**Sayed Saleem**

**Senior Data Architect**

**Cell no: 732-225-1110**

**­­­­­­­ Email id: vishnu@skilltune.com**

**BACKGROUND SUMMARY & CAREER OBJECTIVE**

Around 17 years hands on experience in IT industry, predominantly into backend system, databases, and big data. working in various environments & platforms across varied domains. Having excellent command over written and verbal communication skills as required for coordination and have requisite interpersonal skills to thrive in a team environment. Looking for an opportunity to showcase my skills in a challenging work environment.

# Professional Summary

* Around 17 years of Development and quality assurance experience, in a variety of industries and environments.
* 5 years’ experience in Big Data, Hadoop, NoSQL, NewSQL and Spark
* Expert in the Polyglot Persistence and multimodal Databases.
* Good exposer to open-source databases.
* Expert in Key/Value, Document store, Column and Graph Databases.

# Good Experience in lambda and kappa ARCHITECTURE AND expert in creating the Data lakes.

* Very good exposure in OLAP and OLTP.
* Experienced in managing on-shore and off-shore teams, which includes hiring, mentoring, and in dealing with performance appraisals of team members.
* Extensive experience in SDLC, STLC process development and implementation.
* Good communications and relationship building skills having experience of working onsite at client locations in Europe and USA in the capacity of onsite coordinator.
* Experienced in working across various domains viz. Finance, Insurance, and HealthCare, Gaming, Mobile device and CRM. Environments that include web, client/server platforms.
* Well versed in improving software quality, meet customer expectations and reduce defects by supporting through proven software quality practices and implementing practical and effective process improvement methods.
* Capable of handling multiple projects & teams at the same time
* Good Experience as a Tech / Project Lead

# CERTIFICATIONS

* Certified ScrumMaster® (CSM)
* Cloudera Certified Developer for Apache Hadoop (CCDH)

# EDUCATION

* Master of Computer Application (MCA)

# TECHNICAL SKILLS

* **No SQL Databases: Cassandra**, Mongo DB, neo4j, Couchbase, CouchDB, InfluxDB, ScyllaDB, Redis, RiakDB, HBase, Druid, HyperDex, ClickHouseDB, Snowflake.
* **Database Engines:** RocksDB, InnoDB, LevelDB, BerkeleyDB
* **NewSQL:** VoltDB, NUODB, CreteDB, MemSQL DB, Etc.
* **Multimodal and multi store Databases:** ArrongoDB, OrientDB, MarkLogic
* **Cloud NoSQL Database**: Azure CosmosDB, Amazon DynamoDB, IBM Couldent
* **Big Data Ecosystems**: Hadoop, Map Reduce, HDFS, HBase, Zookeeper, Hive, Pig, Sqoop, Oozie, Flume,
* **Streaming:** Spark Streaming, Flink
* **Cloud:** AWS, Azure, Google cloud
* **Programming and Scripting Languages**: Java, Python and R
* **Databases**: Oracle, SQL Server, MySQL, PostgreSQL
* **Platforms**: Windows, Linux

# PROFESSIONAL EXPERIENCE

**Client**: **Walgreens** **Aug 2019 to Till Date**

**Role: Senior Data Architect**

**Technologies: Cassandra, MongoDB, MariaDB, CosmosDB, Elastic search, Spark, Kafka, Snowflake**

**Location: Chicago, IL**

**Description:**

All the projects in digital platform of Walgreens moving a major transformation from the existing on premises technologies to cloud and microservices. Complete migration of databases from Oracle ATG to NoSQL Databases on all the projects with in the digital and rest of the Walgreens.

## Responsibilities:

* Provide inputs to define and execute strategic roadmap for enterprise data architecture by identifying the current landscape and future business goals.
* Help establish best practice like standards and guidelines for design & development, deployment, support and analytics and mining.
* Data Storage and Computation Frameworks, Performance Optimizations
* Design conceptual and logical data models and flowcharts
* Deployed a fully operational data solution into production on Snowflake DW with technologies such as Microsoft Azure, Azure DevOps, Azure Data Factory, Azure Monitor.
* Understand SD/SN architectures and have deep experience using MPP systems.
* Utilized Snow SQL for command line, DDL, DML, and query syntax.
* Built modern approaches to automating data pipelines.
* Leveraged code repositories and continuous integration tools.
* Demonstrated experience with predictive analytics.
* Coordinate with the Data Science department to identify future needs and requirements.
* Currently working on the Snowflake migration from the existing platform.
* Evaluate Snowflake Design considerations for any change in the application.
* Build the Logical and Physical data model for snowflake as per the changes required.
* Define roles, privileges required to access different database objects.
* Define virtual warehouse sizing for Snowflake for different type of workloads.
* Design and code required Database structures and components.
* Build the Logical and Physical data model for snowflake as per the changes.

