

SHARAT KUMAR H

Email: sharat.kashi@gmail.com

① : +91 998-053-5604

Address: #103, 2nd F, Rama Villas Apt
HSR layout, Bangalore – 560102

Professional Experience: Senior Data Engineer working on Java/Scala and Hadoop technologies with 9 plus years of experience.

TECHNICAL SKILLS

Big Data Frameworks: MapReduce, Hive, Pig, Sqoop, Flume, Oozie, Impala, HDFS, Talend, Hbase and Spark.

Hadoop distribution platform: Cloudera CDH 5.4.8 and MapR 4.1.0.

Languages and frameworks: Java /J2ee, Scala, EJB, Kibana, Elastic search, Struts, Oracle, GWT, SQL, PL/SQL, Axis2, JDBC and JMS (Tibco), SonarQube, Shell scripts.

Graph DB: Neo4J and tiger graph

Web Technologies: HTML, XML, JSP, Servlets, Web Services.

Configuration and build Management: Apache Maven, Tortoise CVS, Tortoise SVN, GIT, Jenkins.

Application and Webserver: WebLogic, Tomcat.

Operating Systems: Windows and Linux.

CERTIFICATION

Cloudera certified developer for 'Designing and Building Big Data Applications'.

EDUCATION

Bachelor of Engineering, Information Science.

P.E.S. Institute of Technology
Visveswaraiah Technological University,
India [Aggregate: 69.41%]
July 2009.

Viveka PU College Kota
Department of Pre-University Education,
Bangalore [89.33%]
April 2005.

SUMMARY

- 9 years of experience in software development includes Analysis, Design and Development in **Java/J2EE** platform.
- Worked in multiple **Bigdata projects** using **Hadoop ecosystem** components like MapReduce, Apache Hive, Apache HBase, Impala, Pig Latin, Flume, Sqoop, Oozie, Talend studio, Apache Spark and Kibana dashboards.
- Exceptional skills in **agile development** and test driven development.
- Developed multiple MapReduce jobs in java for data cleaning and preprocessing.
- Moving data from HDFS to RDBMS and vice-versa using Sqoop and Analyzing/Transforming data with Spark, Hive and Pig.
- Excellent creative, analytical, problem solving skills and **Cloudera** certified Hadoop developer.
- Worked at a client site in **London** to gather requirements and to provide solution design.

PROFESSIONAL EXPERIENCE

United Health Group | Senior Software Engineer **Bangalore | Nov 2016 – till date**

Tools & Technologies: Talend Studio, Java, Scala, Apache Hive, Hbase, Sqoop, MapReduce, Spark, Kibana, Elastic search, Maven, Shell Scripts, Eclipse, Oracle SQL Developer, SVN, Jenkins, Sonar, Neo4J and tiger graph.

Project background: Data Fabric team is involved in automating the entire ingestion process and enrichment of ingested data. This process involves many core functionalities like touch file creation, move files to working directory, entity instance tracking, hive table and view creation, partition creation, record count check, file duplicate check, schema validation check etc. Once the ingestion succeeds, it creates the snapshots for these data. The analytics and data processing happens on these snapshot data and then provide these reports to customer. We use MapR 5.2.1 version currently in live.

Responsibilities:

- Responsible for requirements gathering, solution design, develop high and low level designs.
- Worked on neo4J graphdb to do the healthcare (member, provider and claims) data analysis. Cypher query is being used for loading and processing data.
- Lead the team in data extraction domain feed project. This basically sqoop the data from Oracle. Later Spark with Scala is being used to convert this data into JSON format and to store it in MapRDB. The Kibana ELK with spark listener is being attached to the spark job where we get the live status in the dashboard.
- Lead the team on migration project, where all the java/mapreduce code is rewritten in scala with spark and changed the file format from text to parquet. This improvement in data ingestion tremendously increased the performance while querying data by ~34 times and reduced disk space by ~87%. The ingestion time was also reduced by 50% because of faster processing and efficient code.
- Worked on design and development of automation of the testing framework for Data Fabric team. This framework will test all the core functionalities for the files ingested and provide the reports. Used Talend Studio to develop the framework.

