|  |
| --- |
| **Karunakar Revunuru**  **Big data Developer/Engineer**  **Karunakar.rr@outlook.com**  **970-614-8039** |

****

SUMMARY

* I have **8+ years** of experience in various IT related technologies, which includes **4 years** of hands-on experience in Big Data technologies. Implementation and extensive working experience in wide array of tools in the Big Data Stack like **HDFS, Yarn, Spark, MapReduce, Hive, Flume, Oozie, Sqoop, Kafka, Zookeeper and HBase** with Cloudera and Hortonworks platform in Financial, Retail and Health-care sector.
* Experience in importing data from existing relational databases (Oracle, MySQL and Teradata) that provide SQL interfaces using **Sqoop.**
* Hands on experience in **Avro, Parquet, ORC** files and Combiners, Counters, Dynamic Partitions, bucketing for best practice and performance improvement, Worked on different Compression Codecs (**GZIP, SNAPPY, BZIP**).
* Designed **HIVE** queries to perform data analysis, data transfer and table design to load data into Hadoop environment.
* Extensive experience on importing and exporting data using stream processing platforms like Flume and **Kafka**
* Experience in data workflow scheduler **Zoo-Keeper** and **Oozie** to manage Hadoop jobs by Direct Acyclic Graph (DAG) of actions with the control flows.
* Integrated BI tool like Tableau with Impala and analyzed the data.
* Experience with NoSQL databases like **HBase, Cassandra and MongoDB.**
* Experience in collecting the log data from different sources (webservers and social media) using **Flume, Kafka** andstoringin **HDFS** to perform the MapReduce jobs.
* Involved in converting Hive/SQL queries into Spark transformations using Spark RRD's and Scala.
* Implemented Spark using Scala and utilized data frames and Spark SQL API for faster processing of Data.
* Capable at using **AWS utilities** such as **EMR, S3** and Cloud watch to run and monitor Hadoop/Spark jobs on AWS.
* Knowledge in creating dashboards with the help of business intelligence tool such as Tableau.
* Involved in Agile methodologies, daily scrum meetings, spring planning.

EDUCATIONAL QUALIFICATION

* **Bachelor of Technology** in **Electronics and Communication Engineering** from JNTUA,2010.

CERTIFICATIONS

**Cloudera Certified Spark and Hadoop Developer**

**IBM Certified Developer -Apache Spark**

TECHNICAL SKILLS

|  |  |
| --- | --- |
| **Bigdata Ecosystem** | HDFS and Map Reduce, Hive, Impala, HUE, Oozie,  Zookeeper, Apache Spark, Apache STORM, Apache Kafka, Pig, Sqoop, Flume |
| **NoSQL Databases** | HBase, Cassandra, and MongoDB |
| **Hadoop Distributions** | Cloudera, Hortonworks |
| **Programming languages** | Java, SCALA, Pig Latin, HiveQL. |
| **Scripting Languages** | Shell Scripting |
| **Databases** | MySQL, oracle, Teradata, DB2 |
| **Build Tools** | Maven, Ant |
| **Reporting Tool** | Tableau |
| **Version control Tools** | SVN, Git, GitHub |
| **Cloud** | AWS, Azure |
| **App/Web servers** | WebSphere, WebLogic |
| **Operating Systems** | WINDOWS 10/8 |
| **Development IDEs** | Eclipse IDE, Python(IDLE)  Packages Microsoft Office, putty, MS Visual Studio. |

PROFESSIONAL EXPERIENCE

**Role: Big Data Developer/Engineer May 2017– Till Now**

**Client: DaVita, Denver, CO.**

**Responsibilities:**

* Developed **Sqoop** scripts for importing and exporting data into HDFS and Hive.
* Used **Zookeeper** to store offsets of messages consumed for a specific topic and partition by a specific Consumer Group in **Kafka**.
* Extract Real time feed using **Kafka** and **Spark Streaming** and convert it to RDD and process data in the form of Data Frame and save the data as Parquet format in HDFS.
* Used **Spark** and **Spark-SQL** to read the parquet data and create the tables in hive using the Scala API.
* Improving performance and optimizing of existing algorithms in Hadoop using Spark Context, Spark-SQL, Data Frames and Pair RDD's.
* Involved in creating **Hive** Tables, loading with data and writing Hive queries which will invoke

and run MapReduce jobs in the backend.

* Used **Spark Streaming** APIs to perform transformations and actions on the fly for building common learner data model which gets the data from **Kafka** in Near real time and used Apache NIFI to ingest persist it to **HBase**.
* Worked with **HBase** in creating HBase tables to load large sets of semi structured data coming from various sources.
* Hands on experience on AWS infrastructure services Amazon Simple Storage Service (**Amazon S3**) and Amazon Elastic Compute Cloud (**Amazon EC2**).
* load consumer response data in **AWS S3** bucket into Hive external tables in HDFS location to serve as feed for **tableau** dashboards.
* Worked in **Agile** and used **JIRA** for maintain the stories about project.

**Environment:** Hadoop, Map Reduce, Hive, Spark, Oracle, GitHub, Tableau, Unix, Cloudera, Kafka, Sqoop, Scala, NIFI, HBase, amazon EC2, S3.

