

Abhishek Gupta

Contact : +91 9953507179

Email-id : Abhishek.gupta041988@gmail.com

To leverage my knowledge of the Software Development in the IT Sector and play proactive role in the Software Management for providing customer with best IT solutions

Profile summary

- **BigData Engineer** over 10+ years IT experience which involves 7.0+ years on **Big Data, Hadoop Eco System, Spark Framework, Scala, Python** SparkSql, NoSql, Hbase, Data warehouse, ETL , Datastage 9.x, Sql, Unix, Oracle, Db2.
- Expertise in data pipeline development with acquisition phase and writing complex Logics for processing as per KPI and generating enriched data using big data various stacks
- Developed **Sqoop, Spark** Import, Eval, Query, Incremental load functions to extract the desired data from Rdbms onto the Hadoop layer
- Created and updated hive schema, loading and querying data from Hive tables with **partitions, bucketing** and other techniques to increase performance
- Extracted and stored data in **Hive, Hbase** and applied the transformation by **Spark, Scala** and SparkSql In various file format like **Parquet, Orc, text** etc.
- Expertise on **Cloud AWS** using various services EMR, EC2,S3,Glue for storage, processing etc
- Understanding of RFP document and contributed in designing of project technical Architecture
- Experienced on **Sql and Unix shell scripting** for data extraction and Preparation
- Strong knowledge on python libraries(**Numpy,pandas,matplotlib**),statics like central tendency, correlation, hypothesis etc and Algorithm like Regression, classification, Histogram etc
- Experienced as a **ETL Developer** and worked for Development of Application which involves Stages Transformer, Dataset, Aggregator, Xml, Join, Lookup etc
- Worked in **Agile methodology** for the creation of Mapping documents, Technical documentation, process documentation and Implementation
- Knowledge on NiFi, Python, Atlas, Ranger, Kerberos, Spark Streaming, Kafka, Flume
- Worked on Heath Care and Banking/financial function Domain Projects

Technical Skills

- Big Data Skills : Hadoop , Hdfs ,Sqoop, Map Reduce , Hive, Pig, Yarn
: Spark, SparkSql, Scala, Python, Hotonworks, NiFi
- Data warehouse Tools : ETL – IBM Infosphere Datastage 9.x
- Databases : Oracle,DB2,NoSql,HBase
- Programming language : Scala, Python, Unix Shell script, SQL, Pig Latin
- Desktop Applications : Citrix, Lotus notes, IDE Eclipse, Maven, GIT, Jira
- Tools & Utilities : Remedy, QC, Citrix, Putty, Ultra-edit

Experience Details

Project Details #	
Company	American Express, Gurgaon
Project	CDE - SPACE
Tools/Technologies	Spark, Hive, Hdfs ,Unix, Emr, s3,Git, Jira, Event Engine, Jenkins
Description	This project named SPACE is a Incentive program which maintain, support,

	accommodate changes and fixes in the system which calculates the incentives for the sales-persons based on the basis of performance and targets of their signings for various entities on-boarding on Amex platform on multiple line of business. The system is built on top of in-house build up Cornerstone which acts as a centralized platform of large cluster for all Big Data, AI, ML applications across the wide range of Amex systems.
Responsibilities	<ul style="list-style-type: none"> Implementing code enhancement or modification as per the business logic requirement Created and updated Hive tables, loading and querying data in Parquet, text file format Read hive data into spark framework for the Implementation and enrichments based on various rules and KPIs Creating java spring batch and Unix scripts in Xml files for data flow management and scheduling of jobs Performed UIT and SIT for code enhancement or modification as per the business logic requirement Supported BAU activities for smooth business process and data loading incrementally Worked on Jira and Bit bucket for task tracking and repository

Project Details #	
Company	Capgemini, Malaysia
Project	SCB - IDL Usage
Tools/Technologies	Spark, Scala, Hive, Hdfs, Unix shell script, Hql, Bit bucket, Jira
Description	This project is running as part of the B3IDL program stands for Basel3 Intraday Liquidity, to follow international banking regulations to promote stability and reduce excess economy risk. This project provides a set of management reports which records intraday liquidity usage by various business and client including various charges across multiple countries and this will also enable Treasury to optimize Nostro Management for foreign currency exchange and trade transactions and Liquidity
Responsibilities	<ul style="list-style-type: none"> Involved in understanding and Implementing of technical requirements by analyzing mapping and business documents Loading the daily transaction file into staging Hive table partitions Reading Files and Hive tables into Spark with Scala in Rdd and Data frames to apply various enrichments and transformation as per mappingsheet and business logic processes Created and updated Hive tables, loading and querying data in Orc, text file format Loading and managing final output hive tables after enrichment and integration with other system Performed UIT and SIT for code enhancement or modification as per the business logic requirement Worked on Jira and Bitbucket for task tracking and repository

