**Phani Lonavath Mobile No: (949) 344-6766**

**Hadoop Developer Email id: phani.lonavath@live.com**

**PROFESSIONAL SUMMARY:**

* Overall 7 Years of IT experience in Administering, Installing, Configuring and Maintaining Linux and Hadoop clusters
* Around **4** years of experience working on Hadoop eco-system
* Experience in **configuring, installing** and **managingApache,Cloudera, HortonWorks & MapR** Hadoop Distributions
* Extensive Experience in understanding the client’s Big Data business requirements and transform it into Hadoop centric technologies.
* Experience in configuring **Hadoop on Amazon Web Services (AWS).**
* Analyzing the clients existing Hadoop infrastructure and understanding the performance bottlenecks and provide the performance tuning accordingly
* Experience with Installing Hadoop in new servers and rebuild existing servers
* Performed technical proof of concepts
* Experience in using Automation tools like Chef for installing, configuring and maintaining Hadoop clusters
* Strong knowledge with Hadoop cluster connectivity and security.
* Experience in using Cloudera **Manager** for installation and management of Hadoop Cluster
* Experience in using Ambarifor installation and management of Horton Works Hadoop Cluster
* Expertise in writing **Shell scripts** and debugging existing scripts
* Experience in **Performance Management** of Hadoop Cluster
* Experience in working with amazon Redshift.
* Experience in using **Flume** to load log files into **HDFS**
* Expertise in using **Oozie** for configuring job flows
* Developed **Hive Queries** and automated those queries for analyzing on Hourly, Daily and Weekly basis
* Strong troubleshooting and performance tuning skills
* Experience implementing Kerberos for Cloudera and Horton Works Hadoop distribution.
* Coordinating Cluster services through Zookeeper
* Importing and exporting data into **HDFS** and **Hive** using **Sqoop**
* Experience in importing **and exporting the** preprocessed data into the commercial analytic **database**, e.g. **RDBMS**

**Technical Skills:**

|  |  |
| --- | --- |
| **Hadoop Framework** | HDFS, Hive, Pig, Flume, Oozie, Zookeeper, HBase and Sqoop |
| **NoSQL Databases** | Hbase |
| **Programming/Scripting** | C, SQL, PIG LATIN, UNIX shell scripting |
| **Microsoft** | MS Office, MS Project, MS Visio, MS Visual Studio 2008, MS Project |
| **Databases** | MySQL, Oracle, Redshift |
| **Operating Systems** | Linux, Cent OS, Windows |
| **Cluster Management Tools** | Cloudera Manager, Hue. |
| **IDE** | Net Beans, Eclipse, Visual Studio, Microsoft SQL Server, MS Office |

**Education:**

Bachelor of Engineering in Computer Science, India

**Professional Experience:**

**CenteneCorporation,St. Louis, MO Dec 2014 – Present**

**Hadoop Developer**

**Responsibilities:**

* Create, validate and maintain scripts to load data using Sqoop manually.
* Create Oozie workflows and coordinators to automate Sqoop jobs weekly and monthly.
* Worked on reading multiple data formats on HDFS.
* Developed multiple POCs using Scala and deployed on the Yarn cluster
* Analyzed the SQL scripts and designed the solution to implement
* Running reports in Pig and Hive.
* Develop, validate and maintain HiveQL queries.
* Imported and Exported data (from or to) RDBMS to HDFS using SQOOP.
* Running reports in Pig and Hive Queries.
* Analyzing data with Hive, Pig.
* Designed Hive tables to load data to and from external tables.

**Environment**: Cloudera Manager, HDFS, YARN, Hive, Pig, Zookeeper, Oozie, Sqoop, Flume, Spark, Hue, MySQL.

**AT&T, Dallas, TX June 2013 – May 2014**

**Hadoop Consultant.**

* Administered Hadoop clusters using **Cloudera** including activities like deploy Hadoop cluster, add/remove services, add/remove nodes, keep track of jobs, monitor critical parts of the cluster, and configure name-node high availability and schedule.
* Maintenance, Production support, incident/change management, service request management, Troubleshooting and Transform data between RDBMS and HDFS.
* Configured rack awareness using topology scripts in **Cloudera Distribution.**
* Worked with the Business intelligence team in hosting an Amazon Redshift Databases.
* Performed data completeness, correctness, data transformation and data quality testing using SQL
* Managing, reviewing Hadoop log files. Performance support to Developers
* Migrating Application Teams Video Streaming data (VOD), pay per clicks and Analytics teams to Hadoop from Mainframes, FTP server and EDW.
* Data loading to Hadoop and hive using Sqoop from MySQL, oracle and db2.
* Developed scripts to monitor automation jobs and processes required for Hadoop and setup mail service in case of failure.
* Developed scripts and automated data management from end to end and updating b/w all the clusters.
* Developing Oozie Workflows, Coordinators for automation of jobs.

**Environment:** Java (JDK 1.7), Linux, Shell Scripting, Amazon Redshift, SQL server, **Cloudera Hadoop,** Flume, Sqoop, Pig, Hive, Zookeeper and Hbase, Business Objects and Tableau.

**Visa, Foster City, CA Jan 2013 – June 2013**

**Hadoop Consultant**

Visa is an American multinational financial services corporation headquartered in Foster City, California, United States.[3] It facilitates electronic funds transfers throughout the world, most commonly through Visa-branded credit cards and debit cards.[4] Visa does not issue cards, extend credit or set rates and fees for consumers; rather, Visa provides financial institutions with Visa-branded payment products that they then use to offer credit, debit, prepaid and cash-access programs to their customers.