**Role: Big Data Developer/Engineer Jan 2016 – May 2017**

**Client: Arcellor Mittal, Burns Harbor, IN.**

**Responsibilities:**

* Worked with **SQOOP** import and export functionalities to handle large data set transfer between DB2 database and HDFS.
* Created hive queries for extracting data and sending them to clients.
* Transformation and Analysis in **Hive,** Parsing the raw data using Map reduce and **SPARK**.
* Created **SCALA** programs to develop the reports for Business users.
* Worked on ingesting data from different sources, Followed agile methodology during project delivery.
* Proactively involved in ongoing maintenance, support and improvements in Hadoop cluster.
* Experience in working with NoSQL database **HBase** in getting real time data analytics.
* Familiar with **AWS** **Components** like **EC2, S3.**
* Configured various big data workflows to run on the top of Hadoop using Oozie and these   
  workflows comprise of heterogeneous jobs like Pig, Hive, Sqoop and MapReduce.
* Knowledge of Code **Hub** and **GIT**, Worked/Coordinated with Offshore to complete the tasks.

**Environment:** mapR, Hive, SPARK, SCALA, MapReduce, UNIX scripting, HBASE.

**Role: Big Data Developer/Engineer Nov 2014– Jan 2016**

**Client: Visa, Austin, TX.**

**Responsibilities:**

* Implemented technical architecture and developed various Big Data workflows using custom **MapReduce, Hive, Sqoop**.
* Deployed on premise cluster and tuned the cluster for optimal performance for job execution needs and processes large data sets.
* The logs that are stored on **HDFS** are analyzed and the cleaned data is imported into **Hive**   
  warehouse which enabled end business analysts to write Hive queries.
* Used Maven extensively for building jar files of MapReduce programs and deployed to Cluster.
* Assigned the tasks of resolving defects found in testing the new application and existing applications.
* Analyzing the requirements, designing and developing solutions.
* Managing Project team in achieving the project goals including resource allocation, resolving   
  technical issues and mentoring the resources.
* Used Linux (Ubuntu) machine for designing, developing and deploying of Java modules.

**Environment:** MapReduce, Pig, Hive, Sqoop, FLUME, HBase, JDK 1.6, Maven, Linux.

**Role: Big Data Developer/Engineer Oct 2013– Nov 2014**

**Client: Synchrony Financials, Bangalore, India.  
Responsibilities:**

* Developed ad-clicks based data analytics, for keyword analysis and insights.
* Developed the code for Importing and exporting data into HDFS and Hive using Sqoop.
* Designed docs and specs for the near real-time data analytics using **Hadoop and HBase.**
* Crawled public posts from Facebook and tweets.
* Hands on experience in MapReduce jobs with the Data Science team to analyze this data.
* Converted output to structured data and imported to **Tableau** with analytics team.
* Defined problems to look for right data and analyze results to make room for new project.

**Environment:** Hadoop, HBase, HDFS, MapReduce, Flume, Java, Tableau, Cloudera Manager, Amazon EC2.

**Role SQL/PLSQL Developer Sep 2012– Oct 2013**

**Client: Headstrong LLC,** **Bangalore, India.**

**Responsibilities:**

* Responsible for requirements analysis, application design, coding, testing, maintenance and support.
* Created **Stored Procedures, functions, Data base triggers, Packages and SQL Scripts** based on requirements.
* Created complex **SQL** queries using views, sub queries, correlated sub queries.
* Developed **UNIX** shells/scripts to support and maintain the implementation.
* Created Shell Scripts for invoking SQL scripts and scheduled them using crontab.
* Defect management involving discussion with Business, Process Analysts, and team.
* Defect Tracking and Prepare Test Summary Reports

**Environment:** C++, Oracle PL/SQL (MS Visual Studio, SQL Developer), Unix Shell Scripts.

**Role: SQL/PLSQL Developer Jul 2011– Sep 2012**

**Client: VCare Health Care Systems, Bangalore, India.**

**Responsibilities:**

* Responsible for requirements analysis, coding, testing, and maintenance.
* Performed requirements analysis and object-oriented design.
* Created new Tables, Indexes, Synonyms and Sequences needed as per new requirements.
* Implemented complex **SQL**s using joins, sub queries and correlated sub queries.
* Created Shell Scripts for invoking **SQL** scripts and scheduled them using crontab.
* Prepare Unit Test Cases based on Functional Requirements.

**Environment:** Core Java, Oracle PL/SQL (Eclipse, SQL Developer), Unix Shell Script.

**Role: Java Developer Nov 2010 – Jul 2011**

**Client: WellCare, Bangalore, India.**

**Responsibilities:**

* Interaction with business team for detailed specifications on the requirements and issue resolution.
* Developed user interfaces using **HTML, XML, CSS, JSP, Java Script** and Struts Tag Libraries and defined common page layouts using custom tags.
* Wrote and executed efficient SQL queries (**CRUD operations**), JOINs on multiple tables, to create and test sample test data in Oracle Database using Oracle SQL Developer.
* Used **CVS** for check-in, check-out of files to control versions of files.
* Developed Style Sheet to provide dynamism to the pages and extensively involved in unit testing and System testing using **JUnit** and involved in critical bug fixing.

**Environment:** Java, Struts 1.2, Hibernate 3.0, JSP, JavaScript, HTML, XML, Oracle, Eclipse, JBoss.

**Role: Java Developer May 2010 – Nov 2010**

**Client:** **Sears Retailer, Bangalore, India.**

**Responsibilities:**

* Involved in specification analysis and identifying the requirements.
* Involved in preparation of the Code Review Document, Technical Design Document
* Designed the presentation layer by developing the **jsp** pages for the modules
* Developed controllers and JavaBeans encapsulating the business logic
* Used Jasper Reports to provide print preview of Financial Reports and Monthly Statements.
* Carried out integration testing and acceptance testing
* Participated in the team meetings and discussed enhancements, issues and proposed feasible   
  solutions.

**Environment:** Java1.4, J2EE 1.4 [Servlet, JSP], JDBC, XML, Oracle 8i.