

Senior Hadoop Developer/Big Data Solution Architect Senior Hadoop Developer/Big Data Solution Architect Senior Hadoop Developer/Big Data Solution Architect - AVEVA Inc Cumming, GA 11 years of experience in IT industry as a Software developer, Technical lead, Solution Architecture and Project Manager in various domains More than 2 years of client facing experience in Europe (Finland & Belgium) 3+ years of experience on Big Data ecosystems Hadoop, MapReduce, HDFS, Yarn, Hbase, Hive, Pig, Sqoop, Flume, Spark, Kafka, Oozie, Hue, Zookeeper, Ambari, Hortonworks, MapR, AWS Hands on experience in creating real-time data streaming solutions using Apache Spark, Kafka, HBase, Spark Streaming API(pySpark) & Kafka API Experienced working with different file formats - Avro, Sequence and JSON Good understanding on building Big Data/Hadoop applications using AWS Services like Amazon S3, EMRFS, EMR, RDS, Airflow etc Certified Scrum master Hands on experience in Object oriented analysis, design and programming with C++, Java & Python Strong domain knowledge in Telecom, GIS, Oil and Gas, Marine and Electrical Hands on experience in developing the GIS applications & CAD application customization using GeoServer, HTML, CSS, JQuery, AngularJS, JavaScript, AutoCAD, ObjectARX & RealDWG. Hands on experience in SQL and PL/SQL Experience in using Custom distributions like Hortonworks, MapR and Cloudera Excellent in technical & organization skills, communication skills & Collaboration skills. Work Experience Senior Hadoop Developer/Big Data Solution Architect AVEVA Inc - Huntsville, AL March 2014 to March 2017 Description AVEVA provides engineering/plants Software Solutions to industries such as Process Plants, Engineering/Contracting companies such as EPC, Power, Shipbuilding. It is a niche/specialized segment. As now a days many Oil companies are facing problems like Oil is hard to find, Oil is expensive to produce, human safety etc Our objective is to provide a solution using Big Data/Hadoop technologies to overcome the problem reported by customers like Chevron, Shell, and BP etc Responsibilities Closely work with the customers to understand the Challenges and requirements Analyze the customer data to provide a solutions Installed and configured Apache Hadoop MapReduce, HDFS, developed multiple MapReduce jobs in Java for data cleaning and preprocessing Loading the customer data into HDFS using Hadoop commands Processing and

analyze the data using MapReduce jobs Configure Apache Kafka producer and consumer coding part in java to establish connection from data sources and loaded into HDFS Created HIVE tables to store the processed results in tabular format Written PIG scripts to process the HDFS data

Worked on pulling the data from MySQL database into HDFS using Sqoop Developed Web application for Data Analysis in reporting format using HTML, CSS, JQuery Responsible for Hadoop Admin related tasks like debugging, performance fine-tuning and monitoring for day-to-day activities Designed and implemented Apache Spark streaming application using Python(pySpark) and Scala Load and transform large sets(petabytes) of structured, semi-structured and unstructured data Strong expertise in writing and implementing Pig/Hive scripts, UDFs

Experience in defining and coordination of job flows Good experience in NoSQL database like Hbase Responsibility to manage data coming from different sources Responsibility to write and implement python & Scala coding for Spark implementations Responsible to implement the Spark applications on AWS Write a integration tool in python to connect to AVEVA NET Workhub and Dashboard from Hadoop ecosystem Involved in the loading data from Linux file system to HDFS

Responsible for managing systems on AWS platforms Automated all jobs in Linux shell scripting

As a technical leader, ensure that on-time delivery of the agreed delivery Ensuring that the development process is properly focused and controlled Managing risks and issues at the development with in estimated time, escalating to project manager as required Running the daily meetings ensuring they are timely, focused and brief Environment Java, Hadoop, HDFS, Hive, Pig, Flume, Sqoop, Hbase, Oozie, Spark, Kafka, Zookeeper, Ambari, Hue, Python, Linux, Eclipse, Hortonworks, AWS Hadoop Developer EarthLink Inc - Atlanta, GA January 2013 to March 2014

Description: Traditional GIS applications like GeoServer is limited in dealing with big data challenges including versatile data forms, processing, parallel computing and dynamic mapping and visualization. As a part of predictive analysis, we have studied the GIS data and provided solutions

