
Lab: Creating and Managing Kafka Topics

Exercising the Kafka CLI

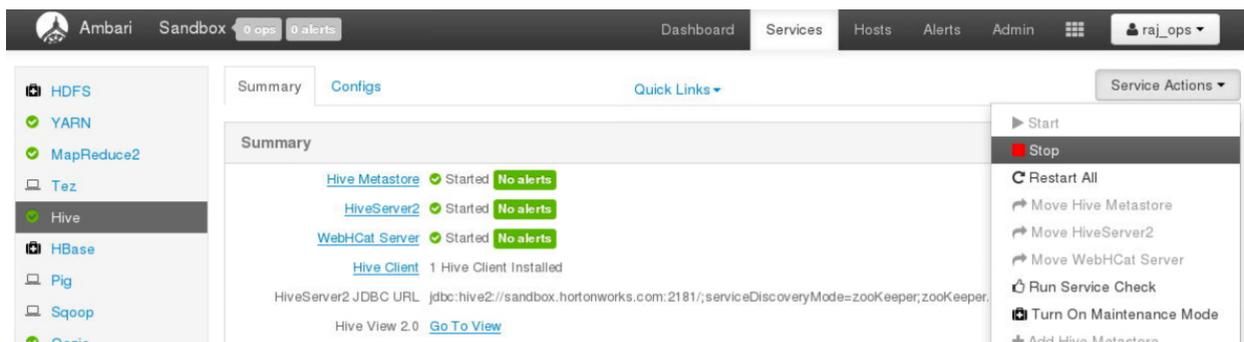
About this Lab

Objective:	Use the Kafka scripts to perform common operations related to Kafka topics
File locations:	n/a
Successful outcome:	Have gained familiarity with common Kafka script operations
Before you begin	n/a
Related lesson:	Kafka Architecture

Start the Kafka Broker

Within the remote desktop, launch the Firefox browser and go to `http://127.0.0.1:8081` and login as user `raj_ops` with password `raj_ops` to **stop the Hive, Oozie, Flume, Spark2 and Zeppelin Notebook services** and **start Kafka, Storm and HBase**.

HINT: Click on the service name on the left-nav and then select Stop or Start appropriately from the Service Actions pull-down menu



The screenshot shows the Ambari web interface. The top navigation bar includes 'Ambari', 'Sandbox', 'Dashboard', 'Services', 'Hosts', 'Alerts', 'Admin', and a user profile 'raj_ops'. The left sidebar lists services: HDFS, YARN, MapReduce2, Tez, Hive (selected), HBase, Pig, Sqoop, and Oozie. The main content area is titled 'Summary' and 'Configs'. It displays a 'Summary' section with three services: 'Hive Metastore' (Started, No alerts), 'HiveServer2' (Started, No alerts), and 'WebHCat Server' (Started, No alerts). Below this, it shows 'Hive Client' (1 Hive Client Installed) and 'HiveServer2 JDBC URL' (jdbc:hive2://sandbox.hortonworks.com:2181/;serviceDiscoveryMode=zooKeeper;zooKeeper...). A 'Service Actions' dropdown menu is open, showing options: Start, Stop (highlighted), Restart All, Move Hive Metastore, Move HiveServer2, Move WebHCat Server, Run Service Check, Turn On Maintenance Mode, and Add Hive Metastore.

You should see successful reporting of five services being stopped and three being started.

Operations	Start Time	Duration	Show: <input type="text" value="All (10)"/>
✓ Start HBase	Today 17:10	68.25 secs	 100% ▶
✓ Start Storm	Today 17:10	13.64 secs	 100% ▶
✓ Start Kafka	Today 17:10	3.23 secs	 100% ▶
✓ Stop Zeppelin Notebook	Today 17:09	15.27 secs	 100% ▶
✓ Stop Spark2	Today 17:08	1.76 secs	 100% ▶
✓ Stop Flume	Today 17:06	686 ms	 100% ▶
✓ Stop Oozie	Today 17:06	6.73 secs	 100% ▶
✓ Stop Hive	Today 17:06	32.26 secs	 100% ▶

Switch to the Appropriate User

The cluster being used is the publicly available HDP Sandbox and this lab will be completed as user `maria_dev`. Ensure you are logged into the sandbox as this user. The following steps assume you started a new Terminal.

```
[root@ip-172-30-0-164 ~]# ssh -p 2222 root@127.0.0.1
root@127.0.0.1's password:
Last login: Thu Jun  1 20:58:25 2017 from 172.17.0.1
[root@sandbox ~]# su - maria_dev
[maria_dev@sandbox ~]$ pwd
/home/maria_dev
[maria_dev@sandbox ~]$
```

Create a Topic

Create a new Kafka topic.

