

## 2. Kafka. Архитектура

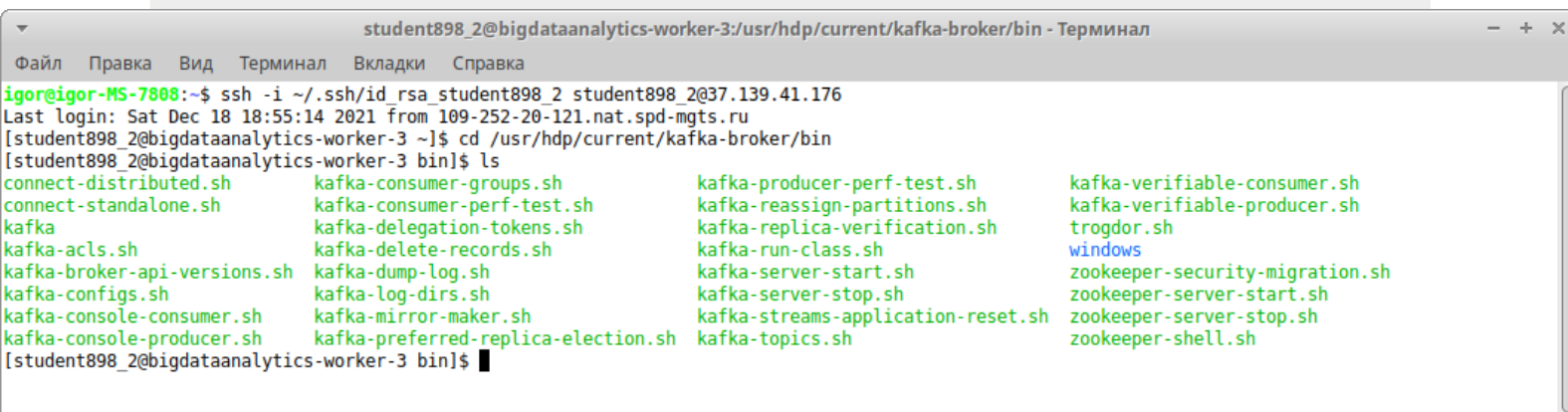
Войти в консоль, повторить команды показанные на лекции, создать свой топик, в одной консоли читать из топика, в другой консоли писать в топик из консоли. Посмотреть как работает очистка очереди. Дополнительное задание, записать в топик данные из файла используя консольные скрипты kafka.

Наполнить очередь объектами, например с помощью консоли, консольной командой записать туда csv. Записать топик с помощью producer csv файл, со структурой и consumer будет читать. Реализовать запись csv файла с помощью скриптов с папки kafka-broker/bin

```
ssh -i ~/.ssh/id_rsa_student898_2 student898_2@37.139.41.176
```

```
cd /usr/hdp/current/kafka-broker/bin
```

