

OBJECTIVE:

To become a successful professional in the field of Information Technology and to work in an innovative and competitive environment providing creative solutions to the challenging problems that the software world is encountering.

PROFILE SUMMARY:

- Having 4+ years of IT experience in Data warehousing environment like **Hadoop, Big Data, Teradata and Informatica 9.x.**
- 2 + Years of **Big Data Ecosystem** experience in ingestion, storage, querying, processing and analysis of big data.
- In depth of understanding/knowledge of Hadoop Architecture and various components such as **HDFS, JobTracker, Task Tracker, Name Node, Data Node and Map Reduce concepts**
- Experience with configuration of Hadoop Ecosystem components: MR, Hive, HBase, Pig, Sqoop, Zookeeper, Oozie
- Experience in optimization of Map reduce algorithm using combiners and reducers to deliver the best results
- Worked on Teradata Database involving scripts such as **Fast load, Multi Load, BTEQ T-Pump.**
- Experience in designing and implementing different types of mappings which includes aggregator, expression, lookup, update, router, sequence generator etc, using Informatica
- Familiarity with (No-SQL) big-data storage concepts
- Followed Test driven development of **Agile and Water Fall Methodology** to produce high quality software.
- Experienced in design and development of **Packages, Cursors, Stored Procedures and Database Triggers** using **PL/SQL** and good understanding of Oracle **data dictionary**
- Worked extensively with Data migration, Data Cleansing, ETL Processes
- Knowledge on HBase NoSQL Database.
- Experience on Yarn architecture and its implementation.
- Ability to learn new technologies.

PROFESSIONAL EXPERIENCE:

Organization	:	cybermate infotek Private Limited
Current Designation	:	Software Engineer
Role	:	Team member
Duration	:	Sept 2011 till date

TECHNICAL SKILLS :

Type	Details
Big Data Ecosystems	HDFS, Hive, Pig, HBase, Sqoop, Hadoop components (NN,JT, TT, ZK), Map Reduce, Oozie, Flume, YARN
Programming Languages	Java
RDBMS	Oracle 8i/9i/10g/11g, , MySQL,, SQL Server
ETL Tools	Informatica 8.x/9.x, Teradata Utilities(Fast load, Multi Load, BTEQ T-Pump)
IDE Tool	Eclipse,
Defect Management Tool	Quality Center
Database	Oracle 10g, SQL Server 2000/2008
Methodologies	Star Schema, Snowflake
Operating System	Windows XP/Windows7 ,Windows server 2003/2008
Testing	ETL Testing and Manual Testing

PROJECT EXPERIENCE:

Project # 1

Title : Trader Joe's Sales DWH

Client : Trader Joe's, Germany

Environment : Hadoop 1.4, Java (jdk1.7), Oracle 11/, SQL server

Duration : Oct-2014 to till date.

Description:

Trader Joe's Electronics Stores, which is one of the organization that sells various numbers of products. Trader Joe's have various stores across Germany, which maintains Sales, Profits with lots of historical data. This project was to create Sales Data Warehouse and generate reports using Reporting services. The aim of sales data Warehouse is to analyze the sales and profit in month wise and year wise. The data gets refreshed daily, weekly, monthly & yearly and is maintained historically. This Data warehouse plays a major role in enabling various stores to view the data at a lowest level and help them to make decision to bring more revenue to company with new policies.

Responsibilities:

- ✓ Involved in **source system analysis, data analysis, data modeling** to ETL (Extract, Transform and Load)
- ✓ Worked extensively in creating **Map Reduce** jobs to power data for search and aggregation
- ✓ Designed a data warehouse using **Hive**
- ✓ Handling **structured** and **unstructured** data and applying ETL processes.
- ✓ Worked extensively with **Sqoop** for importing and exporting the data from HDFS to Relational Database systems/mainframe and vice-versa. Loading data into HDFS.
- ✓ Extensively used **Pig** for data cleansing.
- ✓ Created partitioned tables in **Hive**

