

Big Data - Spark, Hadoop e afins.

Leonardo Gloria

Aula: Streaming



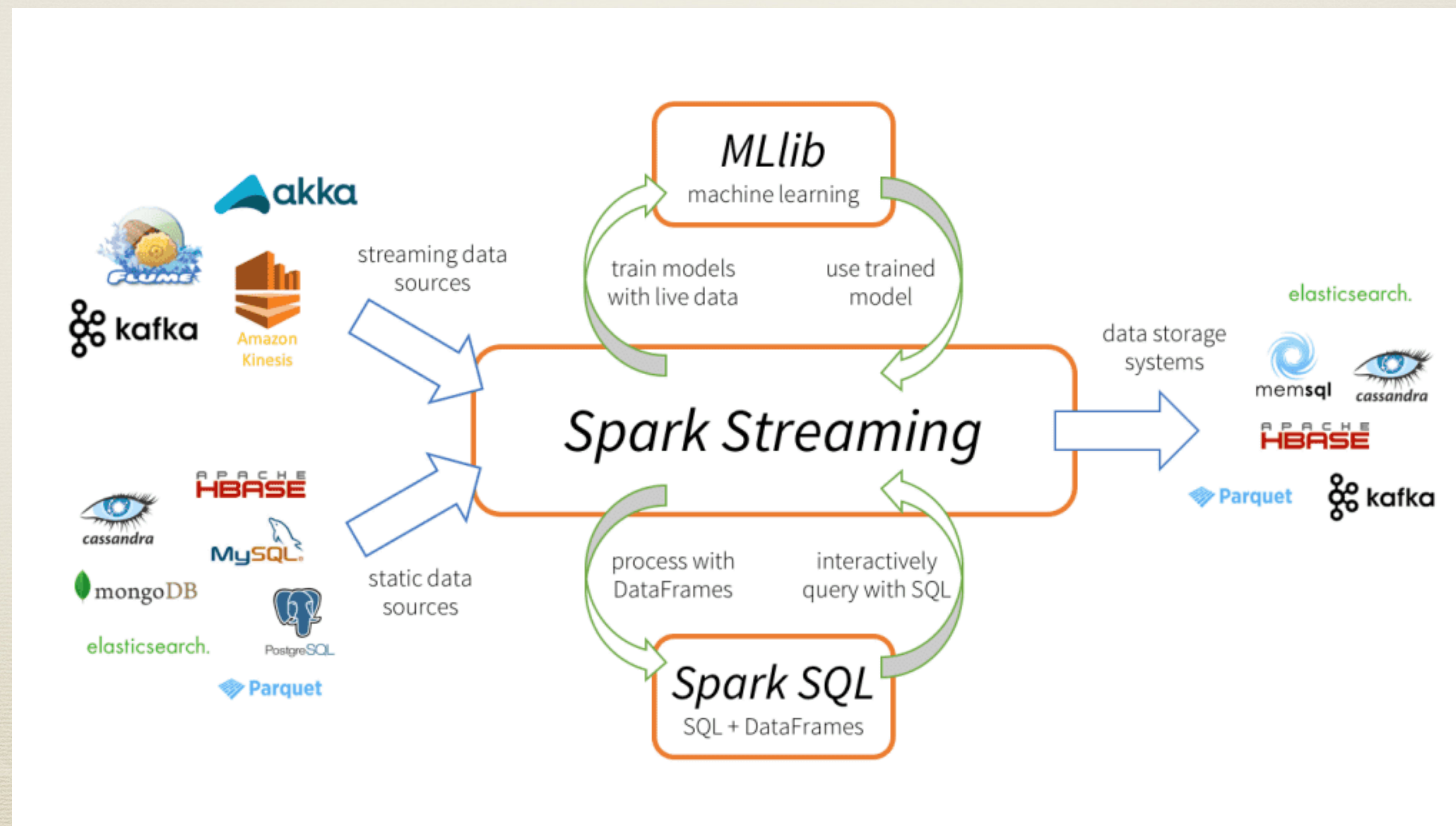
Instituto Infnet

Batch Vs Realtime

Spark Streaming

- * Extensão da API principal do Spark
- * Suporte a processamento de dados em tempo real.
- * Alto desempenho, escalável

Spark Streaming



Spark Streaming

- * Oferece suporte para batch e streaming.
- * Dados em RT podem chegar de diversas fontes: Kafka, Flume, Kinesis, Twitter ou TCP.
- * Dados Estáticos podem chegar de pontos como MongoDB, postgresSql etc.

