```mermaid
graph LR
    API --> PYTHON --> NIFI --> KAFKA
    KAFKA --> FLUME
    FLUME --> HDFS
    HDFS --> HIVE
    HIVE --> HUE
    HIVE --> SPARK
```

API	PYTHON	NIFI	KAFKA

HDFS ← FLUME

HUE ← HIVE → SPARK

Cadastro SPTRANS

Criar login no site da sptrans

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

Python

sudo yum install python34

alias python='/usr/bin/python3.4'

sudo yum install python34-setuptools

sudo easy_install-3.4 pip

sudo pip install requests

Diretórios

cd /home/cloudera
mkdir sptrans

cd sptrans

Programa Python

sptrans.py
No github

shell Python

vi sptrans.sh

/usr/bin/python3.4 /home/cloudera/sptrans/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 ./conf/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

Agrupar linha e as maiores longitudes (X)

Agrupar linha e as menores longitudes (X)

Cruzar maiores e menores

Verificar se existe maior e menor igual (ônibus parado)

Diferença das longitudes