

Spark Streaming

Start spark streaming context:

```
import org.apache.spark.streaming.{Seconds, StreamingContext}
val ssc = new StreamingContext(sc, Seconds(10))
```

Listen to port:

```
val ds = ssc.socketTextStream("localhost", 9999)
```

Listen to dfs changes:

```
val ds = ssc.textFileStream("/user/xyz/streaming-input")
```

Save to file:

```
ds.saveAsTextFiles("/user/xyz/streaming-output")
```

print:

```
ds.print()
```

sliding window:

```
ds.reduceByKeyAndWindow((a: Int, b: Int) => (a + b), Seconds(30), Seconds(10))
```

Tasks

1. Count how many incoming transfers were there for each account in 10 seconds windows.