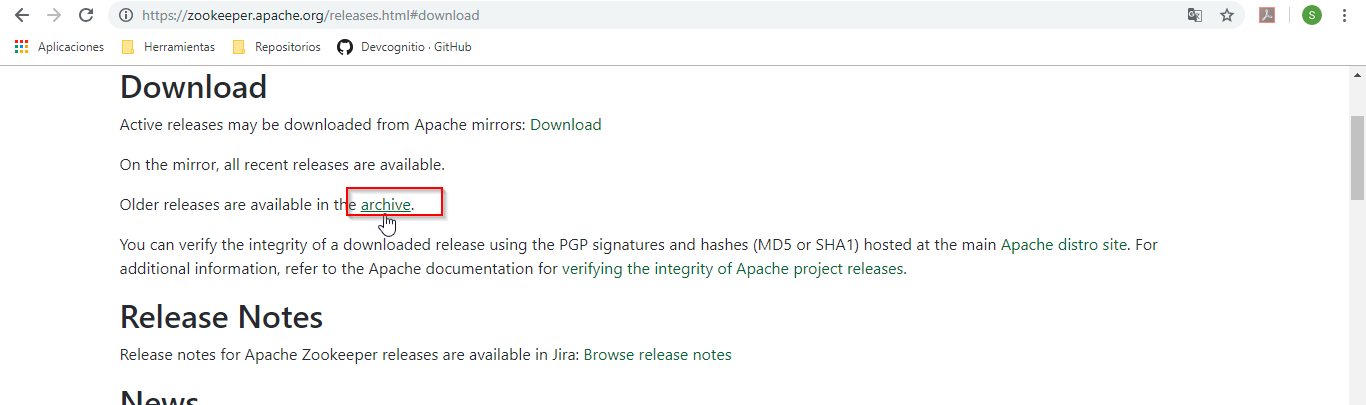
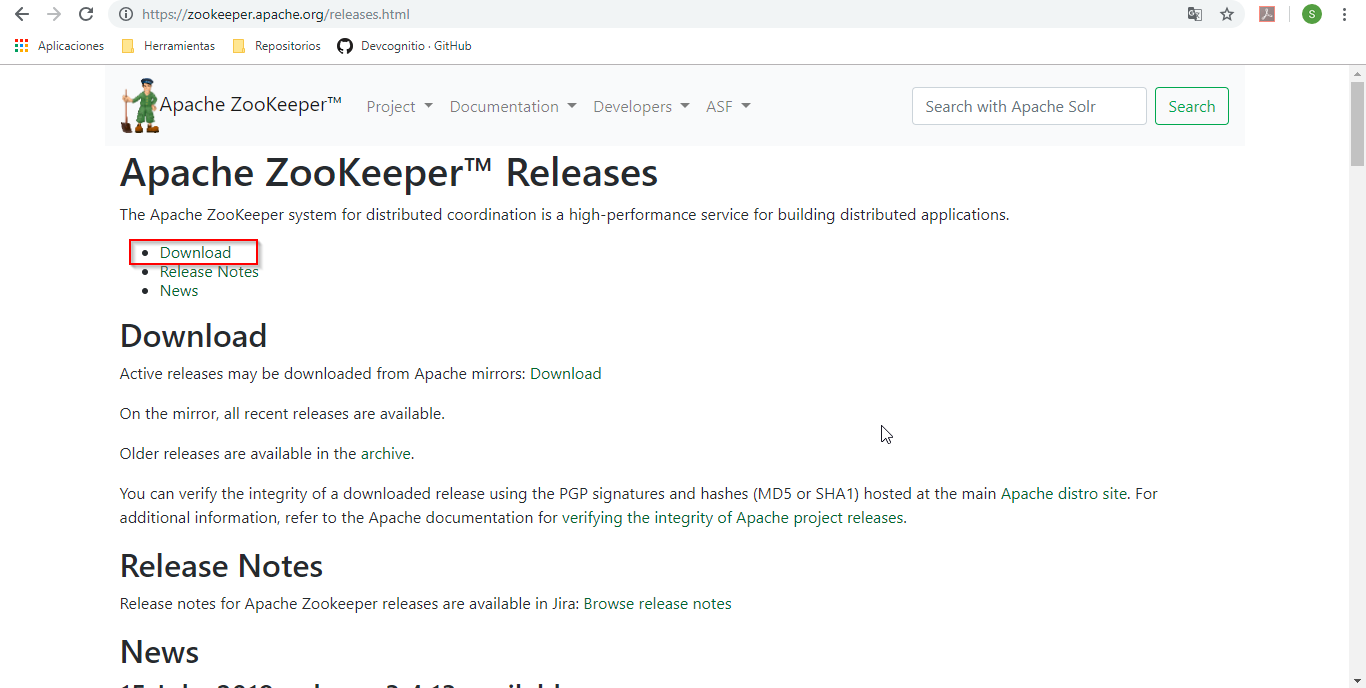
1. Siempre hacer un buen diseño antes de comenzar
2. A mayor número de particiones mayor paralelismo, mejor throughput y más archivos (I/O) y mayor latencia
3. Mantener entre 2000 y 4000 particiones por broker
4. Máximo 20000 particiones en un cluster

