

Bigdata Developer Bigdata Developer Bigdata Developer - Brilliant Edison, NJ Over 5+ Years of professional IT experience in analysis, architectural design, prototyping, development, Integration and testing of applications using Java/J2EE Technologies and Good Working experience in Big Data Technologies. Experience in developing Map Reduce Programs using Apache Hadoop for analyzing the big data as per the requirement. Experienced on major Hadoop ecosystem's projects such as Pig, Hive, HBase and monitoring them with Cloudera Manager. Extensive experience in developing Pig Latin Scripts and using Hive Query Language for data analytics. Hands on experience working on NoSQL databases including HBase, Cassandra and its integration with Hadoop cluster. Experience in implementing Spark, Scala application using higher order functions for both batch and interactive analysis requirement. Good working experience using Sqoop to import data into HDFS from RDBMS and vice-versa. Good knowledge in using job scheduling and monitoring tools like Oozie and Zookeeper. Experience in Hadoop administration activities such as installation and configuration of clusters using Apache, Cloudera and AWS. Experienced in designing, built, and deploying a multitude application utilizing almost all the AWS stack (Including EC2, R53, S3, RDS, DynamoDB, SQS, IAM, and EMR), focussing on high-availability, fault tolerance, and auto-scaling. Developed UML Diagrams for Object Oriented Design: Use Cases, Sequence Diagrams and Class Diagrams using Rational Rose, Visual Paradigm and Visio. Hands on experience in solving software design issues by applying design patterns including Singleton Pattern, Business Delegator Pattern, Controller Pattern, MVC Pattern, Factory Pattern, Abstract Factory Pattern, DAO Pattern and Template Pattern. Developed Web-based applications using Python, Amazon Web Services, jQuery, CSS and Model View control frameworks like Django, Flask and JavaScript. Good experience with design, coding, debug operations, reporting and data analysis utilizing python and using python libraries to speed up development. Hands on experience with Bid Data environment on technologies including Hadoop.

Experienced in creative and effective front-end development using JSP, JavaScript, HTML 5, DHTML, XHTML Ajax and CSS. Good Working experience in using different Spring modules like Spring Core Container Module, Spring Application Context Module, Spring MVC Framework

module, Spring ORM Module in Web applications. Used jQuery to select HTML elements, to manipulate HTML elements and to implement AJAX in Web applications. Used available plug-ins for extension of jQuery functionality. Working knowledge of database such as Oracle10g/11g/12c, Microsoft SQL Server, DB2. Experience in writing numerous test cases using JUnit framework with Selenium. Strong experience in database design, writing complex SQL Queries and Stored Procedures. Experienced in using Version Control Tools like Subversion, Git. Experience in Building, Deploying and Integrating with Ant, Maven. Experience in development of logging standards and mechanism based on Log4J. Strong problem-solving skills, good communication, interpersonal skills and a good team player. Have the motivation to take independent responsibility as well as ability to contribute and be a productive team member.

Work Experience Bigdata Developer Brilliant - Edison, NJ May 2018 to Present Description: Brilliant - Understand concepts and build your problem solving skills with thousands of free problems and examples in math, science, and engineering. Responsibilities: Solid understanding and experience in applying and implementing machine learning algorithms and concepts such as: Classification and Regression, Resampling statistics and bootstrapping using R language Experience in working with Hadoop 2.x version and Spark 2.x (Python and Scala). Extending HIVE and PIG core functionality by using custom User Defined Function's (UDF), User Defined Table-Generating Functions (UDTF) and User Defined Aggregating Functions (UDAF) in Java. Experience in Hadoop Production support tasks by analysing the Application and cluster logs. Implemented Partitioning, Dynamic Partitions, and Buckets in Hive on Avro files to meet the business requirements. Expertise in designing and deployment of Hadoop cluster and different Big Data analytic tools including Pig, Hive, HBase, Oozie, Zookeeper, SQOOP, flume, Spark, Kafka, HBase with MapR Distribution. Assist in upgrading, configuration and maintenance of various Hadoop infrastructures like Pig, Hive, and HBase. Used Spark for interactive queries, processing of streaming data and integration with NoSQL database for huge volume of data. Developed Scala scripts using both Data frames/SQL and RDD/MapReduce in Spark 1.x/2.x for Data Aggregation, queries and writing data back into OLTP system through Sqoop. Used Spark-Streaming APIs to perform necessary transformations

