1. **Создать Flume поток используя Flume сервис**

Сперва в hbase создадим таблицу `Student4_10` с одним семейством колонок `Message`.

Запустим HBASE SHELL

[student4_10@manager ~]$ hbase shell

Создадим таблицу

hbase(main):001:0> create 'Student4_10', 'Message'

0 row(s) in 8.6700 seconds

Посмотрим таблицу через DESCRIBE

hbase(main):001:0> describe 'Student4_10'

Table Student4_10 is ENABLED

Student4_10

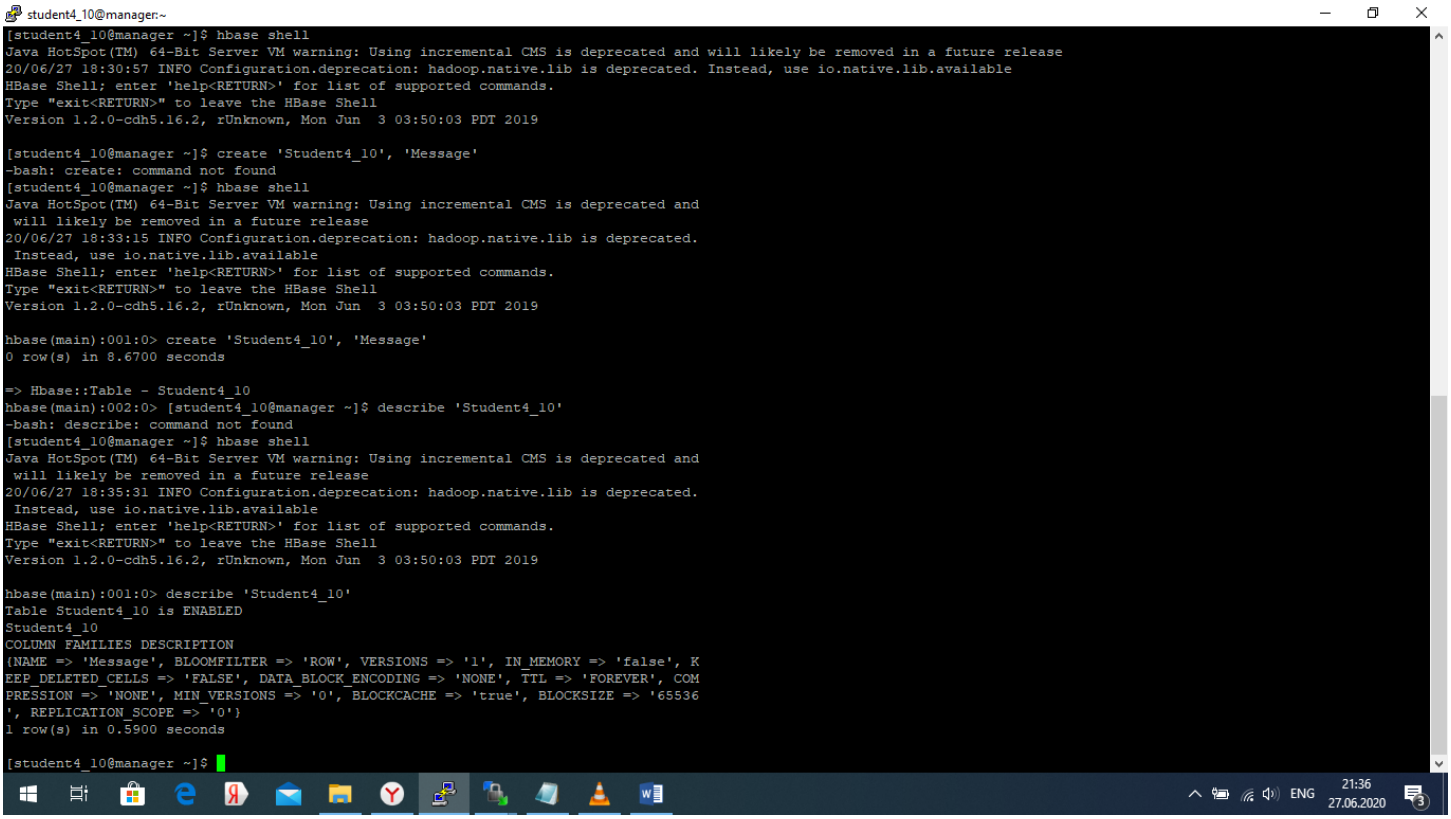
COLUMN FAMILIES DESCRIPTION

{NAME => 'Message', BLOOMFILTER => 'ROW', VERSIONS => '1', IN_MEMORY => 'false', K
EEP_DELETED_CELLS => 'FALSE', DATA_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER',
COM