Project Details #	
Company	LTI, Pune
Project	Nets-Credit Service
Tools/Technologies	Spark, Scala, SparkSql, Hbase, Hive, Sqoop, Hdfs, Ambari, Unix, Sql
Description	Nets Merchant Services has a huge potential to help merchants get more knowledge on how to optimize their business by providing them with data on how, to who and at what times they sell – and how they sell compared to

	competitors. Identify comprehensive list of KPIs for merchants to understand their customers and business better. Few pilot customers from SME and LAKA segments to be on boarded for the KPI's. Enable self service analysis and scenario simulations for Merchants
Responsibilities	<ul style="list-style-type: none"> Involved in understanding of RFP documents and contributed in designing of project technical Architecture Developed Sqoop Import, Eval, Query, Incremental functions to extract the desired data from Rdbms for acquisition phase Reading Files and Hive tables into Spark with Scala in Rdd and Data frames to apply transformation and KPI processes Used Hbase for staging and Integrated with Hive to fetch the Data reside on Hbase and presented through hql query in Hive Created and updated Hive schema, loading and querying data in Parquet, text file format Sql and Unix shell scripting for data extraction and preparation Monitor and management of services and resources by Ambari

Project Details #	
Company	United Health Group
Project	Cirrus Integration
Tools/Technologies	Hadoop, Hive, Sqoop, Hdfs, DB2, Unix, Tivoli, Altova Xml, Datastage 9.x
Description	Cirrus Integration is a HealthCare domain Project where we are fetching data from Multiple Sources and loading into a single repository called Master Data management (MDM). We are extracting all source of Information and generating Xml Files to provide to End User Cirrus. This Information can be share among multiple Applications inside UHG from a single Source. Also It will improve speed of business decision making through advanced analytics and to design and build a system which will provide a single source of truth for UHG.
Responsibilities	<ul style="list-style-type: none"> Developed and used Sqoop Import, Eval, Query, Incremental load functions to extract the desired data Created and updated hive schema, loading and querying data and Hive managed and external tables with a specific partitions Used Pig scripting for data processing and transformation Developed/Modified Datastage jobs, Writing & optimizing Sql and Unix Scripts as per business requirements Migrate the components like Datastage jobs, Unix scripts and Sql scripts using CLM/SVN/manager server Involved in Technical documentation, Mapping Documents, batches, Unit Test Case Document and processes Worked in Agile methodology and Involved in meetings with On-Shore, BA and other streams on daily basis

Project Details #	
Company	HSBC Software development Pvt. Ltd.
Project	FTP Fin-Pro
Tools/Technologies	Datastage 8.X, SQL, Unix, Citrix, IBM-Safr, Quality center
Description	FTP is an acronym of 'Finance Transformation Platform'. FTP is a multi-components Integrated system that provides financial control with a common data repository for regulatory & management reporting. It's deploying 'events based' architecture. it creates warehouse to store the 'reconciled detailed transactional level data to help in reporting.

Responsibilities	<ul style="list-style-type: none">• Defined the Source databases and Target databases in Datastage and Created the Jobs and job sequences• Used most of the stages such as the Transformer, Sequential file, Aggregator, Funnel, joiner, lookup etc• Environment setup for UIT Preparing Input files for self and different components• Defect Tracking & solely responsible for Defect retesting Impact analysis on other modules.• Performed SIT, Performance testing support.• Worked on Utp and Technical specification documents
-------------------------	--

Project Details #	
Company	Tridat Technology Pvt. Ltd. Mumbai
Project	OfficeMax Retailer Supplier
Tools/Technologies	Cobol, JCL, Easytrieve, SQL Spufi, Abend-Aid, XP-editor, Remedy, Db2
Description	OfficeMax is an office supplies retailer founded in 1988 and headquartered in Naperville. Create custom summary bill layout includes various logics on tax calculations and bill formats.
Responsibilities	<ul style="list-style-type: none">• Writing COBOL code from scratch for each and every Customer Layout.• Analyzing and Modification of Existing COBOL programs.• Testing of various components including System Testing, Unit Testing and Integration testing of the products.• Train the new team members with the processes involved.

Educational Qualification

Educational Qualification	Year of Completion	Board/Institution
B.E.	2009	Pune University
12th	2004	C.B.S.E.
10th	2002	Uttar Pradesh

Personal Details

- Name : Abhishek Gupta
- Nationality : Indian
- Marital Status : Married
- Date of Birth : 02 Apr 1988.
- Address : GC-2, Gaur city, Noida, UP
- Linguistic Proficiency : English, Hindi

Declaration

I hereby declare that the particulars given herein are true to the best of my knowledge and belief.

Date – / /2021

Abhishek Gupta