ETL X ELT

Extract
Transform
Load

Extract
Load
Transform

Sockets

Rest Api (HTTP)

Client



Request



Response

Request



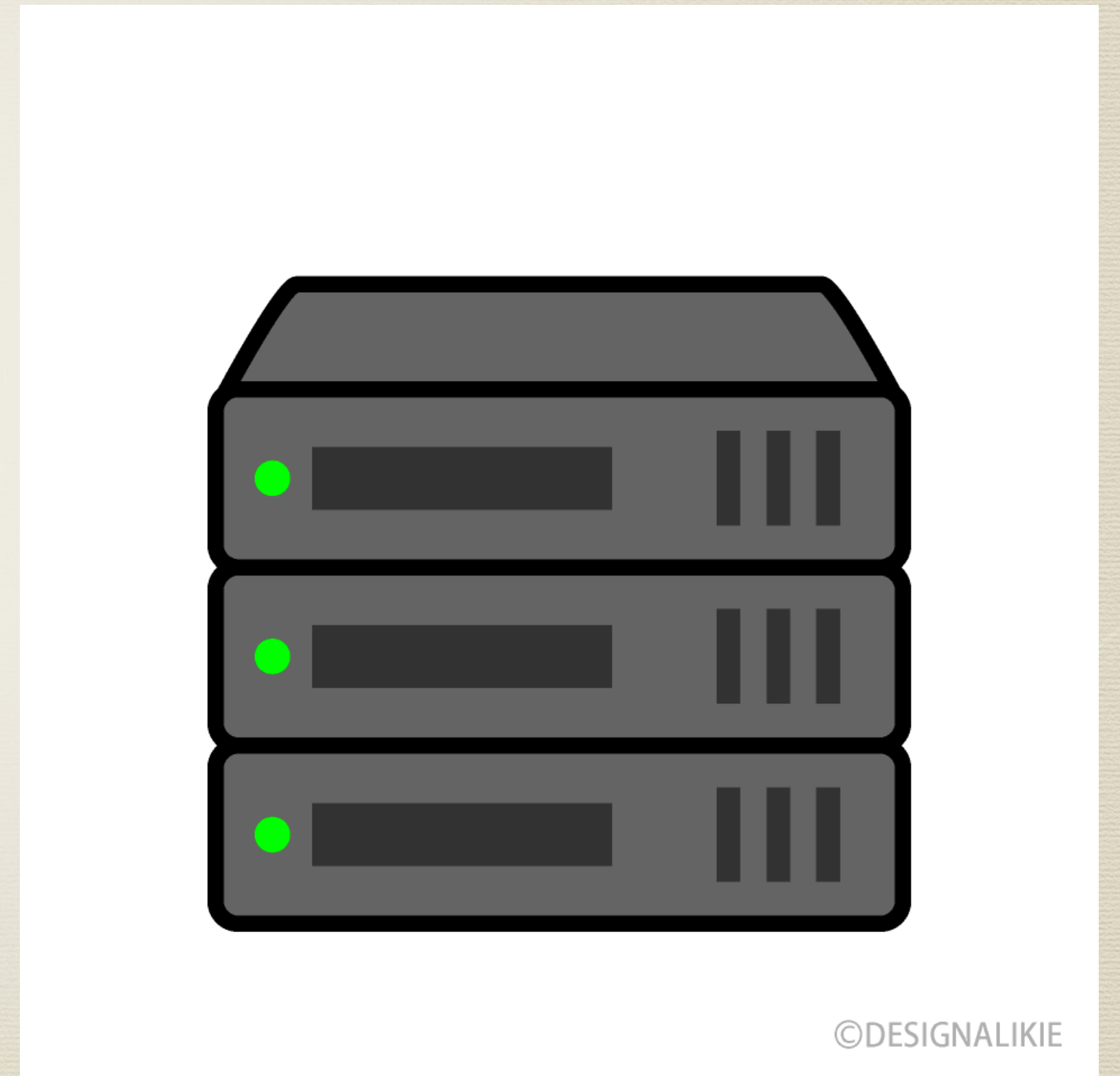
Response

Request



Response