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

List All Topics

List all existing Kafka topics.

```
[maria_dev@sandbox ~]$ /usr/hdp/current/kafka-broker/bin/kafka-topics.sh --
list --zookeeper sandbox.hortonworks.com:2181
ATLAS_HOOK
mariaTopic
[maria_dev@sandbox ~]$
```

Add Data to a Topic

Leverage the console testing harness to produce data to the topic you just created.

```
[maria_dev@sandbox ~]$ /usr/hdp/current/kafka-broker/bin/kafka-console-
producer.sh --broker-list sandbox.hortonworks.com:6667 --topic mariaTopic
Just type some data and then hit the <enter> key
to create another line (aka message) into the topic
Here is a third line. Once done, you can just hit
CTL-C on an empty line to quit the harness
^C[maria_dev@sandbox ~]$
```

Read Data from a Topic

Use the console testing harness to consume data from a topic. Hit CTL-C to stop the script.

```
[maria_dev@sandbox ~]$ /usr/hdp/current/kafka-broker/bin/kafka-console-
consumer.sh --bootstrap-server sandbox.hortonworks.com:6667 --topic
mariaTopic --from-beginning
Just type some data and then hit the <enter> key
to create another line (aka message) into the topic
Here is a third line. Once done, you can just hit
CTL-C on an empty line to quit the harness
^CProcessed a total of 4 messages
[maria_dev@sandbox ~]$
```

Perform Alterations on a Topic

The `--alter` switch on the `kafka-configs.sh` script can allow you to change configurations for individual topics. One use case would be to empty out messages from a topic. There is no such command in Kafka, but messages are guaranteed to be around for a configurable amount of time. Change the retention period to a single second, verify the setting was changed and then see if the messages are removed.

```
[maria_dev@sandbox ~]$ /usr/hdp/current/kafka-broker/bin/kafka-configs.sh --
zookeeper sandbox.hortonworks.com:2181 --alter --entity-type topics --entity-
name mariaTopic --add-config retention.ms=1000
Updated config for entity: topic 'mariaTopic'.
[maria_dev@sandbox ~]$ /usr/hdp/current/kafka-broker/bin/kafka-configs.sh --
zookeeper sandbox.hortonworks.com:2181 --describe --entity-type topics --
entity-name mariaTopic
Configs for topic 'mariaTopic' are retention.ms=1000
[maria_dev@sandbox ~]$ /usr/hdp/current/kafka-broker/bin/kafka-console-
consumer.sh --bootstrap-server sandbox.hortonworks.com:6667 --topic
mariaTopic --from-beginning
Just type some data and then hit the <enter> key
to create another line (aka message) into the topic
Here is a third line. Once done, you can just hit
CTL-C on an empty line to quit the harness
^CProcessed a total of 4 messages
[maria_dev@sandbox ~]$ date
Fri Jun  2 19:47:31 UTC 2017
[maria_dev@sandbox ~]$
```

Note above that the older messages were not removed right away. Also, note below that they were still present a few minutes later. To simulate a real-world scenario of messages continually coming in, add a few more messages from the console producer and then verify that the old messages were removed as well as the new ones just added.

```
[maria_dev@sandbox ~]$ /usr/hdp/current/kafka-broker/bin/kafka-console-
producer.sh --broker-list sandbox.hortonworks.com:6667 --topic mariaTopic
add a few records which will
likely help the expired messages
be deleted (along with these since it is
set to only keep things around for 1 sec!!)
^C[maria_dev@sandbox ~]$
[maria_dev@sandbox ~]$ /usr/hdp/current/kafka-broker/bin/kafka-console-
consumer.sh --bootstrap-server sandbox.hortonworks.com:6667 --topic
mariaTopic --from-beginning
^CProcessed a total of 0 messages
[maria_dev@sandbox ~]$
```

Delete a Topic

Use the CLI tools to remove an existing topic.

```
[maria_dev@sandbox ~]$ /usr/hdp/current/kafka-broker/bin/kafka-topics.sh --
zookeeper sandbox.hortonworks.com:2181 --delete -topic mariaTopic
Topic mariaTopic is marked for deletion.
Note: This will have no impact if delete.topic.enable is not set to true.
```

```
[maria_dev@sandbox ~]$ /usr/hdp/current/kafka-broker/bin/kafka-topics.sh --
list --zookeeper sandbox.hortonworks.com:2181
ATLAS_H00K
__consumer_offsets
lestertester
mariaTopic - marked for deletion
[maria_dev@sandbox ~]$
```

Summary

You should now be able to perform basic Kafka operations via the provided CLI tools.