**Client: Comscore**  **June 2018 to Aug 2019**

**Role: Senior Architect**

**Technologies: Cassandra, MongoDB, Hadoop, Elastic search, Neo4j, Spark, Kafka**

**Location: Reston, VA**

**Description:**

Its BI project. Redesign of the existing Greenplum to Hadoop (HDFS) and Spark. The objective is to convert the storage to Hadoop (HDFS) and converting the PostgreSQL Scripts with Spark. Later this project was switched to Snowflake DB

## *Responsibilities*

* Worked as the senior architect to ComScore project.
  + Choosing the Database for the existing DBBI platform.
  + Worked on Snowflake for the heigh level and low-level design.
* Participate in POC evaluations of technologies and tools.
* Designed Operation wide data lake on snowflake.
* Team of data scientists created cutting-edge approaches to data analytics accessing over 15 separate petabyte data sources managed by separate agencies.
* Created product strategy by collaborating with senior data scientists to define key value offerings and use cases. Identified target prospects and engaged key business decision makers as trusted advisor.
* Managed RFP process.
* Closed two major data analytics client engagements.
* Directed projects through all phases of project lifecycle from client kickoff meeting to final delivery.
* Identified, launched, and coordinated lean six sigma initiatives for business process re-engineering.
* Created project management plan and architecture to implement PaaS solution.

**Client: Delphi** **Feb 2017 to May 2018**

**Role: Architect**

**Technologies:** Cassandra, Mongo DB, neo4j, MySQL, Hadoop, Elasticsearch, Spark, AWS EMR

**Location:** Mountain View, CA

**Description:** This is IoT and Time series project for the car’s sensors. It’s to establish a common platform for predictive maintenance and fraud detection based on the car sensor information. Need a centralized data store for the time series data. A new data pipeline is required for analytical purpose. This project was first designed based on Hadoop eco system, later it has been changed to other NoSQL Databases.

## *Responsibilities*

* Worked as an architect to Movement product.
* Lead the team to develop big data components on Hive, Pig and Spark for enabling Data warehouse capability in Big Data.
* Created different architectural approaches with Cassandra and Elastic search.
* Acted as the main point of contact for the technical discussion between the Moviemento and Valuelabs.
* Designed Cassandra based IoT platform for the Peta Byte Scale workloads.
* Configured EMR and Kafka Cluster.
* Worked on the redesign architecture of Hadoop, HBase, Hive and Apache Phoenix system to Cassandra.
* Designed Analytical Engine and Recommendation Engine.
* Designed Connected Cars and Connected peopled with Graph Database (Neo4j). For the bigger workload designed with other graph databases (Arango DB, Orient DB, Titan DB).
* Created a statistical report conversion of RDBMS to NoSQL.
* Understanding the data nature from different OLTP systems and designing the injection processes for HDFS
* Designing data model on HBase and Hive.
* Configured and Troubleshooting the EMR cluster for the various issues such as:
* View and Monitor a Cluster, connect to the Cluster, terminate a Cluster, Scaling Cluster Resources, cloning a Cluster Using the Console and Submit Work to a Cluster
* What region to run a cluster in, where and how to store data, and how to output results.
* Whether a cluster is long-running or transient, and what software it runs.
* The hardware and networking options that optimize cost, performance, and availability for your application.
* Designed and implemented a database migration to centralize multiple local MySQL database instances to a single Amazon Aurora RDS instance.
* Developed an Analytics and Reporting application in Python to gather database metrics, store them in a OrientDB graph database and produce a daily snapshot report from a Python application.
* Architected a Content Management System database schema for video content.
* Designed and implemented a Social Media Graph Database to support connecting with friends and identifying common interests.

