

Advanced kafka homework.

I created a topic called “hadoop” in kafka:

```
[maria_dev@sandbox-hdp bin]$ ./kafka-topics.sh --create --zookeeper sandbox-hdp.hortonworks.com:2181 --replication-factor 1 --partitions 2 --topic hadoop
Created topic "hadoop".
```

As can be easily noted, it has 2 partitions.

Then I started my producer jar:

```
[maria_dev@sandbox-hdp ~]$ java -jar kafka.jar -topic hadoop -url sandbox-hdp.hortonworks.com:6667 -passStore /home/maria_dev/store.paldb -keywords pavel1,pavel2,pavel3 -howManyMessages 3
Jan 10, 2019 12:38:44 PM com.linkedin.paldb.impl.StoreImpl createReader
INFO: Initialize reader from file store.paldb
Jan 10, 2019 12:38:44 PM com.linkedin.paldb.impl.ReaderImpl <init>
INFO: Opening reader storage
Jan 10, 2019 12:38:44 PM com.linkedin.paldb.impl.StorageReader <init>
INFO: Opening file store.paldb
Jan 10, 2019 12:38:44 PM com.linkedin.paldb.impl.StorageReader <init>
INFO: Storage metadata
```

The producer runs until it has consumed 3 messages. It looks for tweets with hashtags: “#pavel1”, “#pavel2”, “#pavel3”.

I chose those specific tweets, because their hashcodes differ by 1, so it’s easier to guess which partition they will go into.

Now, let’s look at each partition, here’s partition 0:

```
[maria_dev@sandbox-hdp bin]$ ./kafka-console-consumer.sh --bootstrap-server sandbox-hdp.hortonworks.com:6667 --topic hadoop --from-beginning --partition 0
{"created_at":"Thu Jan 10 12:40:51 +0000 2019","id":"1083342906814021632","id_str":"1083342906814021632","text":"#pavel1","source":"\u003ca href=\"http://twitter.com\" rel=\"nofollow\"\u003eTwitter Web Client\u003c/a\u0003e","truncated":false,"in_reply_to_status_id":null,"in_reply_to_status_id_str":null,"in_reply_to_user_id":null,"in_reply_to_user_id_str":null,"in_reply_to_screen_name":null,"user":{"id":"252154948","id_str":"252154948","name":"Paul","screen_name":"PaulOrekhov","location":"Tver, Russia","url":null,"description":null,"translator_type":"none","protected":false,"verified":false,"followers_count":0,"friends_count":2,"listed_count":0,"favourites_count":17,"statuses_count":19,"created_at":"Mon Feb 14 16:14:53 +0000 2011","utc_offset":null,"time_zone":null,"geo_enabled":false,"lang":"en","contributors_enabled":false,"is_translator":false,"profile_background_color":"#C0DEED","profile_background_image_url":"http://abs.twimg.com/images/themes/1/bg.png","profile_background_image_url_https":"https://abs.twimg.com/images/themes/1/bg.png","profile_background_tile":false,"profile_link_color":"#1DA1F2","profile_sidebar_border_color":"#C0DEED","profile_sidebar_fill_color":"#D0E0F6","profile_text_color":"#333333","profile_use_background_image":true,"profile_image_url":"http://abs.twimg.com/sticky/default_profile_images/default_profile_normal.png","profile_image_url_https":"https://abs.twimg.com/sticky/default_profile_images/default_profile_normal.png","default_profile":true,"default_profile_image":false,"following":null,"follow_request_sent":null,"notifications":null,"geo":null,"coordinates":null,"place":null,"contributors":null,"is_quote_status":false,"quote_count":0,"reply_count":0,"retweet_count":0,"favorite_count":0,"entities":{"hashtags":[{"text":"pavel1","indices":[0,7]}],"urls":[],"user_mentions":[],"symbols":[],"favorited":false,"retweeted":false,"filter_level":"low","lang":"und","timestamp_ms":"1547124051179"}
{"created_at":"Thu Jan 10 12:41:38 +0000 2019","id":"1083343070043754496","id_str":"1083343070043754496","text":"#pavel3","source":"\u003ca href=\"http://twitter.com\" rel=\"nofollow\"\u003eTwitter Web Client\u003c/a\u0003e","truncated":false,"in_reply_to_status_id":null,"in_reply_to_status_id_str":null,"in_reply_to_user_id":null,"in_reply_to_user_id_str":null,"in_reply_to_screen_name":null,"user":{"id":"252154948","id_str":"252154948","name":"Paul","screen_name":"PaulOrekhov","location":"Tver, Russia","url":null,"description":null,"translator_type":"none","protected":false,"verified":false,"followers_count":0,"friends_count":2,"listed_count":0,"favourites_count":17,"statuses_count":21,"created_at":"Mon Feb 14 16:14:53 +0000 2011","utc_offset":null,"time_zone":null,"geo_enabled":false,"lang":"en","contributors_enabled":false,"is_translator":false,"profile_background_color":"#C0DEED","profile_background_image_url":"http://abs.twimg.com/images/themes/1/bg.png","profile_background_image_url_https":"https://abs.twimg.com/images/themes/1/bg.png","profile_background_tile":false,"profile_link_color":"#1DA1F2","profile_sidebar_border_color":"#C0DEED","profile_sidebar_fill_color":"#D0E0F6","profile_text_color":"#333333","profile_use_background_image":true,"profile_image_url":"http://abs.twimg.com/sticky/default_profile_images/default_profile_normal.png","profile_image_url_https":"https://abs.twimg.com/sticky/default_profile_images/default_profile_normal.png","default_profile":true,"default_profile_image":false,"following":null,"follow_request_sent":null,"notifications":null,"geo":null,"coordinates":null,"place":null,"contributors":null,"is_quote_status":false,"quote_count":0,"reply_count":0,"retweet_count":0,"favorite_count":0,"entities":{"hashtags":[{"text":"pavel3","indices":[0,7]}],"urls":[],"user_mentions":[],"symbols":[],"favorited":false,"retweeted":false,"filter_level":"low","lang":"und","timestamp_ms":"1547124090096"}
^CProcessed a total of 1 messages
```