PRESSION => 'NONE', MIN_VERSIONS => '0', BLOCKCACHE => 'true', BLOCKSIZE => '65536

', REPLICATION_SCOPE => '0'}

1 row(s) in 0.5900 seconds



Для Source path берём файл /var/log/cron

[student4_10@manager ~]$ ls /var/log/cron

/var/log/cron

Создадим папку для sink в локальной папке

[student4_10@manager ~]$ hdfs dfs -mkdir /home/student4_10/hbase

[student4_10@manager ~]$ hdfs dfs -ls /home/student4_10/

Found 3 items

drwxrwxrwx - flume supergroup 0 2020-06-27 18:03 /home/student4_10/20-06-27

drwxr-xr-x - flume supergroup 0 2020-06-27 18:03 /home/student4_10/flume

drwxr-xr-x - flume supergroup 0 2020-06-27 18:57 /home/student4_10/hbase

Код для Config File

```
# Naming the components on the current agent
Flume4_10.sources = ExecSource
Flume4_10.channels = MemChannel
Flume4_10.sinks = HdfsSink HbaseSink


# Describing/Configuring the source
Flume4_10.sources.ExecSource.type = exec
Flume4_10.sources.ExecSource.command = /bin/tailf /var/log/cron
#Flume4_10.sources.ExecSource.command = /bin/tailf /tmp/myfile
#Flume4_10.sources.ExecSource.command = /bin/tailf /tmp/source4_10
Flume4_10.sources.ExecSource.interceptors = TimestampInterceptor
Flume4_10.sources.ExecSource.interceptors.TimestampInterceptor.type = timestamp

#SINKS

# Describing/Configuring the HDFS sink
Flume4_10.sinks.HdfsSink.type = hdfs
#Flume4_10.sinks.HdfsSink.hdfs.path= /flume/Flume4_10/%y-%m-%d/
Flume4_10.sinks.HdfsSink.hdfs.path= /home/student4_10/hbase/%y-%m-%d/
Flume4_10.sinks.HdfsSink.hdfs.filePrefix = events

# Describing/Configuring the HBASE sink
Flume4_10.sinks.HbaseSink.type        = hbase
Flume4_10.sinks.HbaseSink.table = Student4_10
Flume4_10.sinks.HbaseSink.columnFamily = Message
```

