**Satheesh Reddy L Email :** Pandu.satheeshL@gmail.com

BigData-Hadoop Developer **Mobile : +91-9700917927**

**Career Objective**

**Serving Notice period** and can **join within 30 days**. Working in **BT E-Serv** popularly known as **British Telecom** as a BigData Hadoop Developer. Looking for an intellectually challenging opportunity in **Hadoop development** and **Project Management** that requires Core software experience in all phases of software development life cycle (SDLC) starting from requirements, analysis, architecture, technical design, implementation, performance tuning.

**Experience Summary**

* Having 5 years of experience in **BigData** technology (Spark, Scala, Hive, MapReduce, Sqoop, Pig, Oozie, Shell, Kibana) as a **Developer** and **Project Lead** with multiple teams.
* Working on **CDH** distribution and good command on Hadoop cluster.
* Extracting Structured and Semi-Structured data from **Oracle** and other **File sources** from upstream systems into **HDFS** using scheduled jobs (**Sqoop and shell actions**).
* Worked on **GPG encryption**. Feeds which are coming from upstream systems will be zipped and encrypted using 64 bit algorithm. Need to Decrypt and process into DAA.
* Used **Pig** as ETL to process all our Entities and Extracts. Once done will update Oracle using Sqoop and will send the feeds to Down Stream using **Local Shell actions**.
* Worked on **Oozie** scheduler for unloads and regular jobs (Workflows and Co-ord).
* Used **Hive** as reporting for end user / business to perform ad-hoc analysis. Also used to analyze **Broken Hierarchy** between SAC, CUG, CUS and **organization hierarchy.**
* Implemented Partitions, Bucketing for monthly unloads and **Legacy data** in-order to improve the **performance** of Pig and spark jobs.
* Developed existing pig jobs into **Spark** **Scala** to as part of migration and to reduce the exec timings (Build time has been reduced to 3 days from 15 days)
* **Data Lake** story: Bringing the consumer data from various sources and expose the same to multiple systems. Biggest development and crucial part in all our releases.
* Designed and Developed **Kibana dashboard.** Generating Json files with index for **Logstash.** Having excellent command on **ELK Stack.** Productionised the same**.**
* Working on **CRM** systems as it holds business data and these are destinations.
* Working in **Agile** methodology and chairing the scrum calls as part of Agile and make sure everything is going smooth for each release.

**Workspace Environment**

|  |  |
| --- | --- |
| Big Data Ecosystem | Spark, Scala, HDFS, Map Reduce, Hive, Pig, Sqoop, Oozie, Kibana |
| Programming Languages | Core Java, Pig Latin, Scala |
| Databases | Oracle, CRM systems |
| Scripting Languages | Unix and Shell Scripting |
| Tools | Eclipse, Win SCP, Putty, Webextam, Winmerge |
| Visualization | Kibana, Thoughspot |
| Versioning Tools | SVN, Git(Local), Jira. |

**Employer Details**

|  |  |  |
| --- | --- | --- |
| **Employer Name** | **Duration** | **Role** |
| BT E- serv, Bangalore | Dec-2016 to Till date | Lead & Developer |
| Wipro technologies, Bangalore | Apr-2014 to Nov-2016 | developer |

**Academic Details**

|  |  |  |  |
| --- | --- | --- | --- |
| B.Tech  (Computer science engineering) | Siddhartha Institute of Engineering and technology, Puttur | 76.00% | May-2013 |
| Intermediate  (MPC) | Sri Chaitanya Jr. College, Tirupathi | 94.50% | May-2009 |
| Higher education  (SSC) | APR School, Gyarampalli | 93.00% | May-2007 |

**Project Summary Details**

**Project #3:**

Project Name : CSS OUK (EE Orange)

Client : British Telecom && EE Orange

Technical Environment : HDFS, Spark, Scala, Hive, Sqoop, Oozie, UNIX and shell scripting

Team Members : 14

Duration : January 2018 to till date

Role : Lead and Developer

**Description**: EE is a British mobile network operator, internet service provider and acquired by BT Business. It is having main three data areas specified as Orange, Red and green. Each divided based on importance of data it has. We have multiple vendors for data acquisition. We have a systems where feed is being generated using Abinito and the same will be transferred to PIS for GPG encryption for high security. Once encryption completes it will be transferred to SDEDS, from where we BT will be receiving the data. Once we receive the data from SDEDS the processing will starts from here.