Instalación de herramientas en Windows

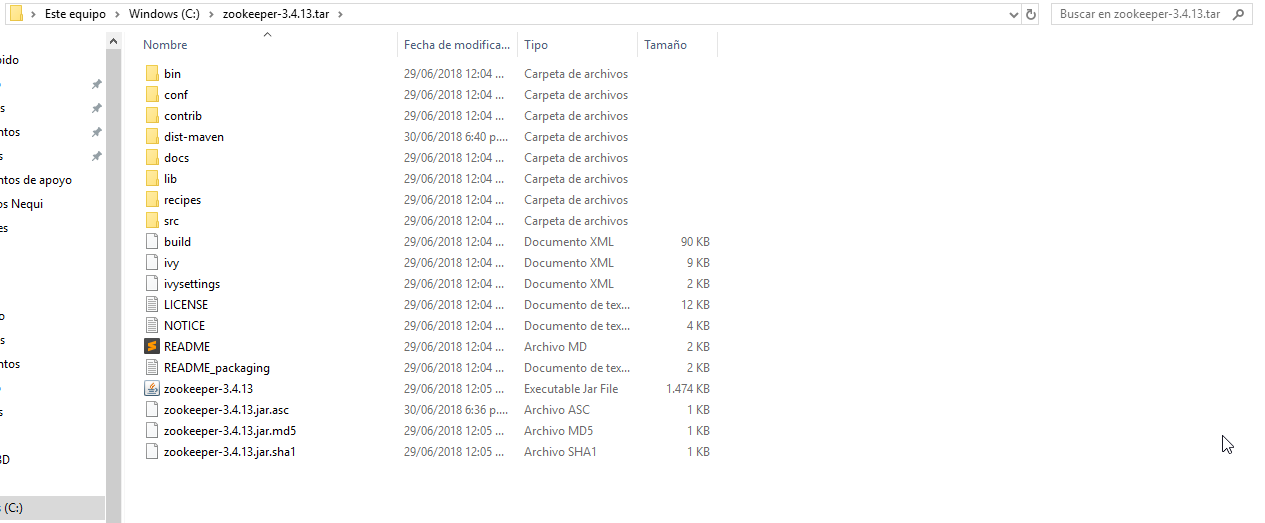
Instalar Zookeeper

<https://zookeeper.apache.org/>

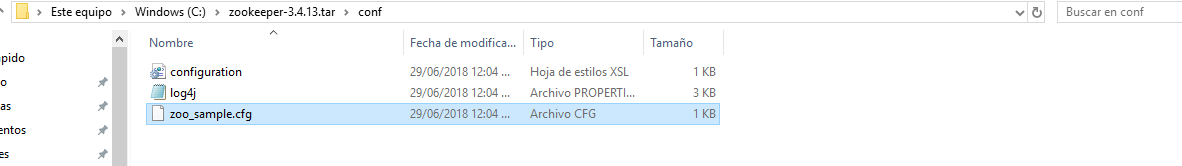




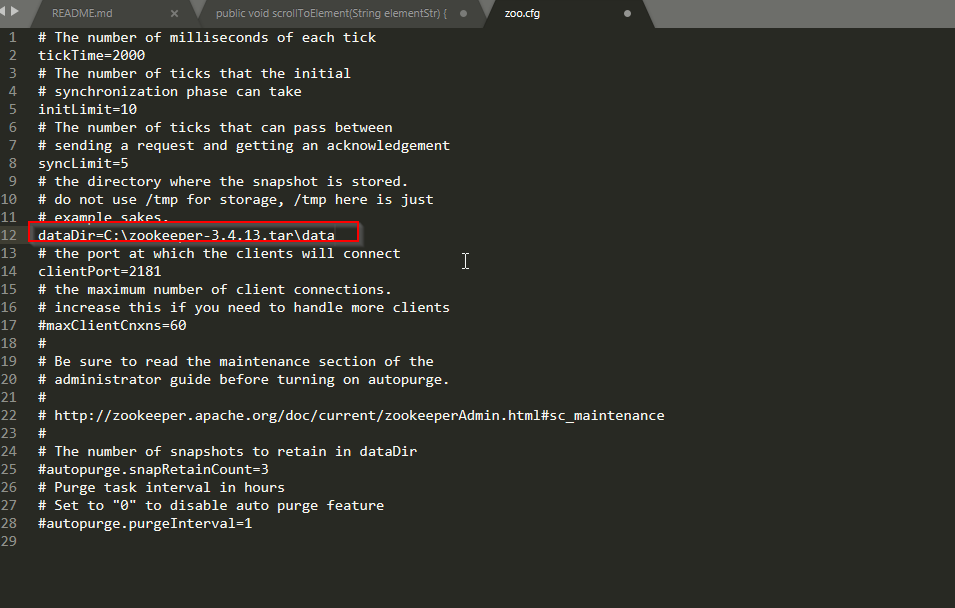
Se debe descomprimir el archivo y llevar la carpeta a una ruta accesible



