|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Sandeep Jarugula** | | | | | | | | |
| **Contact #** | **Alternative contact #** | | **Email Address** | | **Skype** | | | **LinkedIn** |
| 978- 212-9137 | N/A | | [itssandeep.coolguy@gmail.com](mailto:itssandeep.coolguy@gmail.com) | | [itssandeep.coolguy@gmail.com](mailto:itssandeep.coolguy@gmail.com) | | | [www.linkedin.com/in/](http://www.linkedin.com/in/)sandeep-jarugula-496713577/ |
| **Current Location:** | | | Arlington, VA | | | | | |
| **Willing to relocate? If so, please provide the preferred location:** | | | Yes | | | | | |
| **Willing to attend in-person interview?** | | | No | | | | | |
| **INTERVIEW AVAILABILITY**  **(Dates listed must be 48 hours after candidate submission)** | | | | | | | | |
| **Time Slot 1 (Date/Time)** | | | | **Time Slot 2 (Date/Time)** | | **Time Slot 3 (Date/Time)** | | |
| Friday(09-13-2019)/  11:00AM-12:00PM(EST) | | | | Friday(09-13-2019)/  01:00PM-02:00PM(EST) | | Monday(09-16-2019)/  12:00PM-01:00PM(EST) | | |
| **Time Slot 4 (Date/Time)** | | | | **Time Slot 5 (Date/Time)** | |  | | |
| Monday(09-16-2019)/  02:00PM-03:00PM(EST) | | | | Tuesday(09-17-2019)/  01:00PM-02:00PM(EST) | |  | | |
| Availability for new Project: | | | | Immediate | |  | | |
| **Notice Period/ LWD on last project** | | | | 2 week | |  | | |
| **Interviews/ Offers in Pipeline** | | | | No | |  | | |
| **LIST ALL EMPLOYMENT FOR THE PAST 7 YEARS** | | | | | | | | |
| **Dates of Employment**  **(Month/Year – Month/Year)** | | **Parent Employer Name /Project Company Name**  **i.e. ABC Staffing / Project at XYZ Client Name** | | | | | **Location** | |
| Feb’ 18 - Present | | NRECA | | | | | Arlington, VA | |
| Mar’ 16 – Jan’ 18 | | Catalyte | | | | | Chicago, IL | |
| Jan’ 14 – Feb’ 16 | | T - mobile | | | | | Bellevue, WA | |
| **Supplier Synopsis** | | | | | | | | |
| **Mandatory Skills**  **(As listed in JD)** | | **# of Years Experience** | | | **Candidate’s relevant hands-on experience** | | | |
| Spark Hive | | 6+ Years | | | Extensive experience in installing, configuring and using **Big** **Data** **ecosystem** **components** like **Hadoop** **MapReduce, HDFS, Sqoop, Pig, Hive, Impala, Spark** and **Zookeeper** | | | |
| Big Data | | 4 Years | | | Extensive experience in installing, configuring and using Big Data ecosystem components . | | | |
| Python | | 4 Years | | | Developing and implementing **Apache** **NIFI** across various environments, written **QA scripts** in **Python** for **tracking** **files** | | | |
| Pig | | 6+ Years | | | Importing and exporting data into HDFS and HIVE, PIG using Sqoop. | | | |
| **Nice to Have**  **(As listed in JD)** | | **# of Years Experience** | | | **Candidate’s relevant hands-on experience** | | | |
| N/A | | N/A | | | N/A | | | |
|  | |  | | |  | | | |
| **Additional comments:** | | Very Much Interested Consultant. | | | | | | |

