 **Divya kolli**

 **Talend Developer**

 **| Email:** **kollidivya25@gmail.com** **|Phone: 972-987-0149|**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**PROFESSIONAL SUMMARY:**

* **7+ years** of experience in IT Industry involving Software **Analysis, Design, Development, Integration, Maintenance**, Reports, Coding, Bug fixing, Creating Specifications, Production Implementation and Testing.
* Extensively created mappings in Talend using **tMap, tJoin, tReplicate, tParallelize, tConvertType, tflowtoIterate, tAggregate, tSortRow, tFlowMeter, tLogCatcher, tRowGenerator, tNormalize, tDenormalize, tSetGlobalVar, tHashInput, tHashOutput, tJavarow, tAggregateRow, tWarn, tLogCatcher, tMysqlScd, tFilter, tGlobalmap, tDie.**
* Experience using **Talend Integration Suite / Talend Open Studio (6.1/5.x) and 2+ years of experience with Talend Admin Console (TAC).**
* An **expert** in the **ETL Tool Informatica** which includes **components** like **Power Center, Power Exchange, Power Connect, Designer, Workflow Manager, Workflow Monitor, Repository Manager, Repository Server Administration Console**.
* Hands on experience on major components in Hadoop Ecosystem like **Hadoop Map Reduce**, **Spark**, **HDFS, YARN,** **Cassandra**, **Hive**, **Pig**, **HBase**, **Sqoop, Oozie**, **Kafka**.
* Experience in using the Utilities like **Pushdown optimization, CDC techniques**, **Partition and implemented Slowly Changing dimensions Type 1, Type 2** methodology for accessing the full history of accounts and transaction information.
* Converted the **data mart from Logical design to Physical design, defined data types, Constraints, Indexes, generated Schema** in the Database, created Automated scripts, defined storage parameters for the objects in the Database.
* Experience as working with various Relational Databases (RDBMS): **Oracle, MySQL, SQL Server, Complex Flat Files, Datasets, XML and Flat files**.
* Good Knowledge and experience on **Map Reduce** and **Spark** design patterns.
* Experience with distributed systems, large-scale non-relational data stores, **NoSQL map-reduce systems**, data modeling, database performance tuning, and multi-terabyte data warehouses.
* Experience in ingesting log data into **HDFS** using **Kafka**.
* Experience in developing, support and maintenance for the **ETL** (Extract, Transform and Load) processes using **Talend Integration Suite**.
* Experience in submitting Talend jobs for scheduling using **Talend scheduler** which is available in the **Admin Console**.
* Experience in **Spark** to process large stream of data.
* Experienced in running **MapReduce** and **Spark jobs** over YARN.
* Good knowledge and experience in Talend Integration experience with **AWS/Azure**.
* Hands on experience in application development using **Java,** **RDBMS**, and **Linux** shell scripting.
* Well versed in installation, configuration, supporting and managing of Big Data and underlying infrastructure of **Hadoop Cluster**.

**TECHNICAL SKILLS:**

|  |  |
| --- | --- |
| **Big Data Ecosystems** | Spark SQL, Hive SQL, HBase, Sqoop, HDFS, YARN, Kerberos, TD Encryption. |
| **Hadoop Distributions** | Hortonworks 2.4, Hortonworks 2.5, cloudera, |
| **ETL Tools** | Talend Big Data Integration 6.3, Talend Data Integration, Talend Admin console, IBM DataStage, SSIS |
| **NoSQL Databases** | HBase, Cassandra |
| **SQL Databases** | Oracle, MySQL |
| **Programming Languages** | Java, SQL, Pig Latin, Scala |
| **Tools and IDE** | Eclipse, IntelliJ |
| **Defect Management Tools** | HPQC, Jira |
| **Version Control Tools** | Tortoise SVN, GitHub |

