## Для части по SQOOP

Провести импорт таблицы из вашего сервера БД в Hadoop с использованием SQOOP в любых двух вариантах из перечисленных ниже.

* 1. в Hive-таблицу (--hive-import)
  2. в HDFS в формате avro (--as-avrodatafile)
  3. в HDFS в формате sequencefile (--as-sequencefile)

Если у вас нет своего сервера то можно использовать тот Postgres, который я показал на лекции. Пароль expoter\_pass

## Для части по потоковой обработке (Flume)

1. Создать Flume-агент с именем, соответствующим имени своего пользователя (например Flume4\_20)
2. Создать любой Flume поток используя Flume сервис соответствующего номера.
   * Тип источника источник – exeс
   * Тип канала – memory
   * Тип слива – hdfs
3. Убедиться что данные поступают в слив.
4. Создать поверх данных в hdfs таблицу через которую можно просмотреть полученные данные.
5. [Продвинутый вариант] Сделать то-же самое используя несколько сливов в разные места, например в HDFS и в HIve одновременно
6. [Продвинутый вариант] Повторить стандартный пример с выборкой сообщений из Twitter.