Responsibilities Gathered requirements and design the application Installation and Configuration of GeoServer and Hadoop ecosystem tools Installation and Configuration of Tomcat for Development and Testing environment Loading of Shape files into GeoServer Loading the GIS

data into HDFS Used MapReduce framework to identify Delta updates and to refresh persistent files PIG and HIVE extensively used for data analyses Responsible to manage GIS data coming from different sources Developed PIG scripts in the areas where extensive coding needs to be reduced Experienced in MapReduce programs to load the data from application generated log file to Hbase database Write a reusable and tested code in python for Spark Implementation Extensively used Python in Big Data implementations Created Hbase tables to store variable data formats of data coming from several data sources Analyzing the data with Hive, Ping and Hadoop streaming Responsible for monitoring Hadoop cluster, manage and review logs, performance tuning etc Developed Coverage Maps Web application on top of Google Maps using HTML, CSS, JQuery, JavaScript Implemented RESTful scripting for Querying on Shape files Deployed the application in Development and Testing environments Coordinated the Onsite and Offshore teams for smooth implementation Implemented Agile process for this project Give training and mentorship to team members to make them better on the job Review the completed tasks to ascertain compliance with standards Coach all team members and provide necessary advice and guidance Write and forward regular reports to the management Perform regular appraisal of team member's performance to help with improvement Environment HTML, CSS, JavaScript, JQuery, Apache Tomcat, Linux, GeoServer, Eclipse, Java, Hadoop, HDFS, Hive, Pig, Flume, Oozie, Zookeeper, Linux, Hue, Ambari, Hortonworks Team Lead & Techno Functional Architect Proximus Group - Brussels, BE January 2010 to December 2012 Project: Graphical Documentation Automation Description: GDA provides to the NEO division several tools to create or modify the digital graphical documentation of the local loop network. These software's give also several facilities to manage, control and export data from this documentation. During this period below technical & functional enhancements done. ? VBB ? BOSO ? ARK Project ? ARK-III PDF Generator ? GDA 64-Bit Migration ? GDA Migration from 2006 to 2012 Responsibilities Requirements capturing and Analysis of the requirements Functional design and Technical design of the requirements Resource allocation and assigning the tasks to team members Creating the required classes for the development Responsibility of Quality of output to customer. Developed

use cases in VC++, ObjectARX Implemented AutoCAD migration project from AutoCAD 2006 to AutoCAD 2012 Implemented AutoCAD 64-bit project from AutoCAD 32-bit to AutoCAD 64-bit
 Embedded Google Maps into GDA application for interactive selection of area Setting up of Development(Dev), Integration (ITT), acceptance(UAT) environments for application deployment and testing Weekly meetings with customer regarding the health of application and current project activities. Environment SQL, Enterprise Architect, VSS, AutoCAD Map 3D ObjectARX 2006 to 2012 Object ARX Developer ObjectARX - Brussels, BE July 2009 to January 2010 Project: JMSFTB for GDA Description: The main objective of this project is to attach Job Management Application attributes like JMS number, JMS driver and JMS building block number to IVP planned objects like Cables, Distributor etc. JMS application will call a web service which in turn will call a stored procedure in CAS database to collect information related to all planned cables and street cabinets associated with a particular JMS id. Responsibilities Implemented use cases in VC++ and Object ARX. Prepared User-friendly interfaces using MFC. Implementation of various PL/SQL procedures. Used Visual Source safe for doing proper Versioning Control. Unit testing and integration testing. Prepared technical design documentation for use cases. Environment VC++ Developer Neilsoft Limited - Pune, Maharashtra March 2009 to July 2009 Project: Arc+ Software protection and Camera implementation Description: Arc+ Software protection: main aim of the project is to create the evaluation edition and standard edition for the Arc+ software using Aladdin software locking system. Camera Implementation: Main aim of the project is to implement the camera entity like AutoCAD camera feature in the arc+ software. Responsibilities Capturing requirements from client Understanding the requirements and doing functional design Implementation of Technical design for software protection Developed the application using VC++ Implemented Software Protection using Aladdin HASP software's Module test design for the project Unit testing and integration testing. Prepared user-friendly interfaces using MFC. Environment VC++, Aladdin HASP, Codejack xtreme toolkit, Visual Studio, VSS ObjectARX Developer Proximus Group - Brussels, BE December 2007 to February 2009 Project: Genesys-GDA Description: GDA provides to the ANS division several tools to create or modify the digital graphical

documentation of the local loop network. These software's give also several facilities to manage, control and export data from this documentation. The graphical documentation can be defined as: - A description of all the underground cables splices optical fibers and tubes on 'road map' and cable schemes. - An input for exploitation (follow-up of the logical connections) and planning (obvious way to overview the network). - Complementary to the technical documentation (ABR, ITR). - An input for SNS. (The graphical documentation is stored in the Scan & Search (SNS) document management system).

