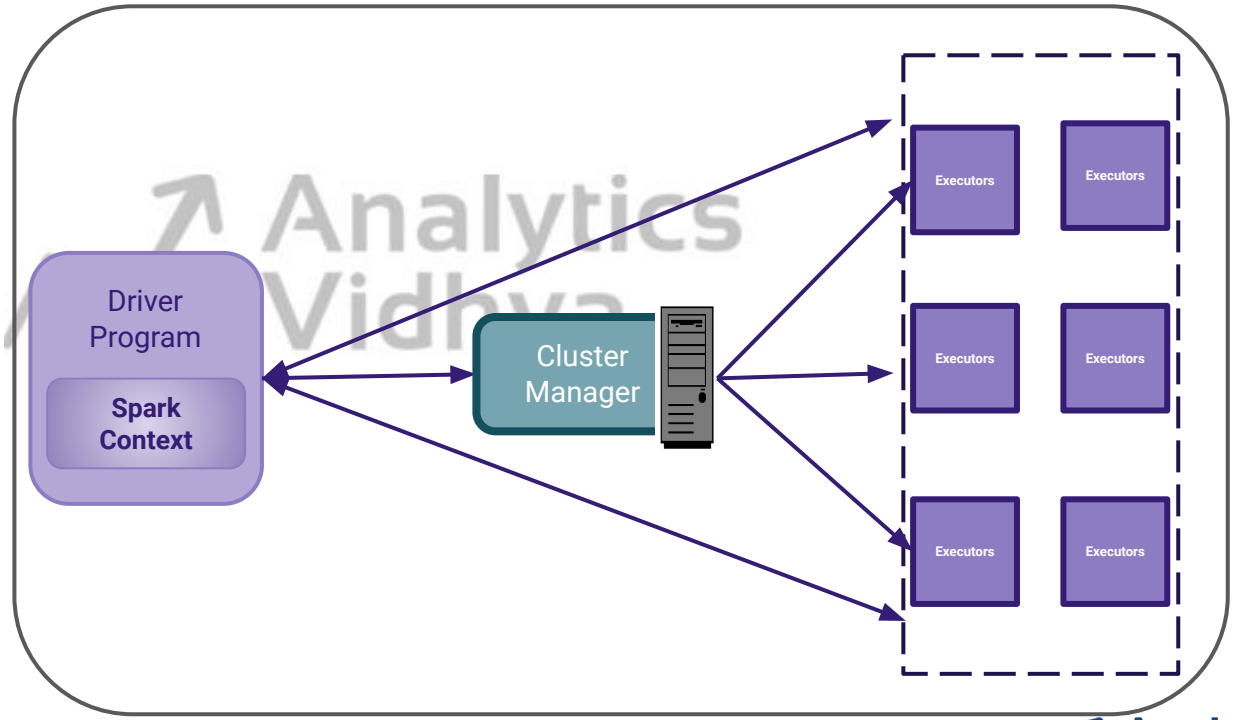
Streaming Context

Spark Context

- Submit jobs
- Ask for resources
- Schedule Tasks
- Send Tasks
- Get Job results



Streaming Context

- StreamingContext is the main entry point for real-time applications
- Connection of the driver program with the Spark engine
- Read data from various real-time sources and convert it to DStream
- Also handles scheduling, transparent to the user

Streaming Context

- StreamingContext is the main entry point for real-time applications
- Connection of the driver program with the Spark engine
- Read data from various real-time sources and convert it to DStream
- Also handles scheduling, transparent to the user

StreamingContext States

1. INITIALIZED
2. ACTIVE
3. STOPPED

Streaming Context

To execute a Spark Streaming application, we need to define the StreamingContext

```
from pyspark import SparkContext
from pyspark.streaming import StreamingContext
sc = SparkContext(master, appName)
ssc = StreamingContext(sc, 1)
```

Streaming Context

To execute a Spark Streaming application, we need to define the StreamingContext

```
from pyspark import SparkContext
from pyspark.streaming import StreamingContext
sc = SparkContext(master, appName)
ssc = StreamingContext(sc, 1)
```

- master is a Spark, Mesos or YARN cluster URL
- to run your code in local mode, use "local[K]" where K>=2 represents the parallelism
- appname is the name of your application
- batch interval time interval (in seconds) of each batch

Thank you

Analytics Vidhya