Server



Streaming API

Client



Request



Event



Event



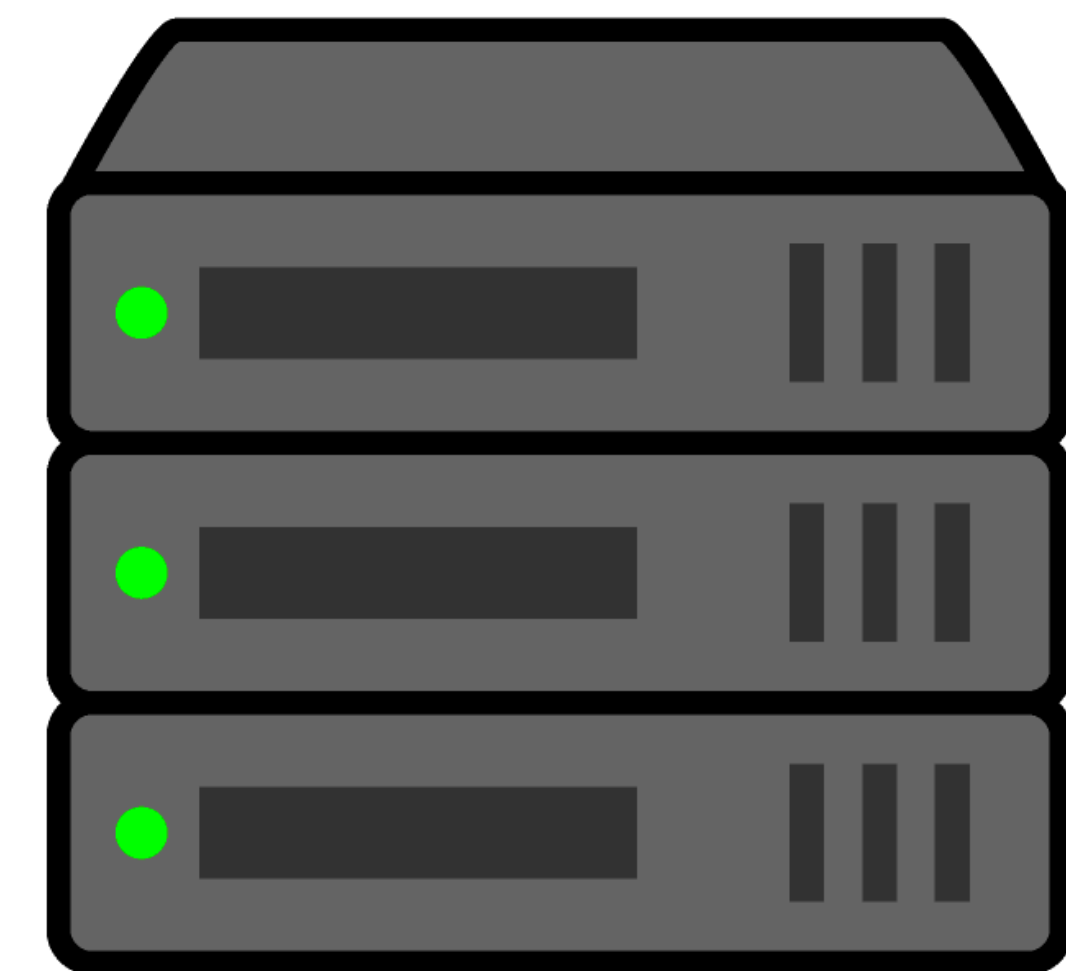