and actions on the fly for building the common learner data model which gets the data from Kafka in near real time and Persists into HBase. Performed advanced procedures like text analytics and processing, using the in-memory computing capabilities of Spark using Scala. Experienced in handling large datasets using Partitions, Spark in Memory capabilities, Broadcasts in Spark, Effective & efficient Joins, Transformations and other during ingestion process itself. Worked on migrating Map Reduce programs into Spark transformations using Spark and Scala. Coordinate with Administration team to enhance Spark Jobs performance by analyzing them. Implemented design patterns in Scala for the Spark application. Developed quality code adhering to Scala coding Standards and best practices. Used Spark API over MapR Hadoop YARN to perform analytics on data in Hive. Explored with the Spark improving the performance and optimization of the existing algorithms in Hadoop using Spark Context, Spark-SQL, Data Frame, Pair RDD's, Spark YARN. Worked on Loading log data into HDFS using Flume, Kafka and performing ETL integrations. Used Reporting Tool Tableau to connect with Hive for generating daily reports of data.

Collaborated with the infrastructure, network, database, application and BA teams to ensure data quality. Worked with different File Formats like TEXTFILE, SEQUENCE FILE, AVROFILE, ORC, and PARQUET for Hive querying and processing. Developed Spark code using Scala and Spark-SQL/Streaming for faster testing and processing of data. Import the data from different sources like HDFS/HBase into Spark RDD. Load the data into Spark RDD and do in memory data Computation to generate the Output response. Experience in Oozie and workflow scheduler to manage Hadoop jobs by Direct Acyclic Graph (DAG) of actions with control flows. Configure Oozie workflow to run multiple Hive and Pig jobs which run independently with time and data availability. Collected and aggregated large amounts of log data using Apache Flume and staging data in HDFS for further analysis. Developed a data pipeline using Kafka and Storm to store data into HDFS. Developed business specific Custom UDF's in Hive, Pig. Responsible for developing Python wrapper scripts which will extract specific date range using Sqoop by passing custom properties required for the workflow. Skilled in using collections in Python for manipulating and looping through different user defined objects. Worked with different kind of compression techniques like

LZO, GZip, Snappy. Worked on various configurations of Oozie bundles for Orchestrating Pig, Hive, Spark, Sqoop Utilized for Apache Nifi data migration from various sources to HDFS destination Worked with the Apache Nifi flow to perform the conversion of Raw XML data into JSON, AVRO Ingested streaming data Apache Nifi with into Kafka Environment: RedHat Linux, MapR, Scala, Python, Rlanguage, Python HDFS Hive, Pig, Sqoop, Flume, Oozie, HBase, Spark Core, Spark SQL, Spark streaming, Kafka, Tableau BigData Developer Eli Lilly&CO - Indianapolis, IN March 2017 to April 2018 Description: Eli Lilly and Company is a global pharmaceutical company headquartered in Indianapolis, Indiana, with offices in 18 countries. Its products are sold in approximately 125 countries. The company was founded in 1876 by, and named after, Col. Eli Lilly, a pharmaceutical chemist and veteran of the American Civil War. Responsibilities: Interact with Solution Architects and Business Analysts to gather requirements and update Solution Architect Document. Handled importing of data from various data sources, performed transformations using Hive, MapReduce, Spark and loaded data into HDFS. Submitted Talend jobs for scheduling using Talend scheduler. Extensively Worked in Agile software development approach using JIRA Leveraged Talend to ingest data into the datalake of datafabric. Unit tested Talend workflows for the correctness of the data. Developed jobs to expose HDFS files to Hive tables and Views depending up on the schema versions. Created Hive tables, partitions and implemented incremental imports to perform ad-hoc queries on structured data. Developed jobs to move inbound files to HDFS file location based on monthly, weekly, daily and hourly partitioning. Optimizing Hive queries, improve performance by configuring Hive Query parameters. Implemented ORC data format for Apache Hive computations to handle the custom business requirements. Imported data from RDBMS (MySQL, Oracle) to HDFS and vice versa using Sqoop (Big Data ETL tool) for Business Intelligence, visualization and report generation. Mapping source to target data and converted data JSON to XML (229 Accord Format) using Talend data mapper. Developed jobs to expose HDFS files to Hive tables and Views depending up on the schema versions. Created Hive tables, partitions and implemented incremental imports to perform ad-hoc queries on structured data. Developed jobs to move inbound files to HDFS file location based on monthly, weekly, daily and