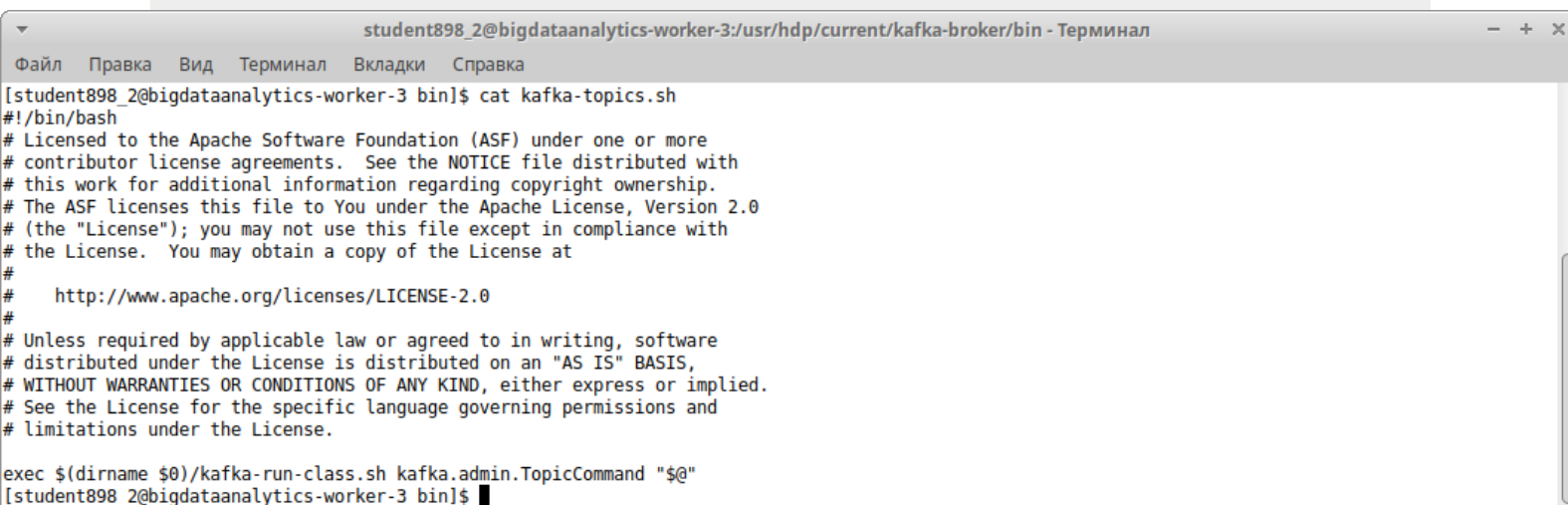
```
ls
```



```
student898_2@bigdataanalytics-worker-3:/usr/hdp/current/kafka-broker/bin - Терминал
Файл  Правка  Вид  Терминал  Вкладки  Справка
igor@igor-MS-7808:~$ ssh -i ~/.ssh/id_rsa_student898_2 student898_2@37.139.41.176
Last login: Sat Dec 18 18:55:14 2021 from 109-252-20-121.nat.spd-mgts.ru
[student898_2@bigdataanalytics-worker-3 ~]$ cd /usr/hdp/current/kafka-broker/bin
[student898_2@bigdataanalytics-worker-3 bin]$ ls
connect-distributed.sh  kafka-consumer-groups.sh  kafka-producer-perf-test.sh  kafka-verifiable-consumer.sh
connect-standalone.sh  kafka-consumer-perf-test.sh  kafka-reassign-partitions.sh  kafka-verifiable-producer.sh
kafka                  kafka-delegation-tokens.sh  kafka-replica-verification.sh  trogdor.sh
kafka-acls.sh          kafka-delete-records.sh     kafka-run-class.sh            windows
kafka-broker-api-versions.sh  kafka-dump-log.sh          kafka-server-start.sh         zookeeper-security-migration.sh
kafka-configs.sh        kafka-log-dirs.sh          kafka-server-stop.sh          zookeeper-server-start.sh
kafka-console-consumer.sh  kafka-mirror-maker.sh     kafka-streams-application-reset.sh  zookeeper-server-stop.sh
kafka-console-producer.sh  kafka-preferred-replica-election.sh  kafka-topics.sh              zookeeper-shell.sh
[student898_2@bigdataanalytics-worker-3 bin]$
```

Смотрим скрипт kafka-topics.sh

```
cat kafka-topics.sh
```



```
student898_2@bigdataanalytics-worker-3:/usr/hdp/current/kafka-broker/bin - Терминал
Файл  Правка  Вид  Терминал  Вкладки  Справка
[student898_2@bigdataanalytics-worker-3 bin]$ cat kafka-topics.sh
#!/bin/bash
# Licensed to the Apache Software Foundation (ASF) under one or more
# contributor license agreements.  See the NOTICE file distributed with
# this work for additional information regarding copyright ownership.
# The ASF licenses this file to You under the Apache License, Version 2.0
# (the "License"); you may not use this file except in compliance with
# the License.  You may obtain a copy of the License at
#
#     http://www.apache.org/licenses/LICENSE-2.0
#
# Unless required by applicable law or agreed to in writing, software
# distributed under the License is distributed on an "AS IS" BASIS,
# WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
# See the License for the specific language governing permissions and
# limitations under the License.
exec $(dirname $0)/kafka-run-class.sh kafka.admin.TopicCommand "$@"
[student898_2@bigdataanalytics-worker-3 bin]$
```

