

Big Data Engineer 实战训练营

拓展阅读

Further Reading & Resources

1. Blogs

- a. Todd Hoff: <http://highscalability.com/>
- b. Martin Fowler: <http://martinfowler.com/>
- c. Martin Kleppmann: <https://martin.kleppmann.com/>
- d. Werner Vogels: <http://www.allthingsdistributed.com/>
- e. Salvatore Sanfilippo: <http://antirez.com/latest/0>
- f. Kyle Kingsbury: <https://aphyr.com/posts>

2. Paper

- a. Apache Zookeeper:
 - i. Google Chubby Paper: <http://static.googleusercontent.com/media/research.google.com/en/archive/chubby-osdi06.pdf>
 - ii. Zookeeper Paper: <http://static.cs.brown.edu/courses/cs227/archives/2012/papers/replication/hunt.pdf>
 - iii. ZAB Protocol: <http://www.tcs.hut.fi/Studies/T-79.5001/reports/2012-deSouzaMedeiros.pdf>
- b. Apache Kafka: <http://research.microsoft.com/en-us/um/people/srikanth/netdb11/netdb11papers/netdb11-final12.pdf>
- c. Apache Cassandra:
 - i. Google GFS Paper: <http://static.googleusercontent.com/media/research.google.com/en/archive/gfs-sosp2003.pdf>
 - ii. Google MapReduce Paper: <http://static.googleusercontent.com/media/research.google.com/en/archive/mapreduce-osdi04.pdf>
 - iii. HDFS Paper: <http://pages.cs.wisc.edu/~akella/CS838/F15/838-CloudPapers/hdfs.pdf>
 - iv. Dynamodb Paper: <http://www.allthingsdistributed.com/files/amazon-dynamo-sosp2007.pdf>

- v. Cassandra Paper: <https://www.cs.cornell.edu/projects/ladis2009/papers/lakshman-ladis2009.pdf>
- d. Apache Spark:
 - i. RDD: <https://www.usenix.org/system/files/conference/nsdi12/nsdi12-final138.pdf>
 - ii. Spark Paper: https://people.csail.mit.edu/matei/papers/2010/hotcloud_spark.pdf
- e. Apache Mesos:
 - i. DRF Algorithms: <https://people.eecs.berkeley.edu/~alig/papers/drf.pdf>
 - ii. Mesos Paper: <https://people.eecs.berkeley.edu/~alig/papers/mesos.pdf>
- 3. Documentation
 - a. Apache Zookeeper:
 - i. Official Documentation: <http://zookeeper.apache.org/doc/trunk/>
 - ii. Tutorial: <http://www.tutorialspoint.com/zookeeper/>
 - iii. Programming Guide: <https://zookeeper.apache.org/doc/trunk/zookeeperTutorial.html>
 - b. Apache Kafka:
 - i. Official Documentation: <http://kafka.apache.org/documentation.html>
 - ii. Tutorial: <http://www.javaworld.com/article/3060078/big-data/big-data-messaging-with-kafka-part-1.html>
 - c. Apache Cassandra:
 - i. Official Documentation: <http://cassandra.apache.org/doc/latest/>
 - d. Apache Spark:
 - i. Official Documentation: <http://spark.apache.org/documentation.html>
 - ii. Programming Guide: <http://spark.apache.org/docs/latest/programming-guide.html>