hourly partitioning. Created Hive External tables and loaded the data in to tables and query data using HQL. Implemented ORC data format for Apache Hive computations to handle the custom business requirements. Implemented a centralized Data Lake in Hadoop with data from various sources. Analyzing and understanding the legacy code and making recommendations on how the new system can be designed. Developed Scala scripts using both Data frames/SQL and RDD/MapReduce in Spark 1.x/2.x for Data Aggregation, queries and writing data back into OLTP system through Sqoop. Used Spark-Streaming APIs to perform necessary transformations and actions on the fly for building the common learner data model which gets the data from Kafka in near real time and Persists into HBase. Performed advanced procedures like text analytics and processing, using the in-memory computing capabilities of Spark using Scala. Stored MapReduce program output in Amazon S3 and developed a script to move the data to RedShift for generating a dashboard using QlikView. Worked on Talend Administrator Console (TAC) for scheduling jobs and adding users

Environment: Hive, MapReduce, Spark, RDBMS, TAC, Data Lake, Amazon S3, RedShift, JSON, XML.

Hadoop Developer Comcast - Herndon, VA May 2016 to February 2017

Description: Comcast Corporation is a global media and technology company with two primary businesses, Comcast Cable and NBCUniversal. Comcast Cable is one of the nation's largest video, high-speed internet, and phone providers to residential customers under the XFINITY brand, and provides these services to businesses.

Responsibilities:

- Provided application demo to the client by designing and developing search engine, report analysis trends, application administration prototype screens using AngularJS, and BootstrapJS.
- Took the ownership of complete application Design of Java part, Hadoop integration
- Apart from the normal requirement gathering, participated in Business meeting with the client to gather security requirements.
- Assisted with the architect to analyze the existing system and future system Prepared design blue prints and application flow documentation
- Experienced in managing and reviewing Hadoop log files Load and transform large sets of structured, semi structured and unstructured data
- Responsible to manage data coming from different sources and application Supported Map Reduce Programs those are running on the cluster
- Responsible in working with Message broker system such as Kafka Extracted data from

mainframes and feed to KAFKA and ingested to HBASE to perform Analytics. Written event driven, link tracking system to capture user events and feed to KAFKA to push it to HBASE. Created MapReduce jobs to extract the contents from HBASE and configured in OOOIE workflow to generate analytical reports. Developed the JAX- RS web services code using apache CXF framework to fetch data from SOLR when user performed the search for documents. Participated in SOLR schema, and ingested data into SOLR for data indexing. Written MapReduce programs to organize the data, and ingest the data to suitable for analytics in client specified format. Hands on experience in writing python scripts to optimize the performance. Implemented Storm builder topologies to perform cleansing operations before moving data into Cassandra. Extracted files from Cassandra through Sqoop and placed in HDFS and processed. Implemented Bloom filters in Cassandra using key space creation. Involved in writing Cassandra CQL statements. Good hands on experience in developing concurrency using spark and Cassandra together. Involved in writing spark applications using Scala. Hands on experience in creating RDDs, transformations and Actions while implementing spark applications. Good knowledge in creating data frames using Spark SQL. Involved in loading data into Cassandra NoSQL Database. Implemented record level atomicity on writes using Cassandra. Written PIG Scripts to query and process the Data sets to figure out the patterns of trends by applying client specific criteria, and configured OOOIE workflows to run the jobs along with the MR jobs. Stored the derived results in HBASE from analysis and make it available to data ingestion for SOLR for indexing data. Involved in integration of java search UI, SOLR and HDFS. Involved in code deployments using continuous integration tool using Jenkins. Documented all the challenges, issues involved to deal with the security system and Implemented best practices. Created Project structures and configurations according to the project architecture and made it available to the junior developer to continue their work. Handled onsite coordinator role to deliver work to offshore. Involved in core reviews and application lead supported activities.

