Saikrishna Cheruvu

Senior Data Engineer | Big Data | PySpark | Hadoop | Presto | Snowflake | SQL | DWH | Python | Unix
Mobile : + 91 8297130002 | Email : odbc.krishna@gmail.com | Linkedin : https://www.linkedin.com/in/saicheruvu/

Technology snapshot :

- Around **10 years of experience** in the IT Industry with specialization in **BigData, Data Warehousing & Business Intelligence** area as **Data Architect.**
- Experience in **Big Data analytics, PySpark, SQOOP, Hive,Ozzie and Impala**.
- In - depth understanding of **Spark** and **Hadoop Architecture** including **HDFS**, **Resource Manager, Node Manager, Name Node, Data Node, Spark Core, Spark SQL, Data Frames** etc.
- Hands on experience in various Big data application phases like data ingestion , data analytics and data visualization.
- Knowledge on **Delta Lake, Presto, Snowflake**.
- Good knowledge on **Numpy**, **Pandas**, **Kolas**.
- Experienced in using **Multiprocessing** and **Multithreading** using python.
- Experience using various Databases such as **Oracle, Netezza, DB2,SQL Server and PostgreSQL**.
- Very well versed in workflow scheduling and monitoring tools such as **Oozie** and **Autosys**.
- Experience in usage of **Hadoop Cloudera distribution**(CDH).
- Experience in transferring data from RDBMS to HDFS and HIVE table using SQOOP
- Experience in **Enterprise Data Warehouse Design, DataMart Design , Data Quality and Data Governance**.
- Experience in **Git, Bitbucket and Ansible tower** pipeline setups and **Hadoop deployments pipeline setup**.
- Involved in **Hadoop Cluster capacity planning**, **performance tuning**, cluster monitoring and troubleshooting.
- Sound knowledge in **UNIX and Python Scripting**.
- Experience in **Base SAS** and data migration from SAS to Hadoop.
- Experience in using REST Services **WebHDFS** and **HttpFS**.
- Involved in the ETL design for Load & Distribution process on Hadoop Data Lake.
- Experience in working with an **Agile** process.
- Having experience maintaining the windows based grid computing application **Moody's Risk Frontier.**

Education

Birla Institute of Technology and Science, Pilani M.Tech, Software systems

Business Analyst at Bank of America
Dec 2012 - Present

Development and Maintenance of ALPS The ALPS COMMON group supports GPS. Data from different source systems and databases such as Commercial, Consumer, ASPEN, CAF, SBL SECBL, and USCC are extracted, filtered, transformed and aggregated and finally loaded into data mart tables which are directly queried by business users and further used to perform downstream segmentation activities through Risk Frontier.
The data from the sources are first extracted and loaded into staging tables after which it is transferred to historical tables. They are loaded from historical tables into the final data mart/reporting tables after necessary transformation and business rules are applied

Roles and Responsibilities :
● Implemented Hadoop data pipeline to prepare the portfolio data marts.
● Performed big data processing using Hadoop,Spark, Hive Sqoop, Oozie, and Impala
● Experience in Importing data from SQL Server, Tetadata,db2,Oracle and Netezza to HDFS, using Sqoop and Spark to load data.
● Successfully handled the Netezza to Hadoop Historical data and Process migration.
● Created an automated tool to import the data from Netezza to Hadoop using Sqoop and Oozie.
● Leveraged WebHDFS and REST API's Created a SAS automation macro to migrate the data from sas to Hdfs.
● As there are lot of transformations required for the legacy systems Extensive performance tuning using explain plans, materialized views, proper distribution of data Building the data marts with dimensions and facts after applying DI/DQs Oracle development for enhancements on ALPS PORTAL application.
● Production Support and enhancements for the Hadoop Batch process.
● Experienced in release management in the production release life cycle using git, ansible.

Liquidity risk Systems (LRS) at Bank of America

Working as a multi-million dollar program to implement liquidity risk management solution for the bank of America treasury and risk management groups to manage the liquidity risk. The Solution would allow them to stress and simulate the liquidity position which will help in managing the liquidity risk more effectively and also adhering to the various regulatory requirements.

Responsible for the setup and deployment initiation through testing (PROD, DEV,QAT/UAT) including risk/issue tracking, financial and budget tracking, defect tracking to ensure remediation on a timely basis.

- Utilized Agile Scrum Methodology to help manage and organize a team with regular code review sessions.
- Extensive experience in Design and Development on OLAP database driven applications.
- Strong Analytical Database Architecture knowledge.
- Experience working in the agile environment and waterfall model.
- Good Knowledge of Hadoop and its related technologies – HDFS, HIVE, PIG, Sqoop, HBase and NoSQL DB.
- Hands on experience in Netezza, Oracle, db2, and Postgres SQL, SAS, Informatica, Autosys.
- Data Analysis, Profiling & Requirement Analysis.
- Performance Tuning, Accessing Very Large Datasets.
- Database preparations for Release cycles and data copy between the host.
- Strong knowledge on nzload,nz migrate and data replications.
- Preparing the Data Marts , FACT and Dimension models.
- Migrating database Code and ETL objects between the lower lanes to production.
- Database and linux environment preparation as part of release cycles.
- Responsible for the Production Issues and resolving as per the process.
- Code reviews and designing and developing the netezza objects and informatica mappings.
- Converted Dataflux rule/jobs to informatica objects.
- Configuring NDM, SFTP and FTP between the environments and integrated the database components.
- Successfully completed Oracle Exadata POC.
- Prepared database capacity planning for netezza database and automated the Capacity report from netezza to html report template.

Software Engineer at IBM India Private Limited
December 2011 - December 2012 (1 year 1 month)
Note: Worked as a contractor on the payroll of Adecco India.
Idea Project - BI ADM Team
Software Engineer at RNF Technologies
2010 - 2011 (2 years)
- Write SQL Queries, Procedures, Functions, Triggers, and Packages using PL/SQL.
- Develop the PL/SQL Code and SQL Queries.
- Impact analysis of the various change requests and enhancements.
- Ensure adherence to quality process and procedures, up to Product satisfies the Customer Requirements. Develop the database programs.
- Worked with Unix and Monitoring Jobs L1 & L2 Support.
- Worked various environments Production, testing, Development.
- Worked with microstrategy client components.