Spark Streaming

This presentation will demonstrate how a local Spark streaming application might be developed.

Stream

Composes a *Data Source ~> Spark Flow ~> Data Sink flow*.

- Source ~ Cassandra, Hdfs, Kafka, Text, …
- Flow ~ Spark Transformations and Actions
- Sink ~ Cassandra, Hdfs, Kafka, Text, …

Technologies

- Scala
- ScalaFX
- Spark
- Kafka
- Cassandra
- Cassandra-Spark Connector

Rationale

- Spark is built with Scala, making Scala the preferred language for Spark development.
- Spark visuals can be built with ScalaFX, a JavaFX wrapper, without using Javascript.
- JavaFX includes Charts, 2D, 3D and Animations.
- JavaFX apps and controls are decorated by CSS.
- Third party JavaFX charts and components exist.
- JavaFX includes a WebView component that allows for the rendering of Javascript content. For instance, D3 could be used to build an exotic chart.

Visual Spark

A ScalaFX application that executes a *Kafka Source ~> Spark Flow ~> Casandra Sink* stream.

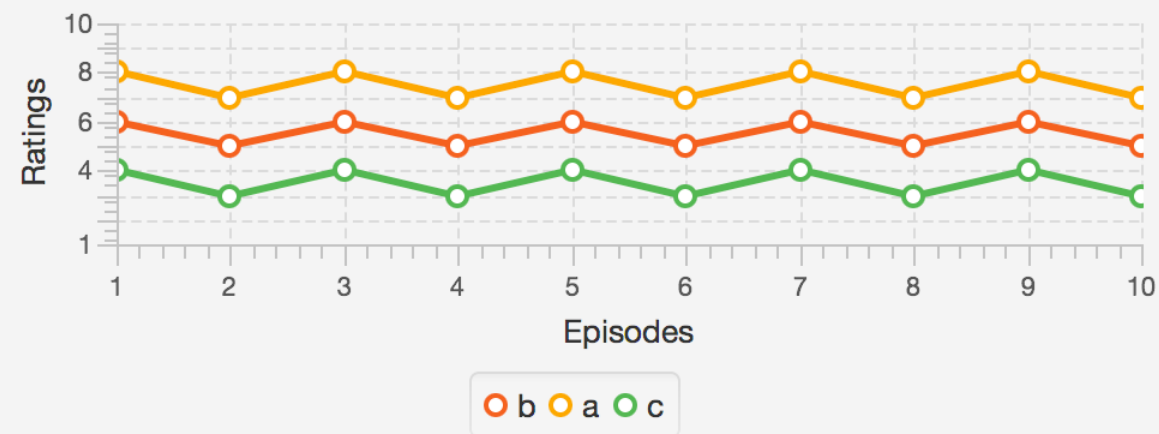The next slide contains a screen shot of Visual Spark after executing and rendering a stream.

Visual Spark

Play

30 messages processed in 22 seconds.

Source

Program	Season	Episode	Rating	
a	1	1	8	
a	1	2	7	
a	1	3	8	
a	1	4	7	
a	1	5	8	
a	1	6	7	
a	1	7	8	
a	1	8	7	

Flow

Episode Ratings

Ratings / Episodes

○ b ○ a ○ c

Sink

Program Ratings

b(33%)   c(21%)

a(45%)

● a(45%) ● b(33%) ● c(21%)