**Summary**

* Overall **9+ years** of working experience as a **Sr. Big Data/Hadoop Developer** in designed and developed various applications like **Big data, Hadoop, Java/J2EE** open-source technologies.
* Experience with all stages of the **SDLC** and **Agile** **Development** **model** right from the requirement gathering to Deployment and production support.
* Good Knowledge in **Amazon AWS concepts** like **EMR** and **EC2** **web** **services** which provides fast and efficient processing
* Extensive experience in installing, configuring and using **Big** **Data** **ecosystem** **components** like **Hadoop** **MapReduce, HDFS, Sqoop, Pig, Hive, Impala, Spark** and **Zookeeper**
* Expertise in using **J2EE** **application** **servers** such as **IBM** **Web** **Sphere,** **JBoss** and **web** **servers** like **Apache** **Tomcat**.
* Experience in different **Hadoop** **distributions** like **Cloudera (CDH3 & CDH4)** and **Horton** **Works** **Distributions (HDP).**
* Experience in developing **web** **services** with **XML** based protocols such as **SOAP, Axis, UDDI** and **WSDL.**
* Solid understanding of **Hadoop MRV1** and **Hadoop MRV2 (or)** **YARN** **Architecture**.
* Good knowledge on various **API** **Managers** like **Azure**.
* Very well verse in writing and deploying **Oozie** **Workflows** and **Coordinators**.
* Highly skill in integrating **Amazon Kinesis** streams with **Spark streaming applications** to build long running real-time applications.
* Good working experience on using **Sqoop** to import data into **HDFS** from **RDBMS** and **vice-versa**.
* Extensive experience in **Extraction, Transformation** and **Loading (ETL)** of data from multiple sources into **Data** **Warehouse** and **Data** **Mart**.
* Hands on experience in configuring and working with **Flume** to load the data from multiple sources directly into **HDFS**.
* Good knowledge in implementing various **data** **processing** **techniques** using **Apache** **HBase** for handling the data and formatting it as required.
* Experience with **Java** **web** **framework** **technologies** like **Apache** **Camel** and **Spring** **Batch**.
* Hands on experience working with **databases** like **Oracle, SQL Server** and **MySQL**.
* Great knowledge of working with **Apache** **Spark** **Streaming** **API** on **Big** **Data** **Distributions** in an active **cluster** **environment**.
* Proficiency in developing secure enterprise **Java** **applications** using **technologies** such as **Maven, Hibernate, XML, HTML, CSS Version Control Systems**
* Developing and implementing **Apache** **NIFI** across various environments, written **QA scripts** in **Python** for **tracking** **files**.
* Excellent understanding of **Hadoop** and underlying **framework** including storage management.
* Good Knowledge and experience of functionality of **NoSQL** **DB** like **Cassandra** and **Mongo DB.**
* Extensive experience in building and deploying applications on **Web/Application** **Servers** like **Web logic, Web sphere**, and **Tomcat**.
* Experience in using **ANT** for building and deploying the projects in servers and also using **Junit** and **log4j** for debugging.
* Have good experience, excellent communication and interpersonal skills which contribute to timely completion of project deliverable well ahead of schedule.

**Technical Skills:**

* **Hadoop/Big Data Technologies:** Hadoop 3.0, HDFS, MapReduce, HBase 1.4, Apache Pig, Hive 2.3, Sqoop 1.4, Apache Impala 2.1, Oozie 4.3, Yarn, Apache Flume 1.8, Kafka 1.1, Zookeeper
* **Cloud Platform:** Amazon AWS, EC2, EC3, MS Azure, Azure SQL Database, Azure SQL Data Warehouse, Azure Analysis Services, HDInsight, Azure Data Lake, Data Factory
* **Hadoop Distributions:** Cloudera , Hortonworks, MapR
* **Programming Language:** Java, Scala, Python 3.6, SQL, PL/SQL, Shell Scripting, Storm 1.0, JSP, Servlets
* **Frameworks:** Spring 5.0.5, Hibernate 5.2, Struts 1.3, JSF, EJB, JMS
* **Web Technologies:** HTML, CSS, JavaScript, JQuery 3.3, Bootstrap 4.1, XML, JSON, AJAX
* **Databases:** Oracle 12c/11g, SQL
* **Operating Systems:** Linux, Unix, Windows 10/8/7
* **IDE and Tools:** Eclipse 4.7, NetBeans 8.2, IntelliJ, Maven
* **NoSQL Databases:** HBase 1.4, Cassandra 3.11, MongoDB, Accumulo
* **Web/Application Server:** Apache Tomcat 9.0.7, JBoss, Web Logic, Web Sphere
* **SDLC Methodologies:** Agile, Waterfall
* **Version Control:** GIT, SVN, CVS

**Education Details:**

B.Tech in Mechanical Engineering from Vignan University, Andhra Pradesh, India (2010).

