|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| venkata  |  | | --- | |  | | hadoopvenkatasai@gmail.com | |  | | 971-800-0023 | |  | | 8+ years of IT EXPERIENCE | |  | | Objective Searching for the opportunity to bring 8 years of programming, technology, and engineering expertise in developing software’s while incorporating critical thinking, problem solving, and leadership. | |  | |  | | --- | | **technologies skill set** **Hadoop Core Services**  HDFS, MapReduce, Hadoop YARN  **Hadoop Data Services**  Apache Hive, Pig, Sqoop, Flume, Kafka  **Hadoop Distributions**  Hortonworks, Cloudera  **Hadoop Operational Services**  Apache Zookeeper, Oozie  **Cloud Computing Services**  AWS (Amazon Web Services), Amazon EC2  **Java & J2EE Technologies**  Java, Servlets, JSP, JDBC, Java Beans  **IDE Tools**  Eclipse, NetBeans, RAD, Jira, Nexus.  **Programming Languages**  C, Java, Unix Shell scripting, Scala  **Operating Systems**  Windows (XP,7,8,10), UNIX, LINUX, Ubuntu, CentOS  **Reporting Tools /ETL Tools**  Tableau, Powerview for Microsoft Excel, Talend  **Databases**  Oracle, MySQL, DB2, Derby, Database (Hbase, Cassandra)  **Web Technologies**  HTML5, JavaScript  **Web Servers**  Web Logic 10.3, Web Sphere 6.1, Apache Tomcat 5.5/6.0  **Environmental Tools** SQL Developer, Win SCP, Putty, SOAP UI, MS Visio, JIRA **Version Control Systems**  CVS, Tortoise, SVN | |

## **Experience**

#### Sr hadoop Developer/Walmart, Bentonville - AR

##### May 2017 – Till Date

* Ingested the data from various data sources (Oracle DB, FTP, Couchbase) into HDFS.
* Developed a Shell Script and Python script for running Sqoop commands, exception Handling and storing the logs while ingesting the Transactional and Electronic Invoice data from Oracle to HDFS.
* Created **Hive** External Tables for the incremental imports into Hive using Ingest, Reconcile, Compact and Purge Strategy.
* Experienced in migrating HiveQL into Impala to minimize query response time.
* Created both external and internal tables in Hive. Worked on Partitions and bucketing.
* Fine Tuning and Productionizing the Teradata SQl queries that are running for long time in a queue.
* Performed Data analysis on the data stored using Spark Scala.
* Case class, object creation and creating data frames, rdds. Applying Transformations and Actions.
* Implemented Avro and parquet data formats for apache Hive computations to handle custom business requirements.
* Worked on AWS to create, manage EC2 instances and Hadoop Clusters. Involved in connecting to target database to get data.
* POC involved in loading data from LINUX file system to AWS S3 and HDFS.
* Using Python, able to connect to the SFTP server and get all the OMS (Order Management System) and Post Tax files into HDFS.
* Developed a Spark application using Scala, which parses the data in HDFS and ingests only the filtered data into Hbase and Solr.
* Created Data frames using Spark SQL and performed a JOIN over large number of tables and the resulting data frame is ingested into Hbase and Solr.
* Created a table in Hbase with Column Families and Column Qualifiers and inserted the data using REST API.
* Using Solr, metadata is stored from which the UI reads the data and projects it to the end-user.
* Understand the Client requirements and designed the best possible approach to meet the Customer Use case.
* Involved in story-driven agile development methodology and actively participated in daily scrum meetings.

