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 ~]\$ Is /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

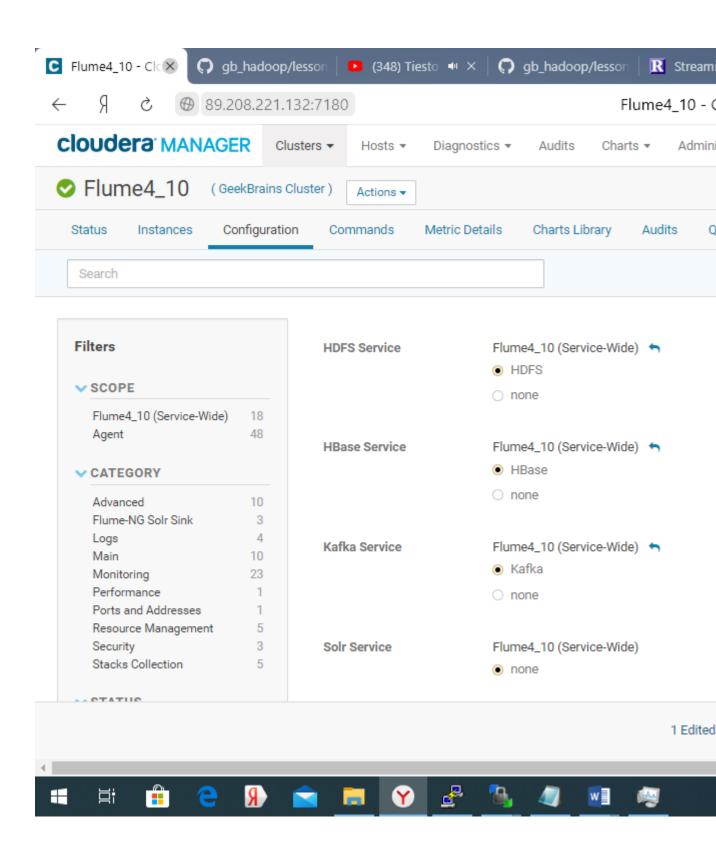
Flume4 10.channels.MemChannel.transactionCapacity = 10

Bind the source and sink to the channel

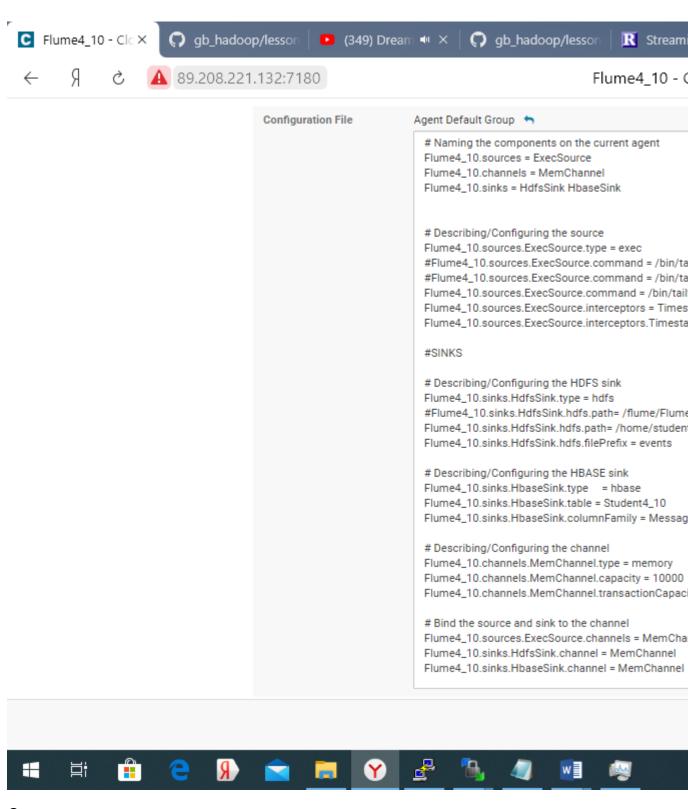
Flume4 10.sources.ExecSource.channels = MemChannel