- ✓ Developed **Pig** Latin scripts to extract the data from the web server output files to load into **HDFS**
- ✓ Developed the **Pig** UDF'S to pre-process the data for analysis.
- ✓ Develop **Hive** queries for the analysts
- ✓ Developed workflow in **Oozie** to automate the tasks of loading the data into HDFS and pre-processing with **Pig** and **MR**.
- ✓ Mentored analyst and test team for writing **Hive** Queries

Project # 2

Title : Aetna Health Insurance care
Client : Aetna, USA
Environment : Hadoop 1.0, Java (jdk1.7), Oracle 11/, SQL server
Role : Team Member.
Duration : May- 2013 to Sept - 2014.

Project Description:

This project deals with US health insurance that provides wide range Consumer-focused health plans for US customers. and the line of coverage is on Medical, dental, vision, life, disability, accident, critical illness plans and programs, the plan may covered for individual and group insurance. The site provides all forms from registration to claims. It provides separate sites for providers, Government & Public Sectors agents & brokers Etc. In keeping with federal and state laws and their responsibility is to protect the privacy of enrollee information.

Responsibilities:-

- ✓ Involved in **source system analysis, data analysis, data modeling** to ETL (Extract, Transform and Load)
- ✓ Worked extensively in creating **Map Reduce** jobs to power data for search and aggregation
- ✓ Designed a data warehouse using **Hive**
- ✓ Handling **structured** and **unstructured** data and applying ETL processes.
- ✓ Worked extensively with **Sqoop** for importing and exporting the data from HDFS to Relational Database systems/mainframe and vice-versa. Loading data into HDFS.
- ✓ Extensively used **Pig** for data cleansing.
- ✓ Created partitioned tables in **Hive**
- ✓ Developed **Pig** Latin scripts to extract the data from the web server output files to load into **HDFS**
- ✓ Developed the **Pig** UDF'S to pre-process the data for analysis.
- ✓ Develop **Hive** queries for the analysts
- ✓ Developed workflow in **Oozie** to automate the tasks of loading the data into HDFS and pre-processing with **Pig** and **MR**.
- ✓ Mentored analyst and test team for writing **Hive** Queries

Project # 3

Project Name : Vodafone Logging and Tracking System (VLTS)
Client : Vodafone
Software tools : Teradata V2R6/12

Other Data bases : Oracle
Environment : UNIX, Windows 2000.
Duration : Oct 2011 - April 2013.

Project Description:

Vodafone logging and tracking system enabled tracking of events within Vodafone's departments, corporate partners and pay as you Talk (PAYT) customer care, including outsourced call centers. This Data warehouse project was implemented with Teradata RDBMS on Unix Server 4400 With MPP architecture. The data is sourced from various applications like, customer information, accounting information, marketing information, postcode information, business info and sales information, which are running on mainframe with Teradata database. This Project provides consistent, reliable and timely customer-centric data that is used for business critical decisions.

Responsibilities:-

- ✓ Maintenance of the **BTEQ scripts** in regard to the Business rules and transformations.
- ✓ Analysis of data through SQL assistant using Teradata SQL.
- ✓ Responsible for overseeing the documents and quality procedures related to the project.
- ✓ Developed mappings in Informatica to load the data from various sources into the Data Warehouse, using different transformations like **Source Qualifier, Expression, Lookup, Aggregate, Update Strategy, Rank and Joiner**.
- ✓ Experienced in handling **slowly changing dimensions**.
- ✓ Employed Workflow Manager Tasks such as **Email, Command task, Event raise, Event wait, Decision, Assignment** to build Workflows and Worklets.

ACADEMICS:

Degree	Institute	Marks	Year Of Passing
M.C.A	Sri Venkateswara university, Tirupathi	74%	2010
B.Sc(Math's)	Sri Venkateswara university, Tirupathi	80%	2007
Intermediate	Board of Intermediate Education(A.P)	79%	2004
10 Th Class	Board of Secondary Education	76%	2002

Thanks
Hari Prasad