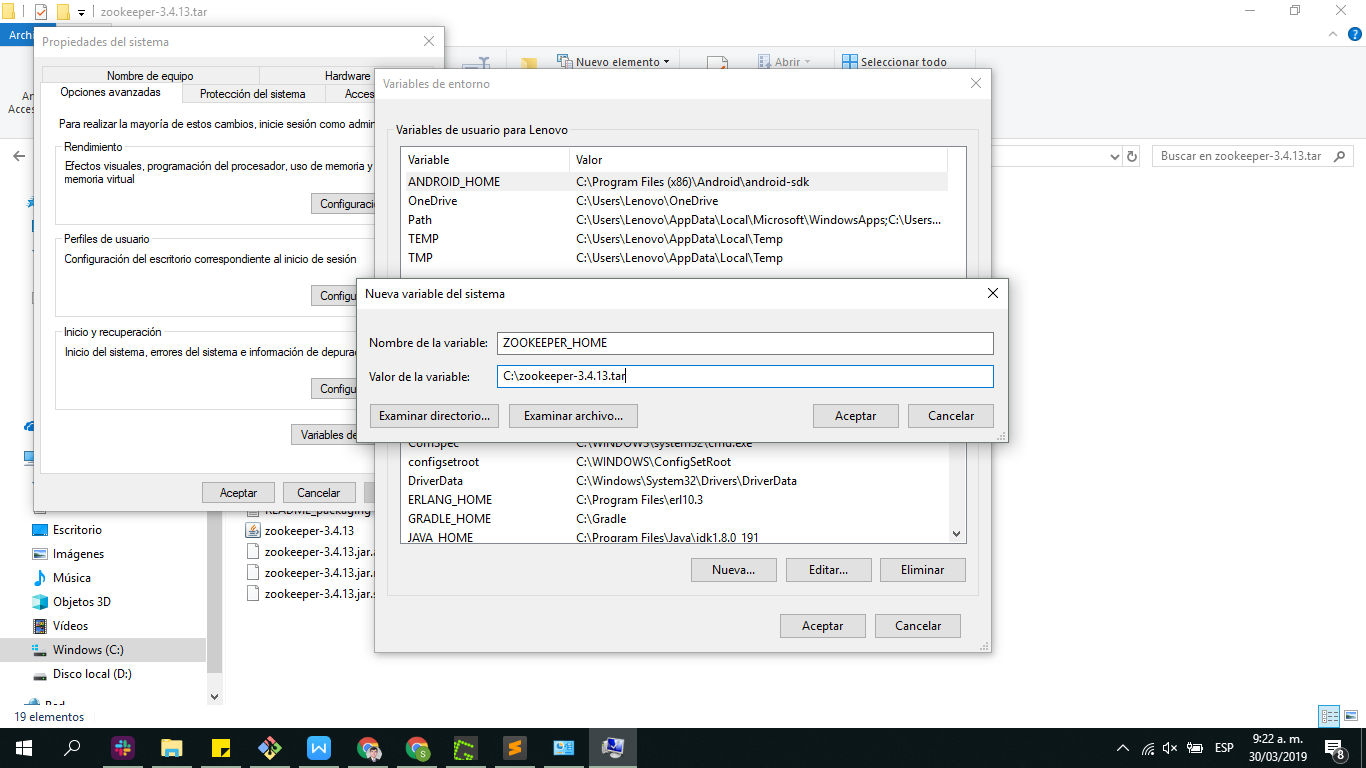
Actualizamos la configuración de Zoo

