|  |  |
| --- | --- |
| Contact Number | 201-688-0542 |
| Alternate Contact Number | N/A |
| Current location of the candidate (City/State) | Phoenix, AZ |
| Email Address | [kiranhadoop55@gmail.com](mailto:kiranhadoop55@gmail.com) |
| Skype ID | [saikiran7149@gmail.com](mailto:saikiran7149@gmail.com) |

**Sai kiran Gandu**

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

**Professional Abridgment:**

* Over **7+** years of professional experience in IT industry with experience in Java and Hadoop Ecosystem.
* Hands on experience in installing, configuring, supporting and managing Hadoop clusters using Cloudera CDH4, CDH5 and Hortonworks HDP 2.2, HDP 2.3 distributions.
* Experience in Hadoop cluster capacity planning, performance tuning and troubleshooting the cluster.
* Experience on Commissioning, Decommissioning, Balancing, and Managing Nodes and tuning server for optimal performance of the cluster.
* Experience in writing the MapReduce programs using Java for analyzing Big Data.
* Good understanding of Hadoop architecture and different daemons of Hadoop clusters which include Name Node and Data Node, Task Tracker, Job Tracker.
* Experience in data ingestion process by using real time integration processes like Apache kafka, Storm and Flume.
* Strong Knowledge in importing streaming data into HDFS using Flume sources, and Flume sinks and transforming the data using Flume interceptors.
* Good knowledge/understanding of NoSQL data bases and hands on work experience in writing applications on NoSQL databases like Cassandra and HBASE.
* Experience working on Linux systems (RHEL 7) and good knowledge on bash shell scripting.
* Experience in installing, configuring and updating servers in RHEL 7 and authenticated the systems using LDAP.
* Experience with version control tolls like GIT, SVN and a Knowledge of BitBucket.
* Development experience with Java such as OOPS Concepts, Strings, IO, Multithreading, Collections and Exception Handling.
* Experience in working with BI technology like Tableau, Data Analysis Packages like Pandas in Python.
* Experienced in working with Amazon Web Services (AWS) using EC2 for computing and S3 as storage mechanism.
* Experience in setting up and managing the clusters on AWS.
* Experienced in the complete life cycle including design, system architecture, development, testing, implementation, Migration and maintenance.
* A very good planning, time management, prioritization and organizational skills.
* Exceptional written, verbal, and interpersonal communication skills.
* Ability to perform business analysis, and desire to improve business processes.
* Ability to effectively work with a variety of organizations, management levels, cultures, and personalities.
* Demonstrated ability to work independently and autonomously.

**Skills:**

Big Data Ecosystems : HDFS, MapReduce, Yarn, Hive, Pig, Sqoop,

Oozie, Spark, Flume, Kafka and Storm

Hadoop distributions : Apache Hadoop, Cloudera, Hortonworks   
Databases : MySQL, Oracle, HBase, Cassandra.

Languages : C, Java, SQL, HQL, Python, Shell, SparkSQL.   
BI/Analytical : SAS, DQ Analyzer, Data Flux, Tableau  
Operating Systems : Windows, RHEL 6 & 7, Unix, Mac OS   
Other Skills : AWS, GIT, SVN, Microsoft Office Suite.

**Experience:**

**Client : American Express, Phoenix AZ Aug 2016 – May 2017**

**Hadoop Consultant**

**Responsibilities:**

* Created and assemble the Foundation layer script to stack the information from social databases

Sql, Oracle, Netezza to Hadoop **Hive** tables utilizing **bash** script.

* Used **Sqoop** to import the verifiable data into hive tables in **text** and **Avro** file formats.
* Developed **Python** Script to transfer the data into target databases Greenplum and Netezza, Hadoop from unified location.
* Implemented partitioning, Bucketing in **Hive** for the better association of the data.
* Implemented Adhoc, every day, week after week, month to month Batch jobs utilizing **Control-M** scheduler.
* Implemented **Apache kafka** for the data ingestion of real time data using **Spark**.
* Wrote the **SQL** queries and utilized as a part of scripting to computerize the email ready warnings to business clients at whatever point their information is stacked into target tables.
* Involved to build up the Python script to peruse the **JSON** objects from the document and store the comes about into TEXT and CSV records.
* Processed text files using various libraries like sed, awk, regex in python.
* Implemented Quality check controls for to prepare the information utilizing **python/Bash** script.
* Involved in the implementation of Hadoop cluster on **AWS** as a part of POC.
* Worked using various Bigdata technologies on the **Hadoop Ecosystem** inHortonworks distribution.
* Experience in managing and reviewing Hadoop **log files.**
* Analyzed large amounts of data sets to determine **optimal** way to aggregate and report on it.
* Exported the analyzed data to the relational databases using Sqoop for visualization and generated financial reports using **Tableau**.
* Experience working with **Git** version control.
* Supported in setting up **QA** environment and updating configurations for implementing scripts with **Pig** and **Sqoop.**
* Working with **Jenkins** for any automation builds which are integrated with **GIT, Docker** as part of infra automation under continuous integration.

**Environment:** Hadoop, MapReduce, Spark, HDFS, Hive, Sqoop, SQL, Tableau, Linux/Unix, Autosys, AWS, EC2, NDM, Hortonworks distribution.

