Criteria	Details		
Candidate Name:	**First Name**	**Middle Name**	**Last Name**
	Mohd Shahid Sagar		
Phone No:	9958066412		
Alternate Number (Mandatory)			
Email ID:	shahidsagar525@gmail.com		
Current Location:	bangalore		
Desired Location:			
Native Location:	delhi ncr		
Total Experience:	4.2		
Relavent Experience:			
Current CTC:	4.5		
Expected CTC:	6		
Hike in Percentage:	**33%**		
Offer in Hand	No		
With Which Comapny			
Company Domain / Size			
Current Offered Salary			
Current Company:	tech mahindra		
Currently on Contract / Permanent:			
Parent Company name if Contract:	pic technology solutions pvt ltd		
Ready to pe on our Payroll:			
Notice Period:	30 days		
Buyout:			
Buyout paid by Candidate?			
Informed that buyout not paid by us?			
Accomodation required?			
Informed Accomodation only 1 Week?			
Reason for Job Change:	looking for delhi ncr		
Is there any gap in employment:			
Reason for Gap:			
Is there any gap in Education:			
Reason for Gap:			
Relocation:			
Reason for Relocation:			
Kids:			
Kids Studying:			
Will relocate with Kids?			
How studies will be managed for Kids:			
Spouse working?			
Will he/she relocate with candidate?			

Md Shahid Sagar
Email: shahidsagar525@gmail.com
Phone: +91 9958066412

Profile Summary

- Having 4 years of overall IT industry experience in Big data and Datawarehouse.
- Around 2.5 years extensive experience in Big Data Hadoop and its components HDFS, Apache Pig, Hive, Sqoop, Flume, MapReduce,Hbase and Oozie.
- Involved in writing the hive, Pig scripts to reduce the job execution time.
- Experience in working with structured and semi-structured data.
- Hands on experience in analyzing Log files for Hadoop and eco system services.
- Good Exposure on Apache Hadoop, PIG Scripting, Distribute Application and HDFS.
- Strong Knowledge of Hadoop and Hive's analytical functions.
- Hands on experience in installation, configuration, supporting and managing Hadoop Clusters using Apache, Cloudera (CDH5.7), Yarn distributions.
- Good experience in using AWS.
- Involved in bench marking Hadoop/HBase cluster file systems various batch jobs and workloads.
- Experience in database design using Stored Procedure, Functions, Triggers and strong experience in writing complex queries for SQL Server.
- Exceptional ability to learn new concepts, Hard working and enthusiastic
- 1.5 years of experience in developing and maintaining the Data Warehouse applications using ETL tools like INFORMATICA Power Center 8.x/9.x to transfer the data from legacy systems to common/ different targets.
- Extensively working as Informatica Developer(Mappings, Sessions, Worklets and Workflows).
- knowledge in DWH concepts, Star schema & Snow flake schema.

Technical Skills

- **BigData Ecosystem:** Hadoop, Hdfs, MapReduce, Hive, Pig, Sqoop, Flume,Hbase, AWS, CDH
- Technologies: Waterfall , Agile.
- ORACLE SQL/PLSQL, UNIX, CORE JAVA.
- INFORMATICA, ODI

Professional Experience

Designation   : Software Developer
Duration   : August 2014 – till date

Project#1   : Big Data Analysis
Position   : Hadoop Developer

Project Description

The main objective of this project is to migrate data from Enterprise Data Warehouse to Hadoop . Data in the warehouse is processed and transformed using sqoop and loaded into HDFS. Later this data is analyzed using Hive that exposes data in HDFS in a distributed Query enabled platform. Sqoop is used to extract data from internal structured data stores and load into HDFS. Data volumes range in Terabytes, which makes it challenging for regular analytics platform

Roles and Responsibilities

- Big DATA/Hadoop Implementation: Responsible for designing/Implementation of hive queries, Indexes views in the HDFS.
- Designed & developed different TD facility in Hadoop, BTQ, Mload, Fload.
- Responsible for developing custom defined UDFs Sqooping data from Teradata DB.
- Created Generic Utility to Convert Tera Data DDL Scripts, statistics to Automated Hive DDL Scripts using java.