Describing/Configuring the channel

Flume4_10.channels.MemChannel.type = memory

Flume4_10.channels.MemChannel.capacity = 10000

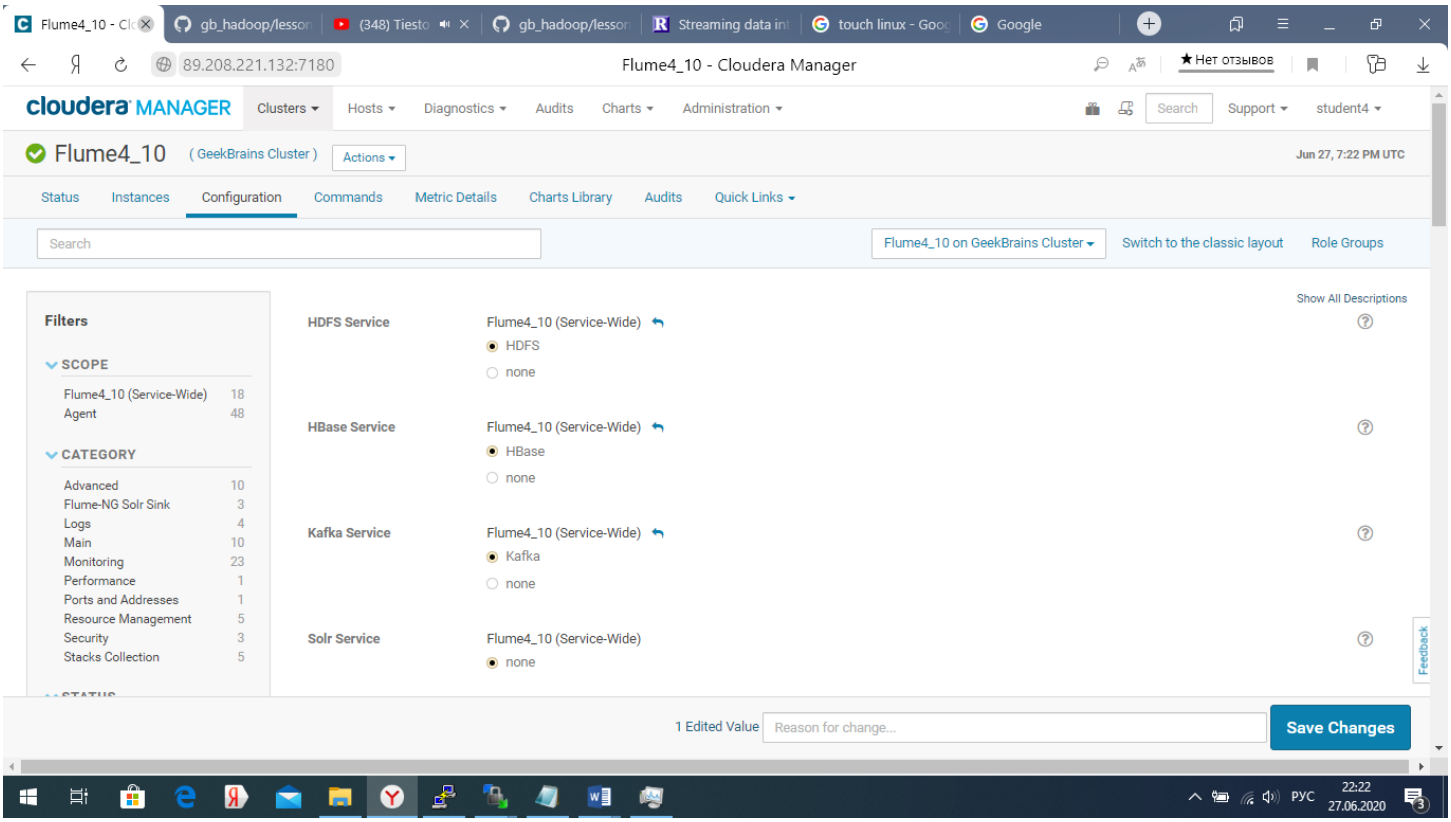
Flume4_10.channels.MemChannel.transactionCapacity = 10