**Client: Target Minneapolis MN Jul 2014 to Aug 2016**

**Role: Hadoop Developer**

**Project: BLAZE - BRMS**

**Responsibilities:**

* Implemented Hadoop Infrastructure and maintained the Hadoop File System using Cloudera ditribution.
* Analyze, develop and implementation of Importing and exporting data into **HDFS using Hive, Sqoop.**
* Designed and developed loading data process from Oracle database into HDFS using Sqoop queries
* Created Hive queries and tables to help LOB identify current credit card trends and take decisions.
* Implemented **Apache Kafka** for the real time data ingestion into HDFS.
* Apply strategies on historical data before promoting them to production.
* **Developed Pig scripts** to parse the raw data, populate staging tables and store the refined data in partitioned DB2 tables for Business analysis.
* Review Hadoop Log files, code reviews and walkthroughs.
* **Developed Pig UDF'S** with **Python** for manipulating the data according to Business Requirements and also worked on developing custom Pig Loaders.
* Used **Spark API** over Hadoop to perform analytics on data in Hive.
* Migrated data from Spark applications with **Tableau** and created visualizations and reports to higher levels of management.
* Worked on Linux Systems with **RHEL 6 & 7** and have a good understanding of the Linux Ecosystem.
* Participate in Requirement gathering, architecture, design, development, test and implementation meetings.
* Understand business requirements, collaborate with stakeholders, discuss on source-target data mapping, design, signoff and implementations.
* Create implementation plan and Detailed Task Schedules
* Implemented **Storm** and **zookeeper** to work in a distributed mode.

**Environment:** Hadoop, Cloudera, MapReduce, HDFS, HBase, SVN, Cloudera, Hive, Spark, Python, Pig, Sqoop, Kafka, Zookeeper, DB2, SQL, Linux, Windows & Microsoft Office.

**JC Penny, Plano, TX Jan 2013 – July 2014**

**Hadoop Admin/Consultant**

**Responsibilities:**

* Handle the installation and configuration of a Hortonworks distribution **Hadoop** cluster.
* Monitored health of all the Processes related to **Name Node HA, HDFS, YARN, HIVE, HBASE, Sqoop.**
* Monitored **disk, Memory, Heap, CPU utilization** on all Master and Slave machines and took necessary measures to keep the cluster up and running on **24/7** basis.
* Setup and support automated Continuous Integration(**CI/CD**) utilizing tools like **Jenkins, Shell Scripts, AWS CLI/API**
* Configured Capacity Scheduler to provide service-level agreements for multiple users of a cluster.
* Wrote Map reduce jobs in **Java**.
* Used **Hive** and created Hive tables and loaded data from Local file system and **HDFS**.
* Created Hive **internal** and **external** tables defined with appropriate static and dynamic **partitions**, intended for efficiency.
* Worked on installing cluster, commissioning & decommissioning of **data nodes.**
* Involved in implementing High Availability and automatic failover infrastructure to overcome single point of failure for **Namenode** utilizing **zookeeper** services.
* Involved in Installing **Kerberos** and Configuring Server, Client Systems to enable Hadoop Security.
* Imported logs from web servers with **Flume** to ingest the data into **HDFS.**
* Involved in collecting metrics for Hadoop clusters using **Ganglia** and **Ambari.**
* Involved in troubleshooting of the **RHEL 6** systems and completed the tickets in time.
* Interacted with developers when we had to deploy new jobs, Jobs throwing exceptions, and data related issues.
* Developed workflow in **Oozie** to automate the tasks of loading the data into **HDFS** and pre-processing with Pig.
* Exported the analyzed data to the relational databases using **Sqoop** for visualization and to generate reports for the **BI** team.
* Responsible for cluster maintenance, adding and removing cluster nodes, cluster monitoring and troubleshooting, manage and review data backups, manage and review Hadoop **log** files.

**Environment:** Hadoop, Sqoop, Flume, Oozie, Jenkins, Git, Hortonworks, Ambari, Linux, HDFS, Pig, Hive, Kerberos, CI/CD, Oracle 10g, MySQL

**Union Bank of Switzerland (UBS), NJ Nov 2011-Dec 2013**

**Java Engineer- Hadoop Administrator**

**Responsibilities:**

* Worked on setting up the Hadoop cluster for the **dev**, **test** and **prod** Environment.
* Configured **Fair** Scheduler to provide fair resources to all the applications across the cluster.
* Worked with application teams to install Hadoop updates, **patches,** version upgrades as required.
* Designing the system and documenting Use Cases, Class diagrams, Sequence diagrams
* Developed Servlets & JSP pages for administrative tools, using custom tags
* Extensively used DOM Parser to Parse the XML document
* Implementing MVC architecture using hibernate value objects and mapping xml files
* Configured JMS and JNDI in Rational Application Developer
* Used commons and log4j logging framework
* Worked on Unit and Integration Testing
* Used JavaScript for client side validations in the JSP and HTML pages
* Performed unit testing, system testing and integration testing
* Developed SQL queries in using complex joins, groups and aggregation.

**Environment:** Java (JDK 1.7), Mapreduce, HDFS, Cloudera distribution, Jsp, log4j, HTML, SQL, Oracle, Linux, Tomcat.

**Virtusa, Hyderabad, India Aug 2009 to Oct 2011**

**Software Engineer- JAVA**

**Responsibilities:**

* Gathered requirements for the project and involved in analysis phase
* Developed quick prototype for the project so as to aid business in deciding the necessary ramifications to the requirements.
* Created UML class and sequence diagrams using Rational rose.
* Designed and created user interface front end screens using **JavaScript**, **HTML** and JSP’s.
* Used strict Object oriented model to separate the roles of the employees and their specifications to achieve extendibility and clarity.
* Developed dynamic content of presentation layer using **JSP**.
* Wrote test cases using **Junit** and coordinated with testing team for integration tests.
* Used **JDBC** to interact with the **Oracle** database for storage and retrieval of information.

**Environment:**Java, Java Scripts, JUNIT, HTML, Web Logic, Tomcat 4.x, Oracle 8i.

**Education:**

B.Tech/ JNTU HYD/2009