VINAY KUMAR MISHRA*(BIG DATA PROFESSIONAL)*

•F-015, BETA-2, GREATER NOIDA •(+91)8178256262 •[innovative.vinkm@gmail.com](mailto:innovative.vinkm@gmail.com),[vinkm98@gmail.com](mailto:vinkm98@gmail.com)

***SUMMARY***

Strong experience in web and distributed Application Development and Production Assistance in handling large volume of data storage and processing. 9+ years of experience in OOPS and functional programming languages like (Java/Scala/Python) along with multiple frameworks like (Spring, Hibernate, JDBC Template etc.) as well as distributed technologies like (HDFS, Spark, Kafka, Elasticsearch etc.)

* Process-oriented, ensuring code quality and performance area improvement.
* Skilled and working experience in **Spark, Kafka, Elasticsearch, TWS,** and **Java/Scala/Python.**
* Outstanding work ethic and team building skills, willing to work overtime and flexible shifts.

***PROFESSIONAL EXPERIENCE***

***Optum* *Global Solutions – (UHG), Noida U.P***.

***Sr. Data Engineer*** ***Sep 2018 – Present***

Worked on several applications like **OEDS** (Optum ERP Data Store), **BRMSDS** (Billing Receivable Management System Data Store), **PEA** (Payment Engine Analytics) and **EPSA** (Electronic Payment Settlement Analytics) in UHG. Each application has its own importance and responsibilities that will serve multiple facilities to business using their capabilities like Distributed Storage, Highly Scaled **MapR** distribution file System, In-memory processing/analytical engine like **SPARK**, Distributed messaging system like **KAFKA** and **Elasticsearch** for search data retrieval with minimum response time SLA like 6 seconds, processing of report data from **Data lake** using **HIVE** tool.

Quarterly performance and spot awards achievement during working with above applications and technologies.

# Xavient Information Systems, Noida, U.P.

***Sr. Software Engineer*** ***Jan 2015 – Aug 2018***

Worked on **Trap-Analyzer (Cable Broadband Network Monitoring and Root Cause detection), Saturn, Mercury, DIP(Data Ingesion Plateform),** projects in OOPS language JAVA, Framework provides a comprehensive programming and configuration model for modern Java-based enterprise applications.

# Ahoy Telecom Pvt. LTD., Noida, UP.

***Sr. Software Engineer*** ***April 2013 – Dec 2014***

Worked on **UAHOY** and **UAD (User Advertisement)** digital ad engine using OOPS language JAVA, Spring, hibernate framework provides a comprehensive programming and configuration model for modern Java-based enterprise applications.

# Sisdon Infosolutions Pvt. LTD., Noida, UP.

***Java Developer*** ***Aug 2011 – MAR 2013***

Worked on **Tennis App** and **Probability Calculator** using OOPS language JAVA, Spring, ORM framework like Hibernate provides a comprehensive programming and configuration model for modern Java-based enterprise applications.

***PROJECTS***

# Project 1# PEA (Payment Engine Analytics) May 2021 – Present

Project **:** EPSA(Electronically Payment Settlement Analytics)

Client : United Health Gruop

Technology : Elasticsearch, HBase, Hive, Spark, Shell Script

Languages : Python, Scala

Role : Team member

Data lake to store UHC payments and claim related data. PEA store and apply some transformation to generate reporting data for business users. Reporting data exposed in form of Hive tables and those are used in Tableau reporting app to publish report for business users every day. Also exposed data through API for PHOW application. In this applications Elasticsearch used to store some key/summary level information for payments and claims. Detailed level information stored in HBase for payments and claims. For coding implementation of this application used Spark, Scala, Unix scripts for storage purposes Elasticsearch, HBase, Hive etc...

# Project 2# EPSA (Electronically Payment Settlement Analytics) Jun 2021 – Present

Project : PEA (Payment Enigne Analytics)

Client : United Health Gruop

Technology : Elasticsearch, HBase, Hive, Spark, Shell Script

Languages : Python, Scala

Role : Team member

Project Description : EPS(Electronically Payment Settlement) is the previously project that handles the whole traffic of the UHC portal as well as batches and other data and query portals.

Existing system has DB2 database that holds the 13 months of Payment data and 3 weeks of the claim data. Older data has to be removed because BD2 has the much traffic that its capacity.

So EPSA(Electronically Payment Settlement Analytics) is the new Bigdata platform that is capable to holds the 10 years of Payment data and 1 years of the claims data, EPSA had the more capable then the existing EPS system. Now EPSA going to replace the existing EPS system very soon. EPSA is still in development phase and we are planning to move this system to production ASAP.

# Project 3# EIS (Enterprise Information Security) Security Audit Logging Jun 2020 – Present

Project : EIS ( Enterpirse Information Security)

Client : United Health Gruop

Technology : Spark, Kafka

Languages : Python

Role : Team member

Within the core UHG environment, Enterprise Information Security (EIS) collects logs in the Enterprise Technology Data Lake (ETDL). Application logs in the ETDL which follow formatting guidance and contain adequate detail can be incorporated into log monitoring, investigation, and retention processes. Applications not utilizing ETDL for these services are responsible for making sure that the repository they use is able to provide security monitoring and log review.