**PROFESSIONAL EXPERIENCE:**

**Wawa Inc, PA Jan 2021 to Present**

**Talend Developer**

**Responsibilities**:

* Worked in the Data Integration Team to perform data and application integration with a goal of moving data more effectively, efficiently and with high performance to assist in business-critical projects coming up with huge data extraction.
* Broad design, development, and testing experience with **Talend and Bigdata Integration Suite** and knowledge in **Performance Tuning** of mappings.
* Deployed and scheduled **Talend** jobs in **Administration console** and monitoring the execution.
* Created separate branches with in the **Talend** repository for **Development, Production and Deployment**.
* Excellent knowledge with **Talend Administration console**, **Talend** **installation**, using Context and global map variables in Talend.
* Created Talend Mappings to populate the data into **dimensions and fact tables**.
* Responsible for developing, support and maintenance for the ETL (**Extract, Transform and Load**) processes using **Talend Integration Suite**.
* Used **Talend** job let and various commonly used **Talend**   transformations components like **tMap, tDie, tConvertType, tFlowMeter, tLogCatcher, tRowGenerator, tSetGlobalVar, tHashInput & tHashOutput**.
* Utilized Big Data components like **tHDFSInput, tHDFSOutput, tPigLoad, tPigFilterRow, tPigFilterColumn, tPigStoreResult, tHiveLoad, tHiveInput, tHbaseInput, tHbaseOutput, tSqoopImport and tSqoopExpor**t.
* Create Hive databases and tables over the **HDFS** data and write **HiveQL** queries on the tables.
* Schedule Hadoop and UNIX jobs using **OOZIE**.
* Responsible for building data model for **ODS/OLAP logical/physical** design.
* Modifies, installs, and prepares technical documentation for system software applications.
* Developed POCs for bulk load options**, web service API with in Talend**.
* Heavily used **Talend** for building **ODS & OLAP** structures, data movements and **XML& JSON** processing.
* Responsible for generating web service **SOAP requests** for larger volume and run through **SOAP service** and load the SOAP response into **POSTGRE** database.
* Set up and managed transactional log shipping, **SQL server Mirroring**, Fail over **clustering** and replication.
* Designed the architecture of **Talend**   jobs in parallel from execution standpoint to reduce the run time.
* Have handled issues related to **cluster start, node failures** and several java specific errors on the system.
* Perform troubleshoot on all tools and maintain multiple servers and provide back up for all files and script management servers.
* Worked Extensively on **Talend** **Admin** Console & Schedule Jobs in Job Conductors, this option is not available in **Talend** **Open Studio**.
* Hands of Experience on many components which are there in the palette to design Jobs & used Context Variables/Groups to Parameterize **Talend** Jobs.
* Worked Extensively on **Talend** **Admin** Console & Schedule Jobs in Job Conductors, this option is not available in **Talend** Open Studio.
* Experience in using Repository Manager for Migration of Source code from Lower to higher environments.
* Created Projects in TAC and Assign appropriate roles to Developers and integrated SVN (Subversion)
* Worked on Custom Component Design and used to have embedded in **Talend**   Studio.
* Used **Talend** **Admin** Console Job conductor to schedule ETL Jobs on daily, weekly, monthly and yearly basis (Cron Trigger).
* Perform validation check and deployment of reports to customer’s staging environment Client, Business objects.

**Environment:**Talend DI , XML files, Flat files, Talend Administrator Console Linux, Unix, Shell Scripting, spark, Pig, Hive, HDFS, Yarn, Hue, Sentry, Oozie.

**Intercontinental Hotels Group, Atlanta,GA Oct 2018 to Dec 2020**

**Talend Developer**

**Responsibilities:**

* Participated in Client Meetings and gathered business requirements and analyzed them.
* **Design, develop, test, implement and support** of **Data Warehousing ETL** using **Talend** and **Hadoop** Technologies.
* Design and Implement ETL processes to import data from and into **Microsoft Azure**.
* Research, analyze and prepare logical and physical data models for new applications and optimize the data structures to enhance data load times and end-user data access response times.
* Create **pig and hive scripts** to process various types of data sets and load them into **data warehouse** built on Hive.
* Develop stored procedures/views in **Snowflake** and use in **Talend** for **loading Dimensions and Facts**.
* Develop merge scripts to UPSERT data into **Snowflake** from an **ETL** source. Created Hive queries that helped market analysts spot emerging trends by comparing fresh data with **EDW** reference tables and historical metrics.
* Created complex mappings in Talend using **tMap, tJoin, tReplicate, tParallelize, tJava, tJavaRow, tJavaFlex, tAggregateRow, tDie, tWarn, tLogCatcher**, etc.
* Created joblets in **Talend** for the processes which can be used in most of the jobs in a project like to **Start job and Commit job**.
* Developed jobs to move inbound files to vendor server location based on monthly, weekly and daily frequency.
* Implemented **Change Data Capture** technology in Talend to load deltas to a Data Warehouse.
* Perform ETL using different sources like **databases, flat files, xml files**.
* Migrated **Snowflake** database to Windows Azure and updating the Connection Strings based on requirement.
* Managed and reviewed Hadoop log files.
* Wrote ETL jobs to read from web APIs using **REST and HTTP calls** and loaded into **HDFS** using java and **Talend**.
* Shared responsibility for administration of **Hadoop**, **Hive** and **Pig** and **Talend**.
* Good experience in reading queue from **JMS**.
* Provided design recommendations and thought leadership to sponsors/stakeholders that improved review processes and resolved technical problems.