Flume4 10.sinks.HdfsSink.channel = MemChannel

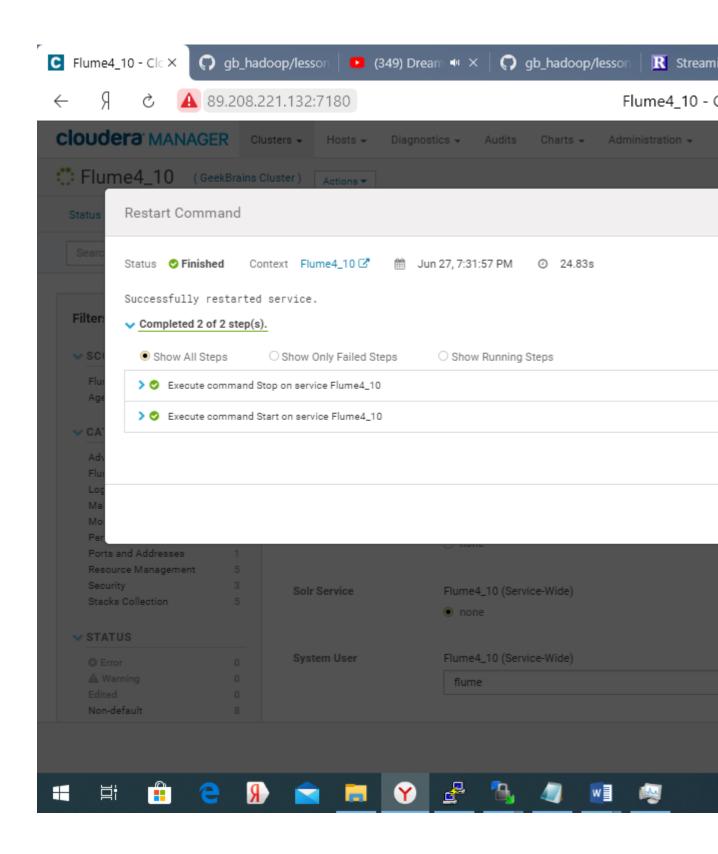
Flume4_10.sinks.HbaseSink.channel = MemChannel



Копируем и вставляем код



Сохраним и запустим



Загрузим данные

Данные загруженные в HDFS

```
[student4 10@manager ~]$ hdfs dfs -ls -r /home/student4 10/hbase/
Found 1 items
drwxr-xr-x - flume supergroup
                                                         0 2020-06-27 22:32 /home/st
[student4_10@manager ~]$ hdfs dfs -ls -r /home/student4_10/hbase/20-
Found 13 items
                3 flume supergroup
                                                      209 2020-06-27 22:32 /home/st
-rw-r--r--
-rw-r--r-- 3 flume supergroup
-rw-r--r-- 3 flume supergroup
                                                     1061 2020-06-27 22:32 /home/st
                                                    1061 2020-06-27 22:32 /home/st
-rw-r--r-- 3 flume supergroup
                                                    1061 2020-06-27 22:32 /home/st
                                                    1061 2020-06-27 22:31 /home/st
                                                    1061 2020-06-27 22:31 /home/st
                                                    1061 2020-06-27 22:31 /home/st
                                                    1061 2020-06-27 22:31 /home/st
-rw-r--r-- 3 flume supergroup
                                                    1061 2020-06-27 22:31 /home/st
                                                    1061 2020-06-27 22:31 /home/st
                                                    1061 2020-06-27 22:31 /home/st
                                                1061 2020-06-27 22:31 /home/st
1061 2020-06-27 22:31 /home/st
[student4 10@manager ~]$
```



























```
student4_10@manager: ~

student2 6 logs
student3 5
student3 8
student3 9
student3 9 db
syslog
27 row(s) in 0.2710 seconds
=> ["Project_S2_3", "Stud3_2", "Student3_10", "Student3_12", "Student3_7:Message", "Student4_10", "Users", "UsersTable", "UsersTable:vi
4", "log table", "mytest:videos", "shlyapka hbase", "student2 5", "s
t3_9_db", "syslog"]
hbase(main):002:0> scan student4 10
NameError: undefined local variable or method `student4 10' for #<Ok
hbase(main):003:0> scan 'Student4 10'
ROW
                                      COLUMN+CELL
 default105c7444-94f7-492d-8817-ef column=Message:pCol, timestamp=15
                                      y) [14601]: finished Oanacron
 6ea6ec6bcd
 default1bb9cff5-0e19-4027-8e60-e2 column=Message:pCol, timestamp=15
                                      y) [22981]: starting Oanacron
 default1f52efc8-f37f-41d5-9d6b-44 column=Message:pCol, timestamp=15
                                      y) [22990]: finished Oanacron
 default2deaadf9-ee99-42a7-89a7-9d column=Message:pCol, timestamp=15
 45a37b70ad
                                      (run-parts /etc/cron.hourly)
 default5c64f5bf-9ee7-4d70-a868-55 column=Message:pCol, timestamp=15
 3b6408452e
                                      (run-parts /etc/cron.hourly)
 default922fdb46-c3c2-45aa-9aaa-41 column=Message:pCol, timestamp=15
                                      run-parts /etc/cron.hourly)
 3eed20f384
 defaultbd6a476d-e424-4ec4-98c1-1b column=Message:pCol, timestamp=15
                                      y) [2925]: starting Oanacron
 ba61918472
 defaultbfc8286e-db17-4c3a-9d07-e2 column=Message:pCol, timestamp=15
                                      y) [2934]: finished Oanacron
 defaultca2128e5-b855-4431-9a27-0a column=Message:pCol, timestamp=15
 23d60eb0fd
                                      y) [14563]: starting Oanacron
                                      column=Message:iCol, timestamp=15
 incRow
10 row(s) in 0.1710 seconds
hbase (main):004:0>
                                      Y
 \blacksquare
      計
                           w
```

Данные загружены в HBASE