```
API → PYTHON → FLUME → KAFKA
                          ↓
HDFS ← NIFI
  ↓
HIVE → SPARK
HUE ← HIVE
```

Cadastro SPTRANS

Criar login no site da sptrans

http://www.sptrans.com.br/desenvolvedores

Python

```
sudo yum install python36

alias python='/usr/bin/python3.6'

sudo yum install python36-setuptools

sudo easy_install-3.6 pip

sudo pip install requests
```

Pre-Req

sudo su -

mkdir /etc/yum.repos.d/bkp

mv /etc/yum.repos.d/cloudera-* /etc/yum.repos.d/bkp

yum clean all

yum check

yum erase apf

yum update ca-certificates

yum upgrade

Python

sudo yum install centos-release-scl

sudo yum install rh-python36

scl enable rh-python36 bash

sudo easy_install-3.6 pip

sudo pip install requests

Diretórios

cd /home/cloudera
mkdir sptrans

cd sptrans

Programa Python

sptrans.py
No github

shell Python

vi sptrans.sh

/opt/rh/rh-python36/root/usr/bin/python sptrans.py

Kafka

cd /home/cloudera/kafka_2.10-0.10.2.1

nohup bin/kafka-server-start.sh config/server.properties &

bin/kafka-topics.sh --create --zookeeper localhost:2181 --replication-factor 1
--partitions 1 --topic sptrans

bin/kafka-topics.sh --list --zookeeper localhost:2181

bin/kafka-console-consumer.sh --bootstrap-server localhost:9092 --topic sptrans

Flume

sptrans.conf
cd apache-flume-1.7.0-bin
conf/
vi sptrans.conf
cd ..

bin/flume-ng agent --conf ./conf --conf-file /home/cloudera/sptrans.conf --name
sptrans -Dflume.root.logger=INFO,console

Hive

```
create external table aula.sptras (
        hora string,
        linha string,
        origem string,
        destino string,
        x string,
        y string)
row format delimited fields terminated by ';'
LOCATION '/aula/sptrans';
```

Hive

```
create external table aula.sptras (
        hora string,
        linha string,
        origem string,
        destino string,
        x string,
        y string)
row format delimited fields terminated by ';'
LOCATION '/aula/sptrans';
```

Análise Spark

/opt/rh/rh-python36/root/usr/bin/python sptrans.py

```
## FLUME  KAFKA
sptrans.sources = source1
sptrans.channels = channel1
sptrans.sinks = sink1

#source para executar a shell e dar o get na api
sptrans.sources.source1.type = exec
sptrans.sources.source1.command = sh /home/cloudera/sptrans.sh
sptrans.sources.source1.channels = channel1
sptrans.sources.source1.restartThrottle = 60000
sptrans.sources.source1.restart =true

#canais em memoria
sptrans.channels.channel1.type = memory
sptrans.channels.channel1.capacity = 100000
sptrans.channels.channel1.transactionCapacity = 100000

#sink para enviar os dados para o Kafka
sptrans.sinks.sink1.channel = channel1
sptrans.sinks.sink1.type = org.apache.flume.sink.kafka.KafkaSink
sptrans.sinks.sink1.kafka.topic = sptrans
sptrans.sinks.sink1.kafka.bootstrap.servers = localhost:9092
sptrans.sinks.sink1.kafka.flumeBatchSize = 200
sptrans.sinks.sink1.kafka.producer.acks = 1
sptrans.sinks.sink1.kafka.producer.linger.ms = 1
```