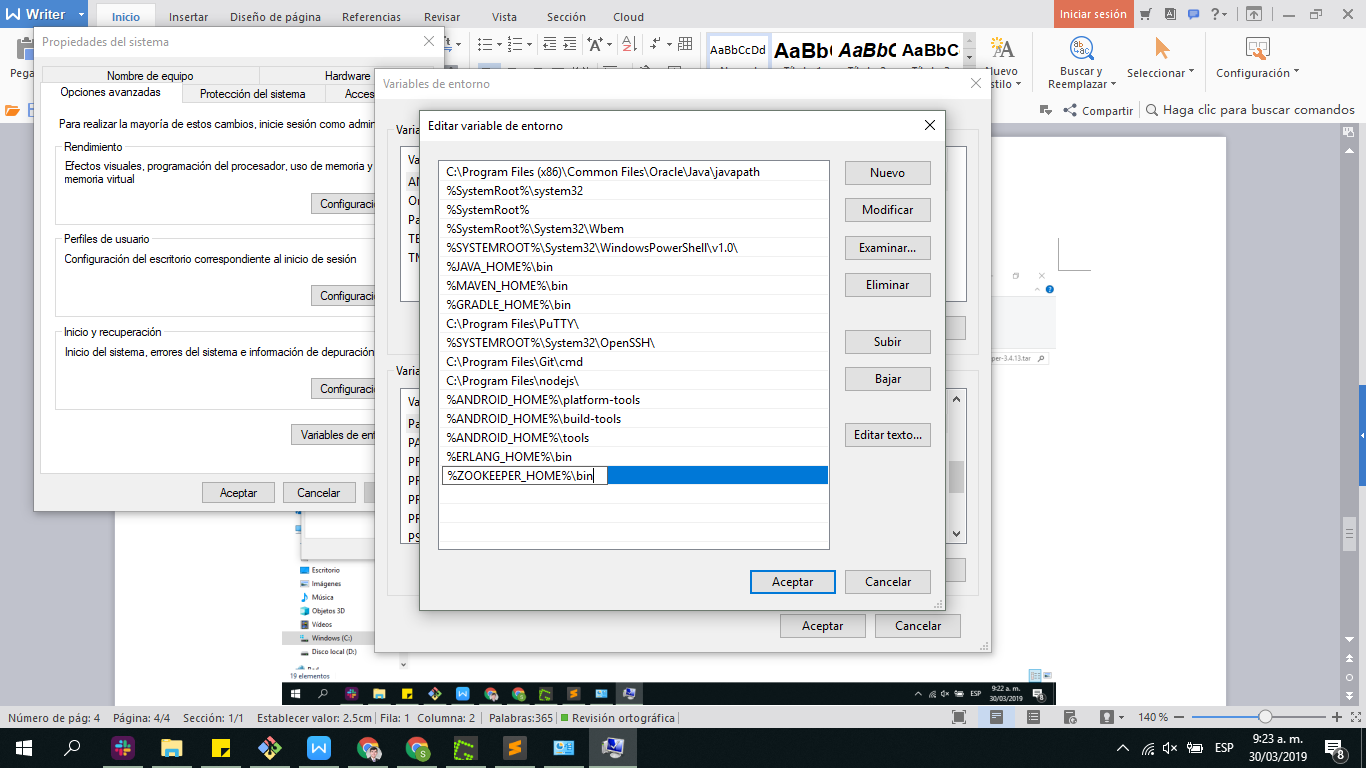


Actualizamos el DataDir para apuntar a la ruta de instalación:

