

Big Data Engineer (GD3111) 5+ years

CA Cupertino

Responsibilities:

- Participate in design and development of Big Data analytical applications
- Design, support and continuously enhance the project code base, continuous integration pipeline, etc.
- Write complex ETL processes and frameworks for analytics and data management
- Implement large-scale near real-time streaming data processing pipelines
- Work inside the team of industry experts on the cutting edge Big Data technologies to develop solutions for deployment at massive scale

Requirements:

- Strong coding experience with Scala, Java, or Python
- In-depth knowledge of Hadoop and Spark, experience with data mining and stream processing technologies (Kafka, Spark Streaming, Akka Streams)
- Understanding of the best practices in data quality and quality engineering
- Experience with version control systems, Git in particular
- Desire and ability for quick learning of new tools and technologies

What will be a plus:

- Knowledge of Unix-based operating systems (bash/ssh/ps/grep etc.)
- Experience with Github-based development processes
- Experience with JVM build systems (SBT, Maven, Gradle)