Technologies and Skills

- Hadoop, MapReduce, Hive, Sqoop, Hdfs, Hue
- Linux, Teradata, Talend

Project#2   : Health care Data Management
Position   : Hadoop Developer

Project Description

Tampa has existing analytic platform "EZAnalytics" which has various performance, scalability and usability issues hence Tampa has identified business need to develop new Analytics platform for their clients. Thus the objective is to migrate the existing Teradata objects and data to Hive, show it to their Clients in the form of Tableau Reports.

Roles and Responsibilities

- Analyzing the requirement to setup a cluster.
- Written the Apache PIG scripts to process the HDFS data.
- Created Hive tables to store the processed results in a tabular format.
- Developed the SQOOP scripts in order to make the interaction between Pig and MySQL Database.
- Created External Hive Table on top of parsed data.
- Importing & Exporting data from RDBMS to HDFS, Hive and HBase

Technologies and Skills

- Primary Skills: Hadoop (HDFS, Map Reduce), Hive, Pig, Sqoop, Flume, Tableau, Informatica
- Languages: Java, HiveQL, Pig Latin, SQL
- Database: Oracle, Hbase, MySQL
- Web server: Apache Tomcat
- Operating Systems: Linux

Project#3   : Data Consolidation
Position   : Hadoop Developer

Project Description

This project involves building a real-time analytics platform based on stored in Enterprise Data Warehouse. Data in the warehouse is processed and transformed using Map Reduce and loaded into HDFS. Later this data is analyzed using Hive that exposes data in HDFS in a distributed Query enabled platform. Sqoop is used to extract data from internal structured data stores and load into HDFS. Data volumes range in Terabytes, which makes it challenging for regular analytics platform.

Roles and Responsibilities:

- Design & Develop solutions using Hadoop to tackle big data, information retrieval & analytics problems.
- Designed & developed different TD facility in Hadoop, BTQ, Mload, Fload.
- Responsible for developing custom defined UDFs Sqooping data from Teradata DB.
- Created Generic Utility to Convert Tera Data DDL Scripts, statistics to Automated Hive DDL Scripts using java.
- Setup Hive with MySQL as a Remote Metastore.
- Automated several jobs for pulling data from FTP server and load data into Hive tables.
- Create & update Hive schema to support a highly-available data warehouse for BI teams.

Technologies and Skills
- Hadoop, MapReduce, Hive, Sqoop, Hdfs
- Linux, Teradata, Talend, MySql

INFORMATICA PROJECT

PTS Metrics Project
Service Area: Technology
Role: ETL Developer; Duration: Nov 2012 – July 2014

PTS is a part of Cardinal Health focusing majorly in manufacturing drugs and other medical products and shipping it to respective clients. Currently PTS has 38 branches and all branches have different source systems. The data warehouse was built to provide business an option to see consolidated operational reports for all branches.

Role: Informatica Developer, Enhancement and Support

Key Responsibilities:

- Analyze source systems and understand source system applications with the help of PTS Business Analysts.
- Create new mapping designs using various tools in Informatica Designer like Source Analyzer, Warehouse Designer, Mapplet Designer and Mapping Designer.
- Develop the mappings using needed Transformations in Informatica tool according to technical specifications
- Created complex mappings that involved implementation of Business Logic to load data in to staging area.
- Used Informatica reusability at various levels of development.
- Developed mappings/sessions using Informatica Power Center 9.x for data loading.

- Performed data manipulations using various Informatica Transformations like Filter, Expression, Lookup (Connected and Un-Connected), Aggregate, Update Strategy, Normalizer, Joiner, Router, Sorter and Union.
- Developed Workflows using task developer, Worklet designer and workflow designer in Workflow manager and monitored the results using workflow monitor.
- Building Reports according to user Requirement.
- Extracted data from Oracle and SQL Server then used oracle for data warehousing.
- Implemented slowly changing dimension methodology for accessing the full history of accounts.
- Optimizing performance tuning at source, target, mapping and session level.

Technical Environment: Informatica 9.x, Oracle, Toad, TWS, Windows and Unix.

Academic Background

- ➤ M.Tech in Computer Engineering(M.D University, Rohtak)
- ➤ B.Tech in Computer Science & Engineering(M.D University, Rohtak)

Personal Details

Present Address	:	**BTM 2nd Stage,Banglore-76**
Permanent Address	:	**Sector-5,Rohini,Delhi-85**

Mohd Shahid Sagar

Bangalore