www.datascienceacademy.com.br

Big Data Real-Time Analytics com Python e Spark

Bibliografia, Referências e Links úteis

O que é Streaming de Dados?
http://datascienceacademy.com.br/blog/o-que-e-streaming-de-dados/

Apache Streaming
http://spark.apache.org/streaming/

Apache Samza
http://samza.apache.org/

Apache Storm
http://storm.apache.org/

Stream Analytix
http://streamanalytix.com/

Spark Streaming + Kinesis Integration
http://spark.apache.org/docs/latest/streaming-kinesis-integration.html

Spark Hub
https://sparkhub.databricks.com/

Apache Spark @Scale: A 60 TB+ production use case from Facebook
https://databricks.com/blog/2016/08/31/apache-spark-scale-a-60-tb-production-use-case.html

PySpark Streaming API
https://spark.apache.org/docs/latest/api/python/pyspark.streaming.html

DataTorrent – Gerenciando mais de 1 Bilhão de Eventos por Segundo em Tempo Real
https://www.infoq.com/br/news/2015/04/datatorrent

By the Numbers: 47 Amazing Uber Statistics
http://expandedramblings.com/index.php/uber-statistics/

TCP/IP
https://pt.wikipedia.org/wiki/TCP/IP

Spark Streming Framework
https://mapr.com/ebooks/spark/06-apache-spark-streaming-framework.html