Responsibilities Understand the requirements and preparing technical design for use cases Implemented use cases in VC++ and Object ARX Prepared User-friendly interfaces using MFC Implementation of various PL/SQL procedures Used Visual Source safe for doing proper Versioning Control Prepared Module test design for use cases Setting up of development(DEV), Integration(ITT), acceptance(UAT) environments for deployment and testing Unit testing and integration testing Prepared documentation for use cases Environment VC++

Developer Neilsoft Limited - Pune, Maharashtra June 2007 to November 2007 **Project:** GIS-CAD System Integration **Description:** Objective of this software application is to exchange the data between Intergraph's GIS system and AutoCAD 2006 in order to have better control on the data flow and flexibility to the designers for editing the drawings. XML will be used as an intermediate medium for data exchange between GIS System and this application. This application has 4 modules. XML2DWG module converts XML document generated by Intergraph GIS system into AutoCAD Drawing. DWG2XML module converts AutoCAD drawing into XML document which is input to Intergraph GIS System. Asset numbering tool assign asset numbers to new features. Design Validation tool validates the AutoCAD drawing before converting into XML.

Responsibilities Understanding the Requirements given by Customer Preparing Technical design for the assigned Use cases Programs are developed Using MSADO15.dll for retrieving data from MSAccess Database. Programs are developed using MSSOAP1.dll for SOAP programming. Used Visual Source safe for doing proper Versioning Control. Unit testing and integration testing.

Environment VC++, ObjectARX, RealDWG, MFC, SOAP, MS Access, Visual Studio, AutoCAD 2006, VSS Java Developer NAPA Oy - Helsinki, FI November 2005 to May 2007 **Project:** Joint

Bulker Project Description: The Joint Bulker project (JBP)Rules, developed by the International Association of Classification Societies (IACS) for both single and double side bulk carriers are expected to come into force for ships contracted for construction on or after 1st April 2006. The purpose of the project is to develop and implement computer software components to support the Joint Bulker Project ? IACS Common Structural Rules for Bulk Carriers checking process. This is to be carried out by modeling the most common structural elements necessary for checking the JBP Rules

Responsibilities Understanding of Specifications and Design Rule Suites development in Java and XML Implementation of test cases using JUnit Framework Used BIRT for implementing Reporting tool.

Environment Java, Junit, BIRT, Eclipse, Star Team SQL, Enterprise Architect, VSS, AutoCAD Map 3D ObjectARX 2006 to 2006 Python & VC++ Developer Neilsoft Limited - Pune, Maharashtra March 2005 to November 2005

Project: CAD Customization Description: The project was for developing the interface software for exporting the 2D sheets created in PDMS to AUTOCAD. After exporting the sheet to ACAD, user was able to add different types of customized annotations to the elements in the sheet. The synchronization was maintained between the 2 CAD systems with the help of COM/SINK function implementation in ACAD and PDMS.

Responsibilities Understanding of Specification and Design Understanding of code Implementation of use cases in VC++, MFC, Python & wxPython Testing the application Defect Fixing

Environment VC++, MFC, Object ARX, Visual Studio, VSS, AutoCAD 2005, AVEVA PDMS, Python, wxPython, Tribon M3 Education BTech in Computer Science and Information Technology Jawaharlal Nehru Technological University - Hyderabad, Andhra Pradesh 2000 to 2004

Additional Information

Technical Skill Set: Big Data Ecosystem Hadoop, MapReduce, HDFS, Yarn, Hbase, Hive, Pig, Flume, Kafka, Oozie, Hue, Spark, Ambari, Zookeeper, Hortonworks, MapR

Programming Languages&IDEs C++, VC++, Java, Python, Scala, Eclipse, Visual Studio

Web Technologies HTML, XML, CSS, JavaScript, JQuery, AngularJS

Methodologies Agile, UML, Design Patterns

APIs ObjectARX (AutoCAD), Real DWG, Google Maps API, MFC

Operating System Windows, Linux

Database Oracle, SQL, PL/SQL

Software Applications AutoCAD, AVEVA NET Workhub & Dashboard, GeoServer, TribonM3

Versioning Control VSS, StarTeam, TFS

Cloud Infrastructure

Amazon Web Services (AWS), EMR, S3, Dynamo DB etc Others Enterprise Architect, Visio & MS
Project

Name: Nicole Bean

Email: joneal@example.org

Phone: (451)438-3290