Event



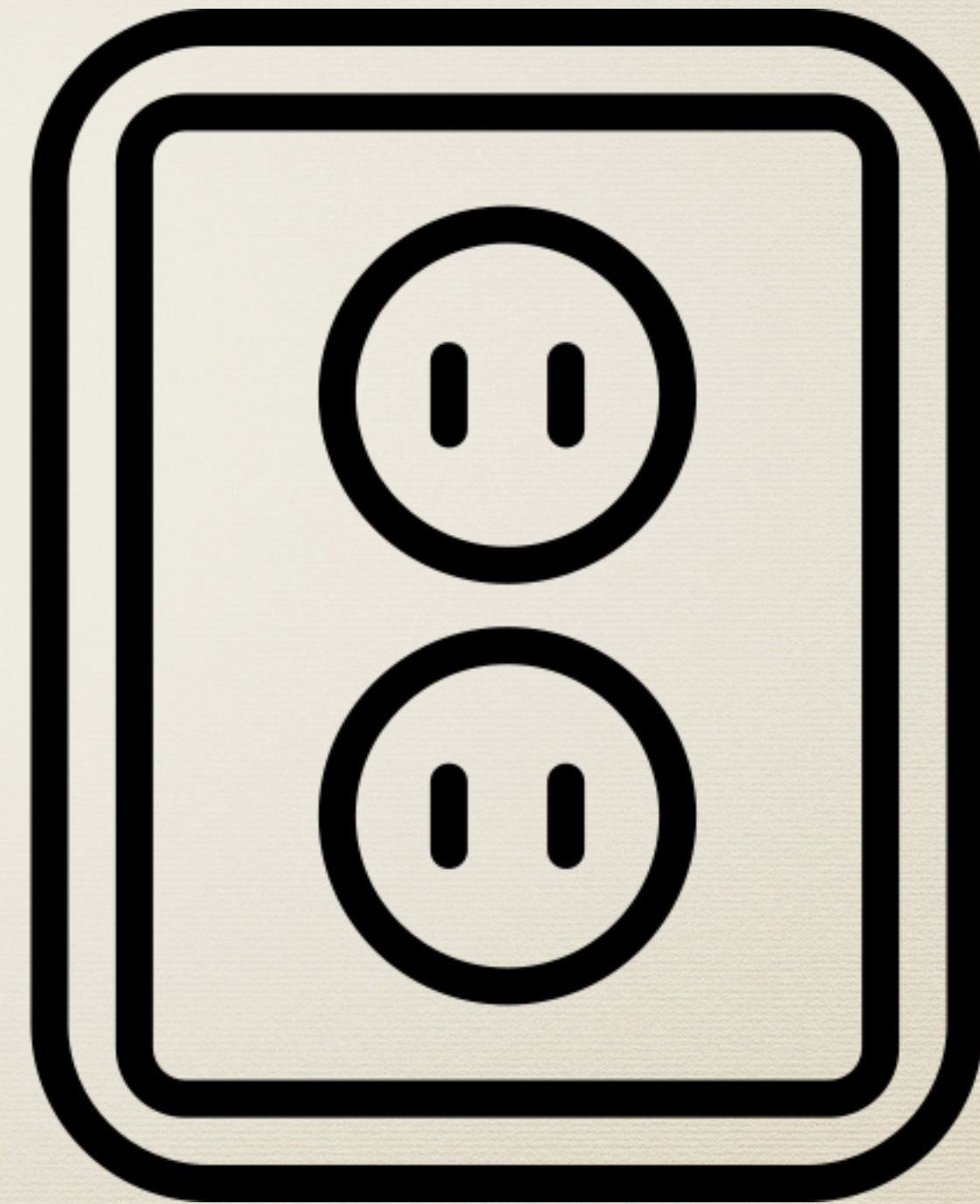
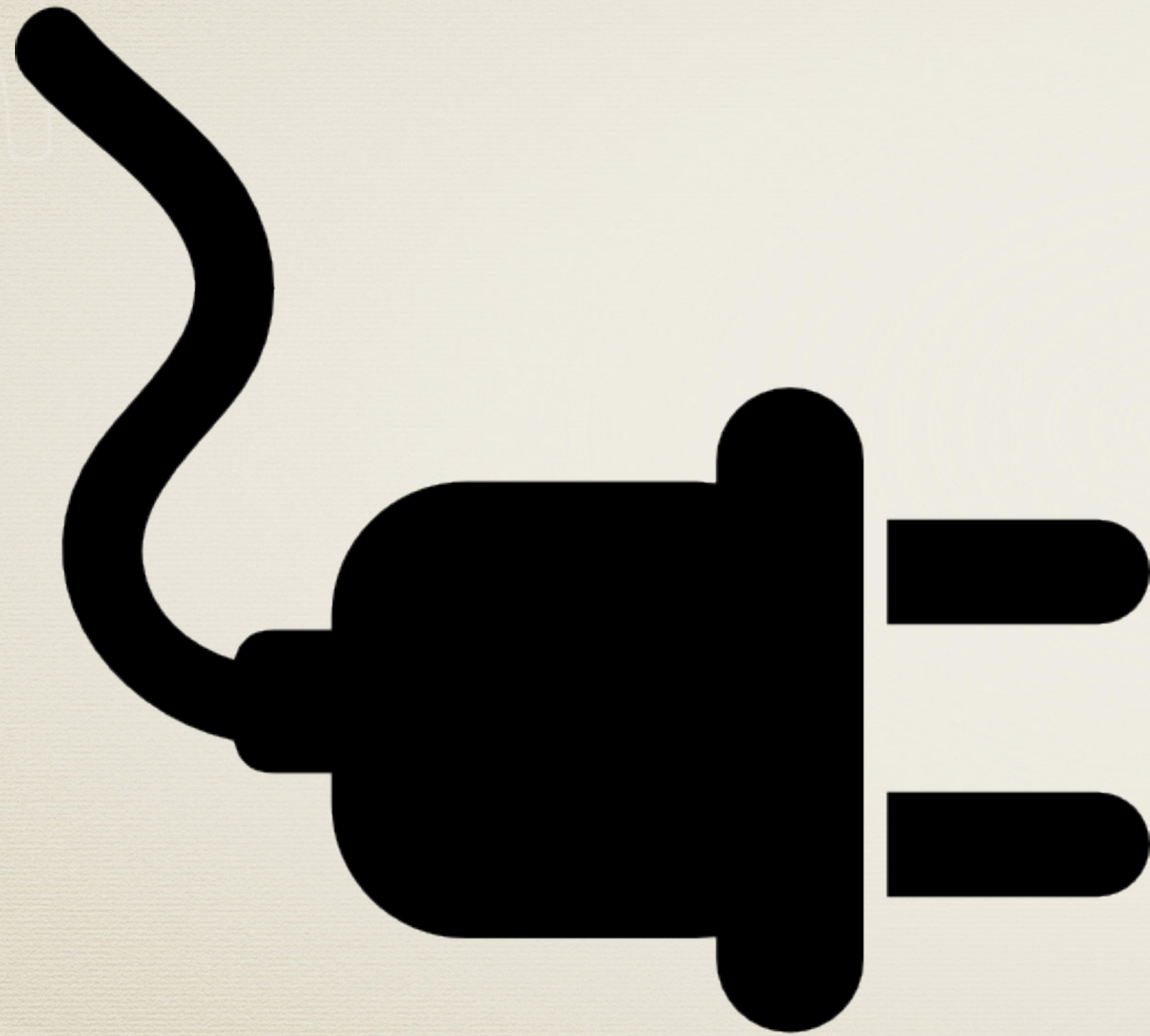
Event