**Environment:** Hadoop, Map Reduce, SQOOP, Splunk, Python, Hortonworks – 2.5.3 ,HDFS, Sqoop, Oracle – 12c, HBase, Spark – 2.0, Solr

## **Experience**

#### SR hadoop developer/Ameritas Life insuaraceCINCINAITI- OH

##### April 2016 –May 2017

* Data Ingestion into the Indie-Data Lake using Open source Hadoop distribution to process Structured, Semi-Structured and Unstructured datasets using Open source.
* Apache tools like FLUME and SQOOP into HIVE environment.
* Expertise in Hive queries, created user defind aggregated function worked on advanced optimization techniqus and have extensive knowledge on joins.
* Migration of ETL process from oracle to hive to test the easy data manipulation using Talend.
* Create hive scripts to extract, transform, load (ETL) and store the data using Talend.
* Developed Sqoop Scripts to extract data from DB2 EDW source databases onto HDFS.
* Worked with Oracle and Teradata for data import/export operations from different data marts.
* We used the most popular streaming tool KAFKA to load the data on Hadoop File system and move the same data to Cassandra NoSQL database.
* Experience in various data transformation and analysis tools like Map Reduce, Pig and Hive to handle files in multiple formats (JSON, Text, XML, Binary, Logs etc.).
* Worked extensively with Data migration, Data cleansing, Data profiling, and ETL Processes features for data warehouses.
* Involved in designing, capacity planning and implementing multi-rack Cassandra cluster.
* Implemented and administered Data center Cassandra Clusters based on the knowledge of architecture and data modeling for Cassandra applications.
* Worked on performance tuning of cluster using Cassandra Configuration file and JVM Parameters.
* Configured internode communication between Cassandra nodes and client using SSL encryption.
* Developed Spark Application by using Python (Pyspark).
* Developed and implemented API services using Python in spark.
* Experience developing Splunk queries and dashboards targeted at understanding application performance and capacity analysis.
* Worked on performing data standardization using PIG scripts.
* Worked in tuning Hive and Pig to improve performance and solved performance issues in Hive and Pig scripts with understanding of Joins, Group and aggregation and how does it translate to Map Reduce jobs.
* Created Partitions, Buckets based on State to further process using Bucket based Hive joins.
* Worked on Performance optimization of Spotfire applications, Spotfire server configuration.
* Extensively worked with Cloudera Distribution Hadoop, CDH 5.x, CDH4.x .
* Worked on NoSQL databases including HBase, MongoDB, and Cassandra.
* Responsible for continuous monitoring and managing Elastic MapReduce cluster through AWS console.
* Implemented Spark using Scala and Spark SQL for faster testing and processing of data.
* Knowledge on handling Hive queries using Spark SQL that integrate with **Spark** environment.
* Worked on migrating MapReduce programs into Spark transformations using Spark and Scala.
* Configured Spark Streaming to receive real time data from the Kafka and store the stream data to HDFS.
* Experience in managing and reviewing Hadoop log files.
* Tested raw data and executed performance scripts using MRUnit.

**Environment:** Hadoop, Hive, Talend, Map Reduce, Pig, SQOOP, Splunk, CDH5, Python,Cloudera Manager CM 5.1.1,HDFS, Pig, DB2, Sqoop, Oozie, Putty, Java.

#### hadoop developer/TIME WARNER CABLE-NC

##### Sept2014- April 2016

* Responsible for building scalable distributed data solutions using Hadoop.
* Written multiple MapReduce programs in Java for Data Analysis.
* Wrote MapReduce job using Pig Latin and Java API.
* Performed performance tuning and troubleshooting of MapReduce jobs by analyzing and reviewing Hadoop log files.
* Collected the logs from the physical machines and the OpenStack controller and integrated into HDFS using Flume.
* Load data from various data sources into HDFS using Kafka.
* Designed and presented plan for POC on impala.
* Experienced in migrating HiveQL into Impala to minimize query response time.
* Implemented Avro and parquet data formats for apache Hive computations to handle custom business requirements.
* Worked on Sequence files, RC files, Map side joins, bucketing, partitioning for Hive performance enhancement and storage improvement.
* Performed extensive Data Mining applications using HIVE.
* Responsible for performing extensive data validation using Hive.
* Sqoop jobs, PIG and Hive scripts were created for data ingestion from relational databases to compare with historical data.
* Utilized Storm for processing large volume of datasets.
* Used Pig as ETL tool to do transformations, event joins, filter and some pre-aggregations.
* Used Visualization tools such as Powerview for excel, Tableau for visualizing and generating reports.
* Setup Hadoop cluster on Amazon EC2 using whirr for POC.
* Implemented test scripts to support test driven development and continuous integration.
* Involved in story-driven agile development methodology and actively participated in daily scrum meetings.