Environment: Java, J2EE, Python, Cassandra, Spring 3.2, MVC, HTML5, CSS, AngularJS, Restful services using CXF web services framework, spring data, SOLR 5.2.1, PIG, HIVE, apache AVRO, Map Reduce, Sqoop Zookeeper, SVN, Jenkins, windows AD, windows KDC, Hortonworks

distribution of Hadoop 2.3, YARN, Ambari Java Developer Magna Quest Technologies Pvt. Ltd-Hyd
March 2015 to January 2016 Description: Magna Quest Technologies limited is a trend-setting and fascinating innovative Enterprise Product-based Solutions Company, which has established its leadership over 15 years, in three broad lines of businesses. Responsibilities: Involved in Analysis, design and coding on Java and J2EE Environment. Implemented struts MVC framework.

Designed, developed and implemented the business logic required for Security presentation controller. Set up the deployment environment on Web Logic Developed system preferences UI screens using JSP and HTML. Developed UI screens using Swing components like JLabel, JTable, JScrollPane, JButtons, JTextFields, etc. Used JDBC to connect to Oracle database and get the results that are required. Designed asynchronous messaging using Java Message Service (JMS). Consumed web services through SOAP protocol. Developed web Components using JSP, Servlets and Server-side components using EJB under J2EE Environment. Designing JSP using Java Beans. Implemented Struts framework 2.0 (Action and Controller classes) for dispatching request to appropriate class Design and implementation of front-end web pages using CSS, DHTML, JavaScript, JSP, HTML, XHTML, JSTL, Ajax and Struts Tag Library. Designed table structure and coded scripts to create tables, indexes, views, sequence, synonyms and database triggers. Involved in writing Database procedures, Triggers, PL/SQL statements for data retrieving. Developed using Web 2.0 to interact with other users and changing the contents of websites. Implemented AOP and IOC concept using UI Spring 2.0 Framework. Developed using Transaction Management API of Spring 2.0 and coordinates transactions for Java objects Generated WSDL files using AXIS2 tool. Developed using CVS as a version controlling tool for managing the module developments. Configured and Tested Application on the IBM Web Sphere App. Server Used Hibernate ORM tools which automate the mapping between SQL databases and objects in Java. Developed using XML XPD, BPEL and XML parsers like DOM, SAX. Developed using XSLT to convert XML documents into XHTML and PDF documents. Written JUnit test cases for Business Objects, and prepared code documentation for future reference and upgrades. Deployed applications using WebSphere Application Server and Used IDE RAD