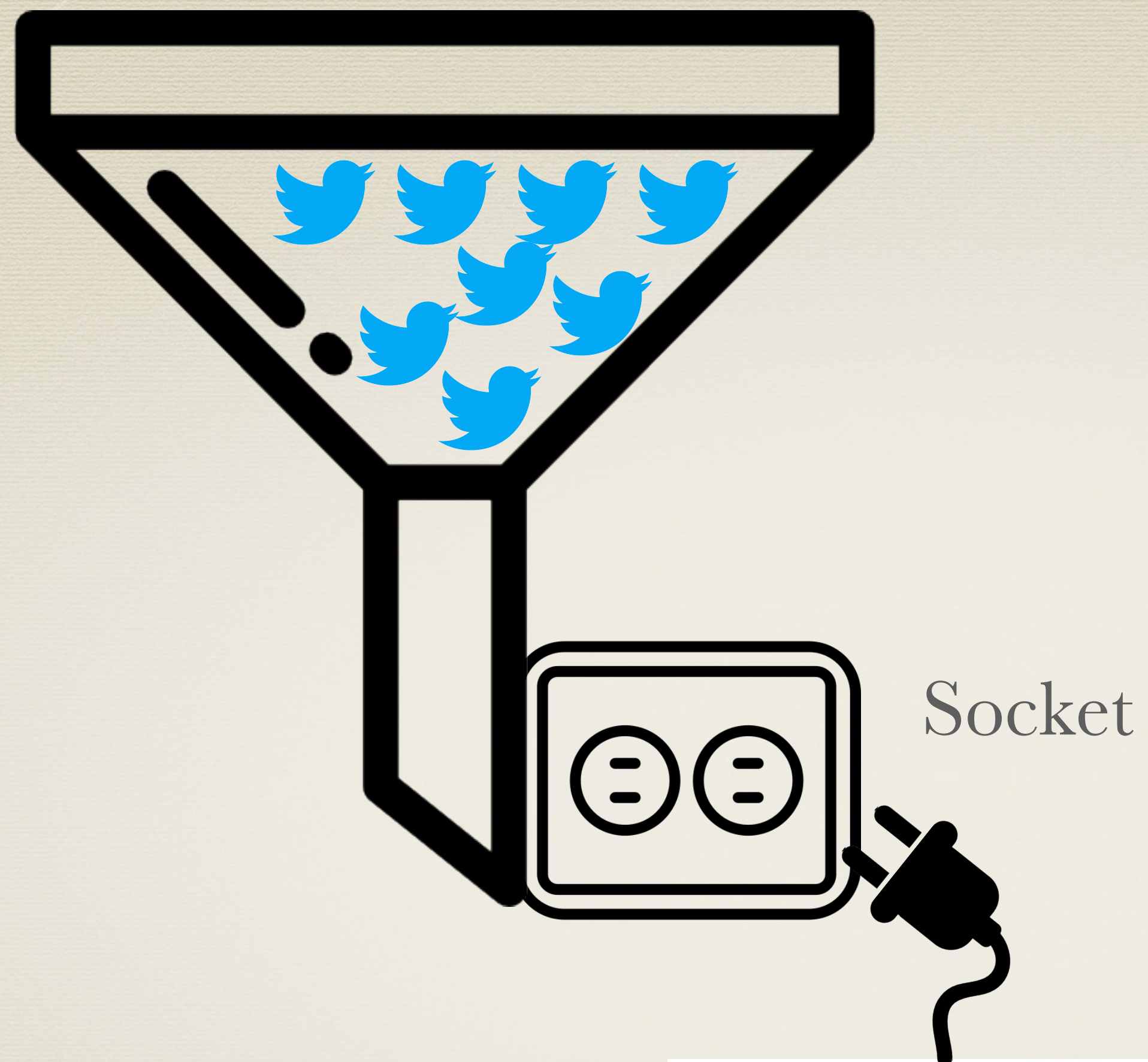


Server



Sockets





Spark
Streaming

Dúvidas?

“Stay hungry, stay foolish.”

—*Steve Jobs*