# Project 4# BRMS-DS (Billing and Revenue Management System – Data Store) Oct 18– May 2021

BRMS-DS (Billing and Revenue Management System-Data Store) is the Big Data platform project that handles the Billing related all the transactions electronically. BRMS is uses the Hive to generate the Billing and revenue reports that make the business growing.

BRMS stores the Oracle transactional data in the multiple level hive tables that helps to create the final reporting table.

Oracle transactional data is sent through Oracle GG(Oracle Golden Gate) and its stored by BRMS DS in multilevel Hive tables.

There are four types of levels of data in Hive:

**L1** is the Level 1 data sent by the Golden gate in the landing zone tables and its having approx. (3000+) tables.

**L2** - is the Hive table that filter out the landing tables and performs the joining on the L1 tables and extracts the specific information from L1 and put the data in L2 tables that having approx. – 500+ tables.

**L2.5** – L.25 is the next layer of the reporting data that will be comes from the L2 tables and this layer having approx. 200+ tables.

**L3** – Level 3 has the final tables that come though the L1, L2, L2.5 and some other aggregated information that helps the reporting system of creates the report of Billing data.

# Project 5# Cable Broadband Network Monitoring and Root Cause Detection Oct 16 – Aug 18

Project : trap-analyzer

Client : Cox Communication Inc. (U.S.)

Technology : Spark-Batch, Spark-Streaming, Kafka, Elasticsearch, Oracle

Languages : Java, Scala

Role : Team member

**Project Description :**

Xavient proposes developing a demonstration solution which will deliver Operational Analytics on the health of the Broadband Network, with specific focus on the access network, Wi-Fi Cable Modems, Amps (Amplifiers), Fiber Nodes, CMTS and fiber links.

This will highlight several 3-4, common use cases where network issues result in disruption to customers broadband internet service, providing the capability to see all modems and their online status via geo-map dashboard UI, as well as alerts at either the modem, Amp, Node or CMTS level, and root cause of an outage in real-time.

The goal of the is to demonstrate the value of a tool which can identify internet delivery issues as they occur, and provide a more accurate probable root cause enabling service techs to more quickly and cost effectively resolve a problem minimizing customer dissatisfaction and support costs.

**Modules of trap-analyzer**

**batch-data-loader :[** Responsible to load snmp records from two different Oracle databases (CRM and EWS) and put it into Elasticserch using Spark-Batch processing,]

* **elasticsearch-utils:** [ This module is responsible to read csv file and put all records in Elasticsearch because elasticsearch gets data from two sources

1. From oracle database and

2. In form of CSV files

These files also need to put in Elasticsearch. So this module is responsible for this task. ]

* **kafka-connect-fs:** [ Responsible to read Cox devices logs and fetch out device traps information using some algorithm provided by **Cox Communication** and send that traps to kafka topic. ]
* **streaming-traps-analyzer :** [This module read topic records and update the devices traps status in Elasticsearch by mac-address.]
* **snmp-maps-rest-client :** [This module is Spring boot rest webservice that is responsible to serves snmp-map records from Elasticsearch on the basing of geo-spatial bounding box query to log\_analyzer-ui.]
* **log\_analyzer-ui :** [This is a google map ui that plots the all received devices traps information on google map using device lat, long and show the helth status as {green, red, amber} by color coding.]
* **trap-simulator :** [This is a simulation ui is responsible to send traps to kafka-topic by consuming snmp-maps-rest-clientrest webservices.]

**Responsibilities**

* Developed code for module **batch-data-loader**
* Developed code for module **streaming-traps-analyzer**
* Developed code for module **snmp-maps-rest-client**
* Developed code for module **trap-simulator**
* Developed code for module **elasticsearch-utils**
* Deployed the all modules on hadoop cluster.

# Project 6# DIP (Data Ingestion Platform): Jan 16 – Oct 16

Project : DIP (Data Ingestion Plateform)

Client : Xavient Internal

Technology : Cassandra, MySQL, Spark, Spark SQL, Kafka, Flume, Sqoop

Languages : Java, Scala

Role : Team member

**Project Description :**

* Xavient setup a CDIP framework in AWS environment.
* Real time data ingestion using CDIP harnessed the powers of Apache Spark to stream data
* The system read files from S3 in (.gz) format and some of the DSP Files
* Perform basic Validations : Null check, Date check, Numeric check
* Save error records in error tables & supports Auditing and Logging
* Support for both Cassandra and Redshirt as Data Store.
* Support source as S3 and Cassandra DB.

**Responsibilities**

* Created a connector using Kafka-Connect API that ingest MySQL data on every insert and put it on kafka topic.
* Created a connector using Kafka-Connect API that ingest Cassandra data on every insert and put it on kafka topic.

# Project 7# Mercury Apr 15 – Dec 15

Project : Mercury

Client : Cox Communication Inc. (U.S.)

Technology : Spark Streaming, Spark-Batch, Kafka, Java, Oracle 10g

Role : Team member

**Project Description :**

Mercury application is a Real time Call Monitoring application which provides site-level data to the management executives and vendors with drill down capabilities. It is an intranet application which will be accessible on COX Network only. Users of the application have to be in COX active directory.

