



Manikanta Borra

2700 ACADIA DR , CORINTH ,TX ,76210.

Ph: +1 206-698-3221 Email: Manikanta2497@gmail.com

TECHNICAL SUMMARY

- Over 2 + years of experience with emphasis on **Bigdata Technologies / Hadoop ecosystem, SQL, Python.**
- Skilled programming in Map-Reduce framework and Hadoop ecosystems.
- Hands on experience in **GCP, Big Query, GCS bucket, G - cloud function, cloud dataflow, Pub/Sub cloud shell, GSUTIL, BQ command line utilities, Data Proc, Stack driver.**
- Have Extensive Experience in IT data analytics projects, Hands on experience in migrating on premise ETLs to **Google Cloud Platform (GCP)** using cloud native tools such as BIG query, Cloud Data Proc. Google Cloud Storage, Composer.
- Good experience with **AWS Databases** such as **RDS(Aurora), Redshift, DynamoDB, Glue Catalog** and **Elastic Cache (Memcached & Redis)**
- Good experience in implementing advanced procedures like text analytics and processing the in-memory computing capabilities with **Apache Impala, Scala.**
- Experience with **NumPy, Matplotlib, Pandas, Seaborn, and Plotly** python libraries.
- Worked on large datasets by using **Pyspark, NumPy** and **pandas.**
- Experience in using **Apache Kafka** for log aggregating.
- Developed a data pipeline using **Kafka** and **Spark Streaming** to store data into HDFS and performed the real-time analytics on the incoming data.
- Vast experience with **Scala** and **Python** In-depth understanding of **MapReduce** and the **Hadoop** Infrastructure Focuses on the big picture with problem-solving.
- Loaded and transformed large sets of structured, semi structured, and unstructured data in various formats like text, zip, **XML** and **JSON.**
- Good understanding of software development methodologies, including **Agile (Scrum).**

SKILL SET:

Hadoop Core Services-	HDFS, Map Reduce, Spark, YARN, Hive, Pig, Scala, Kafka, Flume, Impala, Oozie, Zookeeper
Bigdata Distribution -	Horton works, Cloudera, Amazon EMR, Azure
NO SQL Databases -	HBase, Cassandra, MongoDB
Cloud Computing Tools -	AWS, GCP
Languages -	Python, SQL, Pig Latin, HiveQL, Unix Shell Scripting
Application Servers-	Web Logic, Web Sphere, JBoss, Tomcat, Jetty
Databases	Oracle, MySQL, SQL
Operating Systems	Windows, LINUX

TECHNICAL SKILLS

Programming Skills: Java, Python, C, SQL, TypeScript, XML, Linux, Ruby

Web Technologies: HTML, CSS, JavaScript, react js.

Framework: Spring Boot, React.JS, Hibernate, Spring Data JPA, NET, D3.js, Flask

Other: Apache Kafka, Cloud Computing (AWS, GCP), Spring Security, Oracle, JUnit, Mockito, Jenkins, HTML5/CSS3, AJAX, BootStrap, Splunk, Pandas, Numpy, SciPy, Maven, Gradle, IntelliJ, Git, Agile (Serum), SDLC, Jira, VS Code, xcode, Tableau.

CERTIFICATIONS



- Google cloud certified Professional data engineer.
- Google cloud certified cloud digital leader
- Google cloud certified associate cloud digital leader.

PROFESSIONAL EXPERIENCE:

CLIENT : CVS HEALTH TX, Data engineer

FEB 2023 - CURRENT

- Proficient data engineer with a proven track record in designing, orchestrating, and executing complex Teradata to BigQuery migration workflows.
- Deployed application to GCP using Spinnaker (rpm Involved in complete project Life Cycle, i.e., Design, Implementation, and Unit testing. worked in GCP Data Proc, GCS, Cloud functions, and BigQuery, and used cloud shell SDK in GCP to configure the services Data Proc, Storage, and Bigquery.
- Extensively used agile development methodology and was involved in sprint planning.
- Identify the data to be migrated, including tables, views, and stored procedures, and determine the data volume and its growth rate to plan for BigQuery capacity and performance requirements. Collaborate with stakeholders to define migration goals and priorities.
- Implement data validation checks to verify data consistency between Teradata and BigQuery.
- Developed enterprise-level applications on large-scale distributed systems, ensuring high performance and scalability.
- Proficient in BI concepts with hands-on implementation experience on Cloud databases like Snowflake, Big Query, and Cloudera.
- Strong understanding of object-oriented design, data structures, algorithm design, and problem-solving. Proficient in Python, Scala, and Java.
- Experience in multi-threaded programming, concurrent programming, and synchronization to optimize software performance.
- Developed Big Data applications using Java, Spark, Kafka, enhancing data processing capabilities.
- Hands-on experience with Spark cluster-computing framework & Kubernetes or similar containerization technologies for efficient resource management.
- Experience with Enterprise-ready Cloud Data-Warehousing platforms like Snowflake, ensuring robust data management.
- Designed and developed cloud-based applications including large scale computer services, database services, APIs to design RESTful services, ETL, queues, and notification services.
- Demonstrated knowledge and experience in Software Engineering processes, including test-driven development for quality assurance.
- Skilled at working in tandem with a team of engineers, or alone as required, to deliver projects on time.
- Ability to see the bigger picture impact of decisions, ensuring alignment with business objectives.
- Worked on reading and writing multiple data formats like JSON, ORC, and Parquet on HDFS usingPySpark.
- Worked in Two Migration Projects SAP Hana to Bigquery and on-premises Hadoop to Bigquery

Environment: GCP, Hadoop, pyspark, Kafka, Apache Beam, Big Query, Airflow, Spring 4, Spring MVC, Spring Boot, RESTful web services, Spring Data, JPA, Hibernate 3, SQL, Micro services, JSON, JUnit 4, SVN, Log4J, Jenkins, Tomcat, Jira.

Client: PAYPAL

FEB 2023 - JAN 2024

Role: AWS Data Engineer

Responsibilities:

- Developed ETL data pipelines using Sqoop, Spark, Spark SQL, Scala, and Oozie.
- Used Spark for interactive queries, processing of streaming data and integrated with popular NoSQL databases.



- Experience with AWS Cloud IAM, Data pipeline, EMR, S3, EC2.
- Build data pipelines in airflow in GCP for ETL related jobs using different airflow operators.
- Worked on ETL Migration services by developing and deploying AWS Lambda functions for generating a serverless data pipeline which can be written to Glue Catalog and can be queried from Athena.
- Developed the batch scripts to fetch the data from AWS S3 storage and do require transformations.
- Developed Spark code using Scala and Spark-SQL for faster processing of data.
- Created Oozie workflow engine to run multiple Spark jobs.
- Exploring Spark to improve the performance and optimization of the existing algorithms in Hadoop using Spark-SQL, Data Frame, pair RDD's, Spark YARN.
- Developed Kafka consumer API in Scala for consuming data from Kafka topics.
- Implemented Spark-SQL with various data sources like JSON, Parquet, ORC, and Hive.

EDUCATION:

Governors state University

Master of Science in Applied Computer Sciences

Aug 2021 – Dec 2022

Chicago, Illinois

Jawaharlal Nehru Technological University

Bachelor of Science in Electronics and Communication Engineering

Aug 2015 – Oct 2020

Hyderabad, Telangana