**Professional Summary :**

**NRECA - Arlington, VA**

**Feb’ 18 - Present**

**Sr. Big Data/Hadoop Developer**

**Responsibilities:**

* As a **Big Data/Hadoop Developer,** I worked on **Hadoop eco-systems** including **Hive, HBase, Oozie, Pig,** **Zookeeper, Spark Streaming** and so on with **MapR** **distribution**.
* Worked on **Big Data infrastructure** for batch processing as well as real-time processing. Responsible for building scalable distributed data solutions using **Hadoop**.
* Primarily involved in **Data Migration** process using **Azure** by integrating with **Github** repository and **Jenkins**.
* Utilized **Agile** and **Scrum** **Methodology** to help manage and organize a team of developers with regular code review sessions.
* Hands on experience in installing, configuring, and using Hadoop ecosystem components like **MapReduce, Spark, HDFS, HBase, Oozie, Hive, Sqoop, Pig, and Flume.**
* Involved with all the phases of **Software Development Life Cycle (SDLC) methodologies** throughout the project life cycle.
* Monitored workload, job performance and capacity planning using **Cloudera Manager.**
* Specified the cluster size, allocating Resource pool, Distribution of **Hadoop** by writing the specification texts in **JSON** File format.
* Developed a **JDBC** connection to get the data from **Azure** **SQL** and feed it to a **Spark** **Job**.
* Configured **Spark** streaming to receive real time data from the **Apache Kafka** and store the stream data to **HDFS** using **Scala**.
* Developed the **Sqoop** **scripts** to make the interaction between **Hive** and vertica Database.
* Involved in converting **Hive/SQL** **queries** into Spark transformations using **Spark** **RDDs**, and **Scala**.
* Deployed the application in **Hadoop** **cluster** mode by using **spark** submit **scripts**.
* Queried and analyzed data from **Cassandra** for quick searching, sorting and grouping through **CQL.**
* Worked on the large-scale **Hadoop** **YARN** **cluster** for distributed data processing and analysis using **Spark,** **Hive**.
* Importing and exporting data into HDFS and HIVE, PIG using Sqoop.
* Implemented a distributed messaging queue to integrate with **Cassandra** using **Apache** **Kafka** and **Zookeeper**.
* Implemented various **Hadoop** **Distribution** **environments** such as **Cloudera** and **Hortonworks**.
* Implemented monitoring and established best practices around usage of **elastic** **search**
* Worked on **Apache Nifi** as **ETL** tool for batch processing and real time processing.
* Moved **Relational Database** data using **Sqoop** into **Hive** **Dynamic** partition tables using staging tables.
* Upgraded the **Hadoop** **Cluster** from **CDH3** to **CDH4,** setting up High Availability **Cluster** and integrating **HIVE** with existing applications.
* Captured the data logs from web server into **HDFS** using **Flume** for analysis.
* Involved in developing code to write canonical model **JSON** records from numerous input sources to **Kafka** **Queues**.
* Involved in identifying job dependencies to design **workflow** for **Oozie** **&** **YARN** resource management.
* Used **GitHub** as repository for committing code and retrieving it and **Jenkins** for continuous integration.
* Reviewed the **HDFS** usage and system design for future scalability and fault-tolerance.
* Loaded and transformed large sets of **structured, semi structured** and **unstructured** data in various **formats** like **text, zip, XML** and **JSON**.
* Utilized **Oozie** **workflow** to run **Pig** and **Hive** Jobs Extracted files from **Mongo** **DB** through **Sqoop** and placed in **HDFS** and processed.
* Worked on importing data from **HDFS** to **MYSQL** **database** and vice-versa using **SQOOP**.