Создаем топик. Смотрим какие топики есть в kafka

```
./kafka-topics.sh --create --topic test --zookeeper bigdataanalytics-worker-3:2181 --partitions 3 --replication-factor 2 --config retention.ms=-1
```

```
./kafka-topics.sh
```

```
student898_2@bigdataanalytics-worker-3:/usr/hdp/current/kafka-broker/bin - Терминал
Файл  Правка  Вид  Терминал  Вкладки  Справка

describe. Can also accept a regular
expression except for --create option
--topics-with-overrides      if set when describing topics, only
                             show topics that have overridden
                             configs
--unavailable-partitions     if set when describing topics, only
                             show partitions whose leader is not
                             available
--under-replicated-partitions if set when describing topics, only
                             show under replicated partitions
--zookeeper <String: hosts>  REQUIRED: The connection string for
                             the zookeeper connection in the form
                             host:port. Multiple hosts can be
                             given to allow fail-over.

[student898_2@bigdataanalytics-worker-3 bin]$
```

Обязательным является host zookeeper, посмотрим как он называется, порты

```
cat /etc/hosts
```

```
./kafka-topics.sh --list --zookeeper bigdataanalytics-worker-3:2181
```

```
student898_2@bigdataanalytics-worker-3:/usr/hdp/current/kafka-broker/bin - Терминал
Файл  Правка  Вид  Терминал  Вкладки  Справка

[student898_2@bigdataanalytics-worker-3 bin]$ cat /etc/hosts
127.0.0.1 localhost
10.0.0.21 bigdataanalytics-worker-0.mcs.local bigdataanalytics-worker-0
10.0.0.29 bigdataanalytics-worker-1.mcs.local bigdataanalytics-worker-1
10.0.0.23 bigdataanalytics-worker-2.mcs.local bigdataanalytics-worker-2
10.0.0.6 bigdataanalytics-worker-3.mcs.local bigdataanalytics-worker-3
10.0.0.5 bigdataanalytics-head-0.mcs.local bigdataanalytics-head-0
[student898_2@bigdataanalytics-worker-3 bin]$ ./kafka-topics.sh --list --zookeeper bigdataanalytics-worker-3:2181
__consumer_offsets
incident_event_json
orders_json
sales_unknown
shadrin_iris
test
[student898_2@bigdataanalytics-worker-3 bin]$
```

Изменяем время хранения данных на 1 минуту

```
./kafka-topics.sh --alter --topic test --zookeeper bigdataanalytics-worker-3:2181 --config retention.ms=6000
```

```
student898_2@bigdataanalytics-worker-3:/usr/hdp/current/kafka-broker/bin - Терминал
Файл  Правка  Вид  Терминал  Вкладки  Справка

test
[student898_2@bigdataanalytics-worker-3 bin]$ ./kafka-topics.sh --alter --topic test --zookeeper bigdataanalytics-worker-3:2181 --config retention.ms=6000
WARNING: Altering topic configuration from this script has been deprecated and may be removed in future releases.
        Going forward, please use kafka-configs.sh for this functionality
Updated config for topic "test".
[student898_2@bigdataanalytics-worker-3 bin]$
```

```
./kafka-console-consumer.sh
```

```
student898_2@bigdataanalytics-worker-3:/usr/hdp/current/kafka-broker/bin - Терминал
Файл  Правка  Вид  Терминал  Вкладки  Справка

properties keyed with 'key.
deserializer.' and 'value.
deserializer.' prefixes to configure
their deserializers.
--skip-message-on-error      If there is an error when processing a
                             message, skip it instead of halt.
--timeout-ms <Integer: timeout_ms> If specified, exit if no message is
                             available for consumption for the
                             specified interval.
--topic <String: topic>      The topic id to consume on.
--value-deserializer <String:
                             deserializer for values>
--whitelist <String: whitelist> Whitelist of topics to include for
                             consumption.
[student898_2@bigdataanalytics-worker-3 bin]$
```

