

Ejercicio práctico con Apache Kafka

El ejercicio práctico con Apache Kafka partirá del proyecto final realizado en el curso. El estado de partida es el que se puede encontrar en las últimas lecciones, con el productor de tweets y el consumidor funcionando.

Ahora, el ejercicio consta de tres apartados, con la finalidad de ampliar la aplicación que hemos desarrollado en el curso con las siguientes mejoras:

1. Para completar el primer apartado, la aplicación debe **calcular los hashtags específicos para un idioma elegido**. Por ejemplo, si elegimos idioma español, deberemos aplicar un filtro eficiente para que solamente procese los tweets con idioma “es”.
2. El segundo apartado consiste en calcular los trending topics. **Los trending topics serán los hashtags con el número de ocurrencias superior a un umbral**. Este umbral se podrá especificar vía parametrización. Por ejemplo, si establecemos un umbral de trending topics de 2, la aplicación nos deberá mostrar los tweets cuyos hashtags se hayan repetido más de 2 veces durante la ejecución.
3. El tercer apartado consiste en **integrar con una base de datos** que elijamos, por ejemplo Elasticsearch o MySQL. Para ello, deberemos insertar los resultados de nuestro procesamiento (los tweets) en la base de datos en tiempo real de manera eficiente.

La entrega del ejercicio práctico consistirá en el código fuente con la implementación del consumidor y los tres apartados.