**Environment:** Hadoop 3.0, Hive 2.3, HBase 1.2, Oozie 4.3, Pig 0.17, Zookeeper 3.4, Spark 2.4, Azure, Jenkins 2.15, Agile, JSON, Scala 2.12, Apache Kafka 2.0, Cassandra 3.11, MYSQL 8.0, Sqoop 1.4

**Catalyte - Chicago, IL**

**Mar’ 16 – Jan’ 18**

**Hadoop Developer**

**Responsibilities:**

* Worked on analyzing **Hadoop** **cluster** and different **big data analytic** tools including **MapReduce**, **Hive** and **Spark**.
* Installed and configured the **Hadoop cluster** using the **Cloudera's** **CDH distribution** and monitored the cluster performance using the **Cloudera** **Manager**.
* Involved in Installing, Configuring **Hadoop Eco-System, Cloudera Manager** using **CDH3, CDH4** **Distributions**.
* Worked on **S3** **buckets** on **AWS** to store Cloud Formation Templates and worked on **AWS** to create **EC2** instances.
* Used **Sqoop** to import the data from **RDBMS** to **Hadoop** **Distributed** **File** **System (HDFS)** and later analyzed the imported data using **Hadoop** **Components**
* Used **Oozie** **scripts** for deployment of the application and perforce as the secure versioning software.
* Worked on **No-SQL** **databases** like **Cassandra,** **MongoDB** for **POC** purpose in storing images and URIs.
* Involved in creating **Hive** **tables**, then applied **HiveQL** on those tables, this will invoke and run MapReduce jobs automatically.
* Involved in loading and transforming large sets of **Structured, Semi-Structured** and **Unstructured** **data** and analyzed them by running **Hive** **queries** and **Pig** **scripts**
* Highly skilled in integrating **Kafka** with **Spark** **streaming** for high-speed data processing
* Implemented **POC** to migrate **MapReduce** programs into Spark transformations using **Spark** and **Scala**.
* Developed different kind of custom filters and handled pre-defined filters on **HBase** data using **API**.
* Used **Enterprise Data Warehouse database** to store the information and to make it access all over organization.
* Exported of result set from **HIVE** to **MySQL** using **Sqoop** export tool for further processing.
* Collected and aggregating large amounts of log data using **Flume** and staging data in **HDFS** for further analysis.
* Worked with **Avro** **Data** **Serialization** **system** to work with **JSON** data formats.
* Involved in creating **Hive** **tables**, loading with data and writing hive queries which runs internally in **MapReduce** way.
* Worked on **Spark** for in memory commutations and comparing the **Data** **Frames** for optimizing performance.
* Created **web** **services** request-response mappings by importing source and target definition using **WSDL** file.
* Developed custom **UDF's** to generate unique key for the use in **Apache** **pig** transformations.
* Created conversion scripts using **Oracle SQL queries,** functions and stored procedures, test cases and plans before **ETL** migrations.
* Developed **Shell** **scripts** to read files from edge node to ingest into **HDFS** partitions based on the file naming pattern.

**Environment:** Hadoop 3.0, Hive 2.3, Spark 2.4, AWS, Sqoop 1.4, Oozie 4.3, Cassandra 3.11, MongoDB 4.0, Kafka 2.0, MySQL 8.0, MapReduce, Scala 2.12, Pig 0.17

