

# Consumer Benchmark

1. Заходим на наш стенд, подключаться можно к любой ноде в кластере.

Не забудьте заменить плейсхолдер `<ваш номер студента>` на ваш личный номер студента:

```
ssh <ваш номер студента>@sbox.slurm.io
ssh node-{1-3}.<ваш номер студента>.slurm.io
export USER=<ваш номер студента>
```

2. Переходим в папку со скриптами Кафки:

```
cd /opt/kafka_2.13-2.7.0/bin/
```

3. Бенчмарк консюмера запускается скриптом **kafka-consumer-perf-test.sh**.

Давайте рассмотрим его основные конфиги:

```
./kafka-consumer-perf-test.sh --help
```

- **bootstrap-server** указывает адрес брокера к которому будет подключаться консюмер

- **messages** обозначает количество сообщений которые должны быть зафетчены консюмером

- **topic** указывает название топика из которого бенчмарк будет читать сообщения

- **print-metrics** выводит детальные метрики в конце работы бенчмарка

- **threads** и **num-fetch-threads** ничего не делают (!), хотя их и можно найти во многих статьях про бенчмаркинг Кафки (см. [KAFKA-10126](#))

4. Перед тем как производить тест консюмера нам нужно создать топик...

```
./kafka-topics.sh --bootstrap-server node-1.$USER:9092 --topic consumer-benchmark
--replication-factor 3 --partitions 1 --create
```

5. ...а так же записать в него набор сообщений, для этого воспользуемся

бенчмарком продюсера:

```
./kafka-producer-perf-test.sh --producer-props  
bootstrap.servers=node-1.$USER:9092 --topic consumer-benchmark --throughput -1  
--num-records 1000000 --record-size 100
```

6. Теперь давайте запустим бенчмарк консюмера:

```
./kafka-consumer-perf-test.sh --bootstrap-server node-1.$USER:9092 --topic  
consumer-benchmark --messages 1000000 --print-metrics  
start.time, end.time, data.consumed.in.MB, MB.sec, data.consumed.in.nMsg,  
nMsg.sec, rebalance.time.ms, fetch.time.ms, fetch.MB.sec, fetch.nMsg.sec  
2021-03-13 14:23:19:139, 2021-03-13 14:23:25:956, 95.3783, 13.9912, 1000114,  
146708.8162, 1615645400267, -1615645393450, -0.0000, -0.0006
```

Metric Name

Value

```
consumer-coordinator-metrics:assigned-partitions:{client-id=consumer-perf-consume  
r-15509-1} : 0.000
```

```
consumer-coordinator-metrics:commit-latency-avg:{client-id=consumer-perf-consumer  
-15509-1} : 31.500
```

Как только тест завершится вы увидите набор метрик в терминале, основные из которых:

- **start.time, end.time** — время начала и конца теста

- **data.consumed.in.MB** — общий размер всех считанных сообщений

- **MB.sec** — пропускная способность консюмера в МБ/сек (объем считанных данных за секунду)

- **data.consumed.in.nMsg** — общее количество считанных сообщений

- **nMsg.sec** — пропускная способность консюмера в сообщениях/сек

В приведенном примере мы видим, что консюмер вычитал  $\pm 95$  МБ данных с пропускной способностью  $\pm 14$  МБ/сек.

7. Попробуйте создать топик с несколькими партициями и сравнить результат:

```
./kafka-topics.sh --bootstrap-server node-1.$USER:9092 --topic  
consumer-benchmark-multipartition --replication-factor 3 --partitions 3 --create
```

```
./kafka-producer-perf-test.sh --producer-props  
bootstrap.servers=node-1.$USER:9092 --topic consumer-benchmark-multipartition  
--throughput -1 --num-records 1000000 --record-size 100
```

```
./kafka-consumer-perf-test.sh --bootstrap-server node-1.$USER:9092 --topic  
consumer-benchmark-multipartition --messages 1000000
```

```
start.time, end.time, data.consumed.in.MB, MB.sec, data.consumed.in.nMsg,  
nMsg.sec, rebalance.time.ms, fetch.time.ms, fetch.MB.sec, fetch.nMsg.sec  
2021-03-13 14:41:01:504, 2021-03-13 14:41:05:956, 95.3796, 21.4240, 1000128,  
224646.9003, 1615646461976, -1615646457524, -0.0000, -0.0006
```