

Profile

- Amit has 4.8 years of IT experience in Architecture, Analysis, design, development, implementation, maintenance and support with experience in Cloud platforms like AWS.
- Good understanding and knowledge of big data technology, AWS.
- Expert in understanding the data and designing and implementing the enterprise platforms like Data Lake and huge Data warehouses.
- Have over 3+ years of experience as Consultant with very good exposure on big data Technologies like AWS Glue, Athena and Spark.
- Good understanding of databases and hands on working experience in writing applications on databases with performance tuning and optimization knowledge.
- Has extensive experience working on AWS EMR Clusters and building optimized Glue Jobs as per business requirements.
- Developed spark applications with the help of Spark, SPARK SQL and DATA FRAME API.

Technology Synopsis

Programming Language: Python

Karmik nagar, Dhanbad Jharkhand |

Contact: 91- 7079014262 | amitpaulom@gmail.com |

Professional Experience

Brillio Technologies Data Engineer

Oct 2021 – June 2022

Project: ETL for Pharma Majorclient

In this Big Data project we built a AWS Glue transformation for Redshift Spectrum to move the data from AWS S3 and other external destination sources ; enabling data scientists to leverage GCP (Google Cloud Capabilities) environments for performing research , analysis and experimentation. Built automated Glue template / lambda script in EC2 instance to provision a platform based on batch data streaming for specific global partners and stakeholders. From here Data Scientists requests for relevant data post GCP creation in S3 data lake. Built a datapipeline and separated the data transfer files from AWS S3 and enabled all components of ML and BI for research and analysis.

Activities / responsibilities:

- Analyzing the different source systems and knowing how to extract data from each source system.
- Transforming and Loading the data to S3 by the help Apache Spark on Databricks.
- Coordinated with BI team to provide them with necessary data required for reporting.
- Designed and developed complex data pipelines for customer. • Wrote production level code for logging, querying shipping sales.
- Construct ETL to capture productivity and data quality checks Environment: Python, Spark, Glue, S3, Lambda, Cloud Formation, DynamoDB, Code pipeline, Code Build, Pytest, Step Functions, Athena , Snowflake , Autosys , Shell Scripting Technologies: Python, Spark, Aws Glue, Lambda, S3, Athena , Sagemaker , Redshift

Amit Paul
Data Engineer

Computing Framework :
Spark API- RDD & SparkSql

AWS Cloud Services : S3,
Glue, EMR, Athéna, Lambda

Tools/Ide's : Pycharm, Visual
Studio Code, Microsoft Teams

Data Science : Python
Libraries-Pandas

Education

RAJASTHAN UNIVERSITY
AIET, Rajasthan

Mar 2009 - Sep 2012

Bachelor of Computer
Application

Major in Computer Science and
Mathematics

Cumulative GPA: 7.3/10.0

Language : Hindi , English ,
Bengali

Certification and Training:

.NET, ABAP, Python, Spark,
Aws

Award : Excellency in
education during academic
years from 2009 to 2012

Enum Informatics Private Limited

Data Engineer **Oct 2019 – Aug 2021**

Project: ETL for Retail domain client

In this Big Data project we extract data from different sources like mysql, oracle and put them into Aws S3. From here we do the transformation using AWS Glue and drop the data into processed layer. The processed layer data residing on the data lake is being published to the end user by using Athena Views.

Activities / responsibilities:

- Analyzing the different source systems and knowing how to extract data from each source system.
- Transforming and Loading the data to S3 by the help Apache Spark.
- Coordinated with BI team to provide them with necessary data required for reporting.
- Designed and developed complex data pipelines for customer
- Wrote production level code for logging, querying shipping sales
- Construct ETL to capture productivity and data quality checks

Technologies: Python, Spark, Aws Glue, S3, Athena

Enum Informatics Private Limited

Data Engineer **July 2018 – Sept 2019**

Project: ETL for Retail domain client

In this Big Data project we extract data from HANA DB using Informatica and put them into Aws S3. From here we do the transformation using AWS Glue and write the data back to s3. The data set residing on the s3 are being populated to the end user by using tableau.

Activities / responsibilities:

- Written spark code to read the data, transform the data and write the data in s3.
- Coordinated with BI team to provide them with necessary data required for reporting.
- Validated the data set between the source and target systems.
- Given support to the existing pipelines in the production

Technologies: Python, Spark, Aws Glue, S3, Athena

Isquare Soft Technologies, Bangalore

SAP Technical Consultant **Feb 2018 – June 2018**

Project: Sales Distribution Domain

HPE's new high-end storage platform to driving the next wave of the Intelligent Edge and cloud choices. Customer validation , new creation of batch entity and batch number of customer through HPI and HPE tenant.

Created implementation report for Customer inquiry and validation, used interface implementation and standard method PRESET_SALES_AREA for valid additional data under sales header area. Documented and specified it for Implicit Implementation in BADI for Shipping Sales Requisition.

Roles & Responsibilities Undertaken

- Created the ODATA and REST web services report program to extract data in xml format using OOPS and RFC methodologies
- Created class builder report using classic BADI and implemented it in XD01/XD02 to create and fetch customer details based on shipping additional data screen fields and updated KNVV through standard function methods PRESET_SALES_AREA using interface from implementation.
- Performed Unit testing and Integration testing for ALE Idocs which is generated and executed successfully Worked on Conversion Routine to add field constraints in segment for ALE Idocs through mandatory field LZONE with additional data and classification linking.
- Checked and verified partner details in presentation server file
- Cross verified with client to check extracted data migrated to standard field using BDC report and merged it in custom table through field pointers concept.

Infosys Pvt. Ltd.

Testing Engineer

Aug 2012 – Nov 2012

Project: Maintenance of Finacle software for banking domain

Infosys Finacle solutions address the core banking liquidity management, wealth management, analytics, artificial intelligence, and financial institutions to drive business excellence.

Infosys Finacle universal banking solution builds on the success of Finacle to deliver powerful benefits to global banks. The solution makes possible faster launch of new products and services , helping banks realize a 55 percent return on core banking transformation investments and an average improvement of 33 percent in their time to market as banks aim to reinvent their business and navigate the current challenges in the macroeconomic environment, Finacle 11E promises a simplified approach to banking transformation.

Roles & Responsibilities Undertaken

- Cross verified with client to check extracted data migrated to standard field using BDC report and merged it in custom table through field pointers concept.
- Created test cases in excel sheet and written test scenarios for database testing using query on the data on MySQL Server.
- Performed DDL/DML operations on backend databases, triggered stored procedures while database designing.

Technical skills:

Asp.net, Ado.net, MySQL, Sql server, oracle 11g