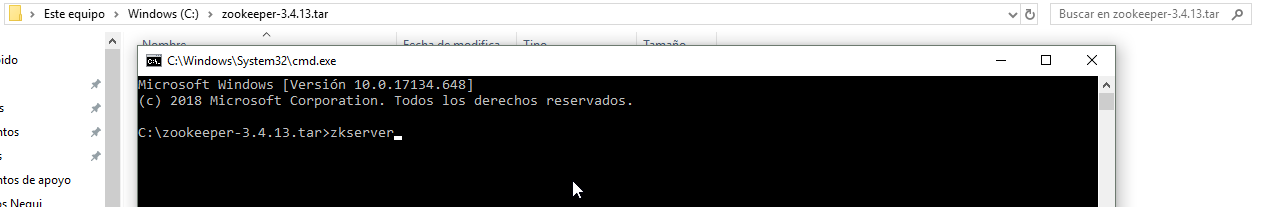


Adicionamos la variable de entorno del zookeper

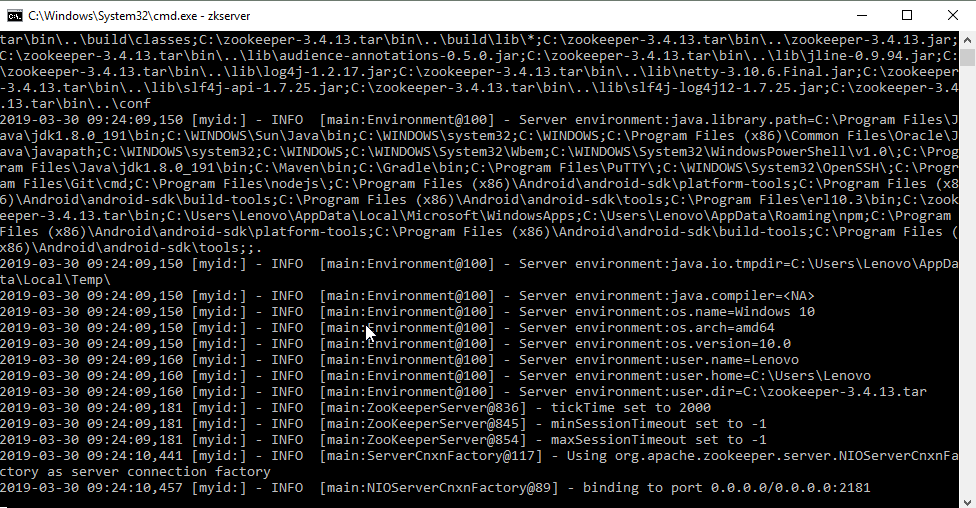




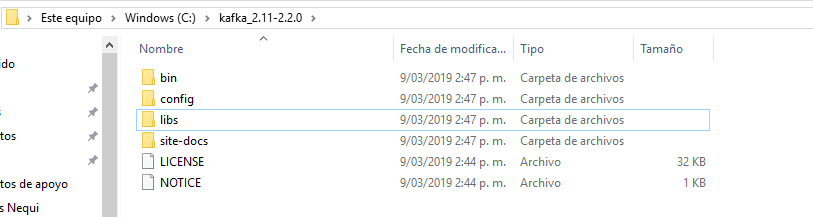
Verificamos la instalación de la aplicación con una línea de comando desde la instalación del zookeeper