**Environment:**Hadoop, Map Reduce, HDFS, Pig, Hive, Sqoop, Flume, Oozie, Java, Linux, Maven, Teradata, Zookeeper, SVN, autosys, Hbase.

#### hadoop developer/american family insurance - wi

##### July 2013 - Sep 2014

* Worked on writing transformer/mapping Map-Reduce pipelines using Java.
* Involved in creating Hive Tables, loading with data and writing Hive queries which will invoke and run Map Reduce jobs in the backend.
* Involved in loading data into **HBase** using **HBase Shell, HBase Client API**, **Pig** and **Sqoop.**
* Designed and implemented Incremental Imports into Hive tables.
* Deployed an **Apache Solr** search engine server to help speed up the search of the government cultural asset.
* Involved in collecting, aggregating and moving data from servers to HDFS using Apache Flume.
* Written Hive jobs to parse the logs and structure them in tabular format to facilitate effective querying on the log data.
* Involved in creating Hive tables, loading with data and writing hive queries that will run internally in map reduce way.
* Experienced in managing and reviewing the Hadoop log files.
* Migrated ETL jobs to Pig scripts do Transformations, even joins and some pre-aggregations before storing the data onto HDFS.
* Implemented the workflows using Apache Oozie framework to automate tasks.
* Worked with Avro Data Serialization system to work with JSON data formats.
* Worked on different file formats like Sequence files, XML files and Map files using Map Reduce Programs.
* Created and maintained Technical documentation for launching HADOOP Clusters and for executing pig Scripts.

**Environment:** Hadoop, Big Data, HDFS, MapReduce, Sqoop, Oozie, Pig, Hive, hbase,Flume, LINUX, Java, Eclipse, Cassandra, Hadoop Distribution of Cloudera., PL/SQL, Windows NT, UNIX Shell Scripting, Putty and Eclipse.

#### sr java/j2ee developer/WELLCARE - FL

##### April 2012- July 2013

* Coordinated with mainframe developers to understand, preserve and migrate legacy application functionality.
* Worked with technologies like Core Java, J2EE, Hibernate, and Oracle.
* Created Framework for Integration Testing (FIT) based test cases for integration testing.
* Tested application functions like handling inputs, processing data, and interaction with database.
* Wrote customization codes for FIT to adapt to the application requirements.
* Implemented the Struts frame work with MVC architecture.
* Created new JSP's for the front end using HTML, Java Script, JQuery, and Ajax.
* Developing JSP pages and configuring the module in the application.
* Developed the presentation layer using JSP, HTML, CSS and client-side validations using JavaScript.
* Involved in designing, creating, reviewing Technical Design Documents.
* Developed DAOs (Data Access Object) using Hibernate as ORM to interact with DBMS - Oracle.
* Worked on MQ for messaging.
* Applied J2EE design patterns like Business Delegate, DAO, Singleton etc.
* Deployed and tested the application using Tomcat web server.
* Using java scripts did client-side validation.
* Involved in developing DAO's using JDBC.
* Involved in coding, code reviews, JUnit testing, Prepared and executed Unit Test Cases.
* JBOSS for application deployment and MySQL for database
* Worked with QA team in preparation and review of test cases.
* JUnit was used for unit testing for the integration testing tool.
* Writing SQL queries to fetch the business data using Oracle as database.
* Identified bugs in the migrated application by running test cases and using Eclipse IDE.
* Worked on JavaScript, JSP, and Servlets as a web application replacement for mainframe front end.
* Provided Hibernate configuration file and mapping files, also involved in the integration of Struts with Hibernate libraries.
* Wrote Ant scripts and handled deployments on Weblogic server at test, stage levels.
* Analyzed, created and proposed remediation measures to fix the bugs in the application.
* Used log4j for logging monitoring errors and exceptions all across the application.
* Analyzed database table relationships to understand legacy application.
* Coordinated with Database Administrators to analyze and optimize the application load on database.
* Identified, analyzed, and re-mediated performance issues in application