**Client: Graphene Med –Tack and trace product**  **July 2015 to Feb 2017**

**Location: SFO, CA, Dubai, and Hyderabad**

**Role: NoSQL SME**

**Technologies:** Cassandra, MongoDB, CouchBase, VoltDB, single store and Hadoop

**Description:** The objective of the product is to collect the bar code information and store in centralized repository. Data need to be collected from the production to till it reaches to the client. This product should support various databases. Based on the client need data should be moved to databases. This project has both TP and AP Data.

## *Responsibilities*

* Define architecture required for POC/role out – Design data architecture (data model) for local implementation or deployment on HDFS distributed multi-node node Cloudera Cluster
* Worked on data modeling and data import and export. Extensive experience in deploying, managing, and developing MongoDB clusters. Experience in creating JavaScript for using DML operation with MongoDB.
* Creation, configuration, and monitoring Shards sets. Analysis of the data to be shared, choosing a shard Key to distribute data evenly. Architecture and Capacity planning for Mongo DB clusters. Implemented scripts for mongo DB import, export, dump and restore.
* Worked on MongoDB database concepts such as locking, transactions, indexes, Shading, replication, schema design. Created multiple databases with sharded collections and choosing shard key based on the requirements. Experience in managing MongoDB environment from availability, performance, and scalability perspectives.
* Worked on creating various types of indexes on different collections to get good performance in Mongo database.
* Worked on creating documents in Mongo database.
* Worked on multiple BSON type objects and used based on requirement.
* The same product is tested with Couchbase as the main database and created new architecture for the same product. In this architecture we have used concept of the auto sync between Databases to ease the synchronization of the data between the databases.
* The same product is tested with Cassandra as the main database and created new architecture for the same product. In this approach we used 3 data center approach writes are directed to first data center and reads are directed 2 and 3 data centers.
* The same product is tested with in memory databases such as VoltDB and single store as the main database and created new architecture for the same product. In this architecture we have used concept called streaming table so that data can be directly stored in the tables rather than using some streaming mechanism.

**Client: Berkadia internal project** **Sep 2014 to June 2015**

**Project NoSQL transformation**

**Location: Red Bank, NJ**

**Role: Lead**

**Technologies:** Cassandra, MongoDB, Couchbase, Hadoop, Hive, Spark

**Description:** The objective of the product is to replace the existing ETL pipeline with the modern technologies.

Upstream data is collected in the DataMart and its processed and sent to the downstream systems and vice versa.

Where the database needs to change from RDBMs to other NoSQL and Hadoop systems. As the volume of the data increased ordinary RDBMS are unable to handle this load to overcome this problem big data tolls are used.

## *Responsibilities*

* Worked as big data and Hadoop Architect
* Created Data lakes for the projects and created analytical pipelines.
* Define architecture required for POC/role out – Design data architecture (data model) for local implementation or deployment on HDFS distributed multi-node node Cloudera Cluster
* Designed Streaming platform using spark to replace the existing ETL.
* Install, Administer, and Optimize NoSQL Databases (MongoDB, Couchbase and Cassandra).

**Client: The Men’s Warehouse and GE**  **Dec 2013 to Aug 2014**

**HIS**

**Location: Hyderabad**

**Role: Lead**

**Technologies:** MySQL Oracle, MS SQL Server, Cassandra, MongoDB, Couchbase, Hadoop, Hive, Spark, Tibco and JMS

**Description:** HIS is an existing product which was developed 10 ago with oracle and Tibco services. HIS collects the log information of all the applications within the organization and send appreciative messages to the respective admin on the Events. This product is to be completely redesigned with the new features which includes the Big Data problem.

Database should be based on the client choice and Tibco service should be replaced with the JMS. This product needs to be customized based on the client needs.

## *Responsibilities*

|  |  |
| --- | --- |
| * Managed a team of 15 members and responsible for the delivery of the project. * Work allocation to offshore team and acting as bridge between onsite, offshore &client. * Responsible for master test plans, environment configurations, application installations and testing schedules. Also, hands on testing of these applications. * Involved in the requirement gathering for the NOSQL (MongoDB and Cassandra) and Hadoop implementation. * Created Test Plan for the Hadoop and BigData implementation. * Analyzed and resolved issues relating to database application performance. * Managed and administered all NoSQL database systems. | |
|  |  |

**Client: GE**  **FEB 2012 to Nov 2013**

**ALOC- Payments**

**Hyderabad**

**Technologies:**  Oracle, Tibco and Java

**Description:** ALOC is GE’s accounting system. ALOC payment system maintains all the information of the Bank Guarantee, letter of credit and other forms of the payments to the clients.

