

Kaska: persistencia de mensajes (parte extra de la práctica)

En este documento se plantea una extensión de la práctica que permite obtener una nota adicional de **2 puntos**, que solo podrá ser obtenida si la parte obligatoria pasa todas las pruebas.

Aviso: antes de afrontar esta práctica optativa, valore el esfuerzo que requiere y asegúrese de que no le perturba la realización de actividades obligatorias de esta u otras asignaturas.

Objetivo de la extensión de la práctica

En la parte final de la práctica obligatoria, se proporciona la funcionalidad que permite que los clientes hagan persistentes en el sistema de ficheros del equipo donde ejecuta el *broker* los *offsets* de los temas a los que están suscritos. Esta funcionalidad hace posible que cuando se reinicia un cliente pueda continuar su trabajo donde lo dejó. Sin embargo, no sirve ante un reinicio del *broker* ya que se pierden los mensajes. En esta extensión, nos planteamos hacer persistentes los mensajes, pero sin que esto apenas repercuta en la eficiencia en la gestión de los mismos gracias al uso de la proyección de archivos (*mmap*).

A continuación, se describe la nueva funcionalidad que hay que añadir a la práctica, que corresponde solo al *broker*, y el modo de operación de la misma.

En primer lugar, dado que se van a hacer persistentes dos tipos de información, vamos a reorganizar la estructura de directorios en el *broker*:

- Dentro del directorio recibido como argumento se creará un directorio *commits* para los *offsets* guardados, tal como se hace en la parte básica, y un directorio *data* para almacenar los mensajes.
- En el directorio *data*, habrá un fichero por cada tema creado con un tamaño de 1GiB (`#define FILE_SIZE 1<<30`). Si tiene alguna limitación en su cuenta use un tamaño menor. Kafka usa esta misma estrategia, pero cuando se llena un fichero, le asigna al tema uno nuevo, denominando segmento a cada uno de ellos.
- El fichero se accederá con la técnica de proyección de archivos asignándole con `ftruncate` el tamaño indicado que hace que el fichero no gaste espacio en disco hasta que no vaya almacenando mensajes.
- Como comprobación de seguridad, se almacenará en los 4 primeros bytes del archivo la secuencia "KASK" (su *magic number*).
- El resto del fichero puede organizarlo como considere oportuno, aunque, lógicamente, almacenará la longitud y contenido de cada mensaje enviado a ese tema.
- El uso de la proyección de archivos permite que el acceso a los mensajes sea a todos los efectos tan eficiente como en la parte básica, no realizándose operaciones de lectura y escritura en el sistema de ficheros. Además, hace posible que el código en el *broker* que realiza el tratamiento de las operaciones de envío y de lectura no se vea afectado prácticamente por el cambio: en la cola almacenaremos la referencia al mensaje, que, en la parte básica, corresponderá al *heap*, mientras que en la extra apuntará a una dirección de la región de memoria asociada al fichero. Nótese que debe seguir cumpliéndose un comportamiento *zero-copy* en el tratamiento de los mensajes.

Hay dos partes principales en esta nueva funcionalidad: lograr que los mensajes sean persistentes y recrear el estado cuando arranca el *broker*.

Con respecto a la persistencia, se debería tener en cuenta lo siguiente:

- Al crear un tema, además de la labor que se realizaba en la parte básica, hay que crear el fichero con el nombre del tema, asignarle el espacio especificado (`FILE_SIZE`) y proyectarlo.
- Se debe incluir al principio el *magic number*.
- Deberá incluir algún tipo de información en el descriptor del tema que permita asociarlo con el fichero. Por ejemplo, se puede almacenar la dirección donde se almacenará el próximo mensaje que se reciba.

- En la recepción de un mensaje, para asegurar el *zero copy* en la gestión de los mensajes, se debe leer directamente del socket en la zona correspondiente de la región asociada al fichero proyectado.

En cuanto a recuperar el estado cuando se reinicia:

- Hay que hacerlo antes de empezar el bucle de servicio.
- Se deben recorrer todas las entradas del directorio tratando solamente aquellas que correspondan a ficheros regulares que contengan el *magic number*.
- Por cada uno de esos ficheros, habrá que abrirlo y proyectarlo creando el tema correspondiente en el mapa e ir identificando en el fichero proyectado los sucesivos mensajes añadiendo una entrada en la cola del mapa por cada uno de los mensajes de manera que en el descriptor se guarde la dirección del mensaje en la región del fichero proyectado.

Entrega de la práctica

Se realizará en la máquina `triqui`, usando el mandato:

```
entrega.sd kaska-extra.2023
```

Este mandato recogerá los mismos ficheros que para la práctica básica.