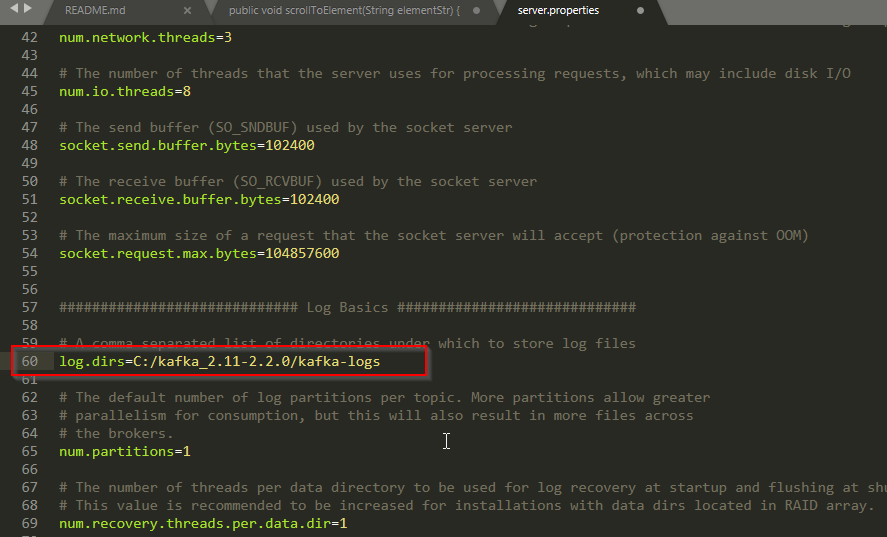
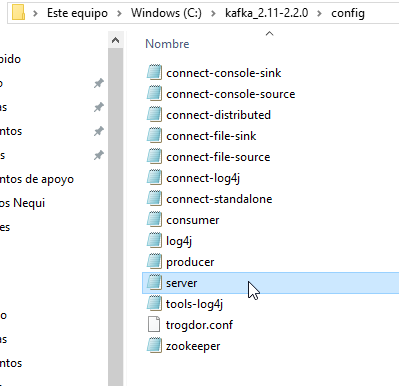
Usando zkserver:



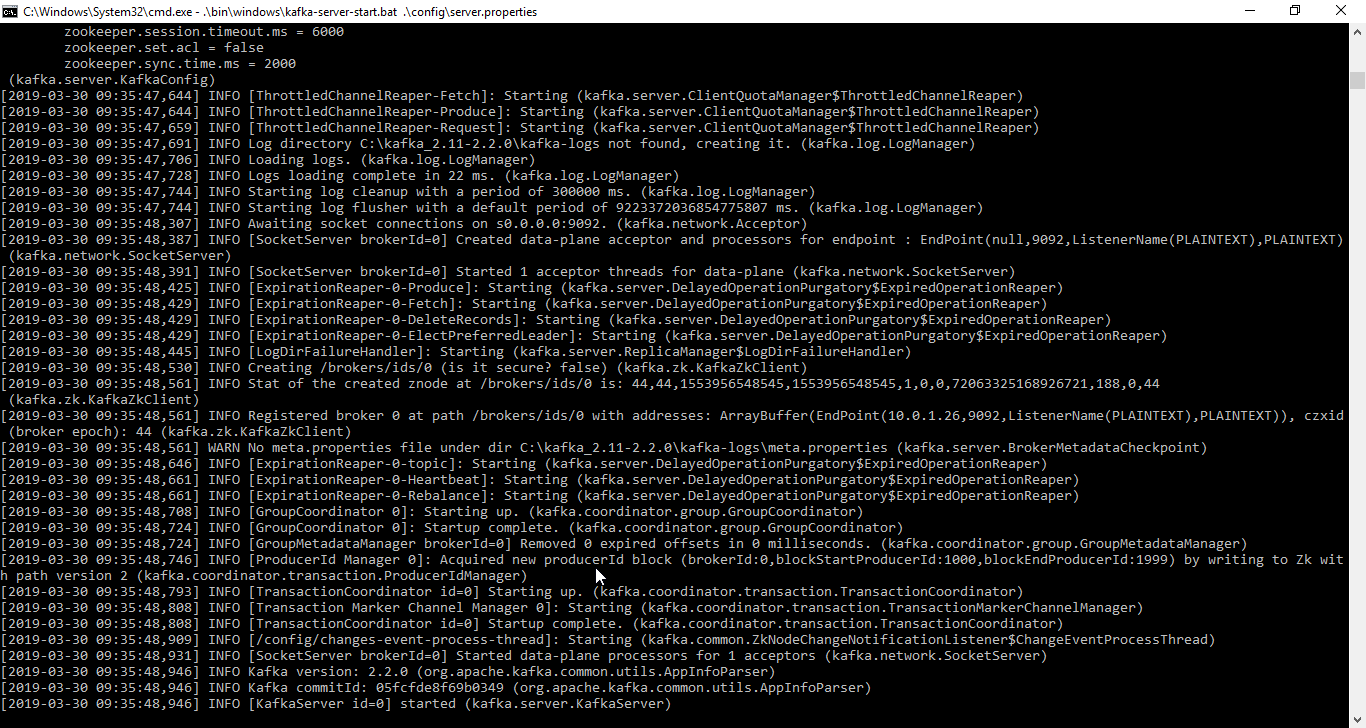
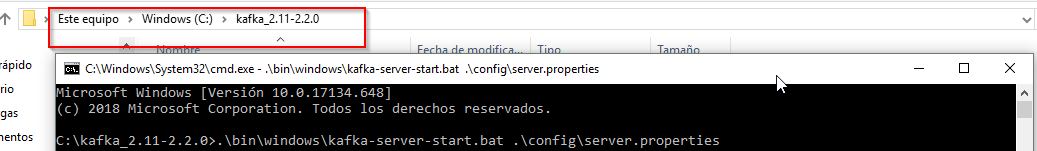
Descomprimir Kafka y llevarlo a una ruta accesible