На этом терминале запускаем чтение сообщений из очереди.

```
./kafka-console-consumer.sh --topic test --bootstrap-server bigdataanalytics-worker-3:6667
```

Открываем терминал в другом окне

В нижнем терминале поставим consumer

В верхнем терминале запускаем producer

```
./kafka-console-producer.sh --topic test --broker-list bigdataanalytics-worker-3:6667
```

Hello world!!!

message 1

message 2

message 3

```
student898_2@bigdataanalytics-worker-3:/usr/hdp/current/kafka-broker/bin - Терминал
Файл  Правка  Вид  Терминал  Вкладки  Справка

igor@igor-MS-7808:~$ ssh -i ~/.ssh/id_rsa_student898_2 student898_2@37.139.41.176
Last login: Tue Dec 21 19:32:50 2021 from 109-252-20-121.nat.spd-mgts.ru
[student898_2@bigdataanalytics-worker-3 ~]$ cd /usr/hdp/current/kafka-broker/bin
[student898_2@bigdataanalytics-worker-3 bin]$ ./kafka-console-producer.sh --topic test --broker-list bigdataanalytics-worker-3:6667
-bash: ./kafka-console-producer.sh: Нет такого файла или каталога
[student898_2@bigdataanalytics-worker-3 bin]$ ./kafka-console-producer.sh --topic test --broker-list bigdataanalytics-worker-3:6667
>Hello world!!!
>Message 1
>Message 2
>message 3
>
```

Открывает третье окно терминала, запустим на нём ещё одного consumer

```
./kafka-console-consumer.sh --topic test --bootstrap-server bigdataanalytics-worker-3:6667
```

message 4

message 5

message 6

```
student898_2@bigdataanalytics-worker-3:/usr/hdp/current/kafka-broker/bin - Терминал
Файл  Правка  Вид  Терминал  Вкладки  Справка
igor@igor-MS-7808:~$ ssh -i ~/.ssh/id_rsa_student898_2 student898_2@37.139.41.176
Last login: Tue Dec 21 19:32:50 2021 from 109-252-20-121.nat.spd-mgts.ru
[student898_2@bigdataanalytics-worker-3 ~]$ cd /usr/hdp/current/kafka-broker/bin
[student898_2@bigdataanalytics-worker-3 bin]$ ./kafka-console-producer.sh --topic test --broker-list bigdataanalytics-worker-3:6667
bash: ./kafka-console-producer.sh: Нет такого файла или каталога
[student898_2@bigdataanalytics-worker-3 bin]$ ./kafka-console-producer.sh --topic test --broker-list bigdataanalytics-worker-3:6667
Hello world!!!
>Message 1
>Message 2
>message 3
>message 4
>message 5
>message 6
█
```

```
student898_2@bigdataanalytics-worker-3:/usr/hdp/current/kafka-broker/bin - Терминал
Файл  Правка  Вид  Терминал  Вкладки  Справка
igor@igor-MS-7808:~$ ssh -i ~/.ssh/id_rsa_student898_2 student898_2@37.139.41.176
Last login: Tue Dec 21 20:01:05 2021 from 109-252-20-121.nat.spd-mgts.ru
[student898_2@bigdataanalytics-worker-3 ~]$ cd /usr/hdp/current/kafka-broker/bin
[student898_2@bigdataanalytics-worker-3 bin]$ ./kafka-console-consumer.sh --topic test --bootstrap-server bigdataanalytics-worker-3:6667
message 4
message 5
message 6
█
```

```
student898_2@bigdataanalytics-worker-3:/usr/hdp/current/kafka-broker/bin - Терминал
Файл  Правка  Вид  Терминал  Вкладки  Справка
--topic <String: topic>          The topic id to consume on.
--value-deserializer <String:  deserializer for values>
--whitelist <String: whitelist> Whitelist of topics to include for
                                consumption.
[student898_2@bigdataanalytics-worker-3 bin]$ ./kafka-console-consumer.sh --topic test --bootstrap-server bigdataanalytics-worker-3:6667
Hello world!!!
Message 1
Message 2
message 3
message 4
message 5
message 6
█
```

