

# Balasubramanyam Peram

## AWS Data Engineer

Minneapolis, Minnesota, 55416

[balasubramanyam9161@gmail.com](mailto:balasubramanyam9161@gmail.com) | (763) 306-1975 | <https://www.linkedin.com/in/balasubramanyam-236874351/>

---

### Summary

AWS Certified Solution Architect with 4 years of experience designing, implementing, and optimizing large-scale data pipelines and processing systems in cloud environments. Strong background in AWS services (S3, Redshift, Lambda, Glue, Athena) and Hadoop ecosystem tools (Hive, HDFS, Spark). Expertise in data ingestion, ETL development, and performance optimization. Skilled in managing cloud and on-premises big data solutions and automating workflows for real-time and batch processing systems.

---

### Technical Skills

- **Cloud Technologies:** AWS (S3, EC2, Lambda, Redshift, Glue, Athena, CloudWatch, DynamoDB)
  - **Big Data Technologies:** Hadoop, HDFS, Hive, Apache Spark, Apache Kafka
  - **Programming Languages:** Python, SQL, Java, Scala, Shell Scripting
  - **ETL Tools:** AWS Glue, Apache Airflow, Apache NiFi
  - **Databases:** MySQL, PostgreSQL, NoSQL (DynamoDB, HBase)
  - **Data Warehousing:** Amazon Redshift, Snowflake, Apache Hive
  - **Data Modeling & Analytics:** Data Lakes, Data Pipelines, Data Visualization (Tableau, Power BI)
  - **Version Control:** Git, GitHub
  - **DevOps Tools:** Docker, Jenkins, Terraform
  - **Operating Systems:** Linux, Windows
- 

### Education

#### Master of Science in Information Technology and Management

Concordia University ST. Paul | May 2023 – Aug 2024 | GPA 4.0

- **Projects:** Real-time data Processing with Apache Kafka and Spark, Hadoop Data Lake Integration with AWS S3
-

## **Professional Experience**

### **AWS Data Engineer**

*Infosys | Nov 2022 – April 2023*

- Designed and implemented end-to-end data pipelines to process large-scale data using AWS services (S3, Glue, Lambda) and Hadoop ecosystem tools (HDFS, Hive, Spark).
- Optimized data workflows, reducing processing time by 30% by integrating Hadoop with AWS Redshift for faster analytics and reporting.
- Built custom data transformation jobs in AWS Glue and Hadoop to handle data from various sources (structured and unstructured).
- Deployed batch processing systems using Apache Spark for large-scale data analysis, improving data processing efficiency by 40%.
- Implemented real-time data streaming solutions using Apache Kafka for data ingestion, processing, and analytics.

### **AWS Data Engineer**

*Logic loop | Oct 2020 – Sept 2022*

- Architected hybrid data solutions using both AWS services and Hadoop for on-premise and cloud data processing, resulting in reduced infrastructure costs by 20%.
- Managed and optimized HDFS-based distributed storage systems, integrating Hadoop with AWS S3 for scalable storage solutions.
- Leveraged Apache Hive for data warehousing and analytics, significantly improving query performance by fine-tuning partitioning and indexing strategies.
- Automated ETL workflows using AWS Glue, reducing manual data processing time and increasing data pipeline reliability.
- Supported large-scale data processing applications in AWS Redshift, ensuring seamless data migration from Hadoop to Redshift for faster querying.

### **Data Engineer Intern**

*Logic loop | Aug 2020 – Oct 2020*

- Worked on Hadoop-based ETL systems to process large datasets from various business applications and load them into Amazon Redshift for reporting.
- Assisted in optimizing MapReduce jobs in Hadoop to handle large datasets, improving job completion time by 35%.
- Supported the development of batch processing pipelines in Apache Spark, enhancing real-time analytics capabilities.

## Cloud Data Analyst Intern

*Planful | April 2019 – June 2019*

- Assisted in building a hybrid cloud architecture integrating Hadoop (HDFS) with AWS S3 for enhanced data storage and retrieval.
  - Developed SQL and Hive queries to perform large-scale data transformations for analytics purposes.
  - Automated routine data ingestion processes with AWS Lambda and Hadoop, significantly reducing manual intervention.
- 

## Certifications

- **AWS Certified Solutions Architect – Associate** | Issued: Sept 2024
- **AI For Everyone** | Issued: April 2022
- **Programming for Everybody (Python)** | Issued: May 2020