#### java/j2ee developer/ ORACLE INDIA PVT. LTD

##### Sep 2010–March 2012

* Responsible for the systems design, architecture, implementation and integration with various technologies like Spring Integration, Web Services, Oracle Advanced Queues and WMQ’s.
* Implemented framework Spring 3.05 and Spring Integration 2.0.5 upgrades.
* Used OSGi container framework to install bundles (modules) developed using Spring and Spring Integration.
* Worked on UI development using JSP on Struts and Spring MVC Frameworks.
* Develop and test PL/SQL stored procedures.
* Developed DAOs (Data Access Object) and DOs (Data Object) using Hibernate as ORM to interact with DBMS - Oracle
* Developed modules that integrate with web services that provide global information
* Used Web services to communicate and integrate within different system.
* Used Log4j for logging the application, log of the running system to trace the errors and certain automated routine functions.
* Build applications using MAVEN.
* Learn and implement upcoming technologies and software updates.
* Participates in high level and low-level design discussions.
* Worked as Web Dynpro Java developer and developed custom applications, and creating the Portal screens.
* Create and access JAX-RS REST and JAX-WS SOAP based Web Services, to access internal and secured applications inside and outside the intranet.
* Responsible for development, testing, deploying and maintaining Spring Integration projects on Virgo (OSGi Container).
* Modified and enhanced the Division of Disability Determinations case processing application.
* Designed and developed functions used by Division of Disability analysts when adjudicating a case and managing caseloads in Java/Oracle environment.
* Diagnoses and resolves application bugs and monitoring logs files across the application.
* Reviews code from other developers to ensure it meets design standards.

#### JAVA developer/ Pratian technologies

##### May 2009 – Sep 2010

* Analysis, design and development of Application based on J2EE using Struts and Hibernate.
* Involved in interacting with the Business Analyst and Architect during the Sprint Planning Sessions.
* Implemented Point to Point JMS queues and MDB's to fetch diagnostic details across various interfaces.
* Worked with WebSphere business integration technologies as WebSphere MQ and Message Broker 7.0 (Middleware tools) on Various Operating systems.
* Perform incident resolution for WebSphere Application Server, WebSphere MQ, IBM Message broker, Process and Portal server.
* Configured WebSphere resources including JDBC providers, JDBC data sources, connection pooling, and JavaMail sessions. Deployed Session and Entity EJBs in WebSphere
* Developed Rich user interface using RIA, HTML, JSP, JSTL, JavaScript, JQuery, CSS, YUI, AUI using Liferay portal.
* Worked on new Portal theme for the website using Liferay and customize for the look and feel.
* Experience in all aspects of Angular JS like "Routing", "modularity", "Dependency injection", "Service calls" and "Custom directives" for development of single page applications.
* Hibernate was used for Object Relational mapping with Oracle database.
* Involved in developing the user interface using Struts tags, core java development involving concurrency/multi-threading, struts-hibernate integration, database operation tasks.
* Integrated Struts and Hibernate ORM framework for persistence and used HibernateDAOSupport with Hibernate Template to access the data.
* Implemented core java functionalities like collections, multi-threading, Exception handling.
* Involved in unit testing using JUnit 4.
* Performed Code optimization and rewriting the database queries to resolve performance related issues in the application.
* Implemented DAO classes which in turn use JDBC to communicate and retrieve the information from DB2 database sitting on Linux/UNIX server.
* Wrote selenium test scripts for regressive testing the application.
* Involved in fixing the production defects for the application.
* Involved in writing SQL, PL/SQL stored procedures using PL/SQL Developer.
* Used Eclipse as IDE for application development.
* Used Weblogic 10.0 for application deployment.
* Supported production deployments and validated the flow of the application after each deployment
* Implemented Action classes in Struts 2.0 framework.
* Used PL/SQL for cresting triggers, packages, procedures and functions.
* Implementation of business logic in Spring Framework with Hibernate in the back-end.
* Wrote custom tags in JSP used to modularize and streamline web page codes.
* Deployed the Application on Web Logic Server