**Roles and Responsibilities:**

* As the data is highly critical, it is being encrypted with 64 bit cryptographic **gpg encryption. Public key** will be shared with PIS which they will be used for encryption.
* Will use Private key to decrypt the same, Unzip it the local and move the files into HDFS post processing store into HDFS (**Using Oozie shell action** to do this)
* Once receive the data into HDFS, process the same using Spark Scala with certain conditions and create **Feeds and Data Lake into CSS Haas.**
* These feeds are being used by all our DS. We have regular ETL jobs in **spark** which are written in **Scala**. After processing storing the same into Hive for other jobs.
* Hive has been used as a reporting and **auditing** purpose.
* **Broken Hierarchy** has been created using Pig in-order to populate missing linkages.
* The same thing has been visualized in **ETL stack (Kibana)** and is self-reflective.
* Used External **Hive** tables on top of the output and used for business analysis.
* Use of **Oozie** workflow to run multiple serial and parallel action using workflows and co-ord.
* Chairing the Scrum calls as part of integration with various vendors as part of Agile.
* **Post Data Lake**, we have design documents received from architects which is in progress of project development phase-2.

**Project #2:**

Project Name : SIM REWRITE (BT E- serv)

Client : BT

Technical Environment : Spark, Scala, HDFS, Pig, Hive, Sqoop, Oozie, Spark, Scala

Team Members : 28

Duration : December 2016 to January 2018

Role : Developer

**Description**: BT is huge telecom retailer which is maintaining billing accounts, address and

Assets of customer groups in CRM systems to maintain customer info. We are working with various vendors to rewrite all existing legacy applications into Hadoop. Developed new unload framework with complete architecture and redesign. Storing all existing data from various sources into Hadoop as a part of data lake story. Working on various technologies In Hadoop to implement all new design and stability. Using agile methodology and chairing Scrum to discuss the status with vendors.

**Roles and Responsibilities:**

* As a part of rewrite, All AI jobs (ETL) has been successfully written into **Spark with Scala** which reduces project cost and increased execution time (15 days to 1 day).
* Developing various **Pig** scripts for ETL processing and scheduled as part of automation.
* Created multiple **hive** jobs to create **audit tables** to identify broken hierarchy.
* Automated the **Unload** framework with redesign and new architecture.
* Using **Sqoop** to import and export data from ORACLE to HDFS and vice-versa.
* Developed **shell** scripts in order to curl data from AI servers and vice versa.
* Created various **Shell** scripts for various monitoring reports (Edge node space, queue size, delta count)—scheduled using Sqoop as shell default jobs disabled.
* Used External **Hive** tables on top of the output and used for business analysis.
* Use of **Oozie** workflow to run multiple serial and parallel action using workflows and co-ord.
* Chairing the Scrum calls as part of integration with various vendors as part of Agile
* Developed **Spark** applications as part of rewrite and done POC before implementing the same.

**Project #1:**

Project Name : Shell FMO Build (Wipro Project)

Client : Shell

Technical Environment : HDFS, Map Reduce, Pig, Hive, Sqoop

Project Duration : September 2014 to till date

Role : Hadoop Developer & Production Support

**Description:** Royal Dutch Shell plc commonly known as Shell is an Anglo-Dutch multinational oil and gas company headquartered in the Netherlands. It is the seventh largest company in the world as of 2016, in terms of revenue, and one of the six oil and gas "super majors". It is having its business unit entire world which is categorized Asia-Pac, Europe, Americas, Australia and Coastline. The data is getting generated from all there areas and is populated in different databases for different regions.

In Shell FMO Build the aim is to merge all these domains and collect data into a single system. The data is stored in Hadoop cluster of 22 nodes (may be increase in future). Data is collected from all existing RDMS systems into HDFS and from different staging locations. The raw data will be ingested into HDFS by using Sqoop from various sources. Data transformations are done by using Pig and SSIS ETL tools. Later on it is moved to hive External tables for reports. Reports are being generated on weekly and monthly basis based on customer requirement.

**Roles and Responsibilities:**

* Moving data files generated from various location into HDFS using Sqoop and Shell scripts.
* Developing Hive and Pig queries for data analysis and reporting.
* Developed PIG scripts to convert the semi structured data to structured data.
* Worked on Hive partitioning and bucketing to improve the performance of system.
* Monitoring Hadoop scripts that will load data into HDFS at regular intervals.
* Writing custom Pig functions as per customer request to clean the data and provide results.
* Monitoring all Pig and Sqoop jobs day to day.
* Submitting the Test case and Test result document to the clients.
* Getting connected with the clients and onshore team to review the code and validation of test results.

**Achievements and Appreciations**

* Rewarded BMPS award for two quarters continuously. Posted in Hal of Fame for six months.
* Awarded with Brilliant performance and promoted within 18 months (Least period in BT).
* Got best team lead award for SIM-Rewrite in BT TSO year end awards.
* Migrated/developed completed AI to Hadoop within span of six months in BT which reduces project cost and saved execution time (Build reduced from 15 days to 3 days).
* Received multiple e-thank cards and appreciations in various all hands meet.
* Two continuous outstanding awards in Wipro for excellent performance in shell.

**Personal Information**

Name : Satheesh Reddy L

Marital Status : Unmarried

Passport : L2858858

Nationality : Indian

Languages known : Hindi, English, Telugu

Notice Period : Serving Notice (30 Days)

**Declaration**

I hereby confirm that the information furnished above is true to the best of my knowledge.

Place: Bangalore (Satheesh Reddy)