## *Responsibilities*

* Managed a team of 5 members and responsible for the delivery of the project.
* Work allocation to offshore team and acting as bridge between onsite, offshore &client.
* Responsible for master test plans, environment configurations, application installations and testing schedules. Also, hands on testing of these applications.
* Configure server-side settings for Cassandra database servers and worked on the Data Modeling of the Use cases.
* Create high availability and load balancing Cassandra cluster.
* Create replica sets for High Availability cluster.
* Monitor deployments for capacity and performance.
* Optimizing the queries.

**Client: Sony**  **Oct 2009 to Jan 2012**

**SCEA (Sony Computer Entertainment America,**

**Chennai, Visakhapatnam**

## Role: Sr Lead

**Technologies:**  Oracle, JAVA, JIVE

**Description:** SCEA isan internal project of Sony. SCEA store the centralized gaming web sites for the gamers.

## *Responsibilities*

* Acted as the main point of contact between many levels within BPA and in SONY including the Business Head, Technical Architect, Project Manager and Business Analyst
* Onsite and offshore coordination
* Given presentation to client on projects handled.
* Prepared test reports and weekly status
* Work allocation to offshore team and acting as bridge between onsite, offshore &client.
* Conducted defect meetings, created test matrices and test lab maintains.
* Conducted Performance testing for the project.

**Client: IGT (International Gaming Technology)** **June 2008 to Sep 2009**

**Leading Casino Service Provider**

**Las Vegas, US**

**Role: Lead and Onsite Coordinator**

**Technologies:**  MSSQL Server, SSIS, SSRS, C# and .Net

**Description:**

## Leading Casino Service Provider is a database upgrade project, where DBA must collect the data from the Casinos and upgrade the databases. In this process DBA must test the upgrade process in the Lab and upgrade the same on the production systems

## *Responsibilities*

* Participated in the Database upgrades. Created clear road maps for the upgrades.
* As a Lead responsible for master test plans, environment configurations, and application installations & testing schedules. Also, having experience of hands-on testing of these applications.
* Managed and maintained relationships with third-party clients to promote and enhance the reach of several marketing efforts for IGT (International Gaming Technology).
* Created Business scenarios and Walkthroughs of Project with client.
* Tested platform updates, hot fixes, and network operating system application upgrades. Also tested line of business application packages on Casino’s Windows platform

**Client: Allianz**  **Sep 2007 to June 2008**

**Minneapolis, US**

**Role: Lead and Onsite Coordinator**

**Technologies:**  MSSQL Server, Oracle, VS10, C# and .Net

**Description:**

In Mass appointment various client’s data is collected and integrated with the Allianz data bases

## *Responsibilities*

* Lead responsible for environment configurations, application installations and testing schedules. Also, hands on testing of these applications.
* Acted as the main point of contact between many levels within HCL and in Allianz including the Business Head, Technical Architect, Project Manager and Business Analyst.
* Conducted several Business scenarios and Walkthroughs of Project.
* Conducted defect meetings, created test matrices and test lab maintains.
* Participated in project estimations.
* Conducted Performance testing for the project using ACT tool.
* Responsible for establishing consistent Dev practices using standard tools across multiple projects.
* Acted as the main point of contact between many levels within HCL and in Allianz including the Business Head, Technical Architect, Project Manager and Business Analyst
* Installed MSSQL server and Oracle Instance.
* Creation Business scenarios and Walkthrough of business scenario with client.
* Wrote an extensive test plan to identify all the paths that a user could follow and all the combinations of events that could occur and leading the test team.

**Client: USSA**  **May 2006 to Aug 2007**

**PML (Probable Maximum Loss)**

**San Antonio, TX, US**

**Role: Lead and Onsite Coordinator**

**Technologies:**  IBM Mainframe, Java

## *Responsibilities*

* Wrote an exhaustive test plan to identify all the paths that a user could follow and all the combinations of events that could occur.
* Creation of Business scenarios and Walkthrough of business scenario with client.
* Perform weekly defect meetings and provide Project direction.
* Perform Onsite coordination works.
* Perform Project estimations.
* Conducted Performance testing for the project.