Intro to spark-submit

Li, Paul

February 25, 2016

What is `spark-submit`?

- `spark-submit` is a script provided in the `bin` directory of your spark installation
- It launches your spark applications onto the spark cluster
- You can explore Sparks documentation on this here:
  http://spark.apache.org/docs/latest/
  submitting-applications.html

spark-submit Usage

Usage:

```
./spark-submit <program> <program parameters>
```

Example:

```
./spark-submit /Users/pli/depot/msis2627/hw/2/hw2.py
/Users/pli/depot/msis2627/hw/2/data.txt
```