Here’s partition 1:

```
[maria_dev@sandbox-hdp bin]$ ./kafka-console-consumer.sh --bootstrap-server sandbox-hdp.hortonworks.com:6667 --topic hadoop --from-beginning --partition 1
{"created_at":"Thu Jan 10 12:41:08 +0000 2019","id":"1083342979484577793","id_str":"1083342979484577793","text":"#pavel2","source":"\u003ca href=\"http://twitter.com\" rel=\"nofollow\"\u003eTwitter Web Client\u003c/a\u0003e","truncated":false,"in_reply_to_status_id":null,"in_reply_to_status_id_str":null,"in_reply_to_user_id":null,"in_reply_to_user_id_str":null,"in_reply_to_screen_name":null,"user":{"id":"252154948","id_str":"252154948","name":"Paul","screen_name":"PaulOrekhov","location":"Tver, Russia","url":null,"description":null,"translator_type":"none","protected":false,"verified":false,"followers_count":0,"friends_count":2,"listed_count":0,"favourites_count":17,"statuses_count":20,"created_at":"Mon Feb 14 16:14:53 +0000 2011","utc_offset":null,"time_zone":null,"geo_enabled":false,"lang":"en","contributors_enabled":false,"is_translator":false,"profile_background_color":"#C0DEED","profile_background_image_url":"http://abs.twimg.com/images/themes/1/bg.png","profile_background_image_url_https":"https://abs.twimg.com/images/themes/1/bg.png","profile_background_tile":false,"profile_link_color":"#1DA1F2","profile_sidebar_border_color":"#C0DEED","profile_sidebar_fill_color":"#D0E0F6","profile_text_color":"#333333","profile_use_background_image":true,"profile_image_url":"http://abs.twimg.com/sticky/default_profile_images/default_profile_normal.png","profile_image_url_https":"https://abs.twimg.com/sticky/default_profile_images/default_profile_normal.png","default_profile":true,"default_profile_image":false,"following":null,"follow_request_sent":null,"notifications":null,"geo":null,"coordinates":null,"place":null,"contributors":null,"is_quote_status":false,"quote_count":0,"reply_count":0,"retweet_count":0,"favorite_count":0,"entities":{"hashtags":[{"text":"pavel2","indices":[0,7]}],"urls":[],"user_mentions":[],"symbols":[],"favorited":false,"retweeted":false,"filter_level":"low","lang":"und","timestamp_ms":"1547124068505"}
^CProcessed a total of 1 messages
```