Mercury is Cox‘s project. Its major aim is to track the customer calls and its function as Sophisticated telephone system – A machine to handle the incoming or outgoing calls, known as an Automated Call Distributor (ACD) – this replaces the switchboard telephones and in itself has embedded program that will handle all the calls made, identify what they are for and route them to the right person to answer the call (provided it has been set up correctly). They can handle and report on thousands of calls per hour!

Many of other types of specialist equipment – Interactive Voice Response (press 1 for this or 2 for that); automated speech recognition which enables a conversation between the computer and the

Customer predictive dialer to make large volumes of outbound telephone calls; resource planning tool to forecast and plan shift patterns to meet expected call volumes etc.

**Responsibilities**

* Requirements gathering and process mapping for the development.
* Written spark streaming jobs to read the feed module kafka topics and put the data into the oracle database.
* Written spark streaming jobs to read the listener module kafka topics and put the data into the oracle database.
* Deploy the spark jobs on the client cluster.

# Project 8# Saturn Jan 15 - Mar 15

Project : Saturn

Client : Cox Communication Inc. (U.S.)

Technology : Java 1.7, Oracle 10g, Spring, JDBC Template

Role : Team member

**Project Description:**

Cox is replacing Aspect & Cisco call centers with Avaya call center solution. “Scout” is a custom-built application currently being used by Call Center supervisors and administrators to make con-figuration changes on Cisco platform. The objective is to replace Scout UI application with Saturn application. Some of existing Scout Application functionality will not be implemented in Saturn, as Avaya Aura will provide separate application screens to perform the same.

Saturn is a new portal application which will be developed using Spring Rest Web Services, Com-positive Application Framework to act as a Contact Centre Administration Utility. It will primarily be the front end (user interface) to configure and administer Cox CPR and IVR call center functionality.

**Cox Post Routing (CPR)** is a custom Solution to facilitate call sharing between Cox call centers and outsource vendors. It has the capability to apply logics over the real-time ACD statistics and exchange call specific data, independent of the telephony switching platforms used by the call sharing parties.

IVR is the Voice Portal Application that manages the flow of every call by playing various mes-sages & prompts for the callers to choose what they intend to do with the call and interacts with CPR at various points in the call flow to make wise suggestions on routing the calls to appropriate teams.

**Responsibilities**

* Requirements gathering and process mapping for the development.
* Developing code for CPR module to configure the schedulers to routing call to multiple destination sites.
* Developing overnight senario to handle US timezone wise scheduler for process routing.
* Developing code for IVR module for various senarios specified by client.
* Move web-methods code to java technology and implements client required enhancements.
* Created and enhanced some stored procedure.

# Project 9# UAD Apr 13 - Dec 14

Project : UAd

Client : Ahoy Telecom Pvt. Ltd.

Technology : Java 1.6, MySQL, Spring, Hibernate

Role : Team member

**Project Description:**

UAD is an Ad Engine; this application software system will be a Digital Advertisement Portal for advertisers and publishers. It acts as a bridge between the advertiser and publisher. This system will be designed to maximize the advertiser’s productivity by providing tools to assist in automating the campaign review and publishing process, which would otherwise have to be performed manually. At the same time, it meets the publisher’s requirements and find suitable campaigns for him just at a button click.

By maximizing the availability of suitable publishers for advertisers and vice versa the system will meet the advertiser’s and publisher’s needs while remaining easy to understand and use. Application supports various types of ads: Advertiser can create three types of Advertisement and Ad-places:

1. Text Ad

2. Banner Ad

3. Video Ad

**Responsibilities**

* Developing Advertiser and Publisher modules dashboards to upload ads and publish ads on their own adplaces.
* Developed code to view Advertiser and Publisher stats reports on dashboard.
* Design table structure and implements code and query to store hourly basis stats in DB.
* Implements video ads dashboard to upload video ads on server and render video ad to user device using WOWZA streaming server.
* Writing code to download stats report in excel format.

# Project 10# Tennis App Aug 11- Apr 13

Project : Tennis App

Client : Sisdon Infosolutions Pvt. Ltd.

Technology : Java 1.6, MySQL, Spring, Hibernate

Role : Team member

**Project Description:**

Tennis App has various portal like Individual, Aggregated filter, etc. which has been developed for UK. It fetches the full description of the individual or multiple tennis player from the website (tennisinsight.com) and stores individually in the database. After that we make various operations on these data to generate sheets and reports.

**Responsibilities:**

* Implementation of functional requirements and developing logic for the same.
* Writing the personalization query for different sort of content.
* Provide support to client for this project

**Educational Qualification**

* Master of Computer Application (MCA) in 2013 from Punjab Technical University.
* Master of Science in Information Technology 2012 from Punjab Technical University.
* Oracle Certified Professional, Java Programmer (OCPJP) Java 1.5 from Oracle.
* Advanced Diploma in computer application ‘A’ Level from DOEACC Society with full time graduation B.A. from Lucknow University.

**(Vinay Kumar Mishra)**