(Rational Application Developer). Environment: Java, J2EE, JSP, Servlets, MVC, Hibernate, Spring 3.0, Web Services, Maven 3.2.x, Eclipse, SOAP, WSDL, Eclipse, jQuery, Java Script, Swings, Oracle, REST API, PL/SQL, Oracle 11g, UNIX. Java Developer Impact Software Solutions-HYD January 2014 to February 2015 Description: IMPACT IT SOLUTIONS is an Expert IT Solutions Company providing all kind of Web Solution like Custom Web development, E-commerce Application Development, Web Application development, Ecommerce Solution Development, Web-based application Development, web & web-based IT solutions to a wide variety of clients across India & rest of the Globe. Responsibilities: Designed & developed the application using Spring Framework Developed class diagrams, sequence and use case diagrams using UML Rational Rose. Designed the application with reusable J2EE design patterns Designed DAO objects for accessing RDBMS Developed web pages using JSP, HTML, DHTML and JSTL Designed and developed a web-based client using Servlets, JSP, Tag Libraries, JavaScript, HTML and XML using Struts Framework. Involved in developing JSP forms. Designed and developed web pages using HTML and JSP. Designed various applets using JBuilder. Designed and developed Servlets to communicate between presentation and business layer. Used EJB as a middleware in developing a three-tier distributed application. Developed Session Beans and Entity beans for business and data process. Used JMS in the project for sending and receiving the messages on the queue. Developed the Servlets for processing the data on the server. Developed views and controllers for client and manager modules using Spring MVC and Spring Core. Used Spring Security for securing the web tier Access. Business logic is implemented using Hibernate. Developed and modified database objects as per the requirements. Involved in Unit integration, bug fixing, acceptance testing with test cases, Code reviews. Interaction with customers and identified System Requirements and developed Software Requirement Specifications. Implemented Java design patterns wherever required. Involved in development, maintenance, implementation and support of the System. Involved in initial project setup and guidelines. Implemented Multi-threading concepts. Developed test cases for Unit testing using JUnit and performed integration and system testing Involved in coding for the presentation layer

using Struts Framework, JSP, AJAX, XML, XSLT and JavaScript Closely worked and supported the creation of database schema objects (tables, stored procedures, and triggers) using Oracle SQL/PLSQL Environment: Java / J2EE, JSP, CSS, JavaScript, AJAX, Hibernate, Spring, XML, EJB, Web Services, SOAP, Eclipse, Rational Rose, HTML, XPATH, XSLT, DOM and JDBC. IntelliJ, Eclipse, NetBeans 2010 to 2012 Development Methodologies Agile/Scrum, UML, Design Patterns, Waterfall. Build Tools Jenkins, Toad, SQL Loader, Maven, ANT, RTC, RSA, Control-M, Oziee, Hue, SOAP UI Reporting Tools MS Office (Word/Excel/Power Point/ Visio/Outlook), Crystal reports XI, SSRS, cognos 7.0/6.0. Databases Microsoft SQL Server 2008,2010/2012, MySQL 4.x/5.x, Oracle 11g, 12c, DB2, Teradata, Netezza Operating Systems All versions of Windows, UNIX, LINUX, Macintosh HD, Sun Solaris Skills Sql, Cassandra, Hdfs, Impala, Mapreduce, Oozie, Sqoop, Hbase, Kafka, Flume, Hadoop, MongoDB, Splunk, Tableau server, C++, Dtd, Fortran, Hadoop, Hbase, Hive Additional Information TECHNICAL SKILLS: Bigdata/Hadoop Technologies Hadoop, HDFS, YARN, MapReduce, Hive, Pig, Impala, Sqoop, Flume, Spark, Kafka, Storm, Drill, Zookeeper and Oozie Languages HTML5, DHTML, WSDL, CSS3, C, C++, XML, R/R Studio, SAS Enterprise Guide, SAS ,R (Caret, Weka, ggplot) , Perl, MATLAB, Mathematica, FORTRAN, DTD, Schemas, JSON, Ajax, Java, Scala, Python (NumPy, SciPy, Pandas), Java Script, Shell Scripting NO SQL Databases Cassandra, HBase, MongoDB, MariaDB Business Intelligence Tools Tableau server, Tableau Reader, Tableau, Splunk, SAP Business Objects, OBIEE, SAP Business Intelligence, QlikView, Amazon Redshift, or Azure Data Warehouse

Name: Kelly Schneider

Email: terri80@example.com

Phone: 2023058997