**Responsibilities:**

* Installed **Hadoop** on clustered Environments on all Environments
* Installed **Cloudera Manager** on **CDH3 clusters**
* Configured the cluster properties to gain the high cluster performance by taking cluster hardware configuration as key criteria
* Implemented the Hadoop Name-node HA services to make the Hadoop services highly available
* By using flume collected web logs from different sources and dumped them into HDFS
* Implemented **Oozie** work-flow for **ETL Process**
* Developed **Hive Scripts** and Temporary Functions for Complex Business Analytics
* Exporting data from RDBMS to **HIVE**, **HDFS** and **HIVE, HDFS** to RDBMS by using **SQOOP**
* Implemented **shell scripts** for **Log-Rolling** day to day processes and made it automated
* Coordinating **FLUME, HBASE** nodes and master using **zookeeper**
* Enabled Kerberos for AD Authentication.
* Commissioned/decommission nodes as needed.
* Streamlined cluster scaling and configuration
* Developed the cron job for storing the Name-node metadata onto the NFS mount directory
* Worked on file system management and monitoring and Capacity planning
* Execute system and disaster recovery processes
* Work with the project and application development teams to implement new business initiatives as they relate to Hadoop.
* Installed and configured operating systems packages.

**Environment:** Cloudera Manager, HDFS, YARN/MRV2, Hive, Pig, Zookeeper, Oozie, Sqoop, Flume, Hue, Teradata and MySQL and Oracle

**NEW Corp, Herndon, VA May 2012 – Jan 2013**

**Hadoop Systems Engineer**

**NEW Corp**, is the leading provider of extended service contracts and product protection programs worldwide. NEW offers customizable, end-to-end customer service solutions for partners in industries such as retail, manufacturing, and satellite TVs. And for more than 26 years, NEW has delivered outstanding results for our partners by delivering exceptional care to their customers. Today, NEW is once again changing the face of customer service delivery with service solutions that enrich the entire consumer ownership experience from beginning to end.

**Responsibilities:**

* Implemented two physical Clusters in the company and installed Hadoop
* Automated Installing Hadoop cluster using **puppet**
* Implemented Name Node Metadata backup using **NFS Mount**
* Manage the day to day operations of the cluster for backup and support
* Used full suite of infrastructure services like **DCHP, PXE, DNS, KICKSTART** and **NFS**
* Implemented automated 24x7 monitoring and escalation infrastructure for Hadoop cluster using Nagios and Ganglia
* Installed and Configured Cassandra on 2-nodes for NOSQL
* Highly involved in operations and troubleshooting Hadoop clusters
* Monitored cluster job performance and capacity planning
* Day to day support for OS Integration and application installations
* Designed and Implemented Security by gathering the user requirements.
* Worked on system and disaster recovery process

**Chamber IT Jan2009 – Dec 2011**

**Performance Engineer**

Chamber IT started it journey with a vision to provide quality software solutions and consulting services for its clients by leveraging the best software industry talent.

* Worked closely with Data Modelers in designing the data model based on the source system tables.
* Configured Physical & BMM layer – import tables, identifying facts, dimensions, created physical tables, join the tables, creating the logical tables and logical columns.
* Worked on **Dimensional Hierarchies (Drill down & Drill up)**, **Level Based & Share Measures** and **Time series Functions** based on business requirements.
* Worked on **Repository Variables** and **Session Variables**.
* Implemented **Security** by creating users, roles, groups, data level security, object level security and time level security.
* Involved in managing appropriate security privileges on subject areas and dashboards per business requirements.
* Created **Alias Tables** for easy reporting and better performance.
* Implemented **Aggregate tables** for better performance.
* Worked on **Answers**, **Dashboards, and Delivers**.
* Developed report with **filters**, **presentation variables**, **prompts**, **column selector**, **view selector**, **charts**, **narrative**, **gauge views**, **pivot table** and other views.
* Created **Dashboards prompts**, **pages**, **links**, **images**, **embedded content**, **text**, **folders** and **guided navigation links**.
* Created pivot tables in many dashboards for providing the ability to rotate rows, columns and section headings to obtain different perspective of the same data.
* Has integrated different reports in to dashboards and has set access permissions accordingly based on user needs.
* Has configured the job manager for working with OBI Delivers, created iBots and scheduled them to display alerts as well as to send the reports to the inbox of users.
* Created different reports like **Standard**, **Dynamic** and **Ad-hoc reports**.
* Performed unit testing and validating the repository.
* Prepared Test scripts, Load Test, Test Data, Test Plan, Test Cases, execute test, validate results, Manage defects and report results
* Used to identify the queries which taking too long and optimize those queries to improve performance
* **Performance Tested SOA based application using Parasoft and LoadRunner Tool**
* Performance Tested application in SaaS model
* **Handled Complex HCF Check in and Checkout Scripts with Custom DLLs provided by EMC**
* Handled View State and Event Validations for .net Scripts
* Independently develop **LoadRunner** test scripts per test specifications/requirements.
* Provide support to the development team in identifying real world use cases and appropriate workflows
* Performs in-depth analysis to isolate points of failure in the application
* Assist in production of testing and capacity certification reports.
* Investigate and troubleshoot performance problems in a lab environment. This will also include analysis of performance problems in a production environment.

**Environment:** OBIEE 10.1.3.3, Informatica Power Center 8.1, **DAC 7.9.X**, Windows Services, Windows 2003 Server UNIX/LINUX, Oracle 9i/10g, SQL, PL/SQL, Toad, LoadRunner, QTP, web -HTML/HTTP, QTP.