Anand Iyer

Big Data Architect

anand.iyer@email.com (412) 555-0165 Pittsburgh, PA linkedin.com/in/anandiyer

SUMMARY

Big data architect with 13 years designing Hadoop, Spark and NoSQL platforms for health insurance, telecom and retail. Owns the architecture of a 2 PB data lake serving 4M members, from Kafka ingestion through Hive, Impala and HBase serving layers. Plans Cloudera upgrades and cloud moves, and turns business needs into designs teams can build.

EXPERIENCE

Big Data Architect

02/2021 - Present

Allegheny Mutual Health

Pittsburgh, PA

Architect for the enterprise data lake of a health insurer with 4M members.

- Own the architecture of a **2 PB** Hadoop and Spark data lake used by 600 analysts and 12 apps.
- Design Kafka ingestion for 45 claims and clinical feeds, landing data in Hive within 5 minutes.
- Lead the Cloudera upgrade of a 300-node cluster with 4 hours of total downtime.
- Set standards for NoSQL databases, with HBase and Cassandra stores serving 8,000 member lookups per second.

Senior Big Data Engineer

05/2016 - 01/2021

Steel City Telecom

Pittsburgh, PA

Senior engineer on a 200-node Hadoop platform for a wireless carrier.

- Built Scala Spark jobs that processed 5 billion network records a day for churn and capacity models.
- Designed Impala data marts that cut analyst query times from 20 minutes to under 30 seconds.
- Moved 300 MapReduce and Pig jobs to Apache Spark, freeing 40% of cluster capacity.
- Created Cassandra tables for device event history covering 7M subscribers.

Hadoop Developer

07/2013 - 04/2016

Monongahela Retail Analytics

Pittsburgh, PA

Developer at a retail analytics firm.

- Wrote Java MapReduce jobs for 2 billion weekly loyalty transactions.
- Built Sqoop and Oozie workflows that imported 90 RDBMS tables each night.
- Tuned Hive queries with partitioning and bucketing, halving report runtimes.
- Designed HBase row keys for a product lookup service handling 2,000 requests per second.

SKILLS

- Hadoop
- Apache Spark
- Hive
- Impala
- HBase
- Cassandra
- Kafka
- Scala
- Java
- NoSQL Databases

EDUCATION

Master of Science

05/2013

Carnegie Mellon University

Pittsburgh, PA

Information Systems Management

CERTIFICATES

AWS Certified Solutions Architect - Professional

08/2023

Amazon Web Services