**Environment:** Talend, HDFS, HBase, MapReduce, Eclipse, XML, JUNIT, Hadoop, Apache Pig, Hive, MS Elastic Search, Web Services.

**Verizon |Richmond, VA Feb 2017 - Sep 2018**

**Sr. ETL/Talend Developer**

**Responsibilities:**

* Involved in **analysis, design, development, and testing** phases and responsible for documenting the technical specifications.
* Process the data and push the valid records to **HDFS**.
* Import data from **MySQL** to **HDFS** using **SQOOP**.
* Tune the **MapReduce**, **PIG** and **Hive** jobs to increase the performance and decrease the execution time of the **jobs**.
* Compress the files downloaded from the servers before storing them in the **cluster** to save cluster resources.
* Write Core java programs to convert the **JSON files to CSV or TSV files** for further processing.
* Optimize already developed long running **MapReduce** and Pig job for better **performance** and **accurate results**.
* Create Hive databases and tables over the **HDFS** data and write **HiveQL** queries on the tables.
* Schedule Hadoop and **UNIX** jobs using **OOZIE** and Developing Scripts and **Batch Job** to **schedule** various Hadoop Program.
* Work with NoSQL databases like **HBase**.
* Write **Pig** and **Hive** **UDFs** for processing and analyzing log files.
* Visualize the complicated data analysis on the dashboards as per the business requirements.
* Integrated **Hive**, **PIG** and **MapReduce** **jobs** with elastic search to publish the metrics to the dashboards.
* Utilized the most used Talend Components such as **tMap, tFilterRow, tAggregateRow, tFileExist, tFileCopy, tFileList, tDie** etc.
* Utilized Big Data components such as **tSqoopExport, tSqoopImport, tHDFSInput, tHDFSOutput, tHiveLoad, tHiveInput, tPigLoad, tPigFilterRow, tPigFilterColumn, tPigStoreResult, , tHbaseInput, tHbaseOutput** along with executing the jobs in Debug mode and also utilizing the tlogrow component to view the sample output.
* Submitted talend jobs for scheduling using Talend scheduler which is available in the Admin Console.
* Deployed talend jobs on various environments including dev, test and production environments.

**Environment**: Hadoop 2x, YARN, HDFS, MapReduce, PIG, HIVE, HBASE, Shell Scripting, Oozie, TALEND, LINUX.

**Kelly Maxson |Hyderabad, IN May 2014 to Jan 2017**

**SQL SERVER/ETL Developer**

**Responsibilities:**

* Performed troubleshooting and maintenance of the reports with any enhancements or changes as needed from time to time.
* Set up report subscriptions as required by BRD and the management.
* Dropped and recreated Indexes in the **SQL Server DS**S while migrating **legacy** **data** to **SQL Server 2005**.
* Used **BCP to transfer data**.
* Created highly complex SSIS packages using various Data transformations like conditional **split, Fuzzy Lookup**, For Each **Loop, Multi Cast, column conversion, Fuzzy Grouping, Script Components**.
* Created **scripts for data validation** and massaging of legacy data in the staging area before moving it to the **DSS**.
* **Schedule** **Jobs** to **run SSIS packages** at night, to feed daily data into the **Decision Support System**.
* Wrote documentation for the **packages**, **scripts** and the jobs created for **DTS** to **SSIS** **migration**.
* Created various reports such as **graphs, charts, matrix, drill down, drill through, parameterized, sub reports** and linked reports etc.
* Generated server-side **T-SQL scripts** for **data manipulation** and **validation** and created various snapshots and materialized views for **remote instances**.
* Modified existing databases by adding/removing tables there by altering referential integrity, primary key constraints and relationships according to requirements.
* Created trace in **SQL Profiler** and used **Database** **engine** tuning advisor for Performance tuning on stored procedures and Scripts.
* Documented the migration process, reports and the **DSS structure and various objects**.
* Worked comfortably with a combination of **Agile** and Spiral methodologies.
* Facilitated meetings between various development teams, DBAs, and Testing teams for timely progress of the migration process.
* Participated in testing during the UAT (QA testing). Mentored and monitored team development and provided regular team status updates to the management.

**Environment:** SQL Server 2005/2000, DTS, SQL SSIS, SSRS, SQL Profiler, Visual Source Safe 2005.