Bind the source and sink to the channel

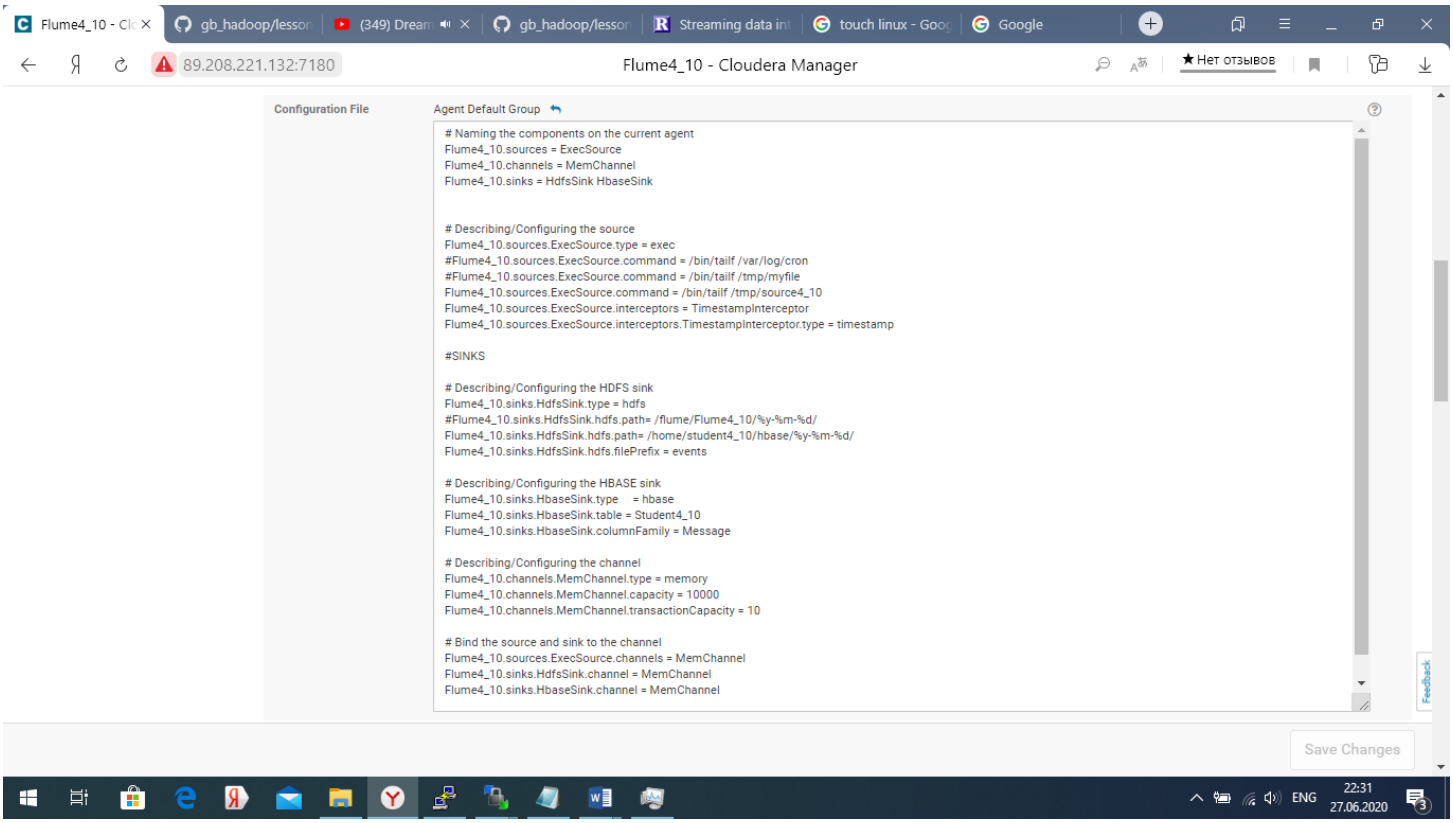
Flume4_10.sources.ExecSource.channels = MemChannel

Flume4_10.sinks.HdfsSink.channel = MemChannel

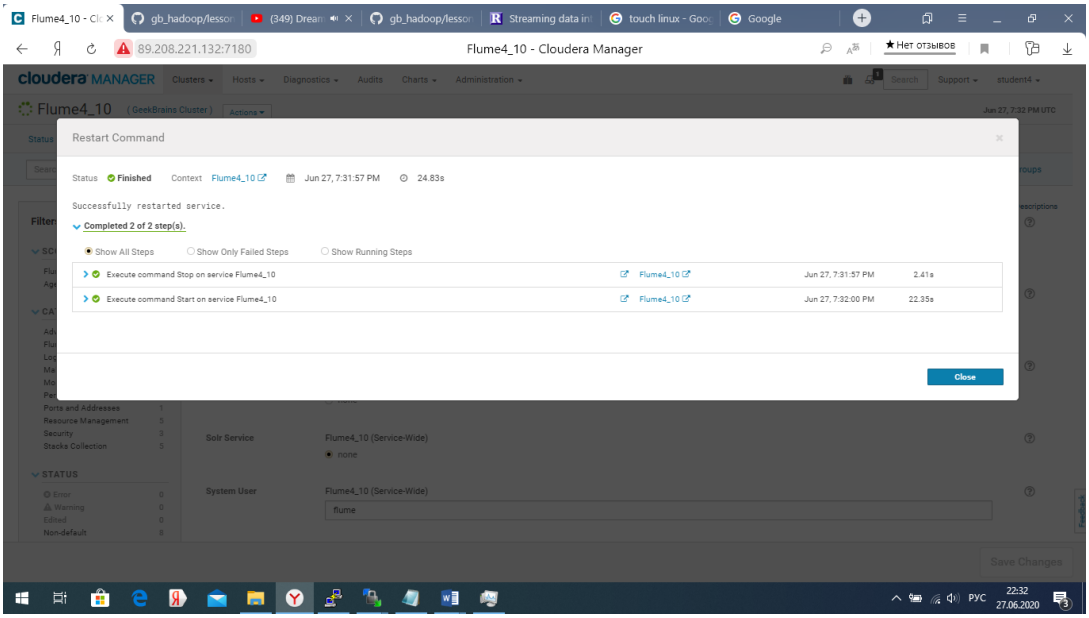
Flume4_10.sinks.HbaseSink.channel = MemChannel