Tools & Technologies: Java, Apache Hive, Pig Latin, Sqoop, MapReduce, Flume, Impala, Oozie, Maven, Eclipse, Oracle SQL Developer, SVN

Project Background: HaaS (Hadoop as a Service) is a Bigdata platform which works on data engineering and data exploitation. HaaS journey begin from data ingestion, data analysis, data processing and finally the processed data will be provided to business. Http fs, Sqoop and Flume is be used to ingest the data into HaaS. Apache Hive, Pig Latin and MapReduce is used for data analysis and data processing. Used Cloudera distributed Hadoop (CDH V 5.4.8) in the live environment.

Responsibilities:

1. Euvat Legislation (Module 1) – **Proposed and implemented a business rule using MapReduce framework** to compute the compliance report to determine if the customer is UK based or not using many evidences like Installation address, Billing address, Sort Code etc. 7 source systems data is ingested into HDFS using sqoop on daily basis & consolidated these data using hive. Used pig script for aggregation and then business rule is applied.
2. Channel Utilization (Module 2) – **Led the delivery of this project.** Objective was to calculate utilization of each trunk in a channel. The CDR(Call Detail Records) data from source is ingested into HDFS every 15 minutes using http-fs framework. Once the files are landed, the hive partition is created and loads these data into hive tables. Then these data is processed and aggregated using hive and the report is exported to oracle system using sqoop.
3. Revenue Assurance (Module 3) – **Worked on importing customer details data from different sources into HDFS using sqoop.** This data is aggregated to compute every customer's actual usage of data, sms, voice calls and broadband utilization by using pig script. Post which the verification is done against the actual bill generated to see if there any revenue is missed.
4. Data Ingestion framework ((Module 4) - **Worked in Design phase of Data ingestion framework.** This framework will automate the entire data ingestion process. Idea is to provide a UI, in which the user just needs to fill some basic details. The end output will be Workflow/Coordinator xml.

Tools & Technologies: Java/J2ee, spring, Hibernate, Axis2, Maven, SQL, PL/SQL, Oracle (DB, App server), JDBC and WebLogic, Tibco BW, Webservice, Eclipse, Oracle SQL Developer, SoapUI, Tibco Designer, Oracle discoverer.

Project Background: The Lufthansa Middleware platform is a service oriented platform (SOA), based on the Airline SOA multi layered architecture, providing middleware and integration services. LH SOA provides synchronous, asynchronous and batch connectivity. It provides event driven mechanisms, notification service capabilities and scheduling services for files. A wide range of interfaces and adapters provide the ability to connect various systems to the LH SOA platform.

Responsibilities:

- Analyzed requirements of a complete module end to end; developed low level design, Complexity estimation, and story point.
- Used AXIS framework for development & deployment of Web Services. Used Castor for mapping XML schema to and from Java objects. Worked on unit testing using JUnit/selenium
- Mentored the internship program.
- Proposed new architecture to replace components with open source tools to simplify and reduce the cost of platform.

Tools & Technologies: Java/J2ee, EJB, Hibernate, SQL, PL/SQL, Oracle, JDBC, GWT, WebLogic, Eclipse, Toad and SOA tools.

Project Background: LIMS(Logical Inventory Management System) –this provides a web service interface for internal and external applications for inventory management which allows users to add, modify or delete records in the logical inventory. It has a corresponding GUI interface to keep a track of these records. The values that can be modified are outlined by Business process and needs.

Responsibilities:

- Worked on java web applications in LIMS project. Used jsp/html for frontend and java, ejb, struts in the backend.
- Developed stored procedures for database operations.
- Innovation ideas like Find Bugs tool for the team.
- Developed project monitoring tool using gwt framework. This is an internal project to track the projects and resource utilization.