Cambiamos en el server.properties el log.dir

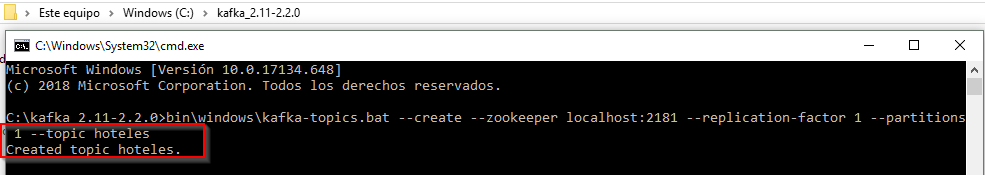


Para iniciar Kafka: .\bin\windows\kafka-server-start.bat .\config\server.properties



Comando para crear tópicos:

bin\windows\kafka-topics.bat --create --zookeeper localhost:2181 --replication-factor 1 --partitions 3 --topic nombretopico



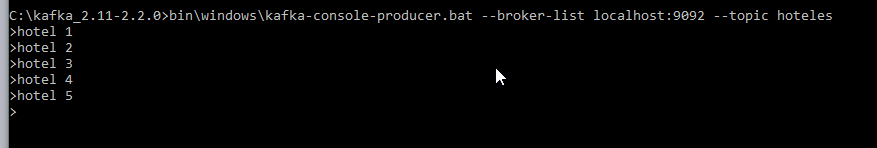
Comando para listar tópicos:

bin\windows\kafka-topics.bat --list --zookeeper localhost:2181



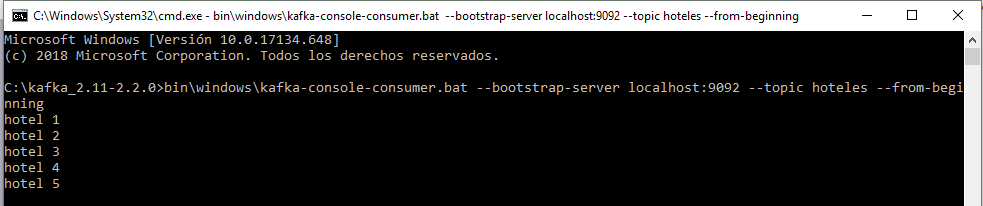
Comando para crear un productor por medio de consola:

bin\windows\kafka-console-producer.bat --broker-list localhost:9092 --topic hoteles



Crear consumidores por medio de consola:

bin\windows\kafka-console-consumer.bat --bootstrap-server localhost:9092 --topic hoteles --from-beginning



Comando para incluir grupos de consumidores

bin\windows\kafka-console-consumer.bat --bootstrap-server localhost:9092 --topic hoteles --from-beginning --group nombregrupo