`./kafka-console-consumer.sh --topic test --bootstrap-server bigdataanalytics-worker-3:6667 --from-beginning`

```
student898_2@bigdataanalytics-worker-3:/usr/hdp/current/kafka-broker/bin - Терминал
Файл  Правка  Вид  Терминал  Вкладки  Справка
igor@igor-MS-7808:~$ ssh -i ~/.ssh/id_rsa_student898_2 student898_2@37.139.41.176
Last login: Tue Dec 21 20:01:05 2021 from 109-252-20-121.nat.spd-mgts.ru
[student898_2@bigdataanalytics-worker-3 ~]$ cd /usr/hdp/current/kafka-broker/bin
[student898_2@bigdataanalytics-worker-3 bin]$ ./kafka-console-consumer.sh --topic test --bootstrap-server bigdataanalytics-worker-3:6667
message 4
message 5
message 6
^CProcessed a total of 3 messages
[student898_2@bigdataanalytics-worker-3 bin]$ ./kafka-console-consumer.sh --topic test --bootstrap-server bigdataanalytics-worker-3:6667 -
-from-beginning
```

./kafka-console-consumer.sh --topic test --from-beginning --bootstrap-server bigdataanalytics-worker-3:6667

```
student898_2@bigdataanalytics-worker-3:/usr/hdp/current/kafka-broker/bin - терминал
Файл  Правка  Вид  Терминал  Вкладки  Справка
Last login: Tue Dec 21 19:32:50 2021 from 109-252-20-121.nat.spd-mgts.ru
[student898_2@bigdataanalytics-worker-3 ~]$ cd /usr/hdp/current/kafka-broker/bin
[student898_2@bigdataanalytics-worker-3 bin]$ ./kafka-console-producer.sh --topic test --broker-list bigdataanalytics-worker-3:6667
-bash: ./kafka-console-producer.sh: Нет такого файла или каталога
[student898_2@bigdataanalytics-worker-3 bin]$ ./kafka-console-producer.sh --topic test --broker-list bigdataanalytics-worker-3:6667
>Hello world!!!
>Message 1
>Message 2
>message 3
>message 4
>message 5
>message 6
>message 7
>
```

```
student898_2@bigdataanalytics-worker-3:/usr/hdp/current/kafka-broker/bin - Терминал
Файл  Правка  Вид  Терминал  Вкладки  Справка
message 6
^CProcessed a total of 3 messages
[student898_2@bigdataanalytics-worker-3 bin]$ ./kafka-console-consumer.sh --topic test --bootstrap-server bigdataanalytics-worker-3:6667
-from-beginning
^CProcessed a total of 0 messages
[student898_2@bigdataanalytics-worker-3 bin]$ ./kafka-console-consumer.sh --topic test --bootstrap-server bigdataanalytics-worker-3:6667
-from-beginning
message 7
^CProcessed a total of 1 messages
[student898_2@bigdataanalytics-worker-3 bin]$ ./kafka-console-consumer.sh --topic test --from-beginning --bootstrap-server bigdataanalyti
s-worker-3:6667
message 7
```

```
student898_2@bigdataanalytics-worker-3:/usr/hdp/current/kafka-broker/bin - Терминал
Файл  Правка  Вид  Терминал  Вкладки  Справка
--value-deserializer <String:
  deserializer for values>
--whitelist <String: whitelist>      Whitelist of topics to include for
                                      consumption.
[student898_2@bigdataanalytics-worker-3 bin]$ ./kafka-console-consumer.sh --topic test --bootstrap-server bigdataanalytics-worker-3:6667
Hello world!!!
Message 1
Message 2
message 3
message 4
message 5
message 6
message 7
```

```
cat kafka-console-producer.sh
```

```
kafka-console-producer.sh --broker-list bigdataanalytics-worker-3:6667 --topic test << dataset.csv
```



[illegible]