**T - mobile - Bellevue, WA**

**Jan’ 14 – Feb’ 16**

**Sr. Spark/Hadoop Developer**

**Responsibilities:**

* Extensively migrated existing architecture to **Spark** **Streaming** to process the live streaming data.
* Involved in writing the shell scripts for exporting log files to **Hadoop** **cluster** through automated process.
* Worked with multiple teams to provision **AWS** **infrastructure** for development and production environments.
* Gathered the **business** **requirements** from the Business Partners and Subject Matter Experts.
* Created custom new columns depending up on the use case while ingesting the data into **Hadoop** **Lake** using **pyspark**.
* Designed and developed applications in **Spark** using **Scala** to compare the performance of **Spark** with **Hive** and **SQL/Oracle.**
* Developed time driven automated **Oozie** **workflows** to schedule **Hadoop** **jobs**.
* Analyzed **Cassandra** **database** and compare it with other open-source **NoSQL** **databases** to find which one of them better suites the current requirement.
* Designed and implemented **map** **reduce** jobs to support distributed processing using **java, Hive,** **Scala** and **Apache Pig.**
* Developed web services using **Restful** **technology** to support **JSON** and **XML**.
* Involved in collecting, aggregating and moving data from servers to **HDFS** using **Apache** **Flume**.
* Involved in converting **Hive/SQL queries** into **Spark** **transformations** using **Spark RDD's.**
* Worked with **NoSQL** **databases** like **HBase** in creating **HBase** tables to load large sets of **semi**-**structured**.
* Worked with **Apache Nifi** for managing the flow of data from sources through automated data flow.
* Created, modified and executed **DDL** and **ETL** **scripts** for **De-normalized** tables to load data into **Hive** and **AWS** **Redshift** **tables**.
* Developed Date, Currency type conversions, and Sales **Spark UDF's** to reduce the load on **Tableau.**
* Managed and review data backups, Manage and review **Hadoop** **log** **files** **Hortonworks** **Cluster**.
* Created continuous integration builds using **Ant, Jenkins,** and **Gradle** and deployed on the **WebSphere** **app** **server**.
* Used **SVN** as **version** **control** **system** for the source code and project documents.
* Extracted files from **CouchDB** through **Sqoop** and placed in **HDFS** and processed.
* Implemented a distributed messaging queue to integrate with **Cassandra** using **Apache** **Kafka** and **Zookeeper**.
* Participated in development/implementation of **Cloudera** **Impala** **Hadoop** **environment**.
* Used Test driven approach for developing the application and Implemented the unit tests using Python **Unit** **Test** **framework**
* Involved in **Ad Hoc** stand up and architecture meetings to set up daily priorities and track the status of work as a part of highly **agile** **work** **environment**.

**Environment:** Spark 2.0, Hadoop 2.3, AWS, pyspark, Scala 1.9, SQL, Oracle 11g, Hive 2.1, Cassandra 3.0, Oozie, MapReduce, java, Apache Pig 0.16, JSON, XML, HBase, Jenkins, Ant, Gradle, Zookeeper 2.9

**GAP - SFO, CA**

**Oct’ 12 – Dec’ 13**

**Java/J2EE Developer**

**Responsibilities:**

* Applied **J2EE** Design Patterns such as **Factory, Singleton,** and **Business delegate,** **DAO, Front** Controller Pattern and **MVC**.
* Designed Rich user Interface Applications using **JavaScript, CSS,** **HTML** and **AJAX** and developed web services by using **SOAP.**
* Implemented modules using **Java APIs, Java collection, Threads**, **XML**, and integrating the modules.
* Successfully installed and configured the **IBM WebSphere Application server** and deployed the business tier components using **EAR** **file**.
* Designed and developed business components using **Session** and **Entity** **Beans** in **EJB**.
* Used **Log4j** for Logging various levels of **information** like **error, info, debug** into the log files.
* Developed integrated applications and light weight component using **spring** **framework** and **IOC** features from **spring** **web** **MVC** to configure application context for spring bean factory.
* Used **JDBC** prepared statements to call from **Servlets** for **database** **access**.
* Used **Angular JS** to connect the **web** **application** to back-end **APIs,** used **RESTFUL** methods to interact with several **API's.**
* Documented all stages of the project, generated documentation using **Java API Documentation** Generator.
* Involved in build/deploy applications using **Maven** and integrated with **CI/CD server Jenkins**.
* Implemented **JQuery** features to develop the dynamic queries to fetch the data from database.
* Involved in developing module for transformation of files across the remote systems using **JSP** and **servlets**.
* Worked on Development bugs assigned in **JIRA** for Sprint following agile process.
* Used **ANT** **scripts** to fetch, build, and deploy application to development environment.
* Extensively used **JavaScript** to provide the users with interactive, Speedy, functional and more usable user interfaces.
* Configured and used **JBoss** **application** **server** for creating Data source and enabling **JNDI** for beans and database lookups.
* Implemented **MVC** **architecture** using **Spark**, **JSP** **&** **Enterprise** **Java** **Beans**.
* Used **AJAX** and **JSON** to make asynchronous calls to the project server to fetch data on the fly.
* Worked on creating a test harness using **POJOs** which would come along with the installer and test the services every time the installer would be run.
* Used **JNDI** to perform lookup services for the various components of the system.
* Deployed the application and tested on **JBoss** **Application** **Server**. Collaborated with Business Analysts during design specifications.
* Developed **Apache** **Camel** middleware routes, **JMS** endpoints, spring service endpoints and used Camel free marker to customize **REST** responses.

