

JACOB PARK

park.jacob.96@gmail.com | <http://jacobpark.me/> | <https://github.com/jparkie>

TECHNICAL SKILLS

Languages and Tools	Java, Gradle, Scala, SBT
Big Data	Kafka, Logstash, Hadoop HDP, Pig, Hive, Spark
Databases	SQLite, Cassandra, Elasticsearch
Microservice Development	Akka, Spray
Mobile Development	Android SDK 4.0 – 5.0
Web Development	Bootstrap, CSS, HTML, JavaScript, Play 2 Framework

WORK EXPERIENCE

Platform Engineer Intern, Paytm Labs, Toronto, Ontario, April 2015 – May 2016

- Implemented SQL over Spark Streaming with Spark SQL and Spark Catalyst for a real-time analytics platform over Kafka, Spark, and Samza.
- Improved bulk loading of +100 GB to 1 TB of data into Cassandra and Elasticsearch by several orders of magnitude against official Spark libraries by utilizing different write paths.
- Utilized the lambda architecture to develop batch and streaming pipelines to write data into Cassandra from HDFS and Kafka through Spark as various event sourced journals and materialized views for fast analytical queries.
- Lead a real-time, fraud-evaluation engine in Akka, Cassandra, and Spray with a cluster manager in Play and an external DSL in Scala; achieving sub-25 milliseconds over 600 rules at 40 000 requests per second.

PROJECTS

Personal Repositories Stars: 468 | Contributed Repositories Stars: 1016

Spark2Elasticsearch – <https://github.com/jparkie/Spark2Elasticsearch>

- Developed a Spark and Scala library to bulk load *DataFrames* into Elasticsearch utilizing the TCP transport instead of the official *elasticsearch-spark*'s REST client.

Spark Cassandra Connector (WIP) – <https://github.com/datastax/spark-cassandra-connector/pull/920>

- Improved bulk loading into Cassandra 2.2-3.0 by atomically writing *SSTables* and streaming them directly to Cassandra nodes through the TCP transport, thus minimizing the write path by skipping the driver.

Akka Persistence Cassandra – <https://github.com/akka/akka-persistence-cassandra>

- Implemented table configurations for *DateTieredCompactionStrategy*, *LeveledCompactionStrategy*, and *SizeTieredCompactionStrategy* for Cassandra 2.1.6 and above.
- Signed *Typesafe, Inc. Individual Contributor License Agreement*.

WhileScalaDsl – <https://github.com/jparkie/WhileScalaDsl>

- Developed an external domain-specific language written in Scala of the WHILE programming language with a context-free-grammar based parser, AST interpreter, and a runtime Java bytecode compiler.

PAID CERTIFICATIONS

HDP Developer: Apache Pig and Hive, *Hortonworks, Toronto, Ontario, July 7 – 10, 2015*

- A 4-day, paid, training course covering the following topics: Hadoop, YARN, HDFS, MapReduce, data ingestion, workflow definition and using Pig and Hive to perform data analytics on Big Data.

PRESENTATIONS

Developing a Real-time Fraud Engine – <http://www.meetup.com/scalator/events/224034596/>

- Presented at Scala Toronto about the development of a real-time, fraud engine with Akka, Cassandra, and Spray with explanations of high-level architecture, low-level design patterns, and an evaluation of various abstractions for parallel processing such as Reactive Streams.

PUBLICATIONS

Professional Scala, 1st Edition, Wrox, June 13, 2016 – ISBN-13: 978-1119267225

- Featured as a technical reviewer of both the rough draft and the final draft of *Professional Scala* written by Janek Bogucki, Alessandro Lacava, Aliaksandr Bedrytski, Matthew de Detrich, and Benjamin Neil.

ACTIVITIES AND INTERESTS

- Technical Speaking (Guest Speaker at Toronto Scala & Typesafe User Group, Attendee at Scala Up North 2015).
- Technical Blogging (Personal Blog, Featured in Cake Solutions' "This week in #Scala").
- Hackathons (Hack The North 2015 Attendee).

AWARDS AND SCHOLARSHIPS

- Awarded Valedictorian Award, Humanitarian Award, Bronze Governor General's Academic Medal.
- Received Schulich Leader Nominee commemorative plaque, Brilliant.org Waterloo Engineering Undergraduate Scholarship Top 10 Candidate.

EDUCATION

Candidate for Bachelor of Applied Science Degree, Computer Engineering (Honours B.A.Sc.),

University of Waterloo, Waterloo, Ontario, September 2015 – April 2020