Копируем и вставляем код

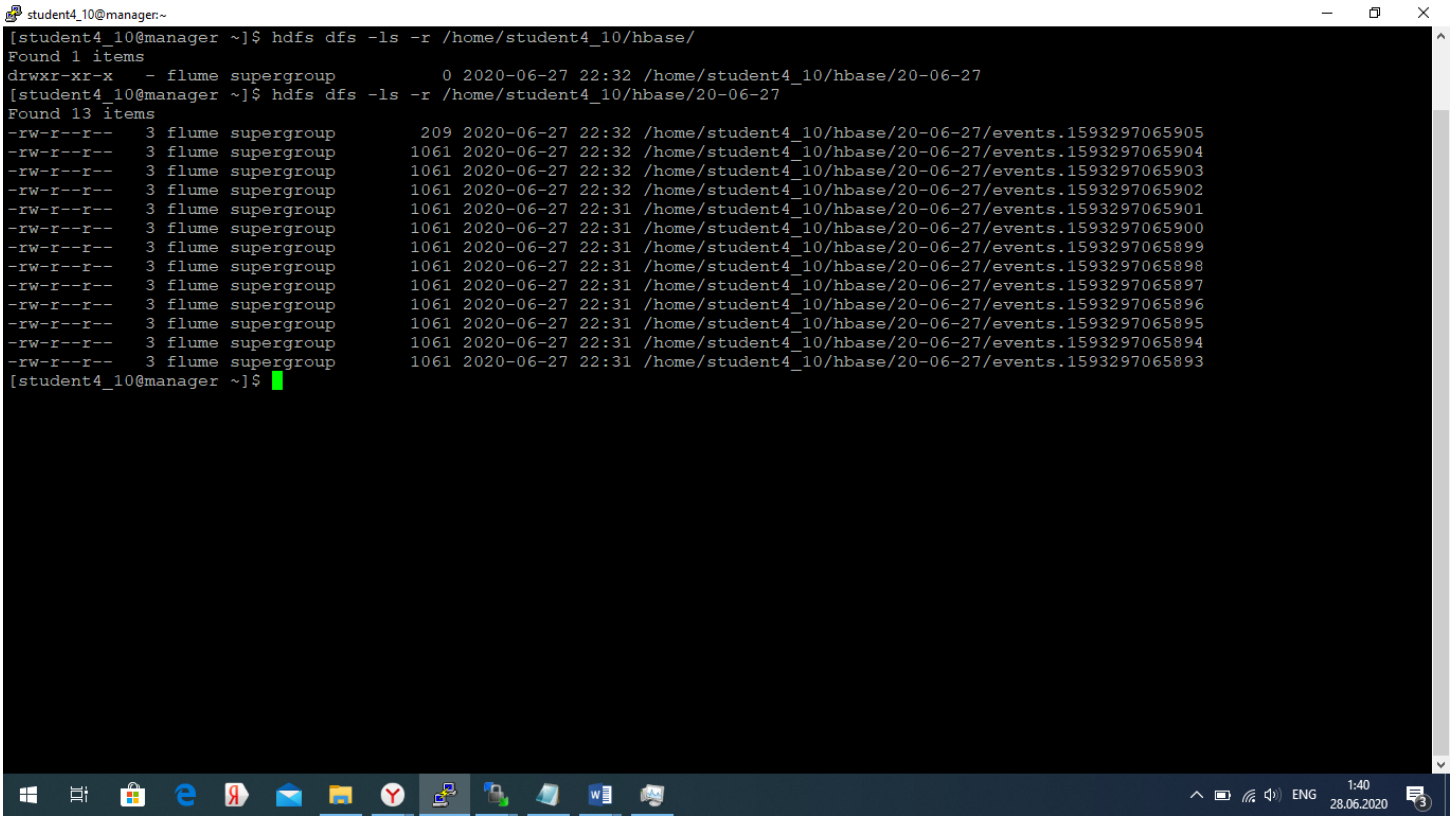


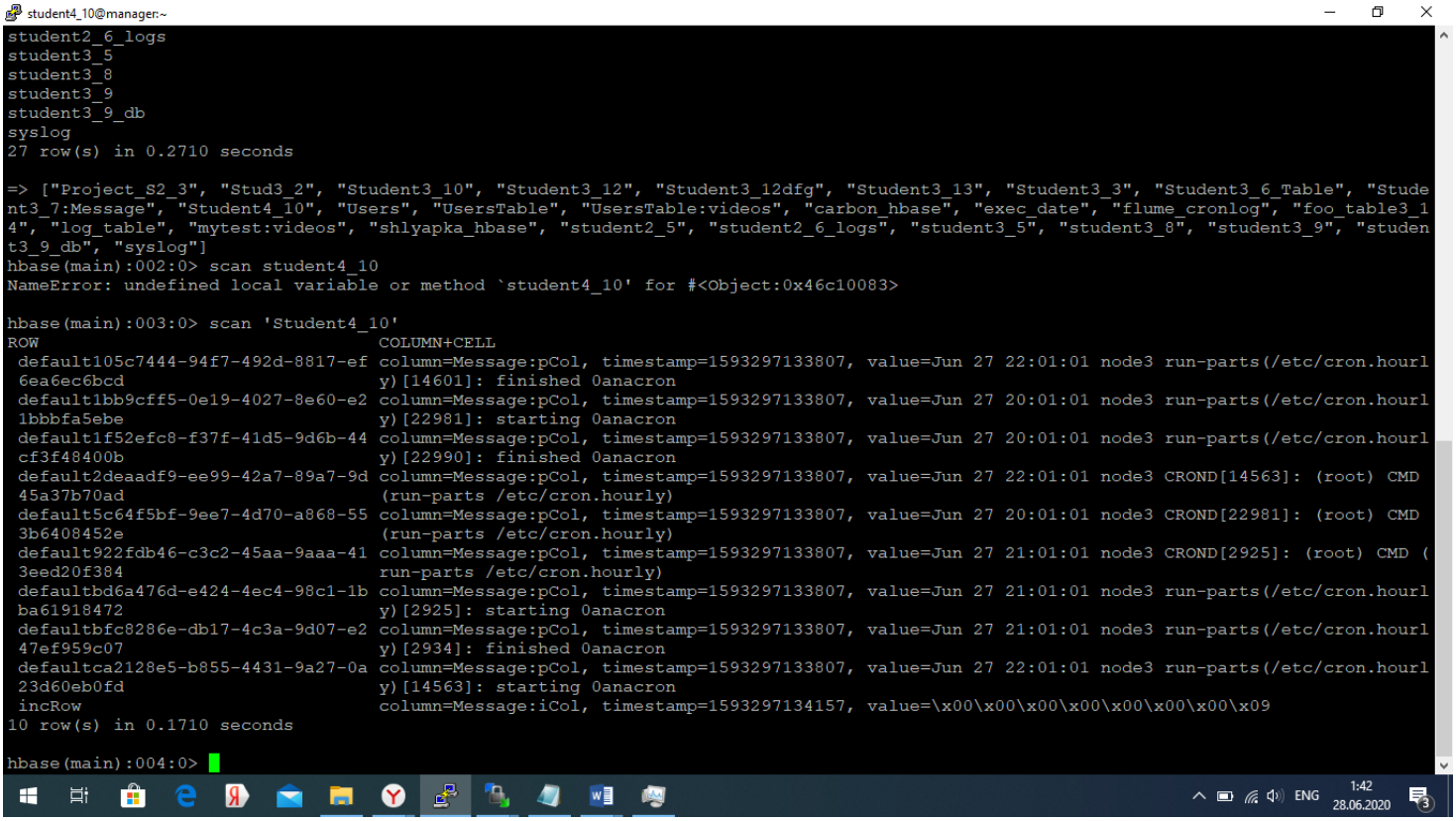
Сохраним и запустим



Загрузим данные

Данные загруженные в HDFS





Данные загружены в HBASE