**Environment:** J2EE, MVC, JavaScript, CSS, HTML, AJAX, Angular JS, Maven, Jenkins, JQuery, ANT, JBoss, AJAX

**ITC Infotech - Bangalore, India**

**Jun’ 10 – Sep’ 12**

**Java Developer**

**Responsibilities:**

* Designed **HTML** prototypes, visual interfaces and interaction of Web-based design.
* Designed and developed a business tiers using **EJBs** and Used **Session** **Beans** to encapsulate the Business Logic.
* Involved in the configuration of **Spring** **MVC** and Integration with **Hibernate**.
* Used **Eclipse IDE** to configure and deploy the application onto **WebLogic** **application** **server** using **Maven** build scripts to automate the build and deployment process
* Designed **CSS** based page layouts that are **cross-browser** compatible and standards-compliant.
* Used **spring** **framework** for Dependency Injection and **JDBC** connectivity.
* Created data source and connection pools in **Web** **Logic** and deployed applications on the server
* Developed **XML** and **XSLT** pages to store and present data to the user using parsers.
* Developed **RESTful** **Web** **services** client to consume **JSON** messages
* Developed a **web application** for developers to generate their own **REST API** calls in, **JAVA,** and **Python.**
* Implemented **business** **logic** with **POJO** using multithreading and design patterns.
* Created test cases for **DAO** **Layer** and service layer using **JUNIT** and bug tracking using JIRA.
* Used **Struts, Front Controller** and **Singleton patterns**, for developing the action and **servlet** **classes,** Involved in
* Built **Web** **pages** that are more user-interactive using **AJAX, JavaScript**, and **React.js.**
* Developed remote integration with third party platforms by using **RESTful** **web** **services**.
* Worked on web-based reporting system with **HTML, JavaScript,** and **JSP**.
* Wrote **REST** **Web** **Services** for sending and getting data from the external interface
* Implemented the application using **Spring Boot Framework** and handles the security using spring security.
* Researched and Executed of **JavaScript** Frameworks, including **Angular JS** and **Node JS**.
* Developed stored procedures and triggers using **PL/SQL** to calculate and update the tables to implement business logic
* **Used Subversion (SVN)** as the configuration management tool to manage the code repository.
* Used **Maven** as the build tool and **Tortoise** **SVN** as the Source version controller
* Created an **XML** configuration file for **Hibernate** for Database connectivity.
* Used **jQuery** for basic animation and end user screen customization purposes.
* Used **GIT** as Version Control System performed Module and Unit Level Testing with **JUnit** and log4j
* Participated in **Unit** **testing** and functionality testing for tracking errors and debugging the code.

**Environment:** HTML, MVC, Hibernate, Eclipse, Maven, CSS, **Java